Prof. Dr. Peter Eichelsbacher

61311
Einführung in die Stochastik

Leseprobe

Fakultät für
Mathematik und Informatik

FernUniversität in Hagen

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, verbleiben bei der FernUniversität. Das Werk darf gemäß § 53 UrhG im privaten Rahmen zum Beispiel durch Kopien oder digitale Speicherung vervielfältigt werden.

Kapitel 15

Der Satz von de Moivre und Laplace

In diesem Kapitel betrachten wir den zentralen Grenzwertsatz für binomialverteilte Zufallsgrößen. Dies ist einer der bedeutendsten Grenzwertsätze der Stochastik. Für unabhängige und identisch verteilte Zufallsgrößen wird das Verhalten der Langzeitmittelwerte untersucht. Beim schwachen Gesetz der großen Zahlen haben wir die Konvergenz der Mittelwerte gegen den gemeinsamen Erwartungswert betrachtet. Ein zentraler Grenzwertsatz beschreibt die asymptotische Größenordnung der Abweichungen der Mittelwerte vom Erwartungswert. Es zeigt sich eine universelle Bedeutung der Normalverteilung, die wir im Rahmen dieses Kurses exemplarisch für Summen von Bernoulli-verteilten Zufallsgrößen studieren. Diese Untersuchung mündet in den Grenzwertsatz von de Moivre und Laplace.

15.1 Der lokale Grenzwertsatz

Zunächst sei daran erinnert, dass die Poisson-Approximation der Binomialverteilung nur sinnvoll ist, wenn n groß und p klein ist. Oft möchte man jedoch $b(k;n,p)$ für einen festen Wert von p und großes n approximieren.

Zunächst betrachten wir $b(n;2n,1/2) = \frac{(2n)!}{(n!)^2}2^{-2n}$. Mithilfe der Stirling-Formel, Satz 7.0.5, folgt

$$b(n;2n,1/2) \sim \frac{1}{\sqrt{\pi n}}$$

(siehe Beispiel 12.0.6) und damit $\lim_{n\to\infty} b(n;2n,1/2) = 0$. Dabei ist $b(n;2n,1/2)$ noch am größten unter den $b(k;2n,1/2)$, wie man aus dem Pascal'schen Dreieck

für die Binomialkoeffizienten ersehen kann. Allgemeiner gilt für ein beliebiges $p \in (0,1)$

$$\lim_{n\to\infty} \max_{0\le k\le n} b(k;n,p) = 0. \tag{15.1}$$

Wir werden dies später beweisen.

Es bezeichne S_n die Anzahl der Erfolge in einem BERNOULLI-Experiment der Länge n und Erfolgswahrscheinlichkeit p. Wir fragen: Wie stark fluktuiert S_n um den Erwartungswert np, in welcher Größenordnung liegen die Abweichungen $S_n - np$ im Limes $n \to \infty$? Wir wissen bereits, dass $S_n - np$ von kleinerer Größenordnung ist als n, denn nach dem Gesetz der großen Zahlen gilt für jedes $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \le \varepsilon\right) = \lim_{n\to\infty} P(|S_n - np| \le \varepsilon n) = 1.$$

Weiter gilt nach (15.1) für jedes $a > 0$:

$$\lim_{n\to\infty} P(|S_n - np| \le a) \le (2a+1) \lim_{n\to\infty} \max\{P(S_n)\} = 0.$$

Um eine hinreichende Anzahl von möglichen Werten für S_n zu „erwischen“, die dann zu einer von 0 verschiedenen Grenzwahrscheinlichkeit führen, muss der zugelassene Bereich mit n unendlich groß werden. Die Größenordnung n aus dem Gesetz der großen Zahlen ist zu großzügig bemessen. Für welche Folgen $(a_n)_n$ positiver reeller Zahlen streben die Wahrscheinlichkeiten $P(|S_n - np| \le a_n)$ weder gegen 0 noch gegen 1? Da wir (zumindest für $p = 1/2$) gesehen haben, dass die größten $b(k;n,p)$ von der Größenordnung $1/\sqrt{n}$ sind, ist es plausibel, wenn sich für die benötigte Anzahl eine Größenordnung von $\sqrt{n}$ ergibt. Dies wird durch das folgende Lemma präzisiert:

15.1.1 Lemma: In obiger Situation gilt

$$\lim_{n\to\infty} P(|S_n - np| \le a_n) = \begin{cases} 1, & \text{falls} \quad a_n/\sqrt{n} \to \infty, \\ 0, & \text{falls} \quad a_n/\sqrt{n} \to 0. \end{cases}$$

Beweis: Mithilfe der TSCHEBYSCHEV-Ungleichung folgt

$$P(|S_n - np| \le a_n) \ge 1 - \frac{np(1-p)}{a_n^2}.$$

Im Fall $a_n/\sqrt{n} \to \infty$ konvergiert die rechte Seite gegen 1, was den ersten Fall beweist. Nun ist für jedes $k \ge 1$

$$b(k;n,p) > b(k-1;n,p)$$

genau dann, wenn $(n-k+1)p > k(1-p)$, also $k < (n+1)p$, gilt. Dies ist in Aufgabe 15.1.2 zu zeigen. Das Histogramm der Binomialverteilung hat also sein Maximum bei $k = k_{n,p} := \lfloor (n+1)p \rfloor$. Damit ist

$$P(|S_n - np| \le a_n) = \sum_{k:|k-np|\le a_n} b(k;n,p) \le (2a_n+1)b(k_{n,p};n,p).$$

