Wirtschaftswissenschaftliches Zentrum
Universität Basel

9

Mathematik 2
Dr. Thomas Zehrt

Zufallsvariablen 3:
Grenzwertsätze

Benötigtes Vorwissen: Der Stoff der Vorlesung „Statistik“ wird als bekannt vorausgesetzt, insbesondere **Kapitel 8 „Zufallsvariablen“** und **Kapitel 10 „Ausblick“**

Inhaltsverzeichnis

1 Stochastische Konvergenz

Sei $(X_i)_{i\in\mathbb{N}} = X_1, X_2, X_3 \ldots$ eine Folge von Zufallsvariablen, die zugehörigen Verteilungsfunktionen seien mit $F_i = F_i(t)$ bezeichnet.

Definition 1.1 *Eine Folge $(X_i)_{i\in\mathbb{N}}$ von Zufallsvariablen konvergiert stochastisch gegen die reelle Zahl a, wenn für beliebige $c > 0$ die folgende Beziehung gilt:*

$$\lim_{i\to\infty} P(|X_i - a| \geq c) \;=\; 0.$$

Äquivalent dazu:

$$\lim_{i\to\infty} P(|X_i - a| < c) \;=\; 1$$

Das ist **keine** Konvergenz im klassischen Sinne und es bedeutet auch nicht, dass X_i gegen a konvergiert. Die zweite Relation könnten wir wie folgt deuten: Für grosse i sind die Werte von X_i mit grosser Wahrscheinlichkeit nahe (Abstand kleiner als c) beim Wert a.

Wir wollen nun untersuchen, was die stochastische Konvergenz für die zugehörigen Verteilungsfunktionen bedeutet. Sei also $(X_i)_{i\in\mathbb{N}}$ eine Folge von Zufallsvariablen so dass für **jedes** $c > 0$

$$\begin{aligned} \lim_{i\to\infty} P(|X_i - a| \geq c) &= 0 \qquad \text{oder} \\ \lim_{i\to\infty} P(X_i - a \leq -c) + \lim_{i\to\infty} P(X_i - a \geq c) &= 0 \end{aligned}$$

gilt. Natürlich muss dann die Wahrscheinlichkeit in beiden Bereichen gegen 0 konvergieren. Wir untersuchen also beide Intervalle:

1.

$$0 \;=\; \lim_{i\to\infty} P(X_i - a \leq -c) \;=\; \lim_{i\to\infty} P(X_i \leq a - c) \;=\; \lim_{i\to\infty} F_i(a-c)$$

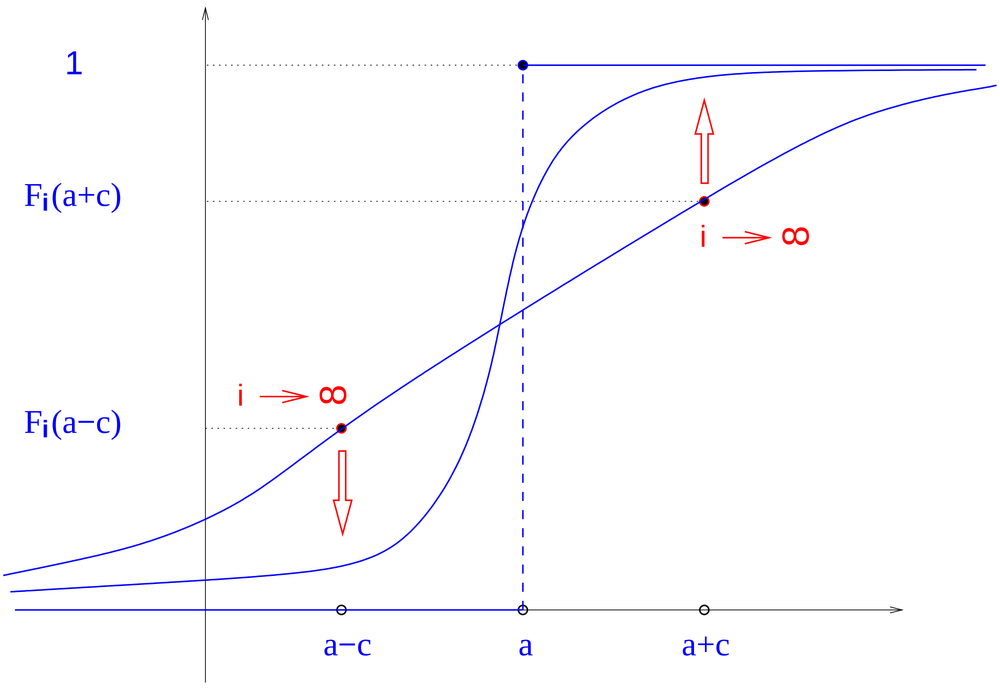
2.

$$\begin{aligned} 0 &= \lim_{i\to\infty} P(X_i - a \geq c) \\ &= \lim_{i\to\infty} (1 - P(X_i < a + c)) \\ &= 1 - \left(\lim_{i\to\infty} F_i(a+c) - \lim_{i\to\infty} \underbrace{P(X_i = a + c)}_{=0} \right) \end{aligned}$$

Für alle $c > 0$ gilt somit:

$$\lim_{i\to\infty} F_i(a-c) \;=\; 0 \qquad \text{und} \qquad \lim_{i\to\infty} F_i(a+c) \;=\; 1$$

Satz 1 *Eine Folge von Zufallsvariablen konvergiert genau dann stochastisch gegen a, wenn die Folge $(F_i(t))_{i\in\mathbb{N}}$ ihrer Verteilungsfunktionen (in jeder Stetigkeitsstelle) gegen die Verteilungsfunktion einer Einpunktverteilung konvergiert.*



Links von a, also an der Stelle $a - c$, gilt $\lim_{i\to\infty} F_i(a-c) = 0$ und rechts von a, also an der Stelle $a + c$, gilt $\lim_{i\to\infty} F_i(a+c) = 1$.

2 Grenzwertsätze

2.1 Das Gesetz der grossen Zahlen

Sei $(X_i)_{i\in\mathbb{N}}$ eine Folge unabhängiger und identisch verteilter Zufallsvariablen mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$. Für das arithmetische Mittel

$$\overline{X}_n \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i \;=\; \frac{1}{n}\,(X_1 + X_2 + \ldots + X_n)$$

von den ersten n dieser Zufallsvariablen gilt

$$E(\overline{X}_n) \;=\; \mu \qquad \text{und} \qquad Var(\overline{X}_n) \;=\; \frac{\sigma^2}{n}$$

Mit der Tschebyschevungleichung folgt dann sofort

$$1 \;\geq\; P\left(\, \left|\overline{X}_n - \mu\right| < c \,\right) \;\geq\; 1 - \frac{Var(\overline{X}_n)}{c^2} \;=\; 1 - \frac{\sigma^2}{n \cdot c^2}$$

Für festes $c > 0$ strebt die rechte Seite für $n \to \infty$ gegen 1.

Satz 2 (Gesetz der grossen Zahlen) *Sei $(X_i)_{i\in\mathbb{N}}$ eine Folge unabhängiger und identisch verteilter Zufallsvariablen mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$. Für das arithmetische Mittel $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ von den ersten n dieser Zufallsvariablen gilt dann*

$$\boxed{\lim_{n\to\infty} P\left(\, \left|\overline{X}_n - \mu\right| < c \,\right) \;=\; 1}$$

d.h. die Folge $(\overline{X}_n)_{n\in\mathbb{N}}$ konvergiert stochastisch gegen μ.

Ist n nur gross genug, so unterscheiden sich $\overline{X}_n$ und μ mit hoher Wahrscheinlichkeit nur um einen (kleinen) Wert $< c$.

