KAPITEL 12

Analytische Methoden

Es seien unabhängige Zufallsvariablen $X_1, \ldots, X_n$ gegeben. Die Verteilungen dieser Zufallsvariablen seien bekannt. Wie bestimmt man dann die Verteilung der Summe $X_1 + \ldots + X_n$? Man kann die Faltungsformel benutzen, jedoch ist diese ziemlich kompliziert, besonders dann, wenn n groß ist. In diesem Kapitel werden wir Methoden einführen, die dieses Problem auf eine viel elegantere Weise lösen, als die Faltungsformel.

12.1. Erzeugende Funktion

Definition 12.1.1. Sei X eine Zufallsvariable mit Werten in $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. Wir schreiben $p_n = \mathbb{P}[X = n]$ für $n \in \mathbb{N}_0$. Dann heißt

$$g_X(z) \overset{def}{=} \mathbb{E}\left[z^X\right] = \sum_{n=0}^{\infty} p_n z^n$$

die *erzeugende Funktion* von X.

Satz 12.1.2. *Die obige Reihe konvergiert absolut für $z \in [-1, 1]$ und es gilt $|g_X(z)| \leq 1$ für alle $z \in [-1, 1]$. Außerdem gilt $g_X(1) = 1$.*

Bemerkung 12.1.3. Die erzeugende Funktion ist also wohldefiniert im Intervall $[-1, 1]$. Es ist bekannt, dass die Summe einer Taylor–Reihe eine unendlich oft differenzierbare Funktion im Konvergenzbereich der Reihe ist. Somit ist $g_X(z)$ unendlich oft differenzierbar. Eigentlich kann man $g_X(z)$ sogar für komplexe Werte von z betrachten, dann ist g_X sogar eine analytische Funktion im Einheitskreis $|z| \leq 1$.

Beweis von Satz 12.1.2. Sei $z \in [-1, 1]$. Um die absolute Konvergenz zu zeigen, betrachten wir die Reihe

$$\sum_{n=0}^{\infty} |p_n z^n| \leq \sum_{n=0}^{\infty} p_n = 1.$$

Somit ist die Reihe absolut konvergent. Es gilt

$$|g_X(z)| = \left|\sum_{n=0}^{\infty} p_n z^n\right| \leq \sum_{n=0}^{\infty} |p_n z^n| \leq \sum_{n=0}^{\infty} p_n = 1.$$

Außerdem gilt $g_X(1) = \sum_{n=0}^{\infty} p_n = 1$. □

Satz 12.1.4. *Für jedes $n \in \mathbb{N}_0$ gilt*

$$\mathbb{P}[X = n] = \frac{g_X^{(n)}(0)}{n!}.$$

Bemerkung 12.1.5. Aus diesem Satz folgt, dass die Verteilung einer Zufallsvariable eindeutig durch die erzeugende Funktion definiert ist. D. h.: sind X und Y zwei Zufallsvariablen mit $g_X(z) = g_Y(z)$ für alle $z \in [-1,1]$, so sind die Verteilungen von X und Y gleich, d.h. es gilt $\mathbb{P}[X = n] = \mathbb{P}[Y = n]$ für alle $n \in \mathbb{N}_0$. Insbesondere sind alle Charakteristika einer Zufallsvariable, wie beispielsweise der Erwartungswert oder die Varianz, in der erzeugenden Funktion versteckt.

Beweis von Satz 12.1.4. Wir können eine konvergente Taylor–Reihe termweise n-mal ableiten:

$$g_X^{(n)}(z) = \left(\sum_{k=0}^{\infty} p_k z^k\right)^{(n)} = \sum_{k=0}^{\infty} p_k (z^k)^{(n)}.$$

Dabei ist

$$(z^k)^{(n)} = \begin{cases} 0, & k < n, \\ n!, & k = n, \\ k(k-1)\dots(k-n+1)z^{k-n}, & k > n. \end{cases}$$

Indem wir nun $z = 0$ einsetzen, erhalten wir $g_X^{(n)}(0) = n! p_n$. □

Warum betrachten wir überhaupt erzeugende Funktionen? Weil sie uns erlauben, die Faltung in ein Produkt umzuwandeln. Das wird im nächsten Satz gezeigt.

Satz 12.1.6. *Seien X, Y unabhängige Zufallsvariablen mit Werten in $\mathbb{N}_0$. Dann gilt:*

$$g_{X+Y}(z) = g_X(z) \cdot g_Y(z), \quad z \in [-1,1].$$

Erster Beweis. Mit der Definition der erzeugenden Funktion erhalten wir

$$g_{X+Y}(z) = \mathbb{E}\left[z^{X+Y}\right] = \mathbb{E}\left[z^X \cdot z^Y\right] = \mathbb{E}\left[z^X\right] \cdot \mathbb{E}\left[z^Y\right] = g_X(z) \cdot g_Y(z),$$

wobei wir benutzt haben, dass X und Y (und somit auch z^X und z^Y) unabhängig sind. □

Zweiter Beweis. Wir wählen die folgende Notation:

$$p_n = \mathbb{P}[X = n] \text{ und } q_n = \mathbb{P}[Y = n].$$

Mit der Faltungsformel erhalten wir dann

$$r_n := \mathbb{P}[X + Y = n] = \sum_{k=0}^{n} p_k q_{n-k}.$$

Nun multiplizieren wir die erzeugenden Funktionen:

$$g_X(z) g_Y(z) = \left(\sum_{k=0}^{\infty} p_k z^k\right)\left(\sum_{l=0}^{\infty} q_l z^l\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} p_k q_{n-k}\right) z^n = \sum_{n=0}^{\infty} r_n z^n = g_{X+Y}(z).$$

□

Beispiel 12.1.7. Sei $X \sim \text{Poi}(\lambda)$. Dann gilt: $g_X(z) = e^{\lambda(z-1)}$.

