3 Grenzwertsätze

3.1 Konvergenz von Zufallsvariablen

Betrachten wir als Beispiel die Zahlen $y \in [0,1)$ in Dezimalform $y = 0, y_1 y_2 y_3 \cdots = \sum_{j=1}^{\infty} y_j 10^{-j}$. Es sei Y_j die diskrete Zufallsvariable $Y_j : [0,1) \to y_j \in \{0,1,\ldots,9\}$, mit $P(Y_j = y) = \frac{1}{10}$ für alle y. Nun setzen wir

$$X_n = \sum_{j=1}^{n} 10^{-j} Y_j$$

und lassen $X_n \to Y = \sum_{j=1}^{\infty} 10^{-j} Y_j$ gehen. Die Y_j sind unabhängig und wir vermuten, dass Y gleichverteilt auf $[0,1)$ ist.

Definition. Es seien $X_1, X_2, \ldots$ reelle Zufallsvariablen auf dem W-Raum $(\Omega, \mathcal{E}, P)$.

1. $X_n \to X$ *fast sicher*: $\Leftrightarrow P(\{\omega \in \Omega : X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega)\}) = 1$, in Zeichen $X_n \xrightarrow{f.s.} X$.

2. $X_n \to X$ *in Wahrscheinlichkeit* (oder stochastisch): $\Leftrightarrow P(|X_n - X| > \varepsilon) \overset{n\to\infty}{\longrightarrow} 0$ für alle $\varepsilon > 0$, in Zeichen $X_n \xrightarrow{i.W.} X$.

3. $X_n \to X$ *in Verteilung*: $\Leftrightarrow P(X_n \le x) \overset{n\to\infty}{\longrightarrow} P(X \le x)$ für alle x, für die $F(x) = P(X \le x)$ stetig ist, in Zeichen $X_n \xrightarrow{i.V.} X$.

Die Interpretation ist: $X_n \xrightarrow{f.s.} X$ falls für große n, $X_n(\omega) \sim X(\omega)$; $X_n \xrightarrow{i.W.} X$, falls für große n, X_n nahe bei X ist.

Satz 3.1. a. $X_n \xrightarrow{f.s.} X_n$ *impliziert* $X_n \xrightarrow{i.W.} X$. *Die Umkehrung gilt nicht allgemein.*
b. $X_n \xrightarrow{i.W.} X_n$ *impliziert* $X_n \xrightarrow{i.V.} X$. *Die Umkehrung gilt nicht allgemein.*

Beweis. a. Es sei o.B.d.A. $X = 0$, ansonsten betrachten wir $X_n - X$. Es sei $A_n = \{\omega \in \Omega : |X_n(\omega)|, |X_{n+1}(\omega)|, \ldots \le \varepsilon\}$, also $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$. Gilt $X_n(\omega \to 0$, so haben wir $|X_n(\omega)| \le \varepsilon$ für $n \ge N$, also $\omega \in A_N$. Es folgt

$$\{\omega \in \Omega : X_n(\omega) \to 0\} \subseteq \bigcup_{n\ge1} A_n \,,$$

also aus Lemma 1.1a) und wegen $X_n \xrightarrow{f.s.} 0$

$$1 = P(\bigcup_{n\geq 1} A_n) = \lim_{n\to\infty} P(A_n)\,.$$

Nun haben wir $\{\omega \in \Omega : |X_n(\omega)| > \varepsilon\} \subseteq \Omega \setminus A_n$, also

$$P(|X_n| > \varepsilon) \leq 1 - P(A_n),$$

und es folgt $\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0$, das heißt $X_n \xrightarrow{i.W.} 0$.

Um zu zeigen, dass $X_n \xrightarrow{i.W.} X \not\Rightarrow X_n \xrightarrow{f.s.} X$ betrachten wir das folgende Gegenbeispiel. Es seien $X_1, X_2, \ldots$ unabhängig auf Ω mit Werten in $\{0,1\}$,

$$P(X_n = 1) = \frac{1}{n},\ P(X_n = 0) = 1 - \frac{1}{n}\,.$$

Wir haben

$$\begin{aligned} P(X_n > \varepsilon) &= \frac{1}{n} \to 0 \quad \text{für } 0 < \varepsilon < 1 \\ P(X_n > \varepsilon) &= 0 \to 0 \quad \text{für } \varepsilon \geq 1\,, \end{aligned}$$

also $X_n \xrightarrow{i.W.} 0$.

Sei nun $A_n = \{\omega \in \Omega : X_k(\omega) = 0 \text{ für } k \geq n\}$. Dann ist wieder $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$, also $P(\bigcup A_n) = \lim P(A_n)$. Nun gilt

$$\begin{aligned} \bigcup_{n\geq 1} A_n &= \{\omega \in \Omega : X_n(\omega) = 0 \text{ für } n \geq n_0\} = \{\omega \in \Omega : X_n(\omega) \to 0\}\,, \\ P(A_n) &= P(X_n = 0 \wedge X_{n+1} = 0 \wedge \ldots) = \prod_{i\geq 0} P(X_{n+i} = 0) \\ &= (1 - \frac{1}{n})(1 - \frac{1}{n+1}) \cdots\,, \end{aligned}$$

also

$$\begin{aligned} P(A_n) &= \lim_{M\to\infty} (1 - \frac{1}{n})(1 - \frac{1}{n+1}) \cdots (1 - \frac{1}{M}) \\ &= \lim_{M\to\infty} (\frac{n-1}{n} \cdot \frac{n}{n+1} \cdots \frac{M-1}{M}) = \lim_{M\to\infty} \frac{n-1}{M} = 0\,, \end{aligned}$$

und somit

$$P(\{\omega \in \Omega : X_n(\omega) \to 0\}) = P(\bigcup_{n\geq 1} A_n) = \lim_{n\to\infty} P(A_n) = 0\,.$$

Demnach gilt nicht $X_n \xrightarrow{f.s.} 0$.