Nach der Stirling-Formel ist nun

$$b(k_{n,p};n,p) \sim \frac{1}{\sqrt{2\pi p(1-p)n}} \left(\frac{np}{k_{n,p}}\right)^{k_{n,p}} \left(\frac{n(1-p)}{n-k_{n,p}}\right)^{n-k_{n,p}}.$$

Wir nutzen nun $|k_{n,p} - np| \le 1$ um zu sehen, dass die rechte Seite durch $C/\sqrt{n}$ für eine Konstante C abgeschätzt werden kann. Zum Beispiel ist der zweite Faktor kleiner gleich $(1+1/k_{n,p})^{k_{n,p}} \le e$. Damit konvergiert $(2a_n+1)b(k_{n,p};n,p)$ im Fall $a_n/\sqrt{n} \to 0$ gegen Null. □

15.1.2 Aufgabe: Zeigen Sie: Für jedes $k \ge 1$ ist

$$b(k;n,p) > b(k-1;n,p)$$

genau dann, wenn $(n-k+1)p > k(1-p)$, also $k < (n+1)p$, gilt.

Also wächst die Abweichung der Summen S_n von np typischerweise mit $\sqrt{n}$. Konvergieren die Wahrscheinlichkeiten $P(|S_n - np| \le c\sqrt{n})$ für fest gewähltes $c > 0$ gegen einen Limes, und wenn ja, wie sieht dieser Limes aus?

Zunächst müssen wir die $b(k;n,p)$ einzeln genauer unter die Lupe nehmen: Es zeigt sich, dass $b(k;n,p)$ durch eine schöne Funktion approximiert werden kann, wenn k in einer Umgebung der Größenordnung $\sqrt{n}$ um den Erwartungswert np liegt. Es ist deshalb günstig, eine Variablentransformation durchzuführen, indem k durch $(k-np)/\sqrt{n}$ ersetzt wird. Tatsächlich dividieren wir durch $\sqrt{np(1-p)}$, was der Standardabweichung der Binomialverteilung entspricht. Wir setzen also

$$x_k := x_k(n,p) := \frac{k-np}{\sqrt{np(1-p)}}. \tag{15.2}$$

x_k hängt natürlich von n und p ab, was wir in der Notation jedoch nicht gesondert betonen. Wir kürzen $1-p$ meist durch q ab.

Der unten stehende Satz besagt, dass $b(k;n,p)$ für große n wie $\frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2}$ ist. Was soll das aber genau heißen?

Zu schreiben

$$\lim_{n\to\infty} \sqrt{2\pi npq}\, b(k;n,p) = e^{-x_k^2/2}$$

ergibt keinen großen Sinn, denn mit $n \to \infty$ und festem k geht $x_k \to -\infty$. Was wir wollen, ist eine Approximation, wenn k in einer Umgebung der Größenordnung $\sqrt{n}$ um np liegt, das heißt x_k in einer Umgebung der Größenordnung 1 um 0. Es zeigt sich, dass in einem solchen Bereich die Konvergenz sogar gleichmäßig gilt. Wir können den Bereich zusätzlich in n größer werden lassen:

15.1.3 Satz: (lokaler Grenzwertsatz) Es seien $0 < p < 1$, $q := 1 - p$ und $(a_n)_{n\in\mathbb{N}}$ eine Folge strikt positiver, reeller Zahlen mit $\lim_{n\to\infty} a_n^3/\sqrt{n} = 0$. Dann gilt mit x_k aus (15.2)

$$\lim_{n\to\infty} \sup_{k:|x_k|\le a_n} \left| \frac{\sqrt{2\pi npq}\, b(k;n,p)}{e^{-x_k^2/2}} - 1 \right| = 0.$$

15.1.4 Bemerkungen: (a) Ist $a_n = A$ eine beliebige, aber feste positive Konstante, so folgt unmittelbar

$$\lim_{n\to\infty} \sup_{k:|x_k|\le A} \left| \frac{\sqrt{2\pi npq}\, b(k;n,p)}{e^{-x_k^2/2}} - 1 \right| = 0.$$

(b) Wir schreiben nachfolgend stets $b(k;n,p) \sim \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2}$ für die obige gleichmäßige Konvergenz und dies auch allgemeiner: Sind $\alpha(k,n)$, $\beta(k,n) > 0$ für $n \in \mathbb{N}_0$, $0 \le k \le n$, so bedeutet (während des unten stehenden Beweises) $\alpha(k,n) \sim \beta(k,n)$, dass für die obige Folge $(a_n)_{n\in\mathbb{N}}$ strikt positiver Zahlen

$$\lim_{n\to\infty} \sup_{k:|x_k|\le a_n} \left| \frac{\alpha(k,n)}{\beta(k,n)} - 1 \right| = 0$$

gilt.

(c) Wir überzeugen uns vom folgenden Sachverhalt, der im Beweis von Satz 15.1.3 mehrfach verwendet wird:

$$\alpha(k,n) \sim \beta(k,n),\ \ \alpha'(k,n) \sim \beta'(k,n) \ \ \Rightarrow\ \ \alpha(k,n)\alpha'(k,n) \sim \beta(k,n)\beta'(k,n). \tag{15.3}$$

Beweis von (15.3)**:** Sei $\alpha := \alpha(k,n), \alpha' := \alpha'(k,n), \beta := \beta(k,n), \beta' := \beta'(k,n)$. Wir verwenden zweimal die Dreiecksungleichung und erhalten

$$\begin{aligned} \left| \frac{\alpha\alpha'}{\beta\beta'} - 1 \right| &\le \left| \frac{\alpha\alpha'}{\beta\beta'} - \frac{\alpha'}{\beta'} \right| + \left| \frac{\alpha'}{\beta'} - 1 \right| \\ &\le \left(\left| \frac{\alpha'}{\beta'} - 1 \right| + 1 \right) \left| \frac{\alpha}{\beta} - 1 \right| + \left| \frac{\alpha'}{\beta'} - 1 \right|. \end{aligned}$$