Oder: Für grosse n gilt mit hoher Wahrscheinlichkeit

$$\mu - c \;<\; \frac{1}{n}\,(X_1 + X_2 + \ldots + X_n) \;<\; \mu + c.$$

Beispiel 2.1 *Die historischen Lottozahlen der letzten 25 Jahre (1302 Ausspielungen) aus Deutschland (7 aus 49) zeigen das Gesetz der grossen Zahlen sehr anschaulich.*

1 187	*2* 194	*3* 194	*4* 178	*5* 176	*6* 187	*7* 175
8 197	*9* 201	*10* 173	*11* 175	*12* 180	*13* 157	*14* 181
15 172	*16* 180	*17* 190	*18* 183	*19* 191	*20* 181	*21* 200
22 191	*23* 186	*24* 177	*25* 196	*26* 200	*27* 180	*28* 167
29 186	*30* 182	*31* 199	*32* 211	*33* 189	*34* 175	*35* 183
36 199	*37* 175	*38* 199	*39* 199	*40* 195	*41* 183	*42* 182
43 189	*44* 181	*45* 188	*46* 191	*47* 175	*48* 195	*49* 207

Jeden Samstag wurden sieben Kugeln (einschliesslich Zusatzzahl) gezogen, es liegt somit ein sehr grosser Stichprobenumfang von $n = 7 \cdot 1302 = 9114$ *vor.*

Genau genommen sind die Ziehungen an einem Tag weder unabhängig noch identisch verteilt, aber wir wollen annehmen, dass beides **fast** *stimmt.*

Das arithmetische Mittel der gezogenen Lottozahlen beträgt

$$\begin{aligned}\overline{X}_{9114} &= \frac{1}{9114}(1 \cdot 187 + 2 \cdot 194 + \ldots + 49 \cdot 207) = 25.2211 \\ Var(\overline{X}_{9114}) &= 200.6512\end{aligned}$$

Die (idealisierte) Wahrscheinlichkeitsverteilung ist die diskrete Gleichverteilung (auf den 49 Zahlen) mit Erwartungswert und Varianz

$$\mu = \sum_{i=1}^{49} i \cdot \frac{1}{49} = \frac{50}{2} = 25 \qquad Var = \frac{49^2 - 1}{12} = 200$$

2.2 Der Satz von Bernoulli

Der folgende Satz ist ein Spezialfall des Gesetzes der grossen Zahlen. Wir starten also mit eine speziellen Folge unabhängiger und identisch verteilter Zufallsvariablen.

Sei A irgend ein Ereignis eines (beliebig oft wiederholbaren) Zufallsexperimentes, welches mit der Wahrscheinlichkeit $p = p(A)$ eintritt und jedes X_i sei Null-Eins-verteilt, d.h.

$$X_i = \begin{cases} 1 \text{ (Ereignis } A \text{ tritt ein im } i\text{-ten Versuch)} & \text{mit } P(X_i = 1) = p \\ 0 \text{ (Ereignis } A \text{ tritt nicht ein im } i\text{-ten Versuch)} & \text{mit } P(X_i = 0) = 1 - p \end{cases}$$

Dann gilt $E(X_i) = p = \mu$ und $Var(X_i) = p(1-p) = \sigma^2$. Die Zufallsvariable

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

stellt hier die relative Häufigkeit des Ereignisses A bei n unabhängigen Wiederholungen dar. Mit der Tschebyschevungleichung folgt wieder

$$\begin{aligned} P(\, |\overline{X}_n - p| < c\,) &\geq 1 - \frac{p(1-p)}{n \cdot c^2} \\ \lim_{n\to\infty} P(\, |\overline{X}_n - p| < c\,) &= 1 \end{aligned}$$

Für grosse n gilt mit hoher Wahrscheinlichkeit

$$p(A) - c < \frac{1}{n} (X_1 + X_2 + \ldots + X_n) < p(A) + c.$$

Satz 3 (Satz von Bernoulli) *Die relative Häufigkeit eines zufälligen Ereignisses A in n unabhängigen Wiederholungen konvergiert stochastisch gegen die Wahrscheinlichkeit $p = p(A)$ des Ereignisses A.*

Beispiel 2.2 *Wir schauen uns an, was das für die Lottozahlen bedeutet. Wir können das arithmetische Mittel bzw. die relative Häufigkeit für jede der Zahlen $k = 1, 2, \ldots, 49$ bestimmen und mit der exakten Wahrscheinlichkeit vergleichen, die für jede Zahl wohl* $1/49 \approx 0.020408$ *sein sollte.*

k	$\frac{\text{Häufigkeit von } k}{9114}$
1	$\frac{187}{9114} \approx 0.020517$
2	$\frac{194}{9114} \approx 0.021286$
$\vdots$	$\vdots$
23	$\frac{186}{9114} \approx 0.020408$
$\vdots$	$\vdots$

<u>Anwendung: Schätzen von $p = p(A)$ durch die relative Häufigkeit $\overline{X}_n$</u>

gegeben: $c > 0$ und Irrtumswahrscheinlichkeit α (Sicherheitswahrscheinlichkeit $1 - \alpha$)

gesucht: Wie gross muss n gewählt werden, so dass

$$P(\ |\overline{X}_n - p| < c\ ) \ \geq \ 1 - \alpha$$

Lösung:

$$P(\ |\overline{X}_n - p| < c\ ) \ \geq \ 1 - \frac{p(1-p)}{n \cdot c^2} \ \geq \ \underbrace{1 - \frac{1/4}{n \cdot c^2} \ \geq \ 1 - \alpha}_{\text{Auflösen nach n}}$$

Die erste Ungleichung ist die Tschebyschevungleichung. Die zweite Ungleichung folgt aus der Feststellung, dass die quadratische Funktion $g(p) = p(1-p) = -p^2 + p$ in $p = 1/2$ ihr globales Maximum annimmt. Auflösen nach der Variablen n liefert uns dann eine Abschätzung der nötigen Stichprobengrösse:

$$n \ \geq \ \frac{1}{4 \cdot \alpha \cdot c^2}.$$

Aufgabe 2.1 *Eine (eventuell nichtfaire) Münze soll n-mal geworfen werden, so dass die relative Häufigkeit für das Ereignis A = „Kopf" mit 98%-iger Wahrscheinlichkeit um weniger als 0.1 von $p = p(A)$ abweicht. Wie gross muss n mindestens sein?*

Lösung:

2.3 Der zentrale Grenzwertsatz

Der zentrale Grenzwertsatz gibt einen Grund für die herausragende Rolle der Normalverteilung, denn er charakterisiert diese als Grenzverteilung von (additiven) Überlagerungen von vielen unabhängigen zufälligen Einzeleffekten.

Sei $X_1, X_2, \ldots$ eine (unendliche) Folge von identisch verteilten und unabhängigen Zufallsvariablen mit $\mu = E(X_i)$ und $\sigma^2 = Var(X_i)$. Wir betrachten:

$$S_n \; = \; \sum_{i=1}^{n} X_i \; = \; X_1 + X_2 + \ldots + X_n$$

Für den Erwartungswert und die Varianz von S_n folgt:

$$\begin{aligned} E(S_n) \; &= \; E\left(\sum_{i=1}^{n} X_i\right) \; = \; \sum_{i=1}^{n} E(X_i) \; = \; n \cdot \mu \\ Var(S_n) \; &= \; Var\left(\sum_{i=1}^{n} X_i\right) \; = \; \sum_{i=1}^{n} Var(X_i) \; = \; n \cdot \sigma^2 \end{aligned}$$

Diese Zufallsvariablen werden nun standardisiert:

$$Y_n \; = \; \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \; = \; \frac{S_n - n \cdot \mu}{\sqrt{n \cdot \sigma^2}}.$$

Y_n heisst auch die <u>standardisierte Summe</u> der $X_1, X_2, \ldots X_n$.

Satz 4 (Zentraler Grenzwertsatz)

$$\boxed{\lim_{n \to \infty} P(Y_n \leq x) \; = \; \Phi(x) \; = \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} \, dt}$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung $N(0,1)$ ist.

Ist eine zufällige Erscheinung additiv aus (möglichts vielen) unabhängigen zufälligen Ereignissen zusammengesetz, so können Wahrscheinlichkeiten **näherungsweise** mit der Funktion Φ bestimmt werden.

Merken sie sich:

$$\boxed{\begin{array}{llll} S_n = \displaystyle\sum_{i=1}^{n} X_i & \sim & N(n \cdot \mu, n \cdot \sigma^2) & \text{für } n \text{ gross} \\[2ex] X_n = \displaystyle\frac{1}{n}\sum_{i=1}^{n} X_i & \sim & N(\mu, \sigma^2/n) & \text{für } n \text{ gross} \\[2ex] Y_n = \dfrac{\displaystyle\sum_{i=1}^{n} X_i - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} & \sim & N(0,1) & \text{für } n \text{ gross} \end{array}}$$

Beispiel 2.3 *Experiment: n-maliger Münzwurf*

Dann gilt $\Omega = \{(\omega_1, \omega_2, \ldots, \omega_n) \,:\, \omega_i \in \{K, Z\}\}$ *mit der Gleichverteilung. Sei*

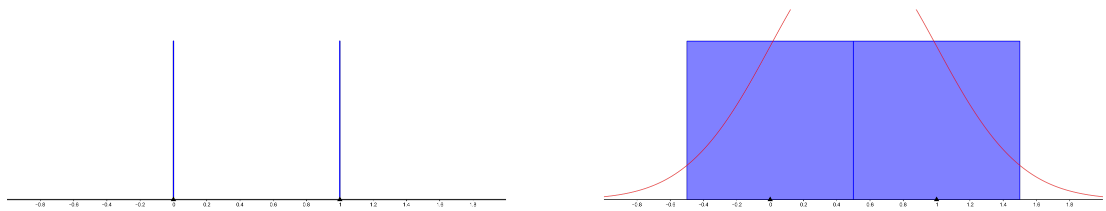
$$X_i = \begin{cases} 0 & \text{beim } i\text{-ten Wurf Kopf} \\ 1 & \text{beim } i\text{-ten Wurf Zahl} \end{cases}$$