b. Es gelte $X_n \xrightarrow{i.W.} X$, mit $F_n(x) = P(X_n \le x)$, $F(x) = P(X \le x)$. Für $\varepsilon > 0$ haben wir

$$\begin{aligned} F_n(x) =\ & P(X_n \le x) = P(X_n \le x \wedge X \le x + \varepsilon) + P(X_n \le x \wedge X > x + \varepsilon) \\ \le\ & F(x+\varepsilon) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Ferner gilt

$$\begin{aligned} F(x-\varepsilon) =\ & P(X \le x - \varepsilon) = P(X \le x - \varepsilon \wedge X_n \le x) + P(X \le x - \varepsilon \wedge X_n > x) \\ \le\ & F_n(x) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Es folgt

$$F(x-\varepsilon) - P(|X_n - x| > \varepsilon) \le F_n(x) \le F(x+\varepsilon) + P(|X_n - x| > \varepsilon).$$

Da $X_n \xrightarrow{i.W.} X$ vorausgesetzt ist, erhalten wir

$$F(x-\varepsilon) \le \liminf F_n(x) \le \limsup F_n(x) \le F(x+\varepsilon).$$

Falls $F(x)$ bei x stetig ist, so gilt mit $\varepsilon \to 0$, $F(x) = \lim\limits_{n\to\infty} F_n(x)$, und es folgt $X_n \xrightarrow{i.V.} X$.

Als Gegenbeispiel für $X_n \xrightarrow{i.V.} X \not\Longrightarrow X_n \xrightarrow{i.W.} X$ betrachten wir Bernoulli Variablen X_n mit Werten in $\{0,1\}$ und

$$P(X_n = 1) = P(X_n = 0) = \frac{1}{2} \quad \text{für alle } n,$$

also $X_1 = X_2 = \cdots = X$. Sei nun $Y = 1 - X$. Wir haben

$$F_n(x) = P(X_n \le x) = \begin{cases} 1 & x \ge 1\,, \\ \frac{1}{2} & 0 \le x < 1\,, \\ 0 & x < 0\,, \end{cases}$$

und ferner $P(Y \le x) = P(X_n \le x)$. Es gilt also $X_n \xrightarrow{i.V.} Y$, aber nicht $X_n \xrightarrow{i.W.} Y$, da $|X_n - Y| = 1$ ist für alle n. □

3.2 Das schwache Gesetz der großen Zahlen

Es seien $X_1, X_2, \ldots$ unabhängige Kopien einer Zufallsvariablen X. Dann legt uns die Intuition nahe, dass der Erwartungswert $\mathrm{E}[X]$ ungefähr gleich $\frac{X_1+\cdots+X_n}{n}$ für große n sein sollte, und das ist auch so.

Für diskrete Zufallsvariable haben wir dies schon im Beispiel nach Satz 2.12 aus der Ungleichung von Tschebyschev begründet. Wir wollen also die Ungleichungen von Markov und Tschebyschev auch für stetige Zufallsvariablen verifizieren.

Satz 3.2. *Sei X stetige Zufallsvariable mit Dichte $f(x)$.*

a. *Ist $X \geq 0$, $a > 0$, so gilt $P(X \geq a) \leq \frac{\mathrm{E}[X]}{a}$. (Markov)*

b. *$P(|X - \mathrm{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}$, $a > 0$. (Tschebyschev)*

Beweis. a. Wir haben

$$\mathrm{E}[X] = \int_0^\infty x f(x)dx = \int_a^\infty x f(x)dx + \underbrace{\int_0^a x f(x)dx}_{\geq 0}$$

$$\geq a \int_a^\infty f(x)dx = aP(X \geq a)\,.$$

b. Mit $Y = |X - \mathrm{E}[X]|$ erhalten wir aus a)

$$P(|X-\mathrm{E}[X]| \geq a) = P(|X-E[X]|^2 \geq a^2) \leq \frac{E[(X-\mathrm{E}[X])^2]}{a^2} = \frac{\mathrm{Var}[X]}{a^2}\,. \qquad \square$$

Satz 3.3 (Schwaches Gesetz der großen Zahlen). *Die reelle Zufallsvariable habe Erwartungswert $\mathrm{E}[X]$ und Varianz $\mathrm{Var}[X]$. Sind $X_1, X_2, \ldots$ unabhängige Kopien von X, so gilt*

$$P(|\frac{X_1+\cdots+X_n}{n} - \mathrm{E}[X]| \geq \varepsilon) \leq \frac{\mathrm{Var}[X]}{n\varepsilon^2} \quad \textit{für alle } \varepsilon > 0\,.$$

Insbesondere gilt für die Folge $\bar{X}_n = \frac{X_1+\cdots+X_n}{n}$, dass $\bar{X}_n \xrightarrow{i.W.} \mathrm{E}[X]$ strebt.

Beweis. Wir haben $E[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \mathrm{E}[X]$, $\mathrm{Var}[\bar{X}_n] = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}[X_i] = \frac{\mathrm{Var}[X]}{n}$, also gilt nach Tschebyschev

$$P(|\bar{X}_n - \mathrm{E}[X]| \geq \varepsilon) \leq \frac{\mathrm{Var}[\bar{X}_n]}{\varepsilon^2} = \frac{\mathrm{Var}[X]}{n\varepsilon^2}. \qquad \square$$

Folgerung 3.4. *Mit den Voraussetzungen des Satzes haben wir*

$$P(|\frac{X_1 + \cdots + X_n}{n} - \mathrm{E}[X]| < t) \geq 1 - \frac{\mathrm{Var}[X]}{nt^2} \quad \textit{für } t > 0.$$

Beispiel. Betrachten wir eine Bernoulli Variable X mit Werten in $\{0,1\}$, $P(X=1) = p$, $P(X=0) = 1-p$, also $\mathrm{E}[X] = p$, $\mathrm{Var}[X] = p(1-p)$. Dann gilt

$$P(|\frac{X_1 + \cdots + X_n}{n} - p| \geq t) \leq \frac{p(1-p)}{nt^2},$$
$$P(|\frac{X_1 + \cdots + X_n}{n} - p| < t) \geq 1 - \frac{p(1-p)}{nt^2}.$$

Die Formeln werden folgendermaßen angewandt.