Beweis. $X \sim \mathrm{Poi}(\lambda)$ heißt, dass $p_n = \mathbb{P}[X = n] = e^{-\lambda}\frac{\lambda^n}{n!}$ für $n \in \mathbb{N}_0$. Somit gilt

$$g_X(z) = \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}z^n = e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda z)^n}{n!} = e^{-\lambda}e^{\lambda z} = e^{\lambda(z-1)}.$$

Beispiel 12.1.8. Seien $X_1 \sim \mathrm{Poi}(\lambda_1)$ und $X_2 \sim \mathrm{Poi}(\lambda_2)$ unabhängige Zufallsvariablen. Dann gilt:

$$X_1 + X_2 \sim \mathrm{Poi}(\lambda_1 + \lambda_2).$$

Beweis. Wir berechnen die Verteilung von $X_1 + X_2$ nicht direkt mit der Faltungsformel, sondern benutzen die erzeugende Funktion. Die erzeugende Funktion von $X_1 + X_2$ ist

$$\begin{aligned} g_{X_1+X_2}(z) &= g_{X_1}(z) \cdot g_{X_2}(z), \quad \text{da } X_1, X_2 \text{ unabhängig} \\ &= e^{\lambda_1(z-1)} \cdot e^{\lambda_2(z-1)}, \quad \text{da } X_i \sim \mathrm{Poi}(\lambda_i) \\ &= e^{(\lambda_1+\lambda_2)(z-1)}. \end{aligned}$$

Auf der rechten Seite erkennen wir die erzeugende Funktion einer $\mathrm{Poi}(\lambda_1 + \lambda_2)$-Verteilung. Die erzeugende Funktion von $X_1 + X_2$ stimmt also mit der erzeugenden Funktion einer $\mathrm{Poi}(\lambda_1 + \lambda_2)$-Verteilung überein. Da nun die erzeugende Funktion einer Zufallsvariable ihre Verteilung eindeutig bestimmt, muss gelten: $X_1 + X_2 \sim \mathrm{Poi}(\lambda_1 + \lambda_2)$. □

Wie zuvor schon erwähnt, steckt die gesamte Verteilung von X, insbesondere der Erwartungswert und die Varianz, in der erzeugenden Funktion. Nun zeigen wir, wo der Erwartungswert und die Varianz versteckt sind.

Definition 12.1.9. Sei X eine Zufallsvariable. Dann heißen die Zahlen

$$\mathbb{E}[X], \mathbb{E}[X^2], \ldots, \mathbb{E}[X^n], \ldots$$

Momente von X.

Satz 12.1.10. *Für jedes $n \in \mathbb{N}_0$ gilt*

$$\mathbb{E}[X(X-1)\ldots(X-n+1)] = g_X^{(n)}(1).$$

Bemerkung 12.1.11. Dies ist an sich kein Moment, aber X^n steckt in diesem Erwartungswert drin und wir können damit letztendlich $\mathbb{E}[X^n]$ rechnerisch isolieren.

Beispiel 12.1.12. Zwei Spezialfälle dieser Formel sind:

$$\begin{aligned} \mathbb{E}X &= g_X'(1), \quad n = 1, \\ \mathbb{E}[X(X-1)] &= g_X''(1), \quad n = 2. \end{aligned}$$

Damit lässt sich nun die Varianz berechnen:

$$\begin{aligned} \operatorname{Var} X &= \mathbb{E}[X^2] - (\mathbb{E}X)^2 \\ &= \mathbb{E}[X(X-1)] + \mathbb{E}X - (\mathbb{E}X)^2 \\ &= g_X''(1) + g_X'(1) - (g_X'(1))^2. \end{aligned}$$

Bemerkung 12.1.13. Die erzeugende Funktion ist im Allgemeinen nur für $|z| \leq 1$ definiert. Daher müssen wir uns die Ableitung der erzeugenden Funktion an den Stellen ± 1 als eine einseitige Ableitung vorstellen.

Beweis von Satz 12.1.10. Wir leiten die Taylor-Reihe $g_X(z)$ n-mal termweise ab:

$$g_X^{(n)}(z) = \sum_{k=0}^{\infty} p_k (z^k)^{(n)} = \sum_{k=n}^{\infty} p_k \cdot k(k-1)\dots(k-n+1)z^{k-n}.$$

Nun setzen wir $z = 1$ ein:

$$g_X^{(n)}(1) = \sum_{k=0}^{\infty} p_k \cdot k(k-1)\dots(k-n+1) = \mathbb{E}[X(X-1)\dots(X-n+1)].$$

Im letzten Schritt wurde die Transformationsformel für den Erwartungswert verwendet. $\square$

Beispiel 12.1.14. Sei $X \sim \operatorname{Poi}(\lambda)$. Wir haben bereits gezeigt, dass $g_X(z) = e^{\lambda(z-1)}$. Es gilt also für jedes $n \in \mathbb{N}$

$$\mathbb{E}[X(X-1)\cdot\ldots\cdot(X-n+1)] = g_X^{(n)}(1) = \left(e^{\lambda(z-1)}\right)^{(n)}\Big|_{z=1} = \lambda^n e^{\lambda(z-1)}\Big|_{z=1} = \lambda^n.$$

Damit kann man den Erwartungswert und die Varianz von X berechnen:

$$\begin{aligned} \mathbb{E}X &= g_X'(1) = \lambda, \\ \operatorname{Var} X &= g_X''(1) + g_X'(1) - (g_X'(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \end{aligned}$$

12.2. Summen mit einer zufälligen Anzahl von Summanden

Beispiel 12.2.1. Einer Versicherung werden N Schäden gemeldet. Dabei sei N eine Zufallsvariable mit Werten in $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Die einzelnen Schadenhöhen seien ebenfalls zufällig und mit $X_1, X_2, \dots$ bezeichnet. Der Gesamtschaden beläuft sich also auf

$$S = X_1 + \ldots + X_N.$$

Dabei besteht die Summe S aus einer zufälligen Anzahl (nämlich N) an Summanden. Die Summanden sind ebenfalls zufällig. Wie bestimmt man die Verteilung von S?

