Wahrscheinlichkeitstheorie und Statistik

9. Woche

Dávid Tóth (BME SZIT)

28., 31. Oktober 2024

Gesetz der großen Zahlen

Beispiel. Nehmen wir an, dass n fairen Würfel unabhängig voneinander geworfen werden, wo n sehr groß ist. Wir bezeichnen mit $X_1, \dots, X_n$ die Augenzahlen der einzelnen Würfel. Wie groß ist ungefähr die Augensumme $X_1 + \cdots + X_n$ aller n Würfel?

Intuitiv geht man davon aus, dass ungefähr $n/6$ Würfel eine 1 zeigen, ungefähr $n/6$ Würfel eine 2 zeigen, usw. Für die Augensumme ergibt sich die Approximation

$$\frac{n}{6} \cdot 1 + \frac{n}{6} \cdot 2 + \frac{n}{6} \cdot 3 + \frac{n}{6} \cdot 4 + \frac{n}{6} \cdot 5 + \frac{n}{6} \cdot 6 = 3{,}5 \cdot n.$$

Gesetz der großen Zahlen

Bemerkungen.

- Die Begriffe "ungefähr" und "groß" sind keine mathematischen Begriffe.
- Die obige Aussage sollte als ein Grenzwert für $n \to \infty$ formuliert werden, nämlich

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = 3{,}5.$$

- Die auf der rechten Seite aufgetauchte Zahl 3,5 kann als der Erwartungswert für die Augenzahl eines Würfels interpretiert werden.

Gesetz der großen Zahlen

Beispiel. Wir betrachten n unabhängige Experimente mit Erfolgswahrscheinlichkeit p. Sei

$$X_i = \begin{cases} 1, & \text{falls Experiment } i \text{ ein Erfolg ist,} \\ 0, & \text{sonst.} \end{cases}$$

Also ist X_i die Indikatorvariable für den Erfolg beim Experiment i.

Die Anzahl der Erfolge in allen n Experimenten ist dann die Zufallsvariable $X_1 + \cdots + X_n$. Wenn n "groß" ist, dann sollte die Anzahl der Erfolge intuitiv "ungefähr" np betragen. Es sollte also gelten, dass

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = p = \mathbb{P}(\text{"Erfolg"}) = \mathbb{E}(X_i).$$

Gesetz der großen Zahlen

- Diese Formeln sind Spezialfälle eines allgemeinen Naturphänomens, das als Gesetz der großen Zahlen bezeichnet wird.
- Um die obigen Beispiele zu verallgemeinern, betrachten wir eine unendliche Folge unabhängiger Zufallsvariablen $X_1, \ldots, X_n, \ldots$, die die gleiche (und ansonsten beliebige) Verteilung besitzen sollen.
- Dann kann man vermuten, dass

$$\lim_{n \to \infty} \frac{X_1 + \cdots + X_n}{n} = \mathbb{E}(X_1),$$

falls der Erwartungswert auf der rechten Seite existiert.

Gesetz der großen Zahlen

Was bedeutet es, dass eine Folge von Zufallsvariablen konvergiert?

Für Funktionen gibt es viele verschiedene (nicht-äquivalente) Konvergenzbegriffe, wir betrachten 3 wichtige Konvergenzarten für Zufallsvariablen:

- *Konvergenz in Wahrscheinlichkeit* oder *stochastische Konvergenz* (sztochasztikus konvergencia)
- *fast sichere Konvergenz* (egy valószínűségű konvergencia)
- *Konvergenz in Verteilung* (eloszlásbeli konvergencia)

Gesetz der großen Zahlen

Definition. Eine Folge von Zufallsvariablen $X_1, X_2, \dots : \Omega \to \mathbb{R}$ *konvergiert in Wahrscheinlichkeit* (oder *stochastisch*) gegen eine Zufallsvariable $X : \Omega \to \mathbb{R}$ wenn für alle $\varepsilon > 0$ gilt:

$$\lim_{n\to\infty} \mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}) = 0.$$

In einer vereinfachten Schreibweise:

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.$$

Wir schreiben dann $X_n \xrightarrow{P} X$.