Es sei p unbekannt. $\bar{X}_n = \frac{X_1+\cdots+X_n}{n}$ gibt die durchschnittliche Anzahl der Erfolge an. Aus $p(1-p) \leq \frac{1}{4}$ für $p \in [0,1]$ folgt

$$P(|\bar{X}_n - p| < t) \geq 1 - \frac{1}{4nt^2}.$$

Die unbekannte Erfolgswahrscheinlichkeit p liegt daher im Intervall $[\bar{x}_n - t, \bar{x}_n + t)]$ mit W-keit $\geq 1 - \frac{1}{4nt^2}$, wobei $\bar{x}_n = \frac{x_1+\cdots+x_n}{n}$ der Durchschnitt der tatsächlich beobachteten Ausgänge ist.

Angenommen wir möchten p bis auf einen Fehler $t = 0{,}01$ mit W-keit 98% bestimmen. Wie oft müssen wir das Experiment durchführen? Aus der Formel sehen wir

$$1 - \frac{1}{4n \cdot 0{,}0001} \geq 0{,}98,$$

und dies ist für $n \geq 125.000$ erfüllt.

Man beachte, dass $1 - \frac{1}{4nt^2}$ stets < 1 ist. Mit Sicherheit kann p also nicht ermittelt werden.

Beispiel. Ein Würfel wird 20 Mal geworfen und zeigt immer 6. Ist der Würfel in Ordnung, das heißt, sind die Zahlen wirklich gleichverteilt?

Es sei X_i die Zufallsvariable mit $X_i = 1$, falls im i-ten Wurf 6 erscheint, $X_i = 0$, falls eine Zahl $\neq 6$ erscheint. Dann ist für $p = \frac{1}{6}$

$$P\left(\left|\underbrace{\frac{X_1 + \cdots + X_{20}}{20}}_{=1} - \frac{1}{6}\right| \geq \frac{5}{6}\right) \leq \frac{5}{36}\frac{1}{20(\frac{5}{6})^2} = \frac{1}{100},$$

also ist der Würfel mit W-keit$\geq 99\%$ nicht in Ordnung.

3.3 Lemma von Borel-Cantelli

Sei ein W-Raum $(\Omega, \mathcal{E}, P)$ gegeben und $A_1, A_2, \ldots$ eine Folge von Ereignissen. Wir definieren

$$A^\infty = \{\omega \in \Omega : \omega \in A_i \text{ für unendlich viele } i\}.$$

Man sieht sofort

$$A^\infty = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k = \bigcap_{n=1}^{\infty} B_n \text{ mit } B_n = \bigcup_{k \geq n} A_k.$$

A^∞ ist also in $\mathcal{E}$. Aus $B_1 \supset B_2 \supset B_3 \supset \ldots$ folgt

$$P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n \to \infty} P(B_n) = P(A^\infty).$$

Beispiel. Wir werfen einen Würfel unendlich oft, und A_n sei das Ereignis, dass im n-ten Wurf 6 geworfen wird. Es sei $\Omega = \{(\omega_1, \omega_2, \ldots)\}$ mit $\omega_i \in \{1, 2, \ldots, 6\}$, also $A^\infty = \{(\omega_1, \omega_2, \ldots) : \omega_i = 6 \text{ für unendlich viele } i\}$.

Satz 3.5 (Lemma von Borel-Cantelli). *Es sei $(\Omega, \mathcal{E}, P)$ gegeben und $A_1, A_2, \ldots$ eine Folge von Ereignissen.*

a. *Ist $\sum_{k \geq 1} P(A_k) < \infty$, so gilt $P(A^\infty) = 0$.*

b. *Sind $A_1, A_2, \ldots$ unabhängig und $\sum_{k \geq 1} P(A_k) = \infty$, so gilt $P(A^\infty) = 1$.*

Beweis. a. Wegen $\sum\limits_{k\geq 1} P(A_k) < \infty$ haben wir $\sum\limits_{k\geq n} P(A_k) \to 0$ für $n \to \infty$. Daraus folgt wegen $P(B_n) \leq \sum\limits_{k\geq n} P(A_k)$, dass $P(B_n) \stackrel{n\to\infty}{\longrightarrow} 0$ geht, also $P(A^\infty) = \lim P(B_n) = 0$.

b. Wir verwenden die Ungleichung $1 - x \leq e^{-x}$ für $0 \leq x \leq 1$. Da die Ereignisse A_n unabhängig sind, so auch die komplementären Ereignisse A_n^c (Lemma 1.5), und wir erhalten

$$P\left(\bigcap_{k=n}^{N} A_k^c\right) = \prod_{k=n}^{N}(1 - P(A_k)) \leq \prod_{k=n}^{N} e^{-P(A_k)} = e^{-\sum\limits_{k=n}^{N} P(A_k)}.$$

Mit $N \to \infty$ ergibt sich $P\left(\bigcap\limits_{k=n}^{N} A_k^c\right) \to 0$, da $\sum\limits_{k\geq 1} P(A_k) = \infty$ ist. Nun ist $\bigcap\limits_{k=n}^{N} A_k^c \supset \bigcap\limits_{k=n}^{N+1} A_k^c \supset \ldots$, also

$$P\left(\bigcap_{k=n}^{\infty} A^c\right) = \lim_{N\to\infty} P\left(\bigcap_{k=n}^{N} A_k^c\right) = 0$$

für alle n, und es folgt

$$P((A^\infty)^c) = P\left(\bigcup_{n=1}^{\infty}\bigcap_{k\geq n} A_k^c\right) \leq \sum_{n=1}^{\infty} P\left(\bigcap_{k\geq n} A_k^c\right) = 0\,,$$

und somit $P(A^\infty) = 1$. □

Bemerkung. In Teil b) kann auf die Unabhängigkeit nicht verzichtet werden. Sei nämlich $A \in \mathcal{E}$, $0 < P(A) < 1$ und $A_n = A$ für alle n. Dann ist $A^\infty = A$, $\sum P(A_n) = \infty$, aber $P(A^\infty) < 1$.

Beispiele. Wir verwenden die folgende Sprechweise: Das Ereignis A tritt *fast sicher* ein, falls $P(A) = 1$ ist.

1. Betrachten wir das Würfelbeispiel von eben, $A_n = \{6 \text{ im } n\text{-ten Wurf}\}$. Die Würfe seien unabhängig, gleichverteilt, dann gilt $P(A_n) = \frac{1}{6}$, also $\sum\limits_{n\geq 1} P(A_n) = \infty$. Die 6 wird somit fast sicher unendlich oft geworfen.

