Kapitel 11

Zentrale Grenzwertsätze

Viele zufällige Größen in der Natur, Wirtschaft und Gesellschaft sind das Ergebnis einer Überlagerung zahlreicher kleiner zufälliger Einflüsse, die weitgehend unabhängig voneinander wirken. So ist der tägliche Schlusskurs einer Aktie das Ergebnis einer i. Allg. großen Zahl von Käufen und Verkäufen, Messergebnisse werden häufig durch zahlreiche Einwirkungen zufälliger Art beeinflusst (Temperatur, Ablesefehler u. a.). Die Wahrscheinlichkeitstheorie widmet sich diesen Fragen, indem sie die Wahrscheinlichkeitsverteilungen der Summe einer großen Anzahl n von einzelnen Zufallsgrößen studiert.
Wie oft in der Mathematik üblich, geht man dabei zum Grenzwert für $n \to \infty$ über, um übersichtliche Resultate zu erzielen. Eine Gruppe entsprechender Sätze, die sogenannten *zentralen Grenzwertsätze*, befasst sich mit Bedingungen an die zugrunde liegenden Zufallsgrößen, unter denen eine *Normalverteilung* im Limes erscheint.

11.1 Lokaler Grenzwertsatz von Moivre-Laplace

Es sei $(X_n, n \geq 1)$ ein Bernoullischema mit dem Parameter $p \in (0,1)$. Dann besitzt $S_n = \sum_{k=1}^{n} X_k$ bekanntlich (siehe Aussage 6.3) eine Binomialverteilung mit den Parametern n und p:

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k} =: b(n,p;k), \quad k = 0,1,\dots,n. \tag{11.1}$$

Es gilt (vgl. Beispiele 4.13 c) und 4.21 c))

$$ES_n = np, \text{ und } D^2 S_n = npq \text{ mit } q = 1 - p.$$

Wir untersuchen, wie sich die Verteilung von S_n bei unbegrenzt wachsendem n verändert. Offenbar wachsen ES_n und $D^2 S_n$ unbeschränkt falls n nach unendlich strebt, und $b(n, p; k)$ konvergiert für $n \to \infty$ bei festen p und k gegen Null. (Beachten Sie $b(n, p; k) \leqq \left(\frac{p}{1-p}\right)^k \cdot \frac{1}{k!} \cdot n^k (1-p)^n$.)

Um dennoch etwas über die asymptotischen Eigenschaften der Binomialverteilung für $n \to \infty$ aussagen zu können, gehen wir zur standardisierten Zufallsgröße S_n^* über:

$$S_n^* = \frac{S_n - ES_n}{\sqrt{D^2 S_n}} = \frac{S_n - np}{\sqrt{npq}}.$$

Diese Zufallsgröße hat die möglichen Werte

$$x_k^{(n)} := \frac{k - np}{\sqrt{npq}},$$

die sie jeweils mit der Wahrscheinlichkeit $b(n, p; k)$ annimmt, $k = 0, 1, \ldots, n$. Die $x_k^{(n)} (k = 0, 1, \ldots, n)$ bilden ein Gitter mit dem Gitterabstand $\triangle_n := (npq)^{-\frac{1}{2}}$, dem kleinsten Gitterpunkt $x_0^{(n)} = -\sqrt{\frac{np}{q}}$ und dem größten $x_n^{(n)} = \sqrt{\frac{nq}{p}}$. Wir führen eine Funktion $\varphi_n(\cdot)$ auf folgende Weise ein:

$$\varphi_n(x) = \frac{b(n, p; k)}{\triangle_n} \quad \text{falls} \quad x \in \left[x_k^{(n)} - \frac{\triangle_n}{2}, x_k^{(n)} + \frac{\triangle_n}{2}\right)$$

$$(k = 0, 1, \ldots, n).$$

$$\varphi_n(x) = 0, \quad \text{falls} \quad x < x_0^{(n)} \quad \text{oder falls} \quad x \geq x_n^{(n)}.$$

φ_n beschreibt ein Säulendiagramm mit $(n + 1)$ senkrechten Säulen der Höhe $\varphi_n(x_k^{(n)})$, der Breite $\triangle_n$ und den Säulenmitten $x_k^{(n)}, k = 0, 1, \ldots, n$.

Die Fläche der k-ten Säule beträgt $b(n, p; k)$ und die Gesamtfläche unter der Oberkante des Säulendiagramms ist gleich Eins.