Satz 12.2.2. *Die folgenden Bedingungen seien erfüllt:*

(1) $N, X_1, X_2, \dots$ *sind unabhängige Zufallsvariablen mit Werten in* $\mathbb{N}_0$.
(2) $X_1, X_2, \dots$ *sind identisch verteilt.*

Dann lässt sich die erzeugende Funktion der Summe $S = X_1 + \ldots + X_N$ *wie folgt berechnen:*

$$g_S(z) = g_N(g_{X_1}(z)), \quad z \in [-1, 1].$$

Bemerkung 12.2.3. Ist $N = n$ konstant (und nicht zufällig), so erhalten wir die Formel $g_S(z) = (g_{X_1}(z))^n$. Diese Formel folgt direkt aus Satz 12.1.6, denn

$$g_S(z) = g_{X_1+\ldots+X_n}(z) = g_{X_1}(z)\cdot\ldots\cdot g_{X_n}(z) = (g_{X_1}(z))^n.$$

Beweis. Wir benutzen die Notation $p_n = \mathbb{P}[N = n]$ für $n \in \mathbb{N}_0$. Nun verwenden wir die Formel von der totalen Wahrscheinlichkeit:

$$\begin{aligned}
g_S(z) &= \sum_{k=0}^{\infty} \mathbb{P}[S = k] \cdot z^k \\
&= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \mathbb{P}[X_1 + \ldots + X_N = k | N = n] \cdot p_n \right) \cdot z^k \\
&= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \mathbb{P}[X_1 + \ldots + X_n = k | N = n] \cdot p_n \right) \cdot z^k.
\end{aligned}$$

Da nun das Ereignis $\{N = n\}$ vom Ereignis $\{X_1 + \ldots + X_n = k\}$ unabhängig ist, erhalten wir

$$\begin{aligned}
g_S(z) &= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \mathbb{P}[X_1 + \ldots + X_n = k] \cdot p_n \right) \cdot z^k \\
&= \sum_{n=0}^{\infty} p_n \cdot \left(\sum_{k=0}^{\infty} z^k \cdot \mathbb{P}[X_1 + \ldots + X_n = k] \right) \\
&= \sum_{n=0}^{\infty} p_n \cdot g_{X_1 + \ldots + X_n}(z).
\end{aligned}$$

Nun sind die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig und identisch verteilt, somit erhalten wir

$$g_S(z) = \sum_{n=0}^{\infty} p_n (g_{X_1}(z))^n = g_N(g_{X_1}(z)).$$

□

Korollar 12.2.4 (Wald-Identität). *Unter den gleichen Voraussetzungen wie in Satz 12.2.2 gilt*

$$\mathbb{E}[X_1 + \ldots + X_N] = \mathbb{E}[N] \cdot \mathbb{E}[X_1].$$

Beweis. Übung. □

Bemerkung 12.2.5. Im Beispiel mit der Versicherung kann man den erwarteten Gesamtschaden berechnen, indem man die erwartete Anzahl an gemeldeten Schäden mit der zu erwarteten jeweiligen Schadenshöhe multipliziert.

12.3. Verzweigungsprozesse

Einen *Verzweigungsprozess* (auch *Galton–Watson Prozess* genannt) kann man sich als ein Modell für eine Kettenreaktion vorstelllen. Hier folgt die Beschreibung dieses Modells. In Generation 0 gibt es 1 Teilchen. Dieses Teilchen erzeugt eine zufällige Anzahl Töchterteilchen, die zu Generation 1 gehören. Jedes Teilchen in Generation 1 erzeugt eine zufällige Anzahl Töchterteilchen, die zu Generation 2 gehören, usw. Dabei wird vorausgesetzt, dass sich alle

Teilchen unabhängig voneinander verhalten. (Die Anzahl der Töchterteilchen, die ein Teilchen produziert, ist also unabhängig davon, wieviele Töchterteilchen die anderen Teilchen produzieren). Wir bezeichnen mit p_k die Wahrscheinlichkeit, dass ein Teilchen k Töchter erzeugt. Wir nehmen an, dass sich diese Wahrscheinlichkeiten nicht von Teilchen zu Teilchen ändern.
Nun geben wir eine präzisere Beschreibung des Modells. Es sei eine Zahlenfolge $p_0, p_1, \ldots$ gegeben, so dass

(1) $p_0, p_1, \ldots \geq 0$.
(2) $p_0 + p_1 + \ldots = 1$.

Es seien $X_{i,n}$, $i \in \mathbb{N}$, $n \in \mathbb{N}_0$, unabhängige identisch verteilte Zufallsvariablen mit $\mathbb{P}[X_{i,n} = k] = p_k$. Dabei soll $X_{i,n}$ die Anzahl der Töchterteilchen des Teilchens i in Generation n bezeichnen. Definiere nun Zufallsvariablen $Z_0, Z_1, \ldots$ induktiv durch $Z_0 = 1$ und

$$Z_{n+1} = \sum_{i=0}^{Z_n} X_{i,n}. \tag{12.3.1}$$

Somit ist Z_n die Anzahl der Teilchen in Generation n. Wie bestimmt man nun die Verteilung von Z_n?