2. Wir haben Urnen $U_1, U_2, \ldots$ gegeben, wobei sich in U_n 1 weiße und $n-1$ rote Kugeln befinden. Das Ereignis A_n sei, dass aus Urne U_n eine weiße Kugel gezogen wird. Wir haben $P(A_n) = \frac{1}{n}$, also $\sum\limits_{n\geq 1} P(A_n) = \infty$ und somit $P(A^\infty) = 1$. Das heißt, fast sicher wird weiß unendlich oft gezogen.

Angenommen in Urne U_n sind 1 weiße und $n^2 - 1$ rote Kugeln. Dann ist $P(A_n) = \frac{1}{n^2}$, $\sum\limits_{n\geq 1} P(A_n) = \sum\limits_{n\geq 1} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$, also wird weiß fast sicher nur endlich oft gezogen.

Das interessante ist, dass keine Verteilung weiß/rot auf die Urnen möglich ist, wo nicht einer dieser beiden Fälle eintritt.

3. Es seien unabhängige Ereignisse $A_1, A_2, \ldots, A_r$ gegeben mit $P(A_i) > 0$ für alle i. Wir betrachten $\Omega_\infty = (\omega_1, \omega_2, \ldots)$. Kommt es oft vor, dass es ein k gibt, mit $\omega_k \in A_1$, $\omega_{k+1} \in A_2, \ldots, \omega_{k+r-1} \in A_r$?

Nehmen wir eine Telefonnummer 8018418. Nun schreiben wir die Ziffern $0, 1, 2, \ldots, 9$ gleichverteilt in einer unendlichen Folge auf. Kommt die Telefonnummer oft in der Folge vor? Dazu definieren wir

$$\begin{aligned}
B_1 &= \{(\omega_n) : \omega_1 \in A_1, \ldots, \omega_r \in A_r\} \\
B_2 &= \{(\omega_n) : \omega_{r+1} \in A_1, \ldots, \omega_{2r} \in A_r\} \\
&\vdots \\
B_m &= \{(\omega_n) : \omega_{(m-1)r+1} \in A_1, \ldots, \omega_{mr} \in A_r\} \\
&\vdots
\end{aligned}$$

Die Ereignisse B_m sind unabhängig mit $P(B_m) = P(A_1) \cdots P(A_r) > 0$ für alle m. Aus $\sum_{m\geq 1} P(B_m) = \infty$ folgt $P(B^\infty) = 1$, und das bedeutet, dass es fast sicher unendlich viele B_m gibt, in denen r aufeinanderfolgende Ergebnisse in $A_1, A_2, \ldots, A_r$ sind. Die Telefonnummer kommt also fast sicher unendlich oft vor.

Eine amüsante Illustration ist der Affe am Computer. Er tippt Buchstaben nach Belieben ein. Nach Borel-Cantelli wird er fast sicher irgendwann Goethes Faust reproduzieren – und das nicht einmal, sondern unendlich oft.

3.4 Das starke Gesetz der großen Zahlen

Es seien wieder $X_1, X_2, \dots$ unabhängige Kopien der reellen Zufallsvariablen X. Wir wollen nun zeigen, dass die Folge $\bar{X}_n = \frac{X_1+\cdots+X_n}{n}$ sogar fast sicher gegen den Erwartungswert konvergiert. In diesem Sinn können wir also sagen, dass der Zufall bei unendlicher Wiederholung verschwindet.

Satz 3.6 (Starkes Gesetz der großen Zahlen). *Für die reelle Zufallsvariable X existieren* $\mathrm{E}[X]$ *und* $\mathrm{Var}[X]$. *Es seien $X_1, X_2, \dots$ unabhängige Kopien von X. Dann gilt*

$$P(\{\omega \in \Omega : \bar{X}_n(\omega) \to \mathrm{E}[X]\}) = 1\,,$$

das heißt $\bar{X}_n \xrightarrow{f.s.} \mathrm{E}[X]$.

Beweis. Wir können wieder $\mathrm{E}[X] = 0$ annehmen, ansonsten verwenden wir die Folge $X_n - E(X)$. Wir gliedern den Beweis in mehrere Schritte.

1. Sei $Y_1, Y_2, \dots$ eine Folge von Zufallsvariablen auf Ω. Wir nennen $\omega \in \Omega$ *gut*, falls $Y_n(\omega) \to 0$. Sei $B_Y = \{\omega \in \Omega : \omega \text{ gut}\}$. Wann gilt $P(B_Y) = 1$? Dazu stellen wir fest:

$$\begin{aligned}
Y_n(\omega) \to 0 &\iff \forall \varepsilon > 0 \text{ gilt } |Y_n(\omega)| < \varepsilon \text{ für } n \geq n_0 \\
&\iff \forall \varepsilon \geq 0 \text{ gilt } |Y_n(\omega)| \geq \varepsilon \text{ für } \textit{endlich} \text{ viele } n \\
&\iff \forall k \in \mathbb{N} \text{ gilt } |Y_n(\omega)| \geq \frac{1}{k} \text{ für } \textit{endlich} \text{ viele } n\,.
\end{aligned}$$

Sei $E_k = \{\omega \in \Omega : |Y_n(\omega)| \geq \frac{1}{k} \text{ für unendlich viele } n\}$, k fest. Dann ist $E_1 \subseteq E_2 \subseteq E_3 \subseteq \dots$, und ferner $B_Y = \bigcap_{k \geq 1} E_k^c$. Somit schließen wir

$$\begin{aligned}
P(B_Y) = 1 &\iff P\left(\bigcap_{k\geq 1} E_k^c\right) = 1 \\
&\iff P\left((\bigcup_{k\geq 1} E_k)^c\right) = 1 \\
&\iff P\left(\bigcup_{k\geq 1} E_k\right) = 0\,.
\end{aligned}$$