Satz 11.1 (Lokaler zentraler Grenzwertsatz von Moivre-Laplace) *Für alle $a > 0$ gilt*

$$\lim_{n\to\infty} \sup_{|x|\le a} |\varphi_n(x) - \varphi(x)| = 0,$$

wobei $\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$

die Dichte der Standard-Normalverteilung auf R_1 *ist:*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{s^2}{2}} \mathrm{ds}, \quad \mathrm{x} \in \mathrm{R}_1.$$

Beweis: siehe z. B. Siraev, I, § 6.
Der Beweis stützt sich im Wesentlichen auf die Stirling'sche Formel der Approximation von Fakultäten $n!$.

Der lokale Grenzwertsatz von Moivre-Laplace wird häufig benutzt, um Wahrscheinlichkeiten der Form $P(k \le S_n \le l)$ näherungsweise zu bestimmen. Es gilt nämlich wegen $S_n^* = \frac{S_n - np}{\sqrt{npq}}$ die Beziehung

$$P(k \le S_n \le l) = \sum_{m=k}^{l} b(n,p;m) =$$

$$P\Big(\frac{k-np}{\sqrt{npq}} \le S_n^* \le \frac{l-np}{\sqrt{npq}}\Big) =$$

$$P\Big(x_k^{(n)} \le S_n^* \le x_l^{(n)}\Big) =$$

$$\int_{x_k^{(n)}-\frac{\triangle_n}{2}}^{x_k^{(n)}+\frac{\triangle_n}{2}} \varphi_n(s)ds \sim \int_{x_k^{(n)}-\frac{\triangle_n}{2}}^{x_k^{(n)}+\frac{\triangle_n}{2}} \varphi(s)ds = \Phi\Big(x_l^{(n)}+\frac{\triangle_n}{2}\Big)-\Phi\big((x_k^{(n)}-\frac{\triangle_n}{2}\big) \qquad (11.2a)$$

Analog erhält man

$$P(k \le S_n < l) \approx \Phi\Big(x_l^{(n)} - \frac{\triangle_n}{2}\Big) - \Phi\big(x_k^{(n)} - \frac{\triangle_n}{2}\big) \qquad (11.2b)$$

$$P(k < S_n \leq l) \approx \Phi\Big(x_l^{(n)} + \frac{\triangle_n}{2}\Big) - \Phi\Big(x_k^{(m)} + \frac{\triangle_n}{2}\Big) \tag{11.2c}$$

$$P(k < S_n < l) \approx \Phi\Big(x_l^{(n)} + \frac{\triangle_n}{2}\Big) - \Phi\Big(x_k^{(n)} - \frac{\triangle_n}{2}\Big). \tag{11.2d}$$

Häufig trifft man auf die folgenden etwas ungenaueren Approximationen, die wir als "grobe Approximation" bezeichnen wollen, im Gegensatz zu der vorhergehenden "feinen Approximation".

$$P(k \leq (<) S_n \leq (<) l) \approx \Phi(x_l^{(n)}) - \Phi(x_k^{(n)}),$$

wobei auf der linken Seite jeweils entweder $\leq$ oder $<$ steht. Sie liefert für größeres n ebenfalls brauchbare Werte.

Beispiel 11.2 16-maliges Werfen einer regulären Münze. Wie groß ist die Wahrscheinlichkeit, dass mindestens sechs und höchstens zehnmal die Zahl oben liegt?

$$n = 16, p = \frac{1}{2}, k = 6, l = 10,\ np = 8, npq = 4$$

1. Exaktes Resultat:

$$P(6 \leq S_{16} \leq 10) = \Big[2\binom{16}{6} + 2\binom{16}{7} + \binom{16}{8}\Big]2^{-16} =$$

$$0,244 + 0,349 + 0,196 = 0,789$$

2. "Grobe" Approximation:

$$P(6 \leq S_{16} \leq 10) = P(-1 \leq S_{16}^* \leq 1) \approx$$

$$\Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0,6826$$

3. "Feine" Approximation:

$$P(6 \le S_{16} \le 10) = P(-1 - \frac{1}{4} \le S^*_{16} \le 1 + \frac{1}{4}\Big) \approx$$

$$\Phi(1,25) - \Phi(-1,25) = 2\Phi(1,25) - 1 = 2 \cdot 0,8944 - 1 = 0,788.$$

Die Approximation ist nicht so gut für $p \neq \frac{1}{2}$. Man berechne sie für $n = 16, p = 0,2$.