Daraus folgt die behauptete Aussage. □

Beweis von Satz 15.1.3: Es gilt nach (15.2)

$$k = np + \sqrt{npq}\, x_k, \;\; n - k = nq - \sqrt{npq}\, x_k,$$

also

$$k \sim np, \;\; n - k \sim nq.$$

Hierbei ist mit $\sim$ die obige gleichmäßige Konvergenz (Bemerkung 15.1.4(b)) gemeint. Mithilfe der Stirling'schen Formel folgt:

$$\begin{aligned} b(k;n,p) &\sim \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, p^k q^{n-k}}{\left(\frac{k}{e}\right)^k \sqrt{2\pi k}\left(\frac{n-k}{e}\right)^{n-k} \sqrt{2\pi(n-k)}} \\ &= \sqrt{\frac{n}{2\pi k(n-k)}}\, \varphi(n,k) \sim \frac{1}{\sqrt{2\pi npq}}\, \varphi(n,k), \end{aligned} \tag{15.4}$$

wobei wir $\varphi(n,k)$ für $(\frac{np}{k})^k (\frac{nq}{n-k})^{n-k}$ schreiben. Man beachte, dass auch hier $\sim$ die obige gleichmäßige Konvergenz bezeichnet. Dass die Stirling-Formel aus Satz 7.0.5 diese Konvergenz impliziert, mache man sich klar. Es ist nun

$$-\ln(\varphi(n,k)) = -nH\Big(\frac{k}{n}\Big|\, p\Big),$$

wobei

$$H(x|\, p) := -x \ln\Big(\frac{x}{p}\Big) - (1-x)\ln\Big(\frac{1-x}{1-p}\Big).$$

Wir kennen diese Funktion schon aus Kapitel 13 (siehe Satz 13.1.2). In Satz 13.1.5 haben wir die Ableitungen von $H(\cdot|p)$ bestimmt. Es gilt $H'(p|\, p) = 0$ und $H''(p|\, p) = 1/p + 1/q = 1/(pq)$. Damit folgt

$$H(x|\, p) = \frac{(x-p)^2}{2pq} + \psi(x-p),$$

wobei wir mit ψ das Restglied in der Taylor-Entwicklung bezeichnen (man denke sich zum Beispiel die Lagrange'sche Form; Joseph-Louis Lagrange (1736–1813)), für die wir also in der Umgebung von p eine Abschätzung

$$|\psi(x-p)| \le c|x-p|^3$$

mit einer geeigneten Konstante c zur Verfügung haben. Wir erhalten somit

$$\left| \ln(\varphi(n,k)) - \frac{n\left(\frac{k}{n} - p\right)^2}{2pq} \right| \le cn \left| \frac{k}{n} - p \right|^3.$$

Man beachte nun, dass aus der Definition der x_k für eine geeignete Konstante $0 < c' < \infty$ folgt

$$\frac{|k-np|^3}{n^2} = c'\frac{|x_k|^3}{\sqrt{n}}.$$

Wählen wir nun ein k mit $|x_k| \le a_n$, so konvergiert die rechte Seite der Ungleichung gegen 0 für $n \to \infty$. Dies folgt unmittelbar aus der Bedingung an die Folge $(a_n)_{n\in\mathbb{N}}$. Da nun aber

$$\frac{n\left(\frac{k}{n}-p\right)^2}{2pq} = \frac{x_k^2}{2},$$

erhalten wir

$$\lim_{n\to\infty} \sup_{k:|x_k|\le a_n} \left| \frac{\varphi(n,k)}{e^{-x_k^2/2}} - 1 \right| = 0.$$

Nach (15.4) ist damit der Satz gezeigt. □

15.1.5 Beispiel: Jemand wirft 1200-mal einen Würfel. Mit welcher Wahrscheinlichkeit hat er genau 200-mal eine 6? Mit welcher Wahrscheinlichkeit 250-mal? Wir berechnen x_k für $k = 200$ bzw. 250, $n = 1200$ und $p = 1/6$. Es gilt $x_{200} = 0$, $x_{250} = \frac{5\sqrt{6}}{\sqrt{10}} \approx 3.873$ sowie

$$b(200; 1200, 1/6) \approx 0.0309019 \quad \text{und} \quad b(250; 1200, 1/6) \approx 0.0000170913.$$

Die sechs Bilder in Abbildung 15.1 (auf Seite 408) illustrieren die Konvergenz der Binomialverteilung gegen die Funktion $\varphi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$. Hier ist jeweils die Funktion $\varphi(x)$ zusammen mit dem skalierten Histogramm

$$f_{n,p}(x) := \begin{cases} \sqrt{np(1-p)}b(k;n,p), & k \in \{0,1,\dots,n\} \text{ mit } |x-x_k| < \frac{1}{2\sqrt{np(1-p)}}, \\ 0, & \text{andernfalls} \end{cases}$$

der Binomialverteilung $b(\cdot\,; n, p)$ gezeichnet. In der linken Spalte der symmetrische Fall mit $p = 1/2$, in der rechten Spalte der asymmetrische Fall $p = 1/5$.