Insbesondere gilt nach Lemma 1.1

$$P(E_k) = 0 \ (\forall k) \Longrightarrow P(B_Y) = 1 .$$

2. Sei $k \geq 1$ fest, $A_{n,k} = \{\omega \in \Omega : |Y_n(\omega)| \geq \frac{1}{k}\}$. Wir erhalten also eine Folge von Ereignissen $A_{1,k}, A_{2,k}, \ldots$ und setzen

$$A_k^\infty = \{\omega \in \Omega : \omega \text{ ist in unendlich vielen } A_{n,k}\} = E_k$$

aus 1). Aus $P(A_k^\infty) = 0$ $(\forall k)$ folgt also $P(B_Y) = 1$. Laut dem Lemma von Borel-Cantelli können wir festhalten

$$\sum_{n\geq 1} P(A_{n,k}) < \infty \ (\forall k) \Longrightarrow P(B_Y) = 1 . \tag{1}$$

3. Sei wieder $k \geq 1$ fest. Wir setzen $Y_n = \bar{X}_n$, dann ist das starke Gesetz genau die Aussage $P(B_Y) = 1$. Wenn wir also zeigen können, dass $\sum_{n\geq 1} P(A_{n,k}) < \infty$ ist für alle k, so sind wir fertig. Nun verwenden wir die Ungleichung von Tschebyschev. Wir haben $A_{n,k} = \{\omega \in \Omega : |\bar{X}_n(\omega)| \geq \frac{1}{k}\}$, und nach Tschebyschev gilt mit $\mathrm{E}[\bar{X}_n] = 0$, $\mathrm{Var}[\bar{X}_n] = \frac{\mathrm{Var}[X]}{n}$

$$P(A_{n.k}) = P(|\bar{X}_n(\omega)| \geq \frac{1}{k}) \leq \frac{\mathrm{Var}[X]}{n} k^2 ,$$

und somit

$$\sum_{n\geq 1} P(A_{n,k}) \leq \mathrm{Var}[X] k^2 \sum_{n\geq 1} \frac{1}{n} .$$

Leider divergiert die harmonische Reihe, also müssen wir die Summe $\sum\limits_{n\geq 1} P(A_{n,k})$ besser abschätzen.

Wir betrachten die Teilfolge $\bar{X}_{n^2}$, $n = 1, 2, \ldots$. Für diese gilt nach Tschebyschev

$$\sum_{n=1}^{\infty} P(A_{n^2,k}) = \sum_{n=1}^{\infty} P(|\bar{X}_{n^2}| \geq \frac{1}{k}) \leq \mathrm{Var}[X] k^2 \sum \frac{1}{n^2} < \infty ,$$

und somit nach (1)

$$P(\{\omega \in \Omega : \bar{X}_{n^2}(\omega) \to 0\}) = 1 . \tag{2}$$

4. Sei $S_\ell = X_1 + \cdots + X_\ell$, und für $m \in \mathbb{N}$ sei $n = n(m)$ mit

$$n^2 \leq m < (n+1)^2 .$$

Wir haben

$$E[S_m - S_{n^2}] = 0, \ \mathrm{Var}[S_m - S_{n^2}] = (m - n^2)\mathrm{Var}[X] .$$

Sei k fest, dann gilt wieder mit Tschebyschev

$$\begin{aligned} &P(\{\omega \in \Omega : |S_m(\omega) - S_{n^2}(\omega)| \geq \frac{n^2}{k}\}) = \\ &P(|S_m - S_{n^2}| \geq \frac{n^2}{k}) \leq \frac{(m-n^2)k^2}{n^4}\mathrm{Var}[X] . \end{aligned}$$

Nun definieren wir die Ereignisse

$$B_{m,k} = \{\omega \in \Omega : \frac{|S_m(\omega) - S_{n^2}(\omega)|}{n^2} \geq \frac{1}{k}\} , \ n = n(m) ,= \lfloor\sqrt{m}\rfloor$$

und haben

$$P(B_{m,k}) \leq \frac{m-n^2}{n^4} k^2 \mathrm{Var}[X] , \ m \geq 1 , \ n = n(m) .$$

Summation über m ergibt

$$\begin{aligned} \sum_{m \geq 1} P(B_{m,k}) &\leq k^2 \mathrm{Var}[X] \sum_{m=1}^{\infty} \frac{m - n(m)^2}{n(m)^4} \\ &= k^2 \mathrm{Var}[X] \sum_{n=1}^{\infty} \sum_{m=n^2}^{(n+1)^2-1} \frac{m-n^2}{n^4} \\ &= k^2 \mathrm{Var}[X] \sum_{n=1}^{\infty} \frac{1+2+\cdots+2n}{n^4} \\ &= k^2 \mathrm{Var}[X] \sum_{n=1}^{\infty} \frac{2n(2n+1)}{2n^4} \\ &= k^2 \mathrm{Var}[X] \sum_{n=1}^{\infty} \frac{2n+1}{n^3} < \infty . \end{aligned}$$

Mit $Y_m = \frac{S_m - S_{n(m)^2}}{n(m)^2}$ haben wir also nach (1)

$$P(\{\omega \in \Omega : \frac{S_m(\omega) - S_{n(m)^2}(\omega)}{n(m)^2} \overset{m\to\infty}{\longrightarrow} 0\}) = 1\,. \tag{3}$$

5. Nun setzen wir die Ergebnisse zusammen. Es sei

$$\begin{aligned} A &= \{\omega : \bar{X}_{n^2}(\omega) \to 0\} \text{ mit } P(A) = 1 \text{ nach (2)} \\ B &= \{\omega : \frac{S_m(\omega) - S_{n(m)^2}(\omega)}{n(m)^2} \to 0\} \text{ mit } P(B) = 1 \text{ nach (3)}\,, \end{aligned}$$

also gilt $P(A\cap B) = P(A)+P(B)-P(A\cup B) \geq 2-1 = 1$, somit $P(A\cap B) = 1$. Für ein $\omega \in A\cap B$ gilt $\bar{X}_{n^2}(\omega) \overset{n\to\infty}{\longrightarrow} 0$, also auch für die Teilfolge $\bar{X}_{n(m)^2}(\omega) \overset{m\to\infty}{\longrightarrow} 0$. Ferner ist

$$\frac{S_m}{n(m)^2} = \frac{S_m - S_{n(m)^2}}{n(m)^2} + \frac{S_{n(m)^2}}{n(m)^2} = \frac{S_m - S_{n(m)^2}}{n(m)^2} + \bar{X}_{n(m)^2}\,.$$