Gesetz der großen Zahlen

Satz. (Schwaches Gesetz der großen Zahlen)
Seien $X_1, X_2, \dots : \Omega \to \mathbb{R}$ unabhängige und identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu$ und $\mathbb{D}^2(X_i) = \sigma^2 < \infty$ für alle $i \in \mathbb{N}$. Dann gilt:

$$\overline{X_n} := \frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} \mu,$$

also

$$\lim_{n \to \infty} \mathbb{P}(\left|\overline{X_n} - \mu\right| \geq \varepsilon) = 0$$

für alle $\varepsilon > 0$.

Ungleichungen von Markow und Tschebyschow

Satz. (Markow-Ungleichung) Sei $X \geq 0$ eine nichtnegative Zufallsvariable. Dann gilt für alle $a > 0$:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Beweis. Wir definieren die Zufallsvariable Y für ein Element $\omega \in \Omega$ folgenderweise:

$$Y(\omega) = \begin{cases} a, & \text{falls } X(\omega) \geq a, \\ 0 & \text{ansonsten.} \end{cases}$$

Also $Y = a \cdot \mathbb{1}_{\{X \geq a\}}$.

Ungleichungen von Markow und Tschebyschow

Beweis. Die Variablen X und $Y = a \cdot \mathbb{1}_{\{X \geq a\}}$ sind nichtnegativ, weiter gilt $X \geq Y$ und deshalb auch

$$\mathbb{E}(X) \geq \mathbb{E}(Y) = a \cdot \mathbb{P}(X \geq a),$$

also

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Ungleichungen von Markow und Tschebyschow

Bemerkung. Die Ungleichung

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

gibt eine nicht triviale Schätzung nur dann, falls $a > \mathbb{E}(X)$ gilt.

Ungleichungen von Markow und Tschebyschow

Korollar. (Tschebyschow-Ungleichung) Sei X eine Zufallsvariable, für die $\mathbb{D}^2(X)$ endlich ist. Dann gilt für alle $a > 0$:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{D}^2(X)}{a^2}.$$

Beweis. Wir wenden die Markow-Ungleichung für die nicht-negative Zufallsvariable $(X - \mathbb{E}(X))^2$ an:

$$\begin{aligned}\mathbb{P}(|X - \mathbb{E}(X)| \geq a) &= \mathbb{P}((X - \mathbb{E}(X))^2 \geq a^2)\\ &\leq \frac{\mathbb{E}\big((X - \mathbb{E}(X))^2\big)}{a^2} = \frac{\mathbb{D}^2(X)}{a^2}.\end{aligned}$$

Ungleichungen von Markow und Tschebyschow

Bemerkung. Wenn wir die Zahl $a = \lambda \cdot \mathbb{D}(X)$ $(\lambda > 1)$ in der Tschebyschow-Ungleichung wählen, dann ergibt sich

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda \cdot \mathbb{D}(X)) \leq \frac{1}{\lambda^2}.$$

Also zeigt die Tschebyschow-Ungleichung, dass die Variable X ihren Wert mit "großer" Wahrscheinlichkeit im Intervall $[\mathbb{E}(X) - \lambda\mathbb{D}(X); \mathbb{E}(X) + \lambda\mathbb{D}(X)]$ annimmt.

Gesetz der großen Zahlen

Satz. (Schwaches Gesetz der großen Zahlen)
Seien $X_1, X_2, \dots : \Omega \to \mathbb{R}$ unabhängige und identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu$ und $\mathbb{D}^2(X_i) = \sigma^2 < \infty$ für alle $i \in \mathbb{N}$. Dann gilt:

$$\overline{X_n} := \frac{X_1 + \cdots + X_n}{n} \xrightarrow[P]{} \mu,$$

also

$$\lim_{n\to\infty} \mathbb{P}(\left|\overline{X_n} - \mu\right| \geq \varepsilon) = 0$$

für alle $\varepsilon > 0$.

Gesetz der großen Zahlen

Beweis. Für die Variable

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n}$$

gilt wegen der Linearität des Erwartungswertes:

$$\mathbb{E}(\overline{X_n}) = \frac{\mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)}{n} = \frac{n\mu}{n} = \mu.$$

Gesetz der großen Zahlen

Beweis. Wir wenden die Tschebyschow-Ungleichung für die Variable $\overline{X_n}$ an:

$$\mathbb{P}(\left|\overline{X_n} - \mu\right| \geq \varepsilon) = \mathbb{P}(\left|\overline{X_n} - \mathbb{E}(\overline{X_n})\right| \geq \varepsilon) \leq \frac{\mathbb{D}^2\left(\overline{X_n}\right)}{\varepsilon^2}.$$

Die Variablen $X_1, \ldots, X_n$ sind unabhängig, also

$$\begin{aligned}
\mathbb{D}^2(\overline{X_n}) &= \mathbb{D}^2\left(\frac{X_1 + \cdots + X_n}{n}\right) \\
&= \frac{1}{n^2}\mathbb{D}^2(X_1 + \cdots + X_n) \\
&= \frac{\mathbb{D}^2(X_1) + \cdots + \mathbb{D}^2(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0,
\end{aligned}$$

wenn $n \to \infty$.