Satz 12.3.1. *Sei $g(t) = \sum_{k=0}^{\infty} p_k z^k$. Dann gilt für die erzeugende Funktion von Z_n:*

$$g_{Z_n}(t) = g(g(\ldots g(t)\ldots)) \qquad \text{(n-mal)}.$$

Beweis. Für $n = 0$ gilt $g_{Z_0}(t) = t$, denn $Z_0 = 1$. Die erzeugende Funktion von $X_{i,n}$ ist g. Wendet man nun Satz 12.2.2 auf die Formel (12.3.1) an, so erhält man

$$g_{Z_{n+1}}(t) = g_{Z_n}(g(t)).$$

Wendet man das induktiv an, so erhält man die Aussage des Satzes. □

Wir können den Erwartungswert von Z_n bestimmen.

Satz 12.3.2. *Sei $\mu := \mathbb{E}[X_{1,0}] = \sum_{k=0}^{\infty} p_k k$. Dann gilt*

$$\mathbb{E}[Z_n] = \mu^n.$$

Bemerkung 12.3.3. Diese Formel ist nicht überraschend. Jedes Teilchen erzeugt im Durchschnitt μ Töchterteilchen für die nächste Generation. Jede Generation ist also im Durchschnitt μ-Mal so groß wie die vorherige. Daher gilt $\mathbb{E}[Z_n] = \mu^n$.

Beweis. Für $n = 0$ ergibt sich nach Definition $\mathbb{E}[Z_0] = 1 = \mu^0$. Wendet man die Wald-Identität auf die Formel (12.3.1) an, so erhält man

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[Z_n] \cdot \mathbb{E}[X_{1,n}] = \mathbb{E}[Z_n] \cdot \mu.$$

Wendet man das induktiv an, so erhält man die Aussage des Satzes. □

Definition 12.3.4. Die *Aussterbewahrscheinlichkeit* q ist definiert durch

$$q = \mathbb{P}[\exists n \in \mathbb{N} : Z_n = 0].$$

Wie lässt sich diese Aussterbewahrscheinlichkeit berechnen? Zunächst einmal gibt es einen trivialen Fall: Ist $p_0 = 0$, so erzeugt jedes Teilchen mindestens ein Töchterteilchen und der Verzweigungsprozess kann nicht aussterben. Somit gilt in diesem Fall $q = 0$. Im nächsten Satz betrachten wir den Fall $p_0 > 0$.

Satz 12.3.5. *Sei $p_0 > 0$. Dann gilt:*

(1) *Ist $\mu < 1$ (der subkritische Fall), so gilt $q = 1$.*
(2) *Ist $\mu = 1$ (der kritische Fall), so gilt $q = 1$.*
(3) *Ist $\mu > 1$ (der superkritische Fall), so ist q die einzige Lösung der Gleichung $g(q) = q$ mit $q < 1$. Es gilt also $0 < q < 1$.*

Bemerkung 12.3.6. Im subkritischen Fall $\mu < 1$ ist jede Generation im Durchschnitt kleiner als die vorherige. Deshalb ist es nicht überraschend, dass der Verzweigungsprozess mit Wahrscheinlichkeit 1 aussirbt.

Bemerkung 12.3.7. Im kritischen Fall $\mu = 1$ ist jede Generation im Durchschnitt genauso groß, wie die vorherige. Es gilt also $\mathbb{E}[Z_n] = 1$ für jedes $n \in \mathbb{N}_0$. Dennoch stirbt der Prozess mit Wahrscheinlichkeit 1 aus. Das kann man sich intuitiv so vorstellen: eine Generation, die im Durchschnitt aus einem Teilchen besteht hat eine gewisse positive Wahrscheinlichkeit, kein einziges Töchterteilchen zu produzieren. Da diese Wahrscheinlichkeit nun zu jedem Zeitpunkt besteht, wird irgendwann tatsächlich kein einziges Töchterteilchen erzeugt und der Prozess stirbt aus. Im kritischen Fall gilt $\lim_{n\to\infty} Z_n = 0$ f.s. (weil Z_n mit Wahrscheinlichkeit 1 ab irgendwann gleich 0 ist). Dabei ist $\mathbb{E}[Z_n] = 1$ für jedes $n \in \mathbb{N}_0$. Hier sehen wir, dass nicht immer $\mathbb{E}[\lim_{n\to\infty} Z_n] = \lim_{n\to\infty} \mathbb{E}[Z_n]$ gelten muss.

Bemerkung 12.3.8. Im superkritischen Fall $\mu > 1$ ist jede Generation im Durchschnitt größer, als die vorherige. Die Aussterbewahrscheinlichkeit ist kleiner als 1. Es sei aber bemerkt, dass auch ein superkritischer Verzweigungsprozess mit positiver Wahrscheinlichkeit aussterben kann. Das geschieht zum Beispiel dann, wenn das allererste Teilchen keine Nachkommen produziert. Das passiert mit Wahrscheinlichkeit p_0, also ist $q \geq p_0$.