Da $\omega \in B$ ist, gilt $\frac{S_m - S_{n(m)^2}}{n(m)^2}(\omega) \longrightarrow 0$ mit $m \to \infty$, und wir erhalten als Ergebnis:

$$\omega \in A \cap B \Longrightarrow \frac{S_m(\omega)}{n(m)^2} \overset{m\to\infty}{\longrightarrow} 0\,.$$

Wegen $m \geq n(m)^2$ gilt ferner für $\omega \in A \cap B$

$$|\bar{X}_m(\omega)| = |\frac{S_m(\omega)}{m}| \leq |\frac{S_m(\omega)}{n(m)^2}| \overset{m\to\infty}{\longrightarrow} 0\,,$$

also

$$\omega \in A \cap B \Longrightarrow \bar{X}_m(\omega) \overset{m\to\infty}{\longrightarrow} 0$$

oder

$$A \cap B \subseteq \{\omega : \bar{X}_m(\omega) \overset{m\to\infty}{\longrightarrow} 0\}\,.$$

Da wie gesehen $P(A \cap B) = 1$ ist, erhalten wir daraus

$$P(\{\omega \in \Omega : \bar{X}_m(\omega) \overset{m\to\infty}{\longrightarrow} 0\}) = 1\,,$$

und der Beweis ist fertig. $\square$

Beispiel. Betrachten wir das Intervall $\Omega = (0,1] \subseteq \mathbb{R}$. Eine Zahl $\omega = 0,y_1y_2y_3\ldots$ heißt *normal*, falls jeder mögliche Block $a_1a_2\ldots a_k \in \{0,1,\ldots,9\}^k$

der Länge k in der Dezimalentwicklung im Limes mit relativer Häufigkeit 10^{-k} vorkommt, und zwar für alle $k \geq 1$. Also, wenn

$$\lim_{n\to\infty} \frac{\#\{a_1 \dots a_k \text{ in } y_1 \dots y_n\}}{n} = \frac{1}{10^k},$$

gilt. Gibt es überhaupt normale Zahlen? Eine besonders schöne Anwendung des starken Gesetzes der großen Zahlen zeigt, dass bei Gleichverteilung auf Ω gilt: $P(\{\omega \in \Omega : \omega \text{ normal}\}) = 1$. Es gibt also überabzählbar viele normale Zahlen.

Wir beweisen die Aussage für $k = 1$, der allgemeine Fall ist nur wenig schwieriger.

Es sei $\omega = 0, y_1(\omega)y_2(\omega)\dots$ die Dezimalentwicklung und $Y_i(\omega) = y_i(\omega)$ die Zufallsvariable (unter Gleichverteilung), $i \geq 1$. wir haben

$$P(Y_1 = b_1 \wedge \dots \wedge Y_m = b_m) = \frac{1}{10^m},$$

da $\{\omega : y_1(\omega) = b_1 \wedge \dots \wedge y_m(\omega) = b_m\}$ ein Intervall der Länge $\frac{1}{10^m}$ ist. Ebenso gilt $P(Y_i = b_i) = \frac{1}{10}$ für alle i, also sind die Variablen Y_i unabhängig.

Nun sei $a \in \{0, 1, \dots, 9\}$ fest vorgegeben, und $X_1, X_2, \dots$ die Folge von Bernoulli Variablen mit

$$X_i(\omega) = \begin{cases} 1 & \text{falls } y_i(\omega) = a \\ 0 & \text{sonst}. \end{cases}$$

Wie eben gesehen ist $P(X_i = 1) = \frac{1}{10}$ für alle i, die Variablen X_i sind somit unabhängige Kopien der Variablen X mit $E[X] = \frac{1}{10}$. Für $\omega \in \Omega$ ist

$$\overline{X}_n(\omega) = \frac{\#\{i : y_i = a, 1 \leq i \leq n\}}{n} = \text{relative Häufigkeit von } a,$$

und das starke Gesetz besagt

$$P(\{\omega : \text{ relative Hüfigkeit von } a \to \frac{1}{10}\}) = 1$$

für alle $a \in \{0, 1, \dots, 9\}$.

3.5 Der zentrale Grenzwertsatz

Der folgende Satz ist der berühmteste Satz der W-Theorie und zeigt die in der Praxis beobachtete zentrale Stellung der Gaußschen Glockenkurve.

Satz 3.7 (Zentraler Grenzwertsatz). *Seien $X_1, X_2, \ldots$ unabhängige Kopien von X mit Erwartungswert $E[X]$ und Varianz* $\mathrm{Var}[X]$. *Dann gilt*

$$\frac{X_1 + \cdots + X_n - n\mathrm{E}[X]}{\sqrt{n}\sqrt{\mathrm{Var}[X]}} \xrightarrow{i.V.} Y \sim N(0,1),$$

das heißt Y ist nach der Standardnormalverteilung verteilt.

Beispiele. 1. Sei X gleichverteilt auf $[0,1]$, $\mathrm{E}[X] = \frac{1}{2}$, $\mathrm{Var}[X] = \frac{1}{12}$. Dann gilt

$$\frac{X_1 + \cdots + X_n - \frac{n}{2}}{\sqrt{n}\sqrt{\frac{1}{12}}} \xrightarrow{i.V.} N(0,1).$$

2. Sei X Bernoulli Variable mit $P(X = 1) = p$, $P(X = 0) = q = 1 - p$, $0 < p < 1$. Dann ist $E[X] = p$, $\mathrm{Var}[X] = pq$, und es folgt

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{npq}} \xrightarrow{i.V.} N(0,1). \tag{1}$$

Damit können wir die Standardnormalverteilung simulieren. Wir werfen eine Münze n Mal, $p = \frac{1}{2}$, dann ist

$$\frac{\text{Anzahl Kopf} - \frac{n}{2}}{\frac{\sqrt{n}}{2}} \xrightarrow{i.V.} N(0,1).$$

Wir werden den zentralen Grenzwertsatz für Bernoulli Variablen beweisen, also die Aussage (1). Dieser Spezialfall heißt Satz von de Moivre-Laplace. Dazu benötigen wir einige Vorbereitungen.