11.2 Der zentrale Grenzwertsatz von Feller-Lévy

Voraussetzung 11.3 Es seien $(X_n, n \ge 1)$ eine Folge unabhängiger, identisch verteilter Zufallsgrößen mit $\sigma^2 := D^2X_1 \in (0,\infty)$ und $S_n := \sum_{k=1}^{n} X_k$.

Insbesondere gilt

$$ES_n = nEX_1, D^2S_n = n\sigma^2. \tag{11.3}$$

Das Gesetz der großen Zahlen besagt, dass die arithmetischen Mittel $M_n = \frac{1}{n}S_n$ P-fast sicher gegen EX_1 konvergieren. Insbesondere streben im vorliegenden Fall auch die Streuungen $D^2M_n = \frac{\sigma^2}{n}$ gegen Null. Daraus folgt, dass die Verteilungen von M_n gegen die ausgeartete Verteilung, die in EX_1 konzentriert ist, konvergieren. Wenn man dagegen S_n zentriert und normiert zu

$$S_n^* = \frac{S_n - ES_n}{\sqrt{D^2S_n}} \quad \text{(Standardisierung)}$$

so hat S_n^* den Erwartungswert Null und die Streuung Eins, und zwar für jedes $n \ge 1$.

Der folgende Grenzwert stellt fest, dass die Verteilungen der S_n^* gegen die Standard-Normalverteilung konvergieren.

Satz 11.4 (Zentraler Grenzwertsatz von Feller-Lévy) *Für die standardisierten Zufallsgrößen* $S_n^*, n \geq 1$ *gilt*

$$\lim_{n\to\infty} \sup_{-\infty\leq a<b\leq\infty} |P(a < S_n^* \leq b) - (\Phi(b) - \Phi(a))| = 0, (2) \qquad (11.4)$$

wobei Φ *die Verteilungsfunktion der Standard-Normalverteilung ist.*

Der Beweis erfolgt mittels des Faltungs- und des Stetigkeitssatzes für charakteristische Funktionen, vgl. auch Übung 12.6.

Die Beweisidee lässt sich folgendermaßen skizzieren:

$$\varphi_{S_n^*}(u) = E\mathrm{e}^{\mathrm{iuS_n^*}} = \mathrm{Ee}^{\mathrm{iu}(\frac{\mathrm{X_1+\ldots+X_n-n}\mu}{\sigma\sqrt{\mathrm{n}}})} =$$

$$E\exp\Big[\frac{iu}{\sigma\sqrt{n}}((X_1-\mu)+\ldots+(X_n-\mu))\Big] =$$

$$\Big[E\exp[\frac{iu}{\sigma\sqrt{n}}(X_1-\mu)]\Big]^n =$$

$$\Big[\varphi_{X_1}\Big(\frac{u}{\sigma\sqrt{n}}\Big)\mathrm{e}^{-\frac{\mathrm{iu}\mu}{\sigma\sqrt{\mathrm{n}}}}\Big]^{\mathrm{n}}, \quad \mathrm{u}\in\mathrm{R}_1.$$

Da nach Voraussetzung $EX_1^2 < \infty$ gilt, ist die charakteristische Funktion φ_{X_1} zweimal stetig differenzierbar, und es gilt

$$\varphi_{X_1}(v) = 1 + i\mu v - \frac{v^2}{2}EX_1^2 + o(v^2)$$

und

$$\mathrm{e}^{iw} = 1 + iw - \frac{w^2}{2} + o(w^2)$$

und man erhält mit $v = \frac{u}{\sigma\sqrt{n}}$ bzw. $w = -\frac{u\mu}{\sigma\sqrt{n}}$ die Beziehung

$$\varphi_{S_n^*}(u) = \Big[1 - \frac{u^2}{2n} + o\Big(\frac{1}{n}\Big)\Big]^n \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-\frac{u^2}{2}},\ u\in R_1.$$

Der Grenzwert ist aber gerade die charakteristische Funktion der Standard-Normalverteilung. Nun ergibt sich die Aussage des Satzes aus dem Stetigkeitssatz für charakteristische Funktionen und der Tatsache, dass die w-Konvergenz von Verteilungsfunktionen F_n gegen eine Verteilungsfunktion F (siehe Bemerkunge nach Stetigkeitssatz 9.6) im Falle, dass F stetig ist, gleichmäßig erfolgt (Übung 11.1).