Alle diese Zufallsvariablen sind unabhängig und gleichverteilt; es gilt $E(X_i) = \frac{1}{2}$ *und* $Var(X_i) = \frac{1}{4}$. *Jedes* X_i *hat die Verteilung* $\left\{\left(0, \frac{1}{2}\right), \left(1, \frac{1}{2}\right)\right\}$. *Nun addieren wir die n Zufallsvariablen* $S_n = X_1 + X_2 + \ldots + X_n$ *und schauen uns die Verteilung dieser neuen Zufallsvariablen an.*

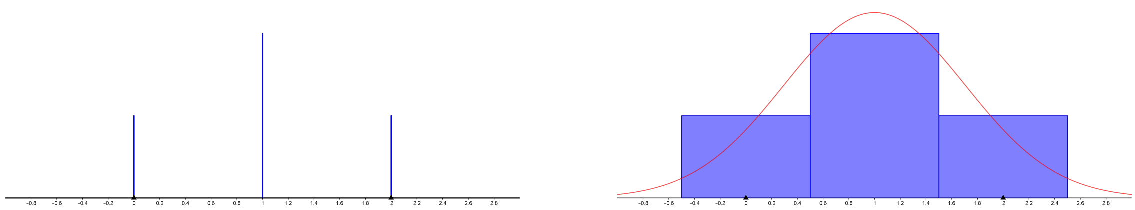
Natürlich gilt für jedes n wie immer bei Summen unabhängiger Zufallsvariablen: $E(S_n) = n/2$ *und* $Var(S_n) = n/4$.

Im folgenden skizzieren wir für einige n die (diskrete) Verteilung von S_n als Stabdiagramm und zusätzlich als Balkendiagramm zusammen mit der Gauss-Glockenkurve mit $\mu = E(S_n) = n/2$ *und* $\sigma^2 = Var(S_n) = n/4$. *Sehen Sie die Aussage des zentralen Grenzwertsatzes?*

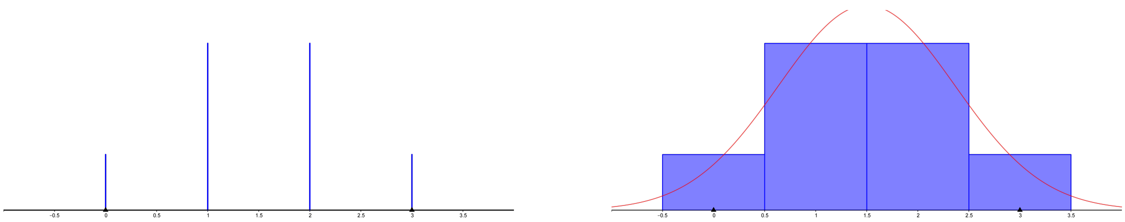
1. $S_1 = X_1$ *Verteilung:* $\left\{\left(0, \frac{1}{2}\right), \left(1, \frac{1}{2}\right)\right\}$



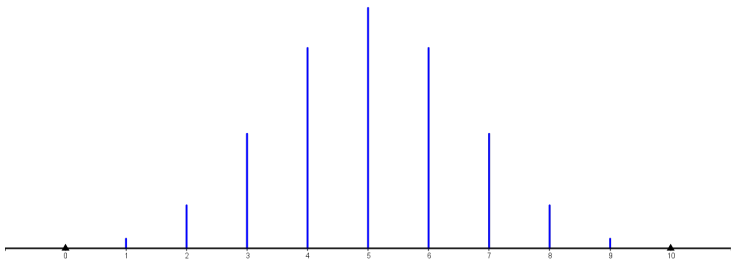
2. $S_2 = X_1 + X_2$, *Verteilung:* $\left\{\left(0, \frac{1}{4}\right), \left(1, \frac{2}{4}\right), \left(2, \frac{1}{4}\right)\right\}$

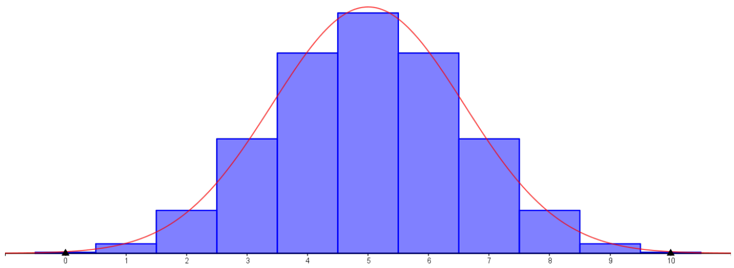


3. $S_3 = X_1 + X_2 + X_3$, *Verteilung:* $\left\{\left(0, \frac{1}{8}\right), \left(1, \frac{3}{8}\right), \left(2, \frac{3}{8}\right), \left(3, \frac{1}{8}\right)\right\}$

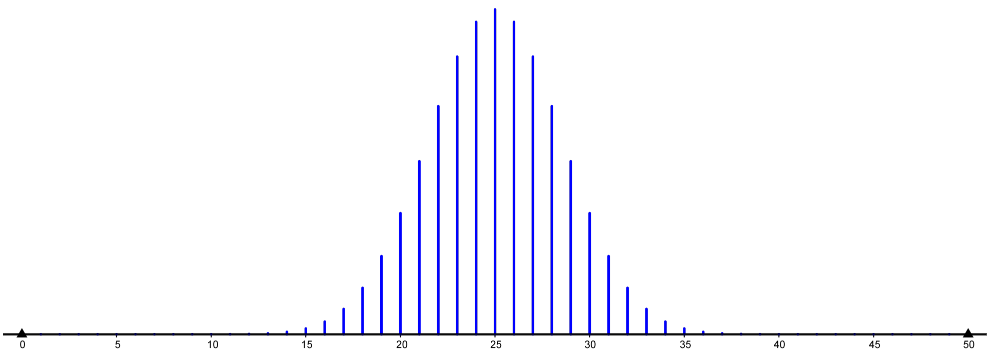


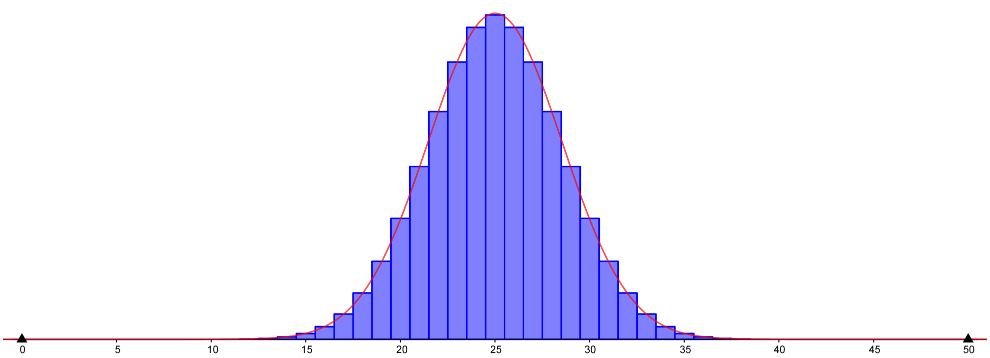
4. $S_{10} = X_1 + \ldots + X_{10}$, *Verteilung:* $\left\{\left(0, \frac{1}{2^{10}}\right), \left(1, \frac{10}{2^{10}}\right), \ldots, \left(10, \frac{1}{2^{10}}\right)\right\}$





5. $S_{50} = X_1 + \ldots + X_{50}$, *Verteilung:*





Eine kombinatorische Überlegung zeigt:

Die Verteilung für $S_n = \sum_{i=1}^{n} X_i$ *ist gegeben durch*

$$\left\{\left(0, \frac{1}{2^n}\right), \left(1, \frac{n}{2^n}\right), \ldots, \left(k, \frac{\binom{n}{k}}{2^n}\right), \ldots, \left(n, \frac{1}{2^n}\right)\right\}.$$