Beweis von Satz 12.3.5. Ist die n-te Generation ausgestorben, so sind auch alle nachfolgenden Generationen leer. Deshalb gelten folgende Inklusionen von Ereignissen:

$$\{Z_1 = 0\} \subset \{Z_2 = 0\} \subset \{Z_3 = 0\} \subset \dots .$$

Es sei q_n die Wahrscheinlichkeit, dass die n-te Generation leer ist:

$$q_n = \mathbb{P}[Z_n = 0] = g_{Z_n}(0) = g\left(g(\dots g(0))\dots\right) \qquad (n\text{-mal}).$$

Dabei ist $g(z) = \sum_{k=0}^{\infty} p_k z^k$ die erzeugende Funktion der Anzahl der Nachkommen eines Teilchens. Mit dem Satz über die Stetigkeit der Wahrscheinlichkeit gilt

$$q = \mathbb{P}\left[\cup_{n=1}^{\infty}\{Z_n = 0\}\right] = \lim_{n\to\infty} \mathbb{P}[Z_n = 0] = \lim_{n\to\infty} g\left(g(\dots g(0))\dots\right) \qquad (n\text{-mal}).$$

Der Rest des Beweises ergibt sich aus dem Bild. □

12.4. Momenterzeugende Funktion (Laplace-Transformierte)

Die erzeugende Funktion hat einen Nachteil: Sie kann nur für Zufallsvariablen, die ganzzahlige, nicht-negative Werte annehmen, definiert werden. Für Zufallsvariablen, die diese Bedingung nicht erfüllen, müssen wir etwas anderes definieren.

Definition 12.4.1. Sei X eine Zufallsvariable mit Werten in $\mathbb{R}$. Dann heißt

$$m_X(t) = \mathbb{E}[e^{tX}] \in (0, \infty], \quad t \in \mathbb{R},$$

die *Laplace-Transformierte* (oder die *momenterzeugende Funktion*) von X .

Beispiel 12.4.2. Sei X eine absolut stetige Zufallsvariable mit Dichte f_X. Dann ist die Laplace-Transformierte gegeben durch

$$m_X(t) = \int_{\mathbb{R}} e^{ty} f_X(y) dy.$$

Bemerkung 12.4.3. Es gilt immer $m_X(0) = 1$.

Bemerkung 12.4.4. Ein Nachteil der Laplace-Transformierten ist, dass sie auch den Wert ∞ annehmen kann. Es sei X z.B. Cauchy-verteilt, d.h. absolut stetig mit Dichte

$$f_X(y) = \frac{1}{\pi} \cdot \frac{1}{1+y^2}, \quad y \in \mathbb{R}.$$

Dann gilt für die Laplace-Transformierte:

$$m_X(t) = \int_{\mathbb{R}} e^{ty} \cdot \frac{1}{\pi} \cdot \frac{1}{1+y^2} dy = \begin{cases} +\infty, & t \neq 0, \\ 1, & t = 0. \end{cases}$$

Die Laplace-Transformierte ist also überall gleich $+\infty$ außer an der Stelle 0. Man kann auch andere Verteilungen konstruieren, die die gleiche Laplace-Transformierte haben. Sei z.B. X' absolut stetig mit Dichte $f_{X'}(t) = \frac{1}{2} y^{-2} \mathbb{1}_{|y|>1}$. Dann gilt $m_X(t) = m_{X'}(t)$ für alle $t \in \mathbb{R}$. Wir sehen, dass die Laplace-Transformierte die Verteilung einer Zufalsvariable nicht eindeutig bestimmt.

Im weiteren werden wir voraussetzen, dass die Laplace-Transformierte endlich in einem Intervall $(-\varepsilon, \varepsilon)$ ist. Im nächsten Satz berechnen wir die Momente von X mithilfe der Laplace-Transformierten.

Satz 12.4.5. *Sei X eine Zufallsvariable mit $m_X(t) < \infty$ für alle $t \in (-\varepsilon, \varepsilon)$ mit $\varepsilon > 0$. Dann gilt:*

$$\mathbb{E}[X^n] = m_X^{(n)}(0).$$

Beispiel 12.4.6. Für $n = 1$ und $n = 2$ ergibt sich

$$m'_X(0) = \mathbb{E}[X], \quad m''_X(0) = \mathbb{E}[X^2].$$

Daraus ergibt sich die Formel für die Varianz:

$$\operatorname{Var} X = m''_X(0) - (m'_X(0))^2.$$

Bemerkung 12.4.7. Dieser Satz erklärt auch die Bezeichnung “momenterzeugende Funktion” für m_X. Die Momente können als Ableitungen der momenterzeugenden Funktion an der Stelle 0 abgelesen werden.

Idee des Beweises von Satz 12.4.5. Wir betrachten die n-te Ableitung von $m_X(t)$:

$$m_X^{(n)}(t) = \left(\mathbb{E}[e^{tX}]\right)^{(n)} = \mathbb{E}\left[\left(e^{tX}\right)^{(n)}\right] = \mathbb{E}[X^n e^{tX}].$$

Setzen wir nun $t = 0$ ein, so erhalten wir $m_X^{(n)}(0) = \mathbb{E}[X^n]$. Allerdings haben wir hier den Erwartungswert mit dem Ableitungsoperator vertauscht. Dieser Schritt bedarf einer Begründung. Daher werden wir den Beweis auf eine andere Weise führen.

Beweis von Satz 12.4.5. Schritt 1. Wir entwickeln die Funktion e^{tX} in eine Taylor-Reihe:

$$e^{tX} = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} = \lim_{m\to\infty} \sum_{k=0}^{m} \frac{(tX)^k}{k!}, \qquad \text{wobei } S_m := \sum_{k=0}^{m} \frac{(tX)^k}{k!}.$$