Der eben angegebene Zentrale Grenzwertsatz ist ein geeignetes Hilfsmittel, um mit guter Näherung Wahrscheinlichkeiten bestimmen zu können, die im Zusammenhang mit arithmetischen Mitteln unabhängiger, identisch verteilter Zufallsgrößen stehen.

Wir werden dafür einige Beispiele angeben. Sie stützen sich alle auf folgende Näherungsgleichung:

Auf Grund des Zentralen Grenzwertsatzes gilt

$$F_{S_n^*}(x) := P(S_n^* \leq x) \approx \Phi(x)\ , \quad x \in R_1 \tag{11.5}$$

Wir werden im Folgenden diese Näherung verwenden, die in Anwendungsfällen meist für nicht allzu große n genügend genau erfüllt ist. (Zur genauen Konvergenzgeschwindigkeit siehe die Ungleichung 11.4.)

Insbesondere folgen die häufig nützlichen Formeln:

$$P(S_n \leq x\sigma\sqrt{n} + n\mu) \approx \Phi(x),\ x \in R_1, \tag{11.6}$$

$$P(S_n \leq y) \approx \Phi\Big(\frac{y - n\mu}{\sigma\sqrt{n}}\Big) \quad , \quad y \in R_1, \tag{11.7}$$

$$P\Big(|\frac{S_n}{n} - \mu| > c\Big) \approx 2\Big(1 - \Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big)\Big),\ c > 0, \tag{11.8}$$

$$P\big(|\frac{S_n}{n} - \mu| \leq c\big) \approx 2\Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big) - 1,\ c > 0. \tag{11.9}$$

Die Werte der Standard-Normalverteilungsfunktion Φ entnimmt man einer entsprechenden Tabelle. (Erinnert sei an die Voraussetzung (11.3).

Im Folgenden geben wir einige Anwendungen dieser Formeln an. Dabei setzen wir voraus, dass $(X_n, n \geq 1)$ den Voraussetzungen 11.3 des Satzes von Feller-Lévy genügt und definieren wie gehabt $S_n = \sum\limits_{k=1}^{n} X_k$.

a) Mit welcher Wahrscheinlichkeit weicht das arithmetische Mittel $\frac{S_n}{n}$ um mehr als c vom Erwartungswert μ ab?

   Antwort: Wegen (11.7) mit der Wahrscheinlichkeit

$$P\Big(|\frac{S_n}{n} - \mu| > c\Big) \approx 2(1 - \Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big)\Big).$$

   Mit welcher Wahrscheinlichkeit überdeckt das (zufällige) Intervall $\Big(\frac{S_n}{n} - c, \frac{S_n}{n} + c\Big)$ den Erwartungswert μ?

   Antwort: Mit der Wahrscheinlichkeit (siehe (11.8))

$$P\Big(\frac{S_n}{n} - c \leq \mu \leq \frac{S_n}{n} + c\Big) = P\Big(|\frac{S_n}{n} - \mu| \leq c\Big) =$$

$$1 - P\Big(|\frac{S_n}{n} - \mu| > c\Big) \approx 2\Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big) - 1$$

b) Es seien $\alpha \in (0,1)$ und n vorgegeben. Wie groß muss man c mindestens wählen, damit gilt

$$P\Big(|\frac{S_n}{n} - \mu| > c\Big) \leq \alpha?$$

   Antwort: Wegen (11.7) wählt man c mindestens so groß, dass $2\Big(1 - \Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big)\Big) \leq \alpha$ erfüllt ist. Das bedeutet

$$c \geqq q_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},$$

wobei q_p das p-Quantil der Standard-Normalverteilung bezeichnet. Siehe Definition 3.35 und Aussage 3.38 sowie die Normalverteilungstabelle.