Wir können noch den versprochenen Beweis von (15.1) nachholen:

Beweis von (15.1)**:** Da die Standardabweichung von S_n gleich $\sqrt{npq}$ ist, so folgt aus der TSCHEBYSCHEV-Ungleichung, dass für $a > 0$ $P(|S_n - np| \ge a\sqrt{npq}) \le 1/a^2$ gilt. Da die Summe nichtnegativer Zahlen mindestens so groß wie jedes ihrer Glieder ist, folgt $b(k; n, p) \le 1/a^2$ für $|x_k| \ge a$. Andererseits ergibt sich aus Satz 15.1.3:

$$\lim_{n\to\infty} \sup_{k:|x_k|\le a} b(k;n,p) = 0.$$

Da a beliebig groß gewählt werden kann, folgt (15.1). □

Wie üblich muss hier bemerkt werden, dass ein reines Limesresultat für die Güte einer Approximation wie in obigem Rechenbeispiel zunächst natürlich gar nichts aussagt. Gefragt sind konkrete Abschätzungen des Fehlers. Dies ist ein technisch aufwendiges Feld, in das wir in diesem Kurs nicht eintreten werden. Wir werden dennoch am Ende des Kapitels eine Abschätzung der Konvergenzgeschwindigkeit bekanntgeben.

15.2 Der Satz von de Moivre und Laplace

Da die Wahrscheinlichkeiten $b(k;n,p)$ für große n somit alle klein sind, ist man meist eher daran interessiert, Summen dieser Wahrscheinlichkeiten zu approximieren.

15.2.1 Satz: (von de Moivre und Laplace) Für beliebige reelle Zahlen a und b mit $a < b$ gilt:

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx. \tag{15.5}$$

Beweis: Sei $k \in \{0,\dots,n\}$. Dann ist $\{S_n = k\} = \{(S_n - np)/\sqrt{npq} = x_k\}$. Also ist die links stehende Wahrscheinlichkeit gleich

$$\sum_{k:a\le x_k\le b} P(S_n = k) = \sum_{k:a\le x_k\le b} b(k;n,p).$$

Wir setzen nun für jeden Summanden auf der rechten Seite seinen in Satz 15.1.3 angegebenen asymptotischen Wert ein und berücksichtigen, dass $x_{k+1} - x_k = \frac{1}{\sqrt{npq}}$ ist. Die Summe dieser Größen nennen wir R_n:

$$R_n := \frac{1}{\sqrt{2\pi}} \sum_{k:a\le x_k\le b} e^{-x_k^2/2}(x_{k+1} - x_k).$$

Unter Verwendung der Gleichmäßigkeit der Konvergenz in $\{k : |x_k| \le A\}$ für jedes A sieht man sofort, dass der Quotient von $P\Big(a \le \frac{S_n-np}{\sqrt{npq}} \le b\Big)$ und dem oben stehenden Ausdruck R_n gegen 1 konvergiert. Das heißt, es existiert eine Nullfolge $(\varepsilon_n)_{n\in\mathbb{N}}$, $\varepsilon_n > 0$, mit

$$R_n(1-\varepsilon_n) \le P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right) \le R_n(1+\varepsilon_n). \tag{15.6}$$

Die Zuordnung zwischen k und x_k ist bijektiv. Wenn k von 0 bis n läuft, dann variiert x_k im Intervall $[-\sqrt{np/q}, \sqrt{nq/p}]$ mit der Schrittweite $x_{k+1} - x_k = 1/\sqrt{npq}$.

Für hinreichend große n umfaßt dieses Intervall das gegebene Intervall $[a, b]$, und die in $[a, b]$ fallenden Punkte x_k teilen dieses in Teilintervalle derselben Länge $1/\sqrt{npq}$. Wenn nun der kleinste und der größte Wert von k mit $a \le x_k \le b$ gleich j bzw. l ist, dann ist

$$x_{j-1} < a \le x_j < x_{j+1} < \cdots < x_{l-1} < x_l \le b < x_{l+1}$$

und die obige Summe läßt sich schreiben als

$$\sum_{k=j}^{l} \varphi(x_k)(x_{k+1} - x_k),$$

wobei $\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ ist. Das ist eine sogenannte Riemann'sche Summe für das bestimmte Riemann-Integral $\int_a^b \varphi(x)\,dx := \lim_{n\to\infty} R_n$. Üblicherweise gehören jedoch in der Lehrbuchliteratur beim Riemann-Integral die Endpunkte a und b mit zu den Stützstellen der Funktion. Doch macht das für $n \to \infty$ und eine dadurch immer feiner werdende Unterteilung keinen Unterschied. Somit konvergiert R_n gegen das Integral in der Behauptung des Satzes. Dieser folgt nun sofort mit (15.6). □

Abraham de Moivre (1667–1754) veröffentlichte dieses Ergebnis in seiner *„Doctrine of Chances“* 1714. Laplace erweiterte das Ergebnis und wies dessen Bedeutung in seiner *„Théorie analytique des probabilités“* 1812 nach. Es handelt sich um den zuerst bekanntgewordenen Spezialfall des *Zentralen Grenzwertsatzes*.

Die Funktion $x \mapsto \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ heißt auch Gauss'sche Glockenkurve, wegen des glockenförmigen Verlaufs ihres Graphen. Gauss hat sich intensiv mit dem zentralen Grenzwertsatz und der Normalverteilung beschäftigt.

Die Integrale $\int_a^b \varphi(x)\,dx$ sind leider nicht in geschlossener Form mithilfe von Polynomen, rationalen Funktionen, Wurzelausdrücken oder elementaren transzendenten Funktionen (wie sin, cos, exp, etc.) darstellbar.