Die Summen S_m können wie folgt abgeschätzt werden:

$$|S_m| = \left|\sum_{k=0}^{m} \frac{(tX)^k}{k!}\right| \leq \sum_{k=0}^{m} \frac{|tX|^k}{k!} = e^{|tX|} \leq S,$$

wobei $S = e^{tX} + e^{-tX}$. Für $t \in (-\varepsilon, +\varepsilon)$ gilt $\mathbb{E}S < \infty$ nach Voraussetzung des Satzes. Wir haben gezeigt, dass $\lim_{m\to\infty} S_m = e^{tX}$ fast sicher (sogar sicher) und $|S_m| \leq S$ mit $\mathbb{E}S < \infty$. Daher können rir den Satz von der majorisierten Konvergenz (Maßtheorie) anwenden:

$$\lim_{m\to\infty} \mathbb{E}[S_m] = \mathbb{E}[e^{tX}].$$

Schritt 2. Es gilt für jedes $t \in (-\varepsilon, \varepsilon)$:

$$m_X(t) = \mathbb{E}[e^{tX}] = \lim_{m\to\infty} \mathbb{E}[S_m] = \lim_{m\to\infty} \mathbb{E}\left[\sum_{k=0}^{m} \frac{(tX)^k}{k!}\right] = \lim_{m\to\infty} \sum_{k=0}^{m} \frac{t^k}{k!}\mathbb{E}[X^k] = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}[X^k].$$

Schritt 3. Dort, wo eine Taylor–Reihe konvergiert, kann man sie beliebig oft termweise ableiten. Es ergibt sich

$$m_X^{(n)}(t) = \sum_{k=n}^{\infty} \frac{k(k-1)\ldots(k-n+1)}{k!}\, t^{k-n}\mathbb{E}[X^k].$$

Setzt man in diese Formel $t = 0$ ein, so verschwinden alle Summanden bis auf den ersten (mit $k = n$). Daraus ergibt sich $m_X^{(n)}(0) = \mathbb{E}[X^n]$. □

Beispiel 12.4.8. Sei $X \sim \text{Poi}(\lambda)$, d.h. $\mathbb{P}[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}$ für $k \in \mathbb{N}_0$. Für die Laplace-Transformierte von X erhalten wir

$$m_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk}\mathbb{P}[X = k] = \sum_{k=0}^{\infty} e^{tk} \cdot e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t\lambda)^k}{k!} = e^{-\lambda} \cdot e^{e^t\lambda} = e^{\lambda(e^t-1)}.$$

Damit lässt sich der Erwartungswert berechnen:

$$\mathbb{E}X = m_X'(0) = \left(e^{\lambda(e^t-1)}\right)'|_{t=0} = \lambda e^t e^{\lambda(e^t-1)}|_{t=0} = \lambda.$$

Satz 12.4.9. *Seien X, Y unabhängige Zufallsvariablen. Dann gilt:*

$$m_{X+Y}(t) = m_X(t) \cdot m_Y(t).$$

Beweis. Mit der Definition der Laplace-Transformierten erhalten wir:

$$m_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX} \cdot e^{tY}] = \mathbb{E}[e^{tX}] \cdot \mathbb{E}[e^{tY}] = m_X(t) \cdot m_Y(t).$$

Dabei haben wir benutzt, dass X und Y (und somit auch e^{tX} und e^{tY}) unabhängig sind. □

12.5. Charakteristische Funktion (Fourier-Transformierte)

Ein Nachteil der Laplace-Transformierten ist, dass sie auch unendlich sein kann (siehe das Beispiel mit der Cauchy–Verteilung). Außerdem ist die Verteilung einer Zufallsvariable nicht immer eindeutig durch die Laplace-Transformierte festgelegt. Wir werden deshalb eine andere Transformation betrachten, die sich von der Laplace-Transformation dadurch unterscheidet, dass man im Exponenten e^{tX} noch die Komplexe Zahl $i = \sqrt{-1}$ hinzunimmt. Damit erreicht man, dass die Transformierte immer endlich ist.

Definition 12.5.1. Sei X eine Zufallsvariable. Dann heißt

$$\varphi_X(t) = \mathbb{E}[e^{itX}] \in \mathbb{C}, \quad t \in \mathbb{R},$$

die *charakteristische Funktion* (oder die *Fourier-Transformierte*) von X.

Bemerkung 12.5.2. Der Erwartungswert von e^{itX} ist

$$\mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX) + i\sin(tX)] = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)].$$

Die charakteristische Funktion $\varphi_X(t)$ existiert also für jedes X und jedes $t \in \mathbb{R}$, denn $|\cos(tX)| \leq 1$ und $|\sin(tX)| \leq 1$.

Bemerkung 12.5.3. Ist X diskret, mit den Werten $y_1, y_2, \ldots$ und den dazugehörigen Wahrscheinlichkeiten $p_1, p_2, \ldots$, so gilt für die charakteristische Funktion von X:

$$\varphi_X(t) = \mathbb{E}\left[e^{itX}\right] = \sum_{k=1}^{\infty} e^{ity_k} p_k.$$

Ist X absolut stetig mit Dichte f_X, so gilt

$$\varphi_X(t) = \mathbb{E}\left[e^{itX}\right] = \int_{-\infty}^{\infty} e^{ity} f_X(y) dy.$$

Beispiel 12.5.4. Nimmt X nur die zwei Werte $+1$ und -1 mit einer Wahrscheinlichkeit von jeweils $1/2$ an, so ist die charakteristische Funktion von X gegeben durch

$$\varphi_X(t) = \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \cos(t).$$

Satz 12.5.5. *Sei X eine Zufallsvariable. Dann hat die charakteristische Funktion φ_X die folgenden Eigenschaften:*

(1) $\varphi_X(0) = 1$ *und* $|\varphi_X(t)| \leq 1$ *für alle* $t \in \mathbb{R}$.
(2) $\varphi_X(-t) = \overline{\varphi_X(t)}$ *für alle* $t \in \mathbb{R}$.
(3) *Die Funktion* φ_X *ist gleichmäßig stetig auf* $\mathbb{R}$.