Mit der Wahrscheinlichkeit $1-\alpha$ gilt dann für die Beziehung $c = q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$$n\mu - q_{1-\frac{\alpha}{2}} \cdot \sigma\sqrt{n} \leq S_n \leq n\mu + q_{1-\frac{\alpha}{2}} \cdot \sigma \cdot \sqrt{n}, \quad \text{d.h.}$$

$$\mu - q_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \frac{S_n}{n} \leq \mu + q_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

c) $\alpha \in (0,1)$ und $c > 0$ seien gegeben. Wie groß sollte man n mindestens wählen, damit gilt:

$$P\Big(|\frac{S_n}{n} - \mu| \leq c\Big) \geq 1 - \alpha$$

Antwort: $P\Big(|\frac{S_n}{n} - \mu| \leq c\Big) = 1 - P\Big(|\frac{S_n}{n} - \mu| > c\Big) \approx$

$$1 - 2\Big(1 - \Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big)\Big) \geq 1 - \alpha$$

Also sollte man n mindestens so groß wählen, dass

$$\Phi\Big(\frac{c\sqrt{n}}{\sigma}\Big) \leq 1 - \frac{\alpha}{2}$$

gilt, d. h.

$$n \geq \frac{\sigma^2}{c^2} \cdot q^2_{1-\frac{\alpha}{2}} \text{ also}$$

$$n \geq n_0 = \Big[\frac{\sigma^2}{c^2} \cdot q^2_{1-\frac{\alpha}{2}}\Big] + 1,$$

wobei $[z] = max\{k \geq 0 | k \in N, k \leq z\}, (z > 0)$ gesetzt wird.

Um die Konvergenzgeschwindigkeit im zentralen Grenzwertsatz von Lévy-Feller abschätzen zu können, ist folgende Ungleichung nützlich.

Ungleichung 11.5 von Berry-Essen
Unter den Voraussetzungen des Satzes von Feller-Lévy und der Annahme $E|X_1|^3 < \infty$ *gilt:*

$$\sup_x |F_{S_n^*}(x) - \Phi(x)| \leq C \cdot \frac{E|X_1 - \mu|^3}{\sigma^3 \sqrt{n}} \tag{11.10}$$

mit einer Konstanten C*, für die* $(2\pi)^{-\frac{1}{2}} < C < 0,8$ *gilt.*

Die Konvergenzordnung $n^{-\frac{1}{2}}$ kann im Allgemeinen nicht verbessert werden. (Siehe z. B. Siraev (1988), Kap. III, §4.)

Der Spezialfall der Binomialverteilung

Für den Fall, dass die $(X_n, n \geq 1)$ ein Bernoullischema mit dem Parameter p bilden, gilt natürlich der zentrale Grenzwertsatz von Feller-Lévy und wird aus historischen Gründen als globaler zentraler Grenzwertsatz von Moivre-Laplace bezeichnet.

Mit

$$S_n = \sum_{k=1}^{n} X_k, S_n^* = \frac{S_n - np}{\sqrt{npq}}, n \geq 1$$

gilt also in diesem Fall

$$\lim_{n\to\infty} \sup_{-\infty \leq a < b \leq \infty} |P(a < S_n^* \leq b) - (\Phi(b) - \Phi(a))| = 0.$$

Eine ähnliche Ungleichung wie die von Berry-Essen (11.9) lautet hier

$$\sup_{x \in R_1} |F_{S_n^*}(x) - \Phi(x)| \leq \frac{p^2 + q^2}{\sqrt{npq}} \tag{11.11}$$

(vgl. Siraev, Kap. I, §6).

Bemerkung 11.6 Als "Praxiserfahrung" gibt Henze (2006) in seinem Kap. 27 die Faustregel $npq \geq 9$ als "für praktische Zwecke ausreichend" an.

Ist $npg \geq 9$ nicht erfüllt, n aber nicht zu klein, so ist evtl. der *Poisson'sche Grenzwertsatz* anwendbar (vgl. Übung 6.6):

Ist $p_n \underset{n\to\infty}{\longrightarrow} 0$, mit $np_n \underset{n\to\infty}{\longrightarrow} \lambda > 0$ so gilt für jede $k \geq 0$

$$\lim_{\substack{n\to\infty \\ p_n\to 0 \\ np_n\to\lambda}} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} =: \pi_{\mathrm{k}}(\lambda).$$

Für $p \ll 1, n \gg 1$ und $\lambda := np < 9$ kann man mit der Näherung

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \pi_k(\lambda)$$

rechnen.