Es gilt offenbar für $a < b$

$$\int_a^b \varphi(x)\,dx = \int_{-\infty}^b \varphi(x)\,dx - \int_{-\infty}^a \varphi(x)\,dx = \Phi(b) - \Phi(a),$$

wobei wir

$$\Phi(y) := \int_{-\infty}^y \varphi(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2}\,dx \tag{15.7}$$

gesetzt haben. (Es ist leicht ersichtlich, dass das uneigentliche Integral konvergiert.) Der Grenzwertsatz von de Moivre und Laplace lässt vermuten, dass

$$\int_{-\infty}^{\infty} \varphi(x)\,dx = 1, \tag{15.8}$$

denn wählen wir ein sehr kleines a und ein sehr großes b, so ist die Wahrscheinlichkeit der linken Seite von (15.5) nahe 1. Wir geben einen Beweis: Dazu verwenden wir Satz 15.2.1 und setzen

$$S_n^* := \frac{S_n - np}{\sqrt{npq}}.$$

(Für das Argument hier spielt p keine Rolle; wir könnten $p = 1/2$ nehmen.) Sei $a > 0$. Dann ist

$$1 = P(-a \le S_n^* \le a) + P(|S_n^*| > a).$$

Nach der Tschebyschev-Ungleichung gilt:

$$P(|S_n^*| > a) \le \frac{1}{a^2} \operatorname{Var}(S_n^*) = \frac{1}{a^2}.$$

Nach Satz 15.2.1 gilt

$$\lim_{n\to\infty} P(-a \le S_n^* \le a) = \int_{-a}^{a} \varphi(x)\, dx.$$

Demzufolge ist

$$1 - \frac{1}{a^2} \le \int_{-a}^{a} \varphi(x)\, dx \le 1$$

für jedes $a > 0$, womit (15.8) bewiesen ist. Wir hatten in Beispiel 10.3.5(a) darauf hingewiesen, dass der Beweis von (15.8) nicht einfach ist. Wir haben ihn hier im Rahmen der Stochastik vorgeführt, indem wir maßgeblich die Güte der Stirling-Approximation verwendet haben. Genau diese liefert den Vorfaktor $\frac{1}{\sqrt{2\pi}}$ auf der rechten Seite von (15.5).

15.2.2 Bemerkungen: (a) Wegen $\lim_{n\to\infty} \sup_k P(S_n = k) = 0$ ist es natürlich gleichgültig, ob in der Aussage von (15.5) $\le$ oder $<$ steht.

(b) Es gilt für $a \in \mathbb{R}$:

$$\begin{aligned} \lim_{n\to\infty} P\left(\frac{S_n - np}{\sqrt{npq}} \le a\right) &= \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx, \\ \lim_{n\to\infty} P\left(\frac{S_n - np}{\sqrt{npq}} \ge a\right) &= 1 - \Phi(a). \end{aligned}$$

Beweis: Wir beweisen in (b) die erste Gleichung; die zweite folgt daraus unmittelbar. Wegen der Symmetrie von φ und (15.8) gilt:

$$\Phi(x) = \int_{-\infty}^{x} \varphi(u)\, du = 1 - \int_{x}^{\infty} \varphi(u)\, du = 1 - \int_{-\infty}^{-x} \varphi(u)\, du = 1 - \Phi(-x)$$

für alle $x \in \mathbb{R}$. Wir setzen wieder $S_n^* := \frac{S_n - np}{\sqrt{npq}}$ und wählen $b > 0$ so groß, dass $-b < a$ gilt. Dann ist nach Satz 15.2.1

$$\begin{aligned}
\limsup_{n\to\infty} P(S_n^* \le a) &= \limsup_{n\to\infty} \Big[P(-b \le S_n^* \le a) + P(S_n^* < -b)\Big] \\
&= \limsup_{n\to\infty} \Big[P(-b \le S_n^* \le a) + (1 - P(S_n^* \ge -b))\Big] \\
&\le \limsup_{n\to\infty} \Big[P(-b \le S_n^* \le a) + (1 - P(-b \le S_n^* \le b))\Big] \\
&= \Phi(a) - \Phi(-b) + \Big[1 - \Phi(b) + \Phi(-b)\Big] \\
&= \Phi(a) + \Phi(-b)
\end{aligned}$$

und

$$\begin{aligned}
\liminf_{n\to\infty} P(S_n^* \le a) &\ge \liminf_{n\to\infty} P(-b \le S_n^* \le a) \\
&= \Phi(a) - \Phi(-b).
\end{aligned}$$

Wegen $\Phi(-b) \to 0$ für $b \to \infty$ folgt die gewünschte Aussage. □

Wir hatten bereits gesehen, dass für $x \le 0$ gilt: $\Phi(x) = 1 - \Phi(-x)$. Es gilt die folgende Approximation für $\Phi(t)$ mit $t \ge 0$ (ohne Beweis):

$$\Phi(t) \approx 1 - \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{t^2}{2}\Big) \left(\frac{a_1}{1+bt} + \frac{a_2}{(1+bt)^2} + \frac{a_3}{(1+bt)^3}\right)$$

mit $b = 0.33267$, $a_1 = 0.4361$, $a_2 = -0.1202$ und $a_3 = 0.937$.