Gesetz der großen Zahlen

Definition. Eine Folge von Zufallsvariablen $X_1, X_2, \dots : \Omega \to \mathbb{R}$ konvergiert *fast sicher* (majdnem mindenütt vagy egy valószínűséggel konvergál) gegen eine Zufallsvariable $X : \Omega \to \mathbb{R}$, wenn

$$\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1.$$

In einer vereinfachten Schreibweise:

$$\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1.$$

Wir schreiben dann $X_n \xrightarrow{f.s.} X$.

Gesetz der großen Zahlen

Bemerkungen.

- Damit diese Definition Sinn ergibt, muss man zeigen, dass

$$\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$$

  ein Ereignis ist.
- Wenn $X_n \xrightarrow{f.s.} X$ gilt, dann gilt auch $X_n \xrightarrow{P} X$.
- Aus $X_n \xrightarrow{P} X$ folgt $X_n \xrightarrow{f.s.} X$ nicht.

Gesetz der großen Zahlen

Satz. (Starkes Gesetz der großen Zahlen)
Seien $X_1, X_2, \dots : \Omega \to \mathbb{R}$ unabhängige und identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu < \infty$ für alle $i \in \mathbb{N}$. Dann gilt

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{f.s.} \mu.$$

Gesetz der großen Zahlen

Was passiert in der obigen Situation, falls die Erwartungswerte nicht endlich sind?

Satz. Seien $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ unabhängige und identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \infty$ für alle $i \in \mathbb{N}$. Dann gilt

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} \text{ existiert und ist endlich}\right) = 0.$$

Gesetz der großen Zahlen

Bemerkung. Bei der stochastischen Konvergenz $X_n \xrightarrow{P} X$ kann es passieren, dass man für jedes Element $\omega \in \Omega$ unendlich viele n finden kann, für die die Werte $X_n(\omega)$ und $X(\omega)$ voneinander weit sind, aber das Verhältnis dieser Elemente in Ω konvergiert gegen 0, wenn n gegen ∞ geht.

Bei der fast sicheren Konvergenz $X_n \xrightarrow{f.s.} X$ bilden die Elemente, für die die Werte $X_n(\omega)$ und $X(\omega)$ für eine beliebig große Zahl n voneinander weit sein können, eine Ereignismenge von Wahrscheinlichkeit 0, also ein *fast unmögliches* Ereignis.

Der zentrale Grenzwertsatz

Wir beschäftigen uns jetzt mit der Verteilung der Summe $X_1 + \cdots + X_n$.

Erinnerung. Falls $X_1, \ldots, X_n$ unabhängige Indikatorvariablen von Ereignissen mit derselben Wahrscheinlichkeit p sind, dann ist $X_1 + \cdots + X_n = S_n$ binomialverteilt mit Parametern n und p, und es gilt

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} \leq t\right) = \Phi(t),$$

wo Φ die Verteilungsfunktion der Standardnormalverteilung ist.

D.h. die standardisierte Variable $(S_n - \mathbb{E}(S_n))/\mathbb{D}(S_n)$ ist approximativ standardnormalverteilt.

Der zentrale Grenzwertsatz

Definition. Eine Folge von Zufallsvariablen $X_1, X_2, \ldots$ konvergiert in Verteilung gegen eine Zufallsvariable X, falls

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$$

für alle $t \in S(X)$ gilt, wobei $S(X)$ die Menge der Stetigkeitspunkte der Verteilungsfunktion F_X ist.

Wir schreiben dann $X_n \xrightarrow{d} X$ oder $F_{X_n} \xrightarrow{d} F_X$. Hier steht der Buchstabe "d" für "distribution".