Nun betrachten wir die zugehörigen normalisierten Zufallsvariablen:

$$Y_n = \frac{S_n - n\frac{1}{2}}{\sqrt{n \cdot \frac{1}{4}}} = \frac{\sum_{i=1}^{n} X_i - \frac{n}{2}}{\sqrt{n} \cdot \frac{1}{2}}$$

z.B. hat Y_2 *dann die Verteilung:* $\left\{ \left(-\sqrt{2}, \frac{1}{4}\right), \left(0, \frac{2}{4}\right), \left(\sqrt{2}, \frac{1}{4}\right) \right\}$
Die Verteilung der Zufallsvariablen $Y_n = \frac{\sum_{i=1}^{n} X_i - \frac{n}{2}}{\sqrt{n} \cdot \frac{1}{2}}$ *ist gegeben durch*

$$\left\{ \left(\frac{0 - \frac{n}{2}}{\sqrt{n} \cdot \frac{1}{2}}, \frac{1}{2^n} \right), \ldots, \left(\frac{k - \frac{n}{2}}{\sqrt{n} \cdot \frac{1}{2}}, \frac{\binom{n}{k}}{2^n} \right), \ldots, \left(\frac{n - \frac{n}{2}}{\sqrt{n} \cdot \frac{1}{2}}, \frac{1}{2^n} \right) \right\}$$

und für $n \to \infty$ *gilt* $Y_n \sim N(0,1)$

Beispiel 2.4 Eine *Theorie behauptet, dass die Entwicklung von Aktienkursen auf (informationseffizienten) Märkten einem* <u>*Random-Walk*</u> *folgt:*

$$\underbrace{k_{t+1}}_{\text{Kurs morgen}} = \underbrace{k_t}_{\text{Kurs heute}} + \epsilon_{t+1}$$

wobei die Kursänderungen $\epsilon_{t+1} = k_{t+1} - k_t$ *Zufallsvariablen sind mit* $E(\epsilon_i) = 0$ *und* $Var(\epsilon_i) = \sigma^2$.

Monatliche Kursänderung: $\epsilon_{t+1} + \epsilon_{t+2} + \ldots + \epsilon_{t+n}$ *(z.B.* $n = 25$ *Handelstage pro Monat)*

Zentraler Grenzwertsatz: $\sum_{i=1}^{25} \epsilon_{t+i}$ $, \sim,$ $N(0, 25\sigma^2)$
(die monatliche Kursänderung ist ungefähr normalverteilt mit dem Erwartungswert 0 *und der* 25*-fachen Varianz* $25\sigma^2$*)*
Schlussfolgerungen aus der Random-Walktheorie

- *Jede Kursprognose, die sich auf die vergangene Kursentwicklung stützt (z.B. die Charttechnik) ist unmöglich.*
- *Einzige akzeptable Prognose: Der Kurs bleibt wie er ist!*

3 Approximationen der Binomialverteilung

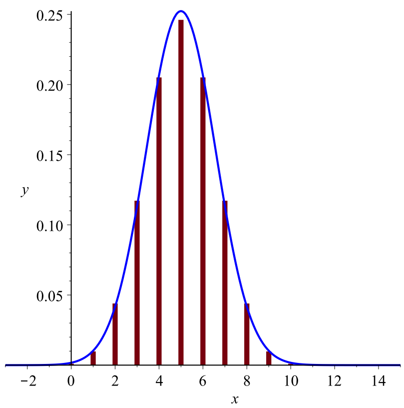
In der Praxis muss man häufig Wahrscheinlichkeiten von binomialverteilten Zufallsvariablen $X \sim B(n; p)$ für grosse n bestimmen. Das ist sehr aufwändig. Hier helfen Approximationen der Binomialverteilung durch Standardverteilungen.

Wir wissen bereits, dass binomialverteilte Zufallsvariablen Summen von identisch zweipunktverteilten und unabhängigen Zufallsvariablen sind. Der zentrale Grenzwertsatz liefert uns somit eine Approximationsmöglichkeit mit Hilfe einer Normalverteilung und vernünftigerweise sollte man die (eindeutig bestimmte) Normalverteilung nutzen, die den selben Erwartungswert und die selbe Varianz wie X hat, d.h. $\mu = np$ und $\sigma = \sqrt{np(1-p)}$.

Beispiel 3.1

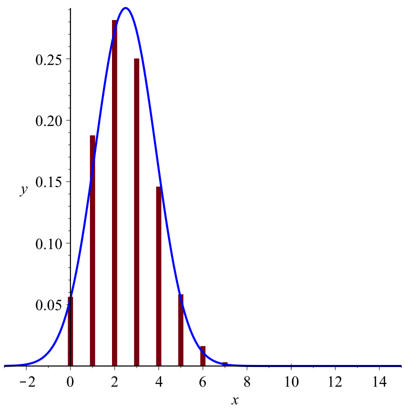
$n = 10$ *und* $p = 0.5 \quad \longrightarrow \quad \mu = n \cdot p = 5$ *und* $\sigma = \sqrt{n \cdot p \cdot (1-p)} \approx 1.58$

In der Skizze sehen Sie die Binomialverteilung $f_{Bi}(x; 10, 0.5)$ *(als rotes Stabdiagramm) und die (blaue) Glockenkurve* $\phi(x; 5, 1.58)$*:*



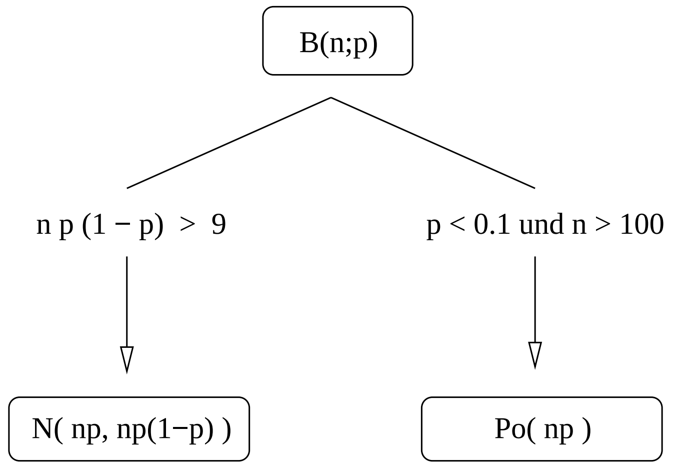
$n = 10$ *und* $p = 0.25 \quad \longrightarrow \quad \mu = n \cdot p = 2.5$ *und* $\sigma = \sqrt{n \cdot p \cdot (1-p)} \approx 1.37$

In der Skizze sehen Sie die Binomialverteilung $f_{Bi}(x; 10, 0.25)$ *(als rotes Stabdiagramm) und die (blaue) Glockenkurve* $\phi(x; 2.5, 1.37)$*:*



Für $n \to \infty$ wäre diese Approximation (immer) perfekt.

Natürlich ist das n stets endlich und die Qualität der Approximation hängt ganz entscheidend auch von p ab. Das ist leicht zu verstehen, denn zunächst einmal sind **alle** Glockenkurven symmetrisch (zur Achse $x = \mu$). Dagegen haben nur die Binomialverteilungen mit $p = 1/2$ eine ähnliche Symmetrie. Für sehr kleine (oder sehr grosse) Werte von p (und relativ kleine Werte n) liefert diese Approximation meist keine guten Resultate. Hier greift man auf eine Approximation durch eine Poissonverteilung zurück.



3.1 Grenzwertsatz von de Moivre und Laplace

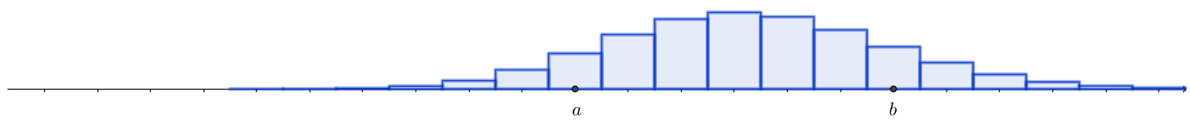
Satz 5 (Grenzwertsatz von de Moivre und Laplace) *Sei $X \sim B(n;p)$ eine binomialverteilte Zufallsvariable, $\mu = np$ und $\sigma = \sqrt{np(1-p)}$. Dann gilt*

$$
\boxed{
\begin{aligned}
f_{Bi}(x;n,p) &= \binom{n}{x} p^x (1-p)^{n-x} &&\approx \phi\left(x; np, \sqrt{np(1-p)}\right) \\
P(a \le X \le b) &= \sum_{x=a}^{b} \binom{n}{x} p^x (1-p)^{n-x} &&\approx \Phi\left(\frac{b-\mu+0.5}{\sigma}\right) - \Phi\left(\frac{a-\mu-0.5}{\sigma}\right)
\end{aligned}
}
$$

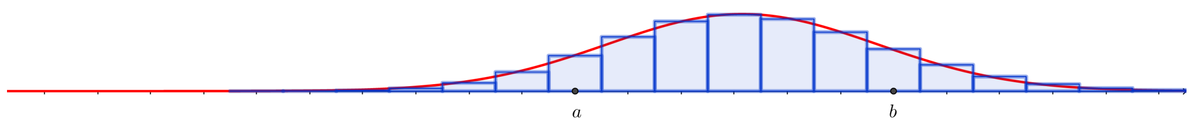
mit hinreichender Genauigkeit, falls $n \cdot p \cdot (1-p) > 9$ gilt.