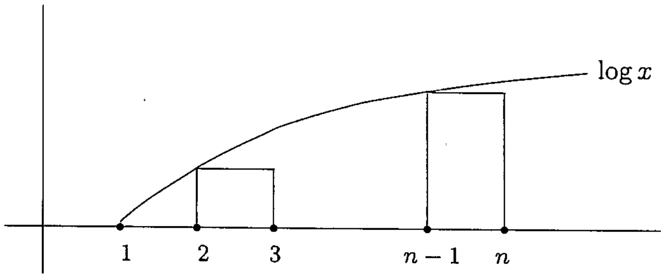
Lemma 3.8 (Stirling Formel). *Wir haben*

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

das heißt

$$\lim_{n\to\infty} \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!} = 1.$$

Wir können uns die Formel plausibel machen. Betrachten wir die Logarithmuskurve:



Durch Ober- und Untersummenbildung erhalten wir

$$\log(n-1)! < \int_1^n \log x dx < \log(n!)$$

und wegen $\int_1^n \log x dx = x \log x - x$

$$\log(n-1)! < n \log n - n + 1 < \log(n!)$$
$$(n-1)! < e(\frac{n}{e})^n < n!\,.$$

Für $n, k,\ n-k$ groß können wir also abschätzen:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} (\frac{n}{k})^k (\frac{n}{n-k})^{n-k}. \qquad (2)$$

Als nächsten wollen wir den Koeffizienten $b(k, n; p)$ abschätzen. Wir gehen in mehreren Schritten vor.

1. Für $b(k, n; p)$ schreiben wir $b_{n,p}(k) = \binom{n}{k} p^k q^{n-k}$, $q = 1-p$. Nun betrachten wir Folgen (k_n) mit $\frac{k_n}{n} \longrightarrow p$. Es muß also auch $k_n \longrightarrow \infty$ gelten, $\frac{n-k_n}{n} = 1 - \frac{k_n}{n} \longrightarrow 1 - p = q$, also auch $n - k_n \longrightarrow \infty$.

Sei $k = k_n$, dann haben wir mit (2)

$$b_{n,p}(k) \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{k(n-k)}} (\frac{pn}{k})^k (\frac{qn}{n-k})^{n-k}, \qquad (3)$$

und mit $\sigma_n = \sqrt{npq}$, $\frac{k}{n} \sim p$, $\frac{n-k}{n} \sim q$

$$\sqrt{\frac{n}{k(n-k)}} \sim \sqrt{\frac{1}{npq}} = \frac{1}{\sigma_n}. \tag{4}$$

2. Als nächstes untersuchen wir das Grenzwertverhalten von

$$f(n,k) = (\frac{pn}{k})^k (\frac{qn}{n-k})^{n-k}.$$

Sei $t_n = \frac{k_n}{n}$, also $t_n \to p$, $\frac{n-k_n}{n} = 1 - \frac{k_n}{n} = 1 - t_n \to q$. Wir setzen kurz $t = \frac{k}{n}$. Dann gilt

$$\begin{aligned} \log f(n,k) &= k \log \frac{p}{t} + (n-k) \log \frac{q}{1-t}, \\ -\log f(n,k) &= k \log \frac{t}{p} + (n-k) \log \frac{1-t}{q} \\ &= n\underbrace{(t \log \frac{t}{p} + (1-t) \log \frac{1-t}{q})}_{g(t)}. \end{aligned}$$

Für die Funktion $g(t)$ sehen wir:
a. $g(p) = 0$,
b. $\begin{aligned} g'(t) &= \log \tfrac{t}{p} + 1 - \log \tfrac{1-t}{q} + (1-t)\tfrac{q}{1-t}(-\tfrac{1}{q}) \\ &= \log \tfrac{t}{p} - \log \tfrac{1-t}{q}, \end{aligned}$
also insbesondere $g'(p) = 0$.
c. $g''(t) = \frac{1}{t} + \frac{1}{1-t}$, $g''(p) = \frac{1}{p} + \frac{1}{q} = \frac{1}{pq}$.
Taylorentwicklung in p ergibt

$$g(t) = \frac{1}{2pq}(t-p)^2 + h(t-p) \text{ mit } |h(t-p)| \le c|t-p|^3.$$

3. Angenommen, es gilt sogar $n(t-p)^3 \to 0$. Dann ist $nh(t-p) \to 0$, und somit

$$|-\log f(n,k) - \frac{n(t-p)^2}{2pq}| = |ng(t) - n\frac{(t-p)^2}{2pq}| = |nh(t-p)| \to 0.$$

Wir setzen nun

$$x(n,k) = \frac{k-np}{\sigma_n}, \tag{5}$$

dann haben wir

$$\frac{n(t-p)^2}{2pq} = \frac{n(\frac{k}{n}-p)^2}{2pq} = \frac{(k-np)^2}{2npq} = \frac{x(n,k)^2}{2},$$

und es folgt

$$|-\log f(n,k) - \frac{x(n,k)^2}{2}| \longrightarrow 0.$$

Daraus folgt

$$e^{-\log f(n,k) - \frac{x(n,k)^2}{2}} \longrightarrow 1,$$

$$\frac{e^{-\frac{x(n,k)^2}{2}}}{f(n,k)} \longrightarrow 1,$$

also

$$f(n,k) \sim e^{-\frac{x(n,k)^2}{2}}. \tag{6}$$

4. Was bedeutet $n(t-p)^3 \to 0$ für die Zahlen $x(n,k)$? Wir haben

$$n(t-p)^3 = n(\frac{k}{n}-p)^3 = \frac{(k-np)^3}{n^2} = \frac{x(n,k)^3\sigma_n^3}{n^2}$$
$$= \frac{x(n,k)^3(pq)^{\frac{3}{2}}}{\sqrt{n}},$$

und sehen

$$n(t-p)^3 \to 0 \Longleftrightarrow \frac{x(n,k_n)^3}{\sqrt{n}} \to 0.$$

5. Fassen wir zusammen: Sei $x(n,k) = \frac{k_n - np}{\sigma_n}$. Falls $\frac{x(n,k_n)^3}{\sqrt{n}} \to 0$ geht, so ist nach (3), (4) und (6)

$$b_{n,p}(k_n) \sim \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_n}e^{-\frac{x(n,k)^2}{2}},$$

und wir erhalten das folgende Resultat.