(4) *Die Funktion φ_X ist positiv semidefinit, d.h., für alle $t_1, \dots, t_n \in \mathbb{R}$ und für alle $c_1, \dots, c_n \in \mathbb{C}$ gilt:*

$$\sum_{k,j=1}^{n} c_k \bar{c}_j \cdot \varphi_X(t_k - t_j) \geq 0.$$

Mit anderen Worten: $(\varphi_X(t_k - t_j))_{k,j=1}^n$ ist eine positiv semidefinite Matrix.

Beweis von 1. Für $t = 0$ gilt $\varphi_X(0) = \mathbb{E}[e^{i \cdot 0 \cdot X}] = \mathbb{E}1 = 1$. Für ein beliebiges $t \in \mathbb{R}$ gilt:

$$|\varphi_X(t)| = \left|\mathbb{E}e^{itX}\right| \leq \mathbb{E}|e^{itX}| = \mathbb{E}1 = 1.$$

Beweis von 2. Wir betrachten die charakteristische Funktion an der Stelle $-t$:

$$\varphi_X(-t) = \mathbb{E}\left[e^{-itX}\right] = \mathbb{E}[\cos(tX) - i\sin(tX)] = \mathbb{E}[\cos(tX)] - i\mathbb{E}[\sin(tX)].$$

Auf der anderen Seite gilt

$$\overline{\varphi_X(t)} = \overline{\mathbb{E}\left[e^{itX}\right]} = \overline{\mathbb{E}[\cos(tX) + i\sin(tX)]} = \overline{\mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)]}.$$

Somit gilt $\varphi_X(-t) = \overline{\varphi_X(t)}$.

Beweis von 3. Seien $t, h \in \mathbb{R}$. Dann gilt

$$|\varphi_X(t+h) - \varphi_X(t)| = \left|\mathbb{E}[e^{i(t+h)X} - e^{itX}]\right| \leq \mathbb{E}\left|e^{itX}(e^{ihX} - 1)\right| = \mathbb{E}\left|e^{ihX} - 1\right| \overset{def}{=} g(h).$$

Nun müssen wir zeigen, dass $\lim_{h \to 0} g(h) = 0$. Zunächst bemerken wir, dass für jeden Ausgang $\omega \in \Omega$: $\lim_{h \to 0} |e^{ihX(\omega)} - 1| = 0$. Somit gilt:

$$|e^{ihX} - 1| \xrightarrow[h \to 0]{f.s.} 0.$$

Außerdem haben wir die Abschätzung $|e^{ihX(\omega)} - 1| \leq 2$. Der Satz von der majorisierten Konvergenz ergibt

$$\lim_{h \to 0} g(h) = \lim_{h \to 0} \mathbb{E}\left|e^{ihX} - 1\right| = \mathbb{E}\left[\lim_{h \to 0}\left|e^{ihX} - 1\right|\right] = \mathbb{E}0 = 0.$$

Es folgt, dass

$$\lim_{h \to 0} \sup_{t \in \mathbb{R}} |\varphi_X(t+h) - \varphi_X(t)| = 0.$$

Also ist φ_X gleichmäßig stetig.

Beweis von 4. Für beliebige $t_1, \dots, t_n$ und $c_1, \dots, c_n \in \mathbb{C}$ gilt

$$\begin{aligned}
\sum_{k,j=1}^{n} c_k \bar{c}_j \varphi_X(t_k - t_j) &= \sum_{k,j=1}^{n} c_k \bar{c}_j \mathbb{E}\left[e^{i(t_k - t_j)X}\right] \\
&= \mathbb{E}\left[\sum_{k,j=1}^{n} c_k \bar{c}_j e^{i(t_k - t_j)X}\right] \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{n} c_k e^{it_k X}\right) \cdot \overline{\left(\sum_{j=1}^{n} c_j e^{it_j X}\right)}\right]
\end{aligned}$$

Mit der Bezeichnung $A := \sum_{k=1}^n c_k e^{it_k X}$ erhalten wir somit

$$\sum_{k,j=1}^n c_k \bar{c}_j \varphi_X(t_k - t_j) = \mathbb{E}[A\bar{A}] = \mathbb{E}[|A|^2] \geq 0,$$

denn $|A|^2 \geq 0$. $\square$

Es gilt auch eine Umkehrung des soeben bewiesenen Satzes.

Satz 12.5.6 (Satz von Bochner). *Sei $\varphi : \mathbb{R} \to \mathbb{C}$ eine Funktion mit den 4 Eigenschaften aus obigem Satz. Dann gibt es eine Zufallsvariable X mit $\varphi_X = \varphi$.*

Ohne Beweis.

Im folgenden Lemma berechnen wir die charakteristische Funktion einer linearen Transformation.

Lemma 12.5.7. *Seien $a, b \in \mathbb{R}$ und X eine Zufallsvariable. Dann gilt:*

$$\varphi_{aX+b}(t) = e^{ibt}\varphi_X(at).$$

Beweis. Wir benutzen die Definition der charakteristischen Funktion:

$$\varphi_{aX+b}(t) = \mathbb{E}[e^{it(aX+b)}] = \mathbb{E}[e^{itb} \cdot e^{i(at)X}] = e^{itb} \cdot \mathbb{E}[e^{i(at)X}] = e^{itb} \cdot \varphi_X(at).$$

$\square$

Satz 12.5.8. *Sei $X \sim N(\mu, \sigma^2)$. Dann gilt*

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

Bemerkung 12.5.9. Spezialfall: Für $X \sim N(0,1)$ gilt $\varphi_X(t) = e^{-\frac{t^2}{2}}$. Somit stimmt in diesem Fall die charakteristische Funktion mit der Dichte bis auf einen Vorfaktor überein.