Bemerkung

- Die Ausgangsfrage ist also, wie gross die Wahrscheinlichkeit ist, dass eine binomialverteilte Zufallsvariable $X \sim B(n,p)$ einen Wert zwischen a und b annimmt. Natürlich gilt $E(X) = np$ und $Var(X) = np(1-p)$. Wir interessieren uns also für die Wahrscheinlichkeit P (Rechtecksflächeninhalt) zwischen a und b, also $P = P(a \le X \le b)$.



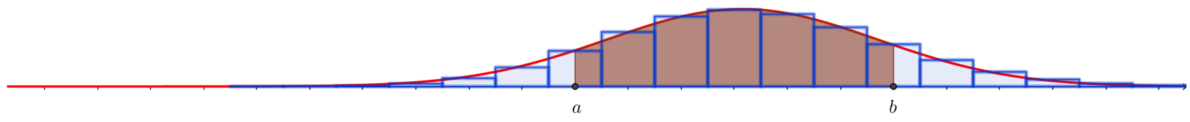
- Wir approximieren die Binomialverteilung durch die Gauss-Glockenkurve $\phi\left(x; np, \sqrt{np(1-p)}\right)$.



- Wir glauben, dass die gesuchte Wahrscheinlichkeit dem Flächeninhalt unter der speziell gewählten Gauss-Glocke recht gut entspricht.

$$P \approx \int_a^b \phi\left(x; np, \sqrt{np(1-p)}\right)\, dx$$

Eventuell könnte man hier auch ahnen, dass es besser wäre, von $a - 0.5$ bis $b + 0.5$ zu integrieren, denn in diesem Fall messen wir die beiden Rechtecke, die in den Punkten a und b sitzen, komplett mit. Also nicht nur eine Hälfte.

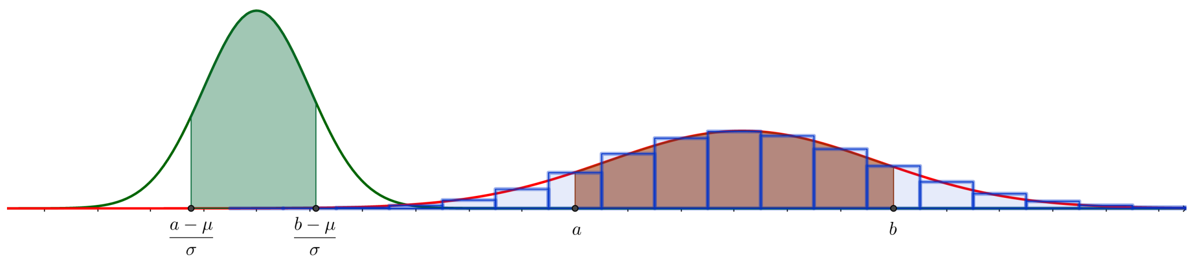


- Wir transformieren die Grenzen $a \to \frac{a-\mu}{\sigma}$ und $b \to \frac{b-\mu}{\sigma}$ so, dass man den gesuchten Flächeninhalt unter der Standardglockenkurve zwischen den transformierten Grenzen finden kann und berechnen diesen mit der Formelsammlung.

$$P \approx \int_a^b \phi\left(x; np, \sqrt{np(1-p)}\right)\, dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \phi(x)\, dx = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

bzw.

$$P \approx \int_{\frac{a-\mu-0.5}{\sigma}}^{\frac{b-\mu+0.5}{\sigma}} \phi(x)\, dx = \Phi\left(\frac{b-\mu+0.5}{\sigma}\right) - \Phi\left(\frac{a-\mu-0.5}{\sigma}\right).$$



Aufgabe 3.1 *Bestimmen Sie approximativ die Zahl*

$$P \;=\; \sum_{k=100}^{150} \binom{1000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} .$$

Lösung:

$$\begin{aligned} P \;&\approx\; \Phi\left(\frac{150-166.67+0.5}{11.78}\right) - \Phi\left(\frac{100-166.67-0.5}{11.78}\right) \;=\; \Phi(-1.37)-\Phi(-5.70) \\ &=\; 0.085 \end{aligned}$$

Ein exakteres Resultat (Geogebra) ist 0.0837.

3.2 Poissonscher Grenzwertsatz

Satz 6 *Für eine Folge von binomialverteilten Zufallsvariablen $X_n \sim B(n;p)$ wobei $p = p(n)$ mit $n \cdot p(n) = \lambda$ (d.h. zu grossen n gehören kleine p) gilt für alle $k = 0, 1, 2, \ldots$:*

$$\boxed{\lim_{n\to\infty} P(X_n = k) = \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} \quad = \quad \frac{\lambda^k}{k!}\, e^{-\lambda}}$$

Beweis: Wir setzten $n \cdot p = \lambda$ bzw. $p = \frac{\lambda}{n}$ und erhalten

$$\lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} \quad = \quad \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

Für den dritten und vierten Faktor gilt sicher:

$$\begin{aligned} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n &= e^{-\lambda} \\ \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} &= 1 \end{aligned}$$

Weiterhin:

$$\begin{aligned} \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k &= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \frac{\lambda^k}{n^k} \\ &= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \\ &= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)}{n \cdot n \cdot n \cdots n} \\ &= \frac{\lambda^k}{k!} \end{aligned}$$

Zusammen folgt die Behauptung.

□

In der Praxis verwendet man diesen Grenzwertsatz zur näherungsweisen Berechnung der Einzelwahrscheinlichkeiten der Binomialverteilung bei grossem n und kleinem p.

<u>Praktische Approximationsregel</u> Für grosse n (es sollte $n > 100$ gelten) und kleine p (es sollte $p < 0.1$ gelten) gilt für $\lambda = n \cdot p$

$$\boxed{f_{Bi}(x; n, p) \;=\; \binom{n}{k} p^k (1-p)^{n-k} \;\approx\; \frac{\lambda^k}{k!}\, e^{-\lambda}}$$

4 *Beweisen wir den zentralen Grenzwertsatz!*

4.1 Konkrete Problemstellung

Natürlich sind die Zeichnungen in den vorhergehenden Kapiteln recht überzeugend und wir mögen glauben, dass (grosse) Summen von identisch verteilten und unabhängigen Zufallsvariablen angenähert normalverteilt sind und sogar gegen Normalverteilungen konvergieren. Aber Sie müssen und sollen hier nichts glauben! Deshalb will ich hier einen möglichen Beweis des zentralen Grenzwertsatzes skizzieren. Mit den uns zur Verfügung stehenden Mitteln werden wir leider auch nicht jedes Details des Beweises ausführen können, d.h. auch hier werden sie wohl einiges glauben müssen. Teile des Beweises führen auch in ganz neue Bereiche der Mathematik, insbesondere in die komplexe **Funktionentheorie**, in die **Theorie der erzeugenden Funktionen** und in die **Fourier-Analysis**. Trotzdem hoffe ich, dass dieser Abschnitt das Verständnis vertieft und sie eventuell neugierig (auf mehr Mathematik) macht.

Beweise, in denen es um Konvergenz geht, sind meistens meistens technisch recht anspruchsvoll. Auch müssten wir zunächst klären, was wir hier überhaupt unter Konvergenz verstehen wollen. Die Mathematik hält einige Konvergenzbegriffe bereit: punktweise Konvergenz, gleichmässige Konvergenz, lokal gleichmässige Konvergenz, Konvergenz im quadratischen Mittel, stochastische Konvergenz, Konvergenz in Verteilung, ... Also konkretisieren wir:

Definition 4.1 *Eine Folge (Z_n) von Zufallsvariablen konvergiert in Verteilung gegen eine Zufallsvariable Z, wenn*

$$\lim_{n\to\infty} P(Z_n \leq x) \quad = \quad P(Z \leq x)$$

für alle x gilt, in denen die Verteilungsfunktion $F(x) = P(Z \leq x)$ stetig ist (nur relevant, falls Z einer diskreten Verteilung folgt).

Nun können wir den zentralen Grenzwertsatz nochmals ausführlich formulieren:

Satz 7 (Zentralen Grenzwertsatz) *Sei (X_n) eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit*

$$\begin{aligned} -\infty < \quad & E(X_n) = \mu && < \infty \\ 0 < \quad & Var(X_n) = \sigma^2 && < \infty. \end{aligned}$$

Dann konvergiert die Folge (Y_n) der standardisierten Summen der X_n, d.h.

$$Y_n \quad = \quad \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}},$$

in Verteilung gegen die standardnormalverteilte Zufallsvariable $Y \sim N(0,1)$.