Lemma 3.9. *Sind* (α_n), (β_n) *Folgen mit* $\frac{x(n,\alpha_n)^3}{\sqrt{n}}, \frac{x(n,\beta_n)^3}{\sqrt{n}} \to 0$, *so ist*

$$b_{n,p}(k_n) \sim \frac{e^{-\frac{x(n,k_n)^2}{2}}}{\sqrt{2\pi}\sigma_n} \quad \textit{für } \alpha_n \le k_n \le \beta_n.$$

Satz 3.10 (Satz von de Moivre Laplace). *Es sei X eine Bernoulli Variable mit $0 < p < 1$, und $X_1, X_2, \dots$, unabhängige Kopien von X. Dann gilt für die zentrierte Variable*

$$S_n^* = \frac{X_1 + \cdots + X_n - np}{\sqrt{npq}}, \quad q = 1 - p,$$

$$S_n^* \xrightarrow{i.V.} N(0,1),$$

das heißt

$$\lim_{n\to\infty} P(a \le S_n^* \le b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad (a, b \in \mathbb{R}).$$

Beweis. Wir teilen den Beweis wieder in mehrere Schritte ein.
1. Sei $S_n = X_1 + \cdots + X_n$. Dann gilt mit $\sigma_n = \sqrt{npq}$

$$a \le S_n^* \le b \Longleftrightarrow a \le \frac{S_n - np}{\sigma_n} \le b \Longleftrightarrow a\sigma_n + np \le S_n \le b\sigma_n + np.$$

Sei $\alpha_n \in \mathbb{N}$ minimal mit $\alpha_n \ge a\sigma_n + np$ und $\beta_n \in \mathbb{N}$ maximal mit $\beta_n \le b\sigma_n + np$. Dann gilt

$$a \le S_n^* \le b \Longleftrightarrow \alpha_n \le S_n \le \beta_n,$$

somit

$$P(a \le S_n^* \le b) = P(\alpha_n \le S_n \le \beta_n) = \sum_{k=\alpha_n}^{\beta_n} b_{n,p}(k).$$

Wir müssen also zeigen:

$$\lim_{n\to\infty} \sum_{k=\alpha_n}^{\beta_n} b_{n,p}(k) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx. \tag{7}$$

2. Aus $\beta_n \le b\sigma_n + np < \beta_n + 1$ folgt

$$b\sigma_n + np - 1 < \beta_n \le b\sigma_n + np,$$

also gilt $\frac{\beta_n}{n} \to p$ und analog $\frac{\alpha_n}{n} \to p$. Mit $x(n,\beta_n) = \frac{\beta_n - np}{\sigma_n}$ haben wir

$$b - \frac{1}{\sigma_n} < x(n,\beta_n) \leq b$$

und analog

$$a \leq x(n,\alpha_n) < a + \frac{1}{\sigma_n}\,.$$

Die Folgen $(x(n,\beta_n))$ und $(x(n,\alpha_n))$ sind somit beschränkt und es folgt

$$\frac{x(n,\beta_n)^3}{\sqrt{n}} \to 0, \ \frac{x(n,\alpha_n)^3}{\sqrt{n}} \to 0\,.$$

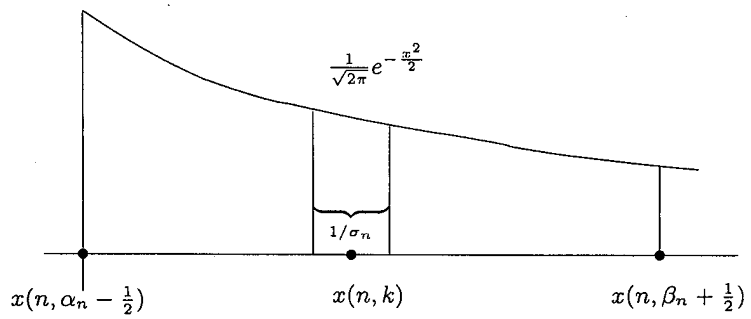
Aus Lemma 3.9 folgt daher

$$b_{n,p}(k) \sim \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{x(n,k)^2}{2}} \text{ für alle } \alpha_n \leq k \leq \beta_n\,.$$

und somit

$$\sum_{k=\alpha_n}^{\beta_n} b_{n,p}(k) \sim \frac{1}{\sqrt{2\pi}\sigma_n} \sum_{k=\alpha_n}^{\beta_n} e^{-\frac{x(n,k)^2}{2}}\,. \tag{8}$$

3. Nun sehen wir uns die rechte Summe in (8) näher an.



Die rechte Summe ist eine Riemannsche Summe mit Intervallen der Länge $\frac{1}{\sigma_n}$, und wir haben

$$x(n,\alpha_n - \frac{1}{2}) = \frac{\alpha_n - \frac{1}{2} - np}{\sigma_n} \to a, \ x(n,\beta_n + \frac{1}{2}) = \frac{\beta_n + \frac{1}{2} - np}{\sigma_n} \to b\,.$$

Es folgt

$$\sum_{k=\alpha_n}^{\beta_n} b_{n,p}(k) \overset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \,,$$

und der Satz ist bewiesen. □

Folgerung 3.11. *Seien $X_1, X_2, \ldots$ unabhängige Kopien der Bernoulli Variablen X mit $0 < p < 1$, $\sigma_n = \sqrt{npq}$. Dann gilt für $S_n = X_1 + \cdots + X_n$ (Anzahl der Erfolge)*

$$\lim_{n\to\infty} P(\alpha \le S_n \le \beta) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-np}{\sigma_n}}^{\frac{\beta-np}{\sigma_n}} e^{-\frac{x^2}{2}} dx = \phi(\frac{\beta - np}{\sigma_n}) - \phi(\frac{\alpha - np}{\sigma_n}) \,,$$

mit $\phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$.

Das letzte Resultat wird in der Praxis zur Abschätzung der Erfolgswahrscheinlichkeit verwendet.

Beispiel. Ein Würfel mit Gleichverteilung wird 600 Mal geworfen. Wie groß ist die W-keit, zwischen 90 und 100 6en zu erhalten? Hier ist $n = 600$, $p = \frac{1}{6}$, $np = 100$, $\sigma_n = \sqrt{600 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 9,13$, $\alpha - np = -10$, $\beta - np = 0$, also

$$P(90 \le S_n \le 100) \sim \phi(0) - \phi(-\frac{10}{9,13}) \sim 0,36 \,.$$