Zahlenbeispiele 11.7

a) In einem Computerprogramm wird nach jeder Operation auf die j-te Dezimale gerundet. Rundungsfehler addieren sich, sind unabhängig und gleichverteilt auf $\left[\frac{-10^{-j}}{2}; \frac{+10^{-j}}{2}\right]$, $n = 10^6$ Operationen werden ausgeführt. Wie groß ist die Wahrscheinlichkeit dafür, dass der absolute Rundungsfehler im Endresultat größer als $c = 500 \cdot 10^{-j}$ ist?

Antwort: Hier sind $X_1, \ldots, X_n$ unabhängig, identisch verteilt, $EX_i = 0, D^2X_i = \frac{10^{-2j}}{12}$.

Auf Grund des Zentralen Grenzwertsatzes von Lévy-Feller ist

$S_n^* = S_n \cdot \sqrt{12} \cdot 10^j$ annähernd Standard-normalverteilt.

Für die gesuchte Wahrscheinlichkeit erhalten wir

$$P(|S_n| > 500 \cdot 10^{-j}) = P(|S_n^*| > \frac{\sqrt{12} \cdot 500}{10^3}) =$$

$$P(|S_n^*| > \sqrt{3}) \approx 2(1 - \Phi(\sqrt{3})) =$$

$$2(1 - 0,9584) = 0,083.$$

Will man dagen eine Schranke, die mit Sicherheit gilt, rechnet man mit dem ungünstigsten Fall, dass alle Fehler gleiches Vorzeichen haben und sich summieren. Dann kann man nur sagen, dass mit Sicherheit

$$S_n \in \Big[-\frac{10^{6-j}}{2}, +\frac{10^{6-j}}{2}\Big]$$

gilt. Das sind Schranken, die weit größer als die vorher bestimmten sind.

b) Ein regulärer Spielwürfel wird 1000mal unabhängig voneinander geworfen. Der Erwartungswert der Augensumme beträgt 3500. In welchem (möglichst kleinem) Intervall $[3500 - c, 3500 + c]$ wird die Augensumme mit der Wahrscheinlichkeit $0,95(0,99$ bzw. Eins) liegen?

Antwort: Die Wahrscheinlichkeit

$$P(|S_{1000} - 3500| \le c) = P\big(|S_{1000}^*| \le \frac{c}{\sigma\sqrt{10^3}}\big) \approx 2\Phi\big(\frac{c}{\sigma\sqrt{10^3}}\big) - 1$$

soll gleich $0,95$ sein. (σ^2 = Streuung der Augenzahl eines Wurfes = $2,917$.)
Daraus folgt $2\Phi\big(\frac{c}{\sigma\sqrt{10^3}}\big) - 1 = 0,95$ also $\Phi\big(\frac{c}{\sigma\sqrt{10^3}}\big) = 0,975$ und

$$c = 92,23 \cdot q_{0,975} = 180,8$$

.

Für $0,99$ an Stelle $0,95$ ergibt sich $c = 237,5$

und für 1 an Stelle $0,95$ erhalten wir $c = 2500$.

c) Wie oft muss man einen Punkt rein zufällig aus dem Einheitsquadrat auswählen, um mit der in Abschnitt 10.4.1 beschriebenen Methode die Zahl $\frac{\pi}{4}$ mit einer approximativen Wahrscheinlichkeit von $0,95$ auf m Stellen genau zu bestimmen?

Antwort: Mit $\alpha = 0,05$ gilt

$$n_0 = \left[\frac{q^2_{1-\frac{\alpha}{2}} \cdot \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{10^{-2m}} \right]$$

$$n_0 = 0,65 \cdot 10^{2m}.$$

11.3 Der zentrale Grenzwertsatz von Lindeberg-Feller

Es seien im Weiteren $(X_n, n \geq 1)$ eine Folge unabhängiger, aber nicht notwendig identisch verteilter Zufallsgrößen, $S_n := \sum_{k=1}^{n} X_k$. Die Verteilungsfunktion von X_n werde mit F_n bezeichnet.

Problem: Unter welchen Bedingungen gibt es Zahlenfolgen (a_n) und (b_n) mit $b_n > 0$, so dass die Verteilungen von $\frac{S_n - a_n}{b_n}$ schwach (d. h. in Verteilung) gegen die Normalverteilung konvergieren?

Ohne weitere Voraussetzungen kann man Konvergenz gegen die Normalverteilung nicht erwarten.