Um den eigentlichen Beweis zu führen, müssen wir noch einen neuen Begriff einführen, die so genannte **charakteristische Funktion** einer Zufallsvariablen. Diese einfache (aber komplexwertige) Funktion enthält alle Informationen über die Zufallsvariable.

Der Beweis verläuft dann wie folgt nicht direkt, sondern über den Umweg der konvergierenden charakteristischen Funktionen:

1. Wir ordnen jedem standardisierten Summanden $Y_1, Y_2, Y_3, \ldots$ (natürlich geschickt) seine charakteristische Funktion $\chi_{Y_1}, \chi_{Y_2}, \chi_{Y_3}, \ldots$ zu.

$$\begin{array}{cccc} Y_1 & Y_2 & Y_3 & \ldots \\ \updownarrow & \updownarrow & \updownarrow & \\ \chi_{Y_1} & \chi_{Y_2} & \chi_{Y_3} & \cdots \end{array}$$

2. Wir zeigen, dass die Folge der charakteristischen Funktionen $\chi_{Y_1}, \chi_{Y_2}, \chi_{Y_3}, \ldots$ gegen die Funktiion $e^{-t^2/2} =: \chi_Y(t)$ konvergiert (genauer gesagt: gleichmässig konvergiert), die charakteristische Funktion der Standardnormalverteilung $Y \sim N(0, 1)$.

$$\begin{array}{cccccc} Y_1 & Y_2 & Y_3 & \ldots & & \\ \updownarrow & \updownarrow & \updownarrow & & & \\ \chi_{Y_1} & \chi_{Y_2} & \chi_{Y_3} & \cdots & \longrightarrow & \chi_{Y \sim N(0,1)} \end{array}$$

3. Wir schlussfolgern, dass deshalb auch die Folge $Y_1, Y_2, Y_3, \ldots$ in Verteilung gegen die Zufallsvariable $Y \sim N(0; 1)$ konvergieren muss.

$$\begin{array}{cccccc} Y_1 & Y_2 & Y_3 & \ldots & \longrightarrow & Y \sim N(0,1) \\ \updownarrow & \updownarrow & \updownarrow & & & \updownarrow \\ \chi_{Y_1} & \chi_{Y_2} & \chi_{Y_3} & \cdots & \longrightarrow & \chi_{Y \sim N(0,1)} \end{array}$$

4.2 Die charakteristische Funktion einer Zufallsvariable

4.2.1 Allgemeines

Es gibt Kenngrössen von Zufallsvariablen, die sich auf Grund von einfachen Rechenregeln als sehr interessant erweisen. Sei X eine Zufallsvariable und s eine (komplexe) Zahl. Dann bilden wir die neue Zufallsvariable e^{sX} und bestimmen ihren Erwartungswert $E(e^{sX})$. Wie kann man diese Zahl deuten, die wir der Zufallsvariablen X zugeordnet haben?

Nutzt man die Reihendarstellung der Exponentialfunktion, erhält man sofort:

$$\begin{aligned} E\left(e^{sX}\right) &= E\left(\sum_{k=0}^{\infty} \frac{1}{k!}\, s^k X^k\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\, s^k E\left(X^k\right) \\ &= \sum_{k=0}^{\infty} \frac{E\left(X^k\right)}{k!}\, s^k . \end{aligned}$$

Das ist eine Potenzreihe in s, deren Koeffizienten im Wesentlichen die Erwartungswerte der Zufallsvariablen $X^0, X^1, X^2, \ldots$, also alle **Momente** von X, sind. Da man die Zahl s stetig variieren kann, erhält man ein ganzes kontinuierliches Spektrum von Kenngrössen $E(e^{sX})$ zur Variablen X.

Insbesondere gilt auch

$$E\left(e^{s(X+Y)}\right) = E\left(e^{sX} e^{sY}\right) = E\left(e^{sX}\right)\, E\left(e^{sY}\right),$$

falls die beiden Zufallsvariablen X und Y, und damit auch die beiden Zufallsvariablen e^{sX} und e^{sY}, stochastisch unabhängig sind.

Tatsächlich hat sich die rein imaginäre Zahl $s = it$ mit $t \in \mathbb{R}$ als besonders interessant erwiesen.

Definition 4.2 *Sei X eine Zufallsvariable. Die charakteristische Funktion χ_X von X ist definiert als*

$$\chi_X(t) = E(e^{itX}) = \begin{cases} \displaystyle\sum_{x_k} P(X = x_k)\, e^{itx_k} & X \text{ diskret mit Werten } x_k \\[2ex] \displaystyle\int_{-\infty}^{\infty} f(x)\, e^{itx} dx & X \text{ stetig mit Dichte } f \end{cases}$$

4.2.2 Diskreter Fall

Für eine diskrete Zufallsvariable X mit Werten $x_1, x_2, \ldots$ gilt also

$$\begin{aligned}\chi_X(t) &= P(X = x_1)\, e^{itx_1} + P(X = x_2)\, e^{itx_2} + \ldots \\ &= \sum_{x_k} P(X = x_k)\, e^{itx_k}\end{aligned}$$

d.h. alle möglichen Ausprägungen mit ihren Wahrscheinlichkeiten sind durch die charakteristische Funktion codiert. Damit ist X eindeutig durch $\chi_X(t)$ bestimmt.

Beispiel 4.1 *Sei $X \sim B(n, p)$ eine binomialverteilte Zufallsvariable. Dann gilt:*

$$\begin{aligned}\chi_X(t) &= \sum_{k=0}^{n} P(X = k)\, e^{itk} \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\, e^{itk} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k (1-p)^{n-k} \\ &= (p\, e^{it} + 1 - p)^n.\end{aligned}$$

In der letzten Zeile haben wir den binomischen Lehrsatz angewendet, der es uns erlaubt, die Summe kompakt darzustellen. Das wird bei diskreten Zufallsvariablen oft gelingen.

Erinnerung (Binomischer Lehrsatz):

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

gilt für alle $a, b \in \mathbb{R}$ (und sogar für alle $a, b \in \mathbb{C}$).

4.2.3 Stetiger Fall

Ist X eine stetige Zufallsvariable mit Dichtefunktion $f = f_X$, so gilt nach Definition

$$\chi_X(t) = \int_{-\infty}^{\infty} f_X(x)\, e^{itx} dx.$$

Mindestens an den Stellen, in denen f stetig ist, gilt umgekehrt:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \chi_X(t)\, e^{-itx} dt$$

Die Wahrscheinlichkeitsdichte (und damit auch die Zufallsvariable selbst) ist also auch hier durch die charakteristische Funktion bestimmt.

Beispiel 4.2 *Für die charakteristische Funktion der standardnomalverteilten Zufallsvariablen $Y \sim N(0,1)$ gilt:*

$$\begin{aligned}
\chi_Y(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, e^{itx} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx-x^2/2}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2-2itx+(-t)^2+t^2)}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}((x-it)^2+t^2)}\, dx \\
&= \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}t^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-it)^2}\, dx \\
&= \ldots \\
&= e^{-\frac{1}{2}t^2}
\end{aligned}$$

Der konkrete und ausführliche Beweis (d.h. der fehlende Teil mit den ...) kann hier leider nicht geführt werden. Dazu müssten wir uns zunächst mit der Theorie der komplex differenzierbaren Funktionen und insbesondere mit Wegintegralen in der komplexen Ebene beschäftigen. Dieses spannende und sehr schöne mathematische Gebiet wird als **Funktionentheorie** *bezeichnet.*

Beispiel 4.3 *Für die charakteristische Funktion der normalverteilten Zufallsvariablen $Y \sim N(\mu, \sigma^2)$ gilt:*

$$\chi_Y(t) = e^{it\mu} e^{-\frac{\sigma^2 t^2}{2}}.$$

Ausblick: Fourier-Analysis Der obige Zusammenhang (stetiger Fall) zwischen den beiden Funktionen f_x und χ_X ist eine Folgerung aus dem Umkehr- und Eindeutigkeitssatz der so genannten Fourier-Analysis. Im Jahre 1807 behauptete der französische Ingenieur *J.B. Fourier* (1786-1830), dass jede T-periodische Funktion f (d.h. $f(x+T) = f(x)$, T bezeichnet die Periodenlänge) durch eine Reihe der Gestalt

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(n\omega t) + b_n \sin(n\omega t)) \quad = \quad \sum_{n=-\infty}^{\infty} c_n \, e^{in\omega t}$$

dargestellt werden kann. Dabei sind die Koeffizienten $a_n, b_n \in \mathbb{R}$ bzw. $c_n \in \mathbb{C}$ natürlich geschickt zu wählen und es gilt $\omega = 2\pi/T$. Die Reihe auf der linken Seite wird als Fourier-Reihe in Sinus-Kosinus Darstellung und die auf der rechten Seite als Fourier-Reihe in komplexer Darstellung bezeichnet.