Der zentrale Grenzwertsatz

Satz. (Der zentrale Grenzwertsatz) Seien $X_1, X_2, \ldots$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu$ und $0 < \mathbb{D}^2(X_i) = \sigma^2 < \infty$. Seien noch $Y \sim N(0;1)$ und $S_n = X_1 + \cdots + X_n$, dann gilt

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} \xrightarrow{d} Y,$$

also

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le t\right) = \Phi(t)$$

für alle $t \in \mathbb{R}$.

Der zentrale Grenzwertsatz

Bemerkung. Wir sahen früher, dass falls die Variablen $X_1, X_2, \dots$ unabhängig und normalverteilt mit Parametern μ und σ^2 sind, dann gilt

$$S_n = X_1 + \cdots + X_n \sim N(n\mu, n\sigma^2),$$

folglich

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim N(0;1).$$

In diesem Fall gilt also die Gleichung

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq t\right) = \Phi(t).$$

Der zentrale Grenzwertsatz ist also eine Verallgemeinerung der obigen Gleichung, aber im Allgemeinfall erhalten wir (nur) einen *Grenzwert*.

Der zentrale Grenzwertsatz

Es gibt solche analogische Ergebnisse, wo die Variablen unabhängig aber nicht notwendig identisch verteilt sind. Dann muss man mehr über die Momente annehmen:

Satz. (Laplace–Lyapunov) Seien $X_1, X_2, \ldots$ unabhängige Zufallsvariablen mit $\mathbb{E}(|X_i - \mu_i|^3) = \rho_i < \infty$, wo $\mathbb{E}(X_i) = \mu_i$. Sei noch $D_n = \left(\sum_{i=1}^n \mathbb{D}^2(X_i)\right)^{1/2} = \mathbb{D}\left(\sum_{i=1}^n X_i\right)$, und nehmen wir an, dass

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n \rho_i}{D_n^3} = 0.$$

Dann gilt

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{D_n} \le t\right) = \Phi(t)$$

für alle $t \in \mathbb{R}$.

Der zentrale Grenzwertsatz

D.h.: die Standardisierung von $\sum_{i=1}^{n} X_i$ ist approximativ standarnormalverteilt, zumindest falls n groß genug ist.

Dieses Ergebnis unterstützt das folgende Prinzip:
die Normalverteilung kann eine gute Approximation sein, wenn eine zufällige Quantität von vielen kleinen, "mehr oder weniger" unabhängigen Faktoren beeinflusst ist.

Der zentrale Grenzwertsatz

Wie schnell ist die Konvergenz im zentralen Grenzwertsatz, d.h.: was für eine Schätzung können wir für den Fehler der Approximation geben?

Satz. (Berry-Esseen) Seien $X_1, X_2, \ldots$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E}(X_i) = \mu$, $0 < \mathbb{D}^2(X_i) = \sigma^2 < \infty$ und $\mathbb{E}(|X_i - \mu|^3) = \rho < \infty$. Dann gibt es eine reelle Zahl $C > 0$, für die die Ungleichung

$$\left| \mathbb{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq t \right) - \Phi(t) \right| < \frac{C\rho}{\sigma^3\sqrt{n}}$$

gilt. (Zum Beispiel $C = 0{,}4748$ ist gut.)

Der zentrale Grenzwertsatz

Beispiel. Sei $X_i = \mathbb{1}_{A_i}$, wo $A_1, \ldots, A_n$ unabhängige Ereignisse mit $\mathbb{P}(A_i) = p$ sind. Dann gelten

$$S_n := X_1 + \cdots + X_n \sim \mathrm{Bin}(n; p),$$
$$\mathbb{E}(X_i) = \mathbb{P}(A_i) = p, \qquad \mathbb{D}^2(X_i) = p(1-p),$$

und

$$\begin{aligned}
\mathbb{E}(|X_i - p|^3) &= (1-p)^3 \cdot \mathbb{P}(X_i = 1) + p^3 \cdot \mathbb{P}(X_i = 0) \\
&= (1-p)^3 p + p^3 (1-p) \\
&= p(1-p)((1-p)^2 + p^2) \\
&= p(1-p)(2p^2 - 2p + 1).
\end{aligned}$$