Beispiel 11.8 Alle X_n seien Cauchyverteilt mit dem Parameter a. Dann ist auch $\frac{S_n}{n}$ Cauchyverteilt mit dem Parameter a. (Beweis mittels charakteristischer Funktionen.) Das heißt für $a_n \equiv 0$ und $b_n = n$ erhalten wir die Konvergenz von $\frac{S_n - a_n}{b_n}$ für $n \to \infty$, aber nicht gegen die Normalverteilung.

Eine wesentliche Rolle bei der Lösung des oben gestellten Problems spielt der folgende Begriff.

Definition 11.9 *Man sagt, die Folge* (X_n) *erfüllt die* Lindeberg-Bedingung (L), *falls gilt* $D^2X_n < \infty, n \geq 1$ *und falls*

$$\lim_{n\to\infty} \frac{1}{D^2S_n} \sum_{k=1}^{n} \int\limits_{\{x:|x-EX_k|\geq\varepsilon\sigma_n\}} |x-EX_k|^2 F_k(dx) = 0 \ \ \forall\varepsilon > 0. \tag{L}$$

Dabei werde $\sigma_n = \sqrt{D^2S_n}$ *gesetzt.*

Falls die Lindeberg-Bedingung (L) gilt, so folgt

$$\lim_{n\to\infty} \max_{1\leq k\leq n} \frac{D^2X_k}{D^2S_n} = 0. \tag{F}$$

Die Eigenschaft (F) wird auch als Feller-Bedingung bezeichnet.

Beweis: Es gilt

$$\frac{D^2X_k}{D^2S_n} \leqq \varepsilon^2 + \frac{1}{D^2S_n} E\Big[(X_k - EX_k)^2 \mathbb{1}_{\{|X_k-EX_k|\geq\varepsilon\sigma_n\}}\Big].$$

Daraus folgt für jedes $\varepsilon > 0$.

$$\max_{1\leq k\leq n} \frac{D^2X_k}{D^2S_n} \leq \varepsilon^2 + \frac{1}{D^2S_n} \sum_{k=1}^{n} \Big[E(X_k - EX_k)^2 \mathbb{1}_{\{|X_k-EX_k|\geq\varepsilon\sigma_n\}}\Big].$$

Aus (L) folgt nunmehr (F). □

Die Feller-Bedingung besagt anschaulich, dass jede der Streuungen $D^2X_k, k = 1, \ldots, n$, für große n verschwindend klein ist im Vergleich zur Streuung D^2S_n der Summe $X_1 + X_2 + \ldots + X_n$.

Aus der Feller-Eigenschaft (F) ergibt sich eine weitere Eigenschaft der Folge (X_n), die man als "Asymptotische Kleinheit der $X_{n,k} := \frac{X_k-EX_k}{\sigma_n}$" bezeichnet:

$$\lim_{n\to\infty} \max_{1\leq k\leq n} P\Big(\frac{|X_k - EX_k|}{\sigma_n} > \varepsilon\Big) = 0. \tag{AK}$$

Der Beweis ergibt sich unmittelbar aus (F) mittels der Tschebyschev'schen Ungleichung:

$$P\Big(\frac{|X_k - EX_k|}{\sigma_n} > \varepsilon\Big) \leq \frac{D^2 X_k}{\sigma_n^2 \cdot \varepsilon^2} \quad , \quad k = 1, \dots, n.$$

Nunmehr haben wir alle Begriffe, um folgenden Satz zu formulieren.

Satz 11.10 (Zentraler Grenzwertsatz von Lindeberg-Feller) *Es sei* $(X_n, n \geq 1)$ *eine Folge unabhängiger Zufallsgrößen über* $(\Omega, \mathfrak{A}, P)$ *mit* $0 < D^2 X_n < \infty$. *Dann sind folgende Aussagen äquivalent:*

1) *Die* $X_{n,k} = \frac{X_k - EX_k}{\sigma_n}, k = 1, \dots, n; n \geq 1$ *mit* $\sigma_n^2 = D^2 S_n$ *sind asymptotisch klein im Sinne von* (AK) *und es gilt*

$$\lim_{n\to\infty} \sup_{-\infty \leq a < b \leq \infty} |P\Big(a < \frac{S_n - ES_n}{\sqrt{D^2 S_n}} \leq b\Big) - (\Phi(b) - \Phi(a))| = 0.$$

b) *Die Lindeberg-Bedingung* (L) *gilt.*

Beweis: Siraev, (1988), Kap. III, 4.