Der Satz 15.2.1 ist eine Präzisierung des Gesetzes der großen Zahlen, welches besagt, dass für jedes $\varepsilon > 0$ $\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \ge \varepsilon\right) = 0$ ist. Letzteres können wir sofort auch aus Satz 15.2.1 herleiten:

$$\begin{aligned}
P\left(\left|\frac{S_n}{n} - p\right| \le \varepsilon\right) &= P\left(-\varepsilon \le \frac{S_n}{n} - p \le \varepsilon\right) \\
&= P\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{pq}} \le \frac{S_n - np}{\sqrt{npq}} \le \frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) \ge P\left(a \le \frac{S_n - np}{\sqrt{npq}} \le b\right),
\end{aligned}$$

sofern n so groß ist, dass $\sqrt{n}\varepsilon/\sqrt{pq} \ge b$ und $-\sqrt{n}\varepsilon/\sqrt{pq} \le a$ sind. Für beliebige Zahlen $a, b \in \mathbb{R}$ ist dies aber für genügend große n der Fall. Somit ist $\lim_{n\to\infty} P(|\frac{S_n}{n} - p| \le \varepsilon) = 1$ für jedes $\varepsilon > 0$.

Dieser Beweis ist natürlich insgesamt wesentlich aufwendiger als der in Kapitel 12 angegebene. Satz 15.2.1 ist jedoch sehr viel informativer als das Gesetz der großen Zahlen.

x	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7356	0.7389	0.7421	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8364	0.8389
1.0	0.8413	0.8437	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8728	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8979	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9146	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9624	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9685	0.9692	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9895	0.9898	0.9901	0.9903	0.9906	0.9908	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9926	0.9928	0.9930	0.9932	0.9934	0.9936
2.5	0.9938	0.9939	0.9941	0.9943	0.9944	0.9946	0.9947	0.9949	0.9950	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9973
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990

Tab. 15.1: Die Verteilungsfunktion $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$ für $x \geq 0$.
Den Wert etwa für $x = 1.75$ findet man in Zeile 1.7 und Spalte 5: $\Phi(1.75) = 0.9599$.

15.2.3 Beispiel: (Ein Anwendungsbeispiel) Eine Fabrik stellt ein Werkstück her mit einer Ausschussrate von 10%. Mit welcher Wahrscheinlichkeit sind unter 400 produzierten mehr als 50 defekt? $n := 400$, $p := 0.1$, $np = 40$, $\sqrt{np(1-p)} = 6$ liefert

$$P(S_n > 50) = P\left(\frac{S_n - np}{\sqrt{np(1-p)}} > \frac{5}{3}\right) \approx 1 - \Phi\left(\frac{5}{3}\right) = \Phi\left(-\frac{5}{3}\right) \approx 0.05.$$

Mit welcher Wahrscheinlichkeit sind zwischen 35 und 45 defekt?

$$\begin{aligned} P(35 \le S_n \le 45) &= P\left(-\frac{5}{6} \le \frac{S_n - np}{\sqrt{np(1-p)}} \le \frac{5}{6}\right) \\ &\approx \Phi\left(\frac{5}{6}\right) - \Phi\left(-\frac{5}{6}\right) = 1 - 2\Phi\left(-\frac{5}{6}\right) \approx 0.6. \end{aligned}$$

Da wir keine Fehlerabschätzungen hergeleitet haben, wissen wir natürlich nicht, wie genau solche Näherungen sind. Die Genauigkeit ist etwas besser, wenn man die Mitte der möglichen Grenzpunkte nimmt; das heißt, im obigen Beispiel schreibt man besser:

$$P(S_n > 50) = P(S_n \ge 50.5) \approx 1 - \Phi\left(\frac{21}{12}\right)$$

$$P(35 \le S_n \le 45) = P(34.5 \le S_n \le 45.5) \approx 1 - 2\Phi\left(\frac{11}{12}\right).$$

Für $n \to \infty$ ist die Korrektur natürlich belanglos, sie ist jedoch immerhin von der Größenordnung $1/\sqrt{n}$.

15.2.4 Aufgabe: Angenommen, bei der Geburt eines Kindes beträgt die Wahrscheinlichkeit für ein Mädchen $p = 0.495$. Mit welcher Wahrscheinlichkeit gibt es in einer Population von 80 Millionen Menschen höchstens 39 Millionen Frauen?

15.2.5 Aufgabe: Unter n Wählern seien S_n Wähler einer Partei P. Wir verwenden S_n/n als einen Schätzwert für die Wahrscheinlichkeit p, dass ein rein zufällig befragter Wähler für die Partei P stimmt. Wie groß muss n sein, damit die Wahrscheinlichkeit eines Irrtums von mehr als 1% nicht größer ist als 0.05?

15.2.6 Aufgabe: Ein Hotel hat 100 Betten. Wie viele Reservierungen darf man akzeptieren, wenn nach Erfahrung eine Reservierung mit Wahrscheinlichkeit 0.2 annulliert wird, und die Wahrscheinlichkeit einer Überbuchung höchstens 0.01 betragen soll?

15.2.7 Aufgabe: Zwei Spieler S und T werfen n mal unabhängig eine Münze. Bei Kopf zahlt T an S einen Euro, im Fall von Zahl umgekehrt S an T. Es sei S_n der Kapitalstand von S nach n Spielen. Bestimmen Sie

$$P\big(-\sqrt{n} \le S_n \le \sqrt{n}\big)$$

approximativ.

Die im vorangegangenen Beispiel erwähnte Korrektur wird *Stetigkeitskorrektur* genannt und begründet sich, wie angedeutet, wie folgt. Mithilfe des Satzes 15.2.1 approximieren wir wie folgt:

$$\begin{aligned} P(k \le S_n \le l) &= P\left(\frac{k-np}{\sqrt{np(1-p)}} \le S_n^* \le \frac{l-np}{\sqrt{np(1-p)}}\right) \\ &\approx \Phi\left(\frac{l-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right) \\ &= \Phi_{x_l} - \Phi_{x_k}. \end{aligned}$$

Der Baustein $P(S_n = s)$ ist im Histogramm der standardisierten Binomialverteilung die Fläche eines Rechtecks mit Mittelpunkt x_s und Grundseite $1/\sqrt{np(1-p)}$. Wenn wir nun dieses Flächenstück bei der Approximation des Histogramms durch ein Integral über φ gut erfassen wollen, sollte die obere Integrationsgrenze nicht x_s sondern $x_s + 1/(2\sqrt{np(1-p)})$ sein. Analog sollte die untere Integrationsgrenze angepasst werden.