Der zentrale Grenzwertsatz

Also

$$\left| \mathbb{P}\left(\frac{S_n - np}{\sqrt{p(1-p)n}} \le t \right) - \Phi(t) \right| < \frac{Cp(1-p)(2p^2 - 2p + 1)}{[p(1-p)]^{3/2}\sqrt{n}}$$

$$= \frac{C(2p^2 - 2p + 1)}{\sqrt{p(1-p)n}}.$$

Für $p = \frac{1}{2}$ haben wir $p(1-p) = 1/4$, $2p^2 - 2p + 1 = 1/2$, d.h.

$$\left| \mathbb{P}\left(\frac{S_n - n/2}{\sqrt{n}/2} \le t \right) - \Phi(t) \right| < \frac{C}{\sqrt{n}} < \frac{1}{2\sqrt{n}}.$$

Der zentrale Grenzwertsatz

Anwendung. Nehmer wir an, dass wir eine nicht (notwendig) reguläre Münze haben, die mit Wahrscheinlichkeit p einen Kopf und mit Wahrscheinlichkeit $1 - p$ eine Zahl gibt.

Wir möchten die Wahrscheinlichkeit p bestimmen.

Werfen wir die Münze n Mal und betrachten wir die relative Häufigkeit p_n der Köpfe (=Anzahl der Köpfe/Anzahl der Würfe).

Das Gesetz der großen Zahlen gibt, dass p_n gegen p konvergiert.

Der zentrale Grenzwertsatz

Die relative Häufigkeit p_n ist nur eine Approximation von p.

Wir möchten p präzis approximieren, z.B. erwarten wir, dass

$$|p_n - p| \leq \lambda$$

für eine kleine positive Zahl $\lambda > 0$ gilt.

Das kann nicht immer garantiert werden, aber wir erawarten dann, dass $|p_n - p| \leq \lambda$ mit großer Wahrscheinlichkeit $1 - \varepsilon$ gilt (wo $\varepsilon > 0$ eine kleine positive reelle Zahl ist):

$$\mathbb{P}(|p_n - p| \leq \lambda) \geq 1 - \varepsilon.$$

Der zentrale Grenzwertsatz

Die relative Häufigkeit p_n kann durch eine Formel

$$\frac{X_1 + \cdots + X_n}{n}$$

ausgedrückt werden. Hier ist X_i die Indikatorvariable des Ereignisses, dass der i-te Wurf einen Kopf gibt. Diese Variablen sind unabhängig und identisch verteilt.

$\mathbb{E}(X_i) = p$ gilt auch, also möchten wir

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mathbb{E}(X_1)\right| \leq \lambda\right) \geq 1 - \varepsilon$$

erreichen.

Der zentrale Grenzwertsatz

Die letzte Wahrscheinlichkeit ist

$$\mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \mathbb{E}(X_1)\right| \leq \lambda\right) = \mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{n}\right| \leq \lambda\right)$$

$$= \mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{\mathbb{D}(X_1)\sqrt{n}}\right| \leq \frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right)$$

$$= \mathbb{P}\left(-\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)} \leq \frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{\mathbb{D}(X_1)\sqrt{n}} \leq \frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right),$$

und nach dem zentralen Grenzwertsatz ist diese Wahrscheinlichkeit ungefähr

$$\approx \Phi\left(\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) - \Phi\left(-\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) = 2\Phi\left(\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) - 1.$$

Der zentrale Grenzwertsatz

Nehmen wir zunächst an, dass diese letzte Approximation gut ist. Um eine gute Schätzung zu erhalten, benötigen wir dann

$$2\Phi\left(\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) - 1 \geq 1 - \varepsilon,$$

also

$$\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)} \geq \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right),$$

denn Φ ist monoton steigend.

Falls λ und ε fix sind, dann gibt eine genügend große Anzahl der Würfe ($=n$) eine gute Schätzung:

$$n \geq \frac{\Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right)^2 \cdot \mathbb{D}^2(X_1)}{\lambda^2} = \frac{\Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right)^2 \cdot p(1-p)}{\lambda^2}$$

muss gelten.

Der zentrale Grenzwertsatz

Wir sahen (Übungsblatt 2., Aufgabe 6.), dass $p(1-p) \leq \dfrac{1}{4}$ immer gilt, also ist es genügend, falls

$$n \geq \frac{\Phi^{-1}\left(1-\frac{\varepsilon}{2}\right)^2}{4\lambda^2}$$

erfüllt wird.