Nach der Klärung, was man überhaupt genau unter dieser Behauptung von Fourier verstehen muss (die obige Reihe sollte wohl in irgendeinem Sinne gegen f konvergieren), konnte *P. L. Dirichlet* im Jahre 1829 einen exakten Beweis der Behauptung Fouriers liefern.

Heute zählt die Fourier-Analysis zu einem wichtigen Hilfsmittel in den Naturwissenschaften, gerade bei der Untersuchung periodischer Phänomene.

Fourier hat in seinen Formeln für die Entwicklung T-periodischer Funktionen in Reihen die Grenzwertbildung $T \to \infty$ versucht, um die Fourier-Analysis auch auf nicht-periodische Funktionen auszudehnen. Dieser Grenzübergang führt dann nicht mehr zu einer unendlichen Reihe als Repräsentant der Funktion, sondern zu einer neuen (Integral)funktion.

Definition 4.3 *Sei $f : \mathbb{R} \to \mathbb{C}$ eine Funktion. Dann heisst*

$$F(\omega) \quad := \quad \mathcal{F}\{f(x)\} \; = \; \int_{-\infty}^{\infty} e^{-i\omega x} \, f(x) \, dx$$

(falls das uneigentliche Integral für alle $\omega \in \mathbb{R}$ existiert) die Fourier-Transformierte bzw. Spektralfunktion von f.

Die inverse Fouriertransformierte (irgend)einer Funktion $F(\omega)$ ist

$$\mathcal{F}^{-1}\{F(\omega)\} \quad := \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega x} \, F(\omega) \, d\omega$$

(falls das uneigentliche Integral für alle $\omega \in \mathbb{R}$ existiert).

Beispiel 4.4 *Die Rechtecksfunktion (die auch eine Wahrscheinlichkeitsdichte ist)*

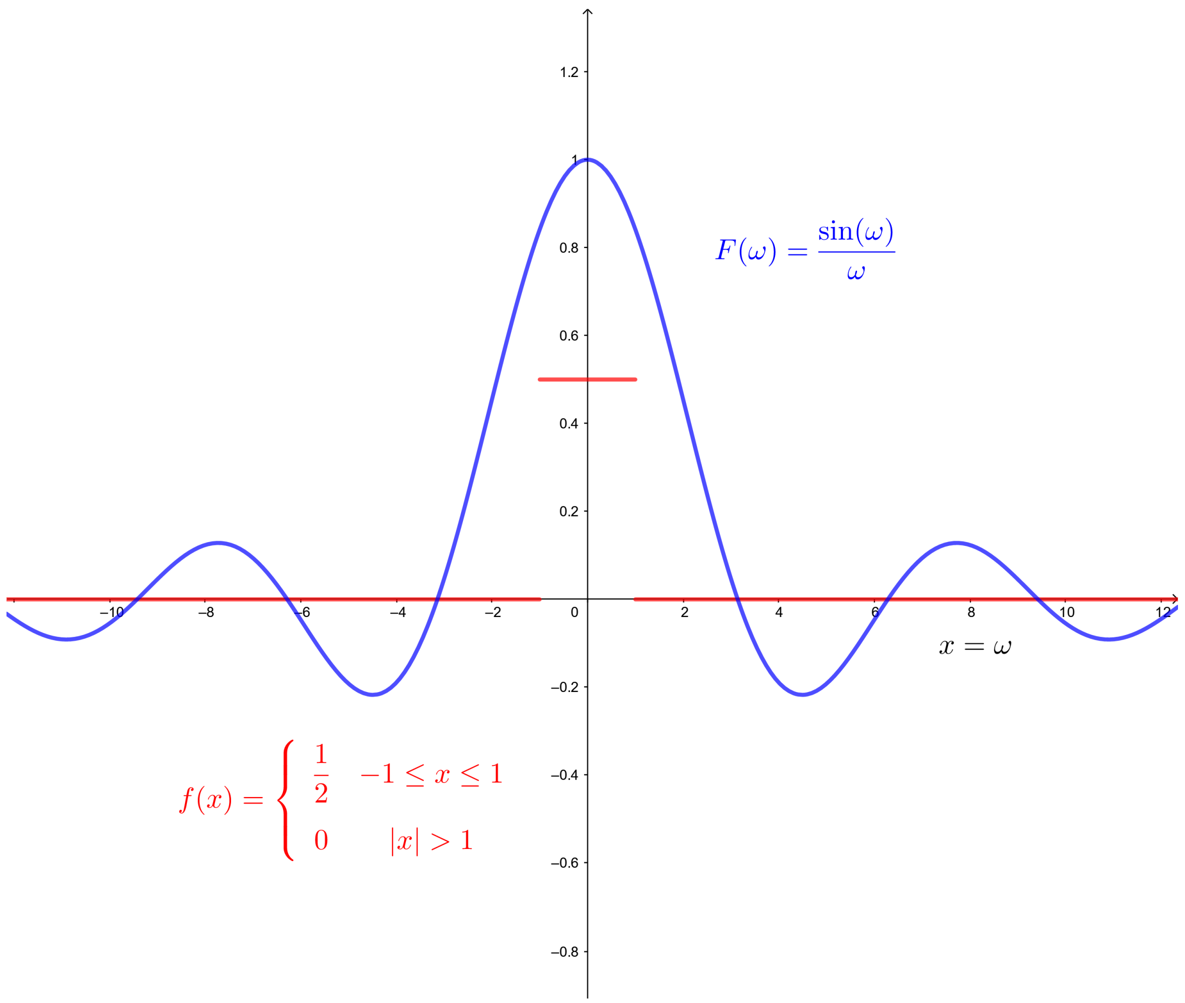
$$f(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & |x| > 1 \end{cases}$$

hat falls $\omega \neq 0$ die Spektralfunktion

$$\begin{aligned} F(\omega) \;=\; \mathcal{F}\{f(x)\} &= \int_{-\infty}^{\infty} e^{-i\omega x}\, f(x)\, dx \;=\; \int_{-1}^{1} e^{-i\omega x}\, \frac{1}{2}\, dx \\ &= \left[\frac{1}{2} - \frac{1}{i\omega} e^{-i\omega x}\right]_{x=-1}^{x=1} \\ &= \frac{1}{\omega}\,\frac{1}{2i}(e^{i\omega} - e^{-i\omega}) \\ &= \frac{1}{\omega}\,\sin(\omega) \end{aligned}$$

Für die letzte Umformung benötigt man das Wissen um die Zusammenhänge zwischen den Winkelfunktionen und der Exponentialfunktion. Tatsächlich gilt für alle komplexen Zahlen (also auch für alle reellen Zahlen) ω die wunderschöne **Euler-Gleichung**

$$\boxed{e^{i\omega} \;=\; \cos(\omega) + i\sin(\omega).}$$



Für die inverse Fourier-Transformierte der Funktion gilt:

$$\begin{aligned}\mathcal{F}^{-1}\left\{\frac{1}{\omega}\,\sin(\omega)\right\} &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega x}\,\frac{\sin(\omega)}{\omega}\,dx \\ &= \ldots \\ &= \begin{cases} \frac{1}{2} & -1 < x < 1 \\ \frac{1}{4} & |x| = 1 \\ 0 & |x| > 1 \end{cases}\end{aligned}$$

Mit Ausnahme der beiden Sprungstellen von f gilt also $\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} = f(x)$, zumindest in diesem Beispiel.

Und tatsächlich kann Folgendes bewiesen werden:

Satz 8 (Umkehr- und Eindeutigkeitssatz) *Ist die Funktion $f : \mathbb{R} \to \mathbb{C}$*

- *absolut integrierbar, d.h. f ist auf jedem endlichen Intervall stückweise stetig mit*

$$\int_{-\infty}^{\infty} |f(x)|\,dx < \infty,$$

- *in jedem endlichen Intervall stückweise differenzierbar und*
- *gilt die so genannte Mittelwerteigenschaft*

$$f(t) = \frac{1}{2}\left[\lim_{\tau\to t-} f(\tau) + \lim_{\tau\to t+} f(\tau)\right]$$

  für alle $t \in \mathbb{R}$

dann ist

$$\boxed{\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} = f(x)}$$

für alle $x \in \mathbb{R}$, d.h. die Korrespondenz $f(x) \leftrightarrow F(\omega)$ ist bijektiv.