In früheren Jahren wurde die Bedeutung von Satz 15.2.1 auch dadurch hervorgehoben, dass für die Binomialverteilung die Berechnung von Bereichswahrscheinlichkeiten der Form $P(k \le S_n \le l)$ sehr kompliziert war. Mit leistungsfähigen Computer-Algebra-Systemen ist dieses numerische Problem heute überwunden. Phantastisch ist hingegen, dass Satz 15.2.1 sich weitgehend verallgemeinern lässt. Der folgende Satz zeigt die *Universalität* der Normalapproximation. Wir zitieren ihn ohne Beweis. Ein allgemeineres Ergebnis wird für gewöhnlich in einer Vorlesung „Wahrscheinlichkeitstheorie“ bewiesen.

15.2.8 Satz: (Zentraler Grenzwertsatz) Für jedes $n \in \mathbb{N}$ seien auf einem W.-Raum (Ω_n, p_n) n unabhängige Zufallsgrößen $X_1^{(n)}, \ldots, X_n^{(n)}$ definiert, die für alle i und n dieselbe Verteilung haben und deren Erwartungswert E und Varianz V existieren (notwendigerweise für alle $X_i^{(n)}$ dieselben). Dann gilt für jedes $a \in \mathbb{R}$

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i^{(n)} - E}{\sqrt{V}} \le a\right) = \Phi(a).$$

Wir sehen also, dass das Verhalten einer Summe unabhängiger und identisch verteilter Zufallsgrößen asymptotisch für $n \to \infty$ nur vom Erwartungswert und von der Varianz, nicht aber von der speziellen Gestalt der Verteilung von X_1 bestimmt wird. Der Begriff *zentraler Grenzwertsatz* beschreibt tatsächlich *nicht* die zentrale Rolle, die dieser Satz in der Stochastik spielt, sondern das Phänomen, dass Wahrscheinlichkeiten zentraler Ereignisse untersucht werden, also solche Ereignisse, die in der Nähe des Erwartungswertes stattfinden.

In Kapitel 3 hatten wir mithilfe der Tschebyschev-Ungleichung für eine Stichprobe $(x_1, \ldots, x_n)$ bewiesen, dass mindestens 50 % der Werte im $2s_x$-Intervall liegen, siehe Satz 3.2.12 und die anschließende Diskussion. Wie ungenau diese Angabe war, können wir nun darstellen. In Satz 15.2.1 sei $a = -k$ und $b = k$. Mit $\mathbb{E}(S_n) = np$ und $\mathbb{V}(S_n) = np(1-p)$ folgt

$$\lim_{n\to\infty} P\Big(\mathbb{E}(S_n) - k\sqrt{\mathbb{V}(S_n)} \le S_n \le \mathbb{E}(S_n) + k\sqrt{\mathbb{V}(S_n)}\Big) = \int_{-k}^{k} \varphi(x)\,dx = 2\Phi(k) - 1.$$

Für $k = 2$ ist $2\Phi(2) - 1 \approx 0.954$, also liegt S_n für große n ungefähr mit Wahrscheinlichkeit 0.954 in den Grenzen $\mathbb{E}(S_n) \pm 2\sqrt{\mathbb{V}(S_n)}$.

15.2.9 Aufgabe: Bestimmen Sie obige Approximation auch für $k = 1$ und $k = 3$.

Zum Abschluss dieses Kapitel wollen wir die beiden verschiedenen Approximationen der Binomialverteilung, die Poisson-Approximation und die Normalapproximation, voneinander abgrenzen. Dazu erinnern wir, dass eine Poisson-Approximation gut ist, wenn np^2 klein ist. Dies hatten wir bewiesen, siehe Satz 14.2.1. Mit etwas mehr Aufwand kann die Fehlerschranke in Satz 14.2.1 sogar noch verbessert werden. Es gilt:

$$\Delta(n,p) \le \min\Big(np^2, p\Big).$$

Man erkennt sofort, dass diese Fehlerabschätzung immer dann die Poisson-Approximation als gut erscheinen lässt, sobald nur p klein ist. Wir beweisen diese Abschätzung nicht. Sie nutzt eine Methode, die auf Charles Stein (geb. 1920) zurückgeht. Stein ist emeritierter Professor an der Stanford University, USA.

Im Fall der Normalapproximation gibt es ebenfalls Fehlerabschätzungen. Wenn man in der Situation von Satz 15.2.8 zusätzlich fordert, dass $\mathbb{E}(|X_1|^3) < \infty$, so kann man beweisen:

$$\sup_{z\in\mathbb{R}} \left| P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i^{(n)} - E}{\sqrt{V}} \le z \right) - \Phi(z) \right| \le 0.8 \frac{\mathbb{E}(|X_1 - \mathbb{E}(X_1)|^3)}{V^{3/2}} \frac{1}{\sqrt{n}}.$$

Dies ist der Satz von Andrew C. Berry und Carl-Gustav Esseen (1918–2001). Wir beweisen diesen Satz nicht. In der allgemeinen Situation wie in Satz

15.2.8 ist dies ein tiefes Resultat der Wahrscheinlichkeitstheorie. Im Fall der Binomialapproximation könnten wir einen direkten Beweis rein technisch gesehen wagen, aber aus Platzgründen verzichten wir an dieser Stelle darauf. Für Bernoulli-verteilte X_i folgt, dass die Normalapproximation der Binomialverteilung gut ist, wenn

$$\frac{p(1-p)^3+(1-p)p^3}{(p(1-p))^{3/2}}\frac{1}{\sqrt{n}}=\frac{p^2+(1-p)^2}{\sqrt{np(1-p)}}$$

klein ist. Da $1/2 \leq p^2+(1-p)^2 \leq 1$ gilt, ist dieser Ausdruck klein, wenn $np(1-p)$ groß ist. Liegt p nahe bei 1, so sind beide Approximationen schlecht.