Für $\lambda = 0{,}001$ und $\varepsilon = 0{,}005$ erhalten wir die untere Schranke

$$n \geq \frac{2{,}807033768^2 \cdot 10^6}{4} \approx 1\,969\,859{,}6.$$

Der zentrale Grenzwertsatz

Wenn auch der Fehler der Approximation im zentralen Grenzwertsatz berücksichtigt wird, dann erhalten wir

$$\mathbb{P}\left(\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mathbb{E}(X_1)\right| \leq \lambda\right) = 2\Phi\left(\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) - 1 + E,$$

wo

$$|E| \leq \frac{(2p^2 - 2p + 1)}{\sqrt{p(1-p)n}}$$

wegen des Satzes von Berry-Esseen gilt.

Der zentrale Grenzwertsatz

Falls n groß genug ist, dann haben wir

$$2\Phi\left(\frac{\lambda\sqrt{n}}{\mathbb{D}(X_1)}\right) - 1 + E \geq 1 - \varepsilon + E.$$

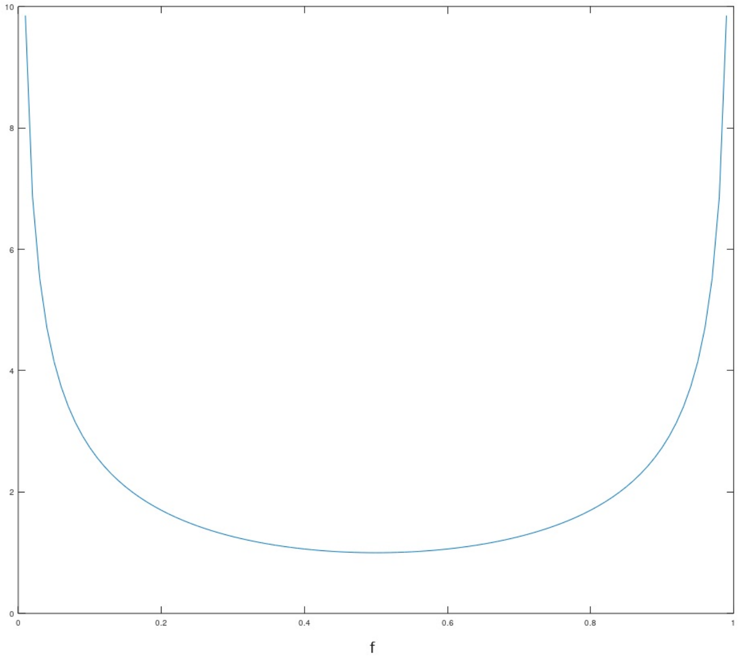
Um eine untere Schranke $\mathbb{P}(|p_n - p| \leq \lambda) \geq 1 - \varepsilon - \delta$ zu erhalten, ist es also genügend, falls

$$\frac{(2p^2 - 2p + 1)}{\sqrt{p(1-p)n}} \leq \delta$$

erfüllt wird. Die Funktion

$$f(p) = \frac{(2p^2 - 2p + 1)}{\sqrt{p(1-p)}}$$

nimmt aber große Werte an, falls p in der Nähe von 0 oder 1 ist:


10
8
6
4
2
0
0
0.2
0.4
0.6
0.8
1
f

Der zentrale Grenzwertsatz

Nehmen wir an, dass $p \in [\frac{1}{4}; \frac{3}{4}]$, dann gilt

$$|E| \leq \frac{5}{2\sqrt{3n}} < 0{,}001,$$

falls $n \geq 1\,969\,860$, und dann

$$\mathbb{P}(|p_n - p| \leq 0{,}001) > 0,994.$$

D.h.: ungefähr 2 Million Würfe liefern mit großer Wahrscheinlichkeit eine ziemlich gute Annäherung von p.

Der zentrale Grenzwertsatz

Bemerkungen.

- Meistens gibt es keine Kraftquelle oder Zeit, so viele Experimente zu machen oder einen so großen Datensatz anzulegen.
- Die obige Rechnung zeigt dann, dass die große Präzisität und/oder die große Sicherheit oft aufgegeben werden müssen, wenn man Verteilungen oder ihre Parameter durch statistische Methoden bestimmen möchte. Aber meistens sind auch schwachere Schätzungen hinreichend.
- Um die Rechnungen zu vereinfachen, betrachtet man oft solche zufällige Quantitäten, die als normalverteilt angenommen werden können. In diesem Fall kann die Schätzung der Fehler im zentralen Grenzwertsatz weggelassen werden.