Beispiele 11.11

a) (X_n) unabhängig, $EX_n \equiv EX_1 = m, D^2 X_n \equiv D^2 X_1 = \sigma^2 \in (0, \infty)$.

   Dann ist die Lindeberg-Bedingung erfüllt, denn es gilt

$$\frac{1}{\sigma_n^2} \sum_1^n \int\limits_{\{x||x-m| \geq \sigma_n \cdot \varepsilon\}} |x - m|^2 dF_1(x) =$$

$$\frac{n}{n\sigma^2} \int\limits_{\{x||x-m| \geq \sqrt{n}\sigma\varepsilon\}} |x - m|^2 dF_1(x) \to 0$$

   wegen $P^{X_1}(\{x||x - m| \geq \sqrt{n}\sigma\varepsilon\}) \leq \frac{D^2 X_1}{n\sigma^2 \cdot \varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0$

b) Für ein $\delta > 0$ sei die folgende *Ljapunov-Bedingung* erfüllt:

$$\frac{1}{\sigma_n^{2+\delta}} \sum_1^n E|X_k - m_k|^{2+\delta} \underset{n\to\infty}{\longrightarrow} 0 \qquad (Ljap).$$

Dann gilt die Lindeberg-Bedingung (L).

Beweis: Für jedes $\varepsilon > 0$ haben wir

$$|X_k - m_k|^{2+\delta} = \int_{R_1} |x - m_k|^{2+\delta} dF_k(x) \geq$$

$$\geq \int_{\{x||x-m_k|\geq\varepsilon\sigma_n\}} |x - m_k|^{2+\delta} dF_k(x) \geq \varepsilon^\delta \sigma_n^\delta \int_{\{x||x-m_k|\geq\varepsilon\sigma_n\}} |x - m_k|^2 dF_k(x)$$

$$\Longrightarrow \frac{1}{\sigma_n^2} \sum_1^n \int_{\{x||x-m_k|\geq\varepsilon\sigma_n\}} (x - m_k)^2 dF_k(x) \leq$$

$$\frac{1}{\varepsilon^\delta \sigma_n^{2+\delta}} \sum_{k=1}^n E|X_k - m_k|^{2+\delta} \to 0$$

Es gibt Folgen (X_n) unabhängiger Zufallsgrößen mit $S_n^* = \frac{\sum_{k=1}^n (X_k - EX_k)}{\sqrt{D^2 S_n}}$ $\xrightarrow{w} N(0,1)$, wo weder (L) gilt noch Asymptotische Kleinheit (AK) vorliegt:

Die Zufallsgrößen $(X_n, n \geq 1)$ seien unabhängig und normalverteilt mit

$$EX_n \equiv 0, \quad D^2 X_1 = 1, D^2 X_n = 2^{n-2}, n \geq 2.$$

Dann ist die Streuung $D^2 S_n$ von $S_n = \sum_{k=1}^n X_k$ gleich $\sum_{k=1}^n D^2 X_k = 2^{n-1}$. Wir setzen wie üblich

$$S_n^* = \frac{1}{\sqrt{D^2 S_n}} \sum_{k=1}^n X_k, \quad n \geq 1.$$

Die Folge $(X_n, n \geq 1)$ genügt nicht der Lindeberg-Bedingung, da insbesondere die Fellereigenschaft (F) nicht gilt:

$$\max_{k=1,\ldots,n} \frac{D^2 X_k}{D^2 S_n} = \max_{k=1,\ldots,n} \frac{2^{k-2}}{2^{n-1}} = \frac{1}{2}.$$

Außerdem sind die $X_{n,k} := \frac{X_k}{\sqrt{D^2 S_n}}$; $\quad k = 1, \ldots, n; n \geq 1$ nicht asymptotisch klein im Sinne von (AK), da für alle $\varepsilon > 0$ und $n \geq 1$ die Gleichung

$$\max_{k=1,\ldots,n} P\Big(\frac{|X_k|}{\sqrt{D^2 S_n}} > \varepsilon\Big) = P\Big(\frac{|X_n|}{\sqrt{2^{n-1}}} > \varepsilon\Big) = 2\big(1 - \Phi(\varepsilon)\big) > 0$$

erfüllt ist.

Andererseits genügt $(X_n, n \geq 1)$ trivialerweise dem zentralen Grenzwertsatz:

S_n^* ist für jedes $n \geq 1$ Standard-normalverteilt.