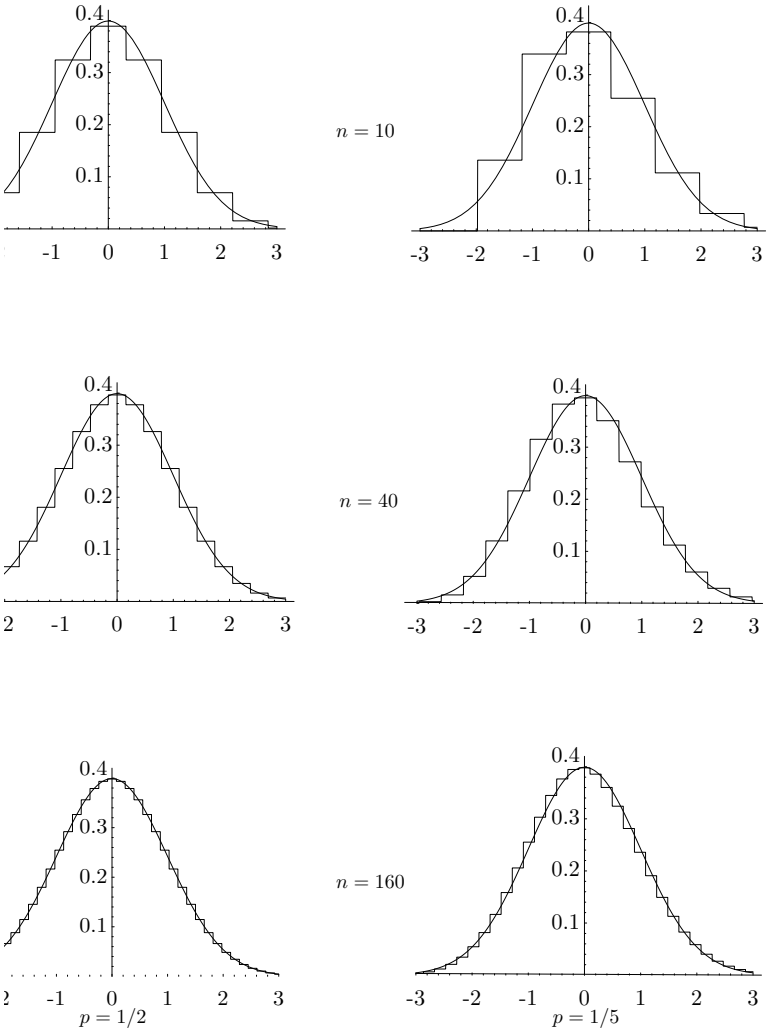

Abb. 15.1: Numerische Illustration von Satz 15.1.3, links für $p = 1/2$ und rechts für $p = 1/5$.

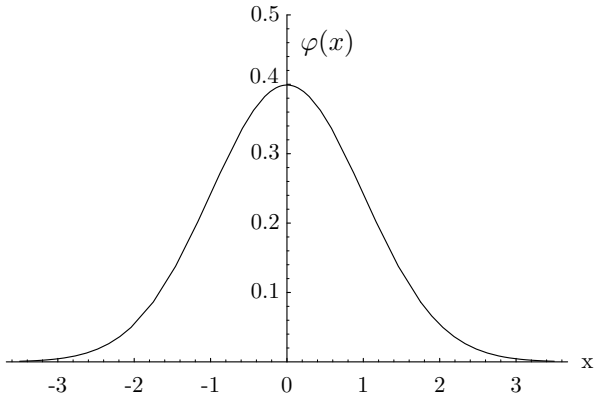

Abb. 15.2: Die Funktion $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

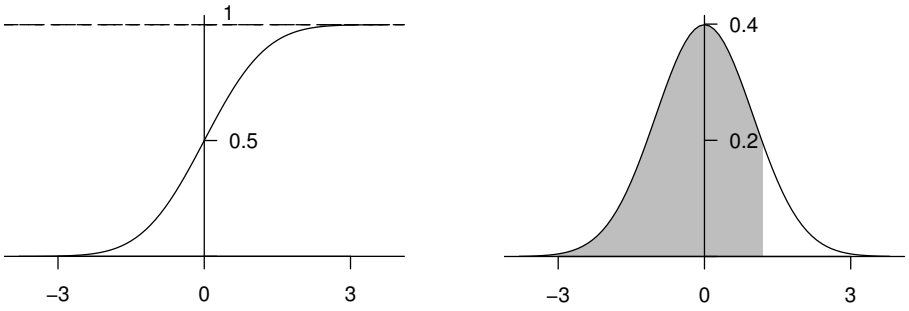

Abb. 15.3: Veranschaulichung der Verteilungsfunktion Φ (links) und ϕ (rechts). Die grau unterlegte Oberfläche unter dem Graphen von ϕ (rechts) entspricht $\Phi(t) = \int_{-\infty}^{t} \phi(u)\, du$.