4.3 Der Beweis

Sicher können wir zunächst den Ausdruck Y_n wie folgt umformen:

$$\begin{aligned} Y_n &= \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} \\ &= \frac{1}{\sqrt{n}} \frac{X_1 - \mu + X_2 - \mu + \ldots + X_n - \mu}{\sigma} \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{X_j - \mu}{\sigma} \\ &= \sum_{j=1}^{n} \frac{U_j}{\sqrt{n}}. \end{aligned}$$

mit den neuen Zufallsvariablen $U_j := \dfrac{X_j - \mu}{\sigma}$. Es ist leicht zu sehen, dass alle U_j standardisiert sind und eine unabhängige Folge (wie die X_j) bilden.
Für die charakteristischen Funktionen der Zufallsvariablen $U_j/\sqrt{n}$ gilt:

$$\begin{aligned} & \chi_{U_j/\sqrt{n}}(t) \\ = \ & E\left(e^{itU_j/\sqrt{n}}\right) \\ = \ & \sum_{k=0}^{\infty} \frac{1}{k!}\, i^k\, t^k\, E\left((U_j/\sqrt{n})^k\right) \\ = \ & \sum_{k=0}^{\infty} \frac{1}{k!}\, i^k\, t^k \left(\frac{1}{\sqrt{n}}\right)^k E\left((U_j)^k\right) \\ = \ & \frac{1}{0!}\, i^0\, t^0 \left(\frac{1}{\sqrt{n}}\right)^0 E\left((U_j)^0\right) + \frac{1}{1!}\, i^1\, t^1 \left(\frac{1}{\sqrt{n}}\right)^1 E\left((U_j)^1\right) + \frac{1}{2!}\, i^2\, t^2 \left(\frac{1}{\sqrt{n}}\right)^2 E\left((U_j)^2\right) \\ & + \sum_{k=3}^{\infty} \frac{1}{k!}\, i^k\, t^k \left(\frac{1}{\sqrt{n}}\right)^k E\left((U_j)^k\right) \\ = \ & 1 + it0 - \frac{1}{2} t^2 \left(\frac{1}{\sqrt{n}}\right)^2 + \sum_{k=3}^{\infty} \frac{1}{k!}\, i^k \left(\frac{t}{\sqrt{n}}\right)^k E\left((U_j)^k\right) \\ = \ & 1 - \frac{t^2}{2n} + \underbrace{\sum_{k=3}^{\infty} \frac{1}{k!}\, i^k \left(\frac{t}{\sqrt{n}}\right)^k E\left((U_j)^k\right)}_{=:R(t/\sqrt{n})} \end{aligned}$$

Beachten Sie, dass jeder Summand im Restglied R der Summe einen Grad von mindestens 3 in t hat. Somit gilt sicher:

$$\lim_{t \to 0} \frac{R(t/\sqrt{n})}{t^2} = 0.$$

Für die charakteristischen Funktionen der Zufallsvariablen Y_n gilt dann:

$$\begin{aligned}
\chi_{Y_n}(t) &= E\left(e^{itY_n}\right) \\
&= E\left(e^{it\sum_{j=1}^n \frac{U_j}{\sqrt{n}}}\right) \\
&= E\left(e^{it\frac{U_1}{\sqrt{n}}}\, e^{it\frac{U_2}{\sqrt{n}}} \cdots e^{it\frac{U_n}{\sqrt{n}}}\right) \\
&= E\left(e^{it\frac{U_1}{\sqrt{n}}}\right)\, E\left(e^{it\frac{U_2}{\sqrt{n}}}\right) \cdots E\left(e^{it\frac{U_n}{\sqrt{n}}}\right) \\
&= \chi_{U_1/\sqrt{n}}(t)\, \chi_{U_2/\sqrt{n}}(t) \cdots \chi_{U_n/\sqrt{n}}(t) \\
&= \left[1 - \frac{t^2/2}{n} + R(t/\sqrt{n})\right]^n,
\end{aligned}$$

wobei wir die Multiplikativität der Erwartungswerte für unabhängige Zufallsvariablen ausgenutzt haben.

Wir wissen bereits aus dem Kurs Mathematik 1, dass

$$\lim_{n\to\infty} \left[1 - \frac{t^2/2}{n}\right]^n = e^{-t^2/2}$$

gilt und das wäre hier auch unser Wunschresultat. Leider stört der Term $R(t/\sqrt{n})$. Glücklicherweise gilt aber auch das folgende Resultat (vgl. Königsberger, Analysis 1):

Satz 9 *Sei w_n eine Zahlenfolge mit $\lim_{n\to\infty} w_n = w$. Dann gilt*

$$\lim_{n\to\infty} \left[1 + \frac{w_n}{n}\right]^n = e^w.$$

Sicher gilt nun zunächst

$$1 - \frac{t^2/2}{n} + R(t/\sqrt{n}) = 1 + \frac{-t^2/2 + nR(t/\sqrt{n})}{n}$$

und

$$\lim_{n\to\infty} \left[-t^2/2 + nR(t/\sqrt{n})\right] = -\frac{t^2}{2},$$

da jeder Summand in $R(t/\sqrt{n})$ einen Term $(1/\sqrt{n})^k = 1/n^{k/2}$ mit $k \geq 3$ enthält.

Damit haben wir nun letztendlich gezeigt, dass

$$\lim_{n\to\infty} \chi_{Y_n}(t) = \lim_{n\to\infty} \left[1 - \frac{t^2/2}{n} + R(t/\sqrt{n})\right]^n = e^{-t^2/2}$$

gilt. Also konvergieren die charakteristischen Funktionen χ_{Y_n} gegen die charakteristische Funktion $\chi_Y(t) = e^{-t^2/2}$ der Standardnormalverteilung $Y \sim N(0,1)$. Daraus folgt, dass auch die Zufallsvariablen Y_n in Verteilung gegen die Zufallsvariable Y konvergieren. □

5 Übungsaufgaben

1. $X_1, \ldots, X_n$ seien unabhängige, identisch verteilte Zufallsvariablen mit Erwartungswert μ und Varianz σ^2. Wir betrachten die Zufallsvariablen

$$\begin{aligned} \overline{X} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\ Z_i &= \frac{X_i - \mu}{\sigma} \quad \text{für } i = 1, 2, \ldots n. \end{aligned}$$

   (a) Bestimmen Sie (nochmals) den Erwartungswert von $\bar{X}$.
   (b) Bestimmen Sie (nochmals) die Varianz von $\bar{X}$.
   (c) Bestimmen Sie den Erwartungswert von Z_i.
   (d) Bestimmen Sie die Varianz von Z_i.

2. Die folgenden Ausdrücke sollen berechnet werden.

$$\begin{aligned} P_1 &= \sum_{k=0}^{100} \binom{500}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{500-k}, \\ P_2 &= \sum_{k=100}^{120} \binom{500}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{500-k}, \\ P_3 &= \sum_{k=120}^{160} \binom{500}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{500-k}. \end{aligned}$$

   (a) Deuten Sie die Zahlen P_1, P_2 und P_3 als Wahrscheinlichkeiten der Form $P(a \leq X \leq b)$ mit einer passenden Zufallsvariablen X. Welche Verteilung hat X?
   (b) Können Sie P_1, P_2 und P_3 ,,exakt,, (z.B. mit Ihrem Taschenrechner) bestimmen? Tun Sie das.
   (c) Bestimmen Sie P_1, P_2 und P_3 mit Hilfe der Approximation
      i. über die Normalverteilung mit Randkorrektur und
      ii. über die Normalverteilung ohne Randkorrektur.

3. Die Wahrscheinlichkeit, dass ein Brennelement in einem Kernreaktor den Bedingungen einer Qualitätsprüfung nicht genügt, beträgt $p = 0.0002$. Wie gross ist die Wahrscheinlichkeit, dass höchstens 2 von 5000, genau eins von 1000 bzw. keines von 100 dieser Brennelemente die Qualitätsbedingungen nicht erfüllen? Rechnen Sie exakt und mit Hilfe der Poissonapproximation.

4. Seien $X_1, X_2, \ldots$ unabhängige und identisch (zweipunkt)verteilte Zufallsvariablen mit

$$X_i = \begin{cases} 0 & \text{mit } P(X_i = 0) = 1/3 \\ 2 & \text{mit } P(X_i = 2) = 2/3 \end{cases}$$

   und sei $S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$.

   (a) Bestimmen Sie die Verteilung der Zufallsvariablen S_3.
   (b) Bestimmen Sie Erwartungswert und Varianz der Zufallsvariablen S_n.
   (c) Was kann man über die Verteilung der Zufallsvariablen S_n für sehr grosse n aussagen? Begründung!