Beweis. Schritt 1. Wir betrachten den Fall, dass $X \sim N(0,1)$. Zuerst berechnen wir die Laplace-Transformierte:

$$\begin{aligned}
m_X(t) &= \int_{-\infty}^{\infty} e^{ty} f_X(y) dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ty} \cdot e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-t)^2} dy \\
&= e^{\frac{t^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \\
&= e^{\frac{t^2}{2}},
\end{aligned}$$

wobei wir die neue Variable $z = y - t$ eingeführt haben.

Schritt 2. Aus der Laplace-Transformierten berechnen wir durch Hinzufügen des Faktors i die charakteristische Funktion:

$$\varphi_X(t) = m_X(it) = e^{\frac{(it)^2}{2}} = e^{-\frac{t^2}{2}}.$$

Schritt 3. Sei nun $X \sim N(\mu, \sigma^2)$ mit beliebigen Parametern μ, σ^2. Dann haben wir die Darstellung $X = \sigma Y + \mu$, wobei $Y \sim N(0,1)$. Mit Lemma 12.5.7 gilt dann:

$$\varphi_X(t) = e^{it\mu} \cdot \varphi_Y(\sigma t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}.$$

$\square$

Satz 12.5.10. *Seien X, Y unabhängige Zufallsvariablen. Dann gilt*

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t).$$

Beweis. Übung. $\square$

Nun berechnen wir Momente mithilfe der charakteristischen Funktion.

Satz 12.5.11. *Sei X eine Zufallsvariable mit $\mathbb{E}|X|^n < \infty$ für ein $n \in \mathbb{N}$. Dann ist die charakteristische Funktion φ_X n-mal stetig differenzierbar und es gilt:*

$$\mathbb{E}[X^n] = i^{-n} \varphi_X^{(n)}(0).$$

Zum Beweis benötigen wir ein Lemma.

Lemma 12.5.12. *Für $y \in \mathbb{R}$ gilt die Ungleichung $|e^{iy} - 1| \leq |y|$.*

Beweis. Für $y \geq 0$ gilt

$$|e^{iy} - 1| = \left| \int_0^y e^{is} ds \right| \leq \int_0^y |e^{is}| ds = y.$$

Für $y \leq 0$ benutzen wir, dass $|e^{iy} - 1| = |e^{-iy} - 1|$. $\square$

Beweis von Satz 12.5.11. Schritt 1. Wir betrachten den Fall $n = 1$. Sei also $\mathbb{E}|X| < \infty$. Wir werden zeigen, dass φ_X stetig differenzierbar ist und dass

$$\varphi_X'(t) = \mathbb{E}[iXe^{itX}]. \tag{12.5.1}$$

Wir stellen den Differenzenquotienten der Funktion φ_X auf:

$$\frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \mathbb{E}\left[e^{itX} \frac{e^{ihX} - 1}{h} \right].$$

Mit der Taylor–Reihe der Exponentialfunktion erhalten wir

$$\lim_{h \to 0} \left(e^{itX} \frac{e^{ihX} - 1}{h} \right) = iXe^{itX}.$$

Außerdem erhalten wir mit Lemma 12.5.12 die Abschätzung

$$\left| e^{itX} \frac{e^{ihX} - 1}{h} \right| = \left| \frac{e^{ihX} - 1}{h} \right| \leq |X|.$$

Wir haben vorausgesetzt, dass $\mathbb{E}|X| < \infty$. Also können wir den Satz von der majorisierten Konvergenz benutzen:

$$\lim_{h \to 0} \frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \mathbb{E}\left[iXe^{itX} \right].$$

Somit ist φ_X differenzierbar und die Formel (12.5.1) gilt. Außerdem ist die Funktion $\mathbb{E}[iXe^{itX}]$ stetig. Der Beweis dafür ist identisch mit dem Beweis, dass φ_X stetig ist.
Nun setzen wir $t = 0$ in (12.5.1) ein: $\varphi'_X(0) = \mathbb{E}[iX]$. Das beweist den Satz für $n = 1$.

SCHRITT 2. Sei nun $n \in \mathbb{N}$ beliebig und $\mathbb{E}|X|^n < \infty$. Wendet man die Methode von Schritt 1 induktiv an, so erhält man die Formel

$$\varphi_X^{(n)}(t) = \mathbb{E}[(iX)^n e^{itX}].$$

Setzt man $t = 0$ ein, so erhält man die Aussage des Satzes. □

BEISPIEL 12.5.13. Sei $X \sim N(0,1)$ standardnormalverteilt. Die Taylor-Entwicklung der charakteristischen Funktion lautet: $\varphi_X(t) = e^{-t^2/2} = 1 - \frac{1}{2}t^2 + \ldots$. Damit gilt

$$\begin{aligned}\varphi'_X(0) &= 0 \Rightarrow \mathbb{E}X = 0,\\ \varphi''_X(0) &= -1 \Rightarrow \mathbb{E}X^2 = i^{-2} \cdot (-1) = 1.\end{aligned}$$

Also gilt für die Varianz von X: $\operatorname{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}X)^2 = 1$.

Wenn die charakteristische Funktion einer Zufallsvariable bekannt ist, dann kann man die Momente dieser Zufallsvariable ausrechnen. Wie kann man aber die ganze Verteilung, also die Verteilungsfunktion der Zufallsvariable, wiederherstellen?

SATZ 12.5.14 (Umkehrformel). *Sei Z eine Zufallsvariable mit charakteristischer Funktion φ_Z und Verteilungsfunktion F_Z. Dann gilt für alle $a, b \in \mathbb{R}$ mit $a < b$ und $\mathbb{P}[Z = a] = \mathbb{P}[Z = b] = 0$ die Formel*

$$\mathbb{P}[a \le Z \le b] = F_Z(b) - F_Z(a) = \frac{1}{2\pi} \lim_{c \to \infty} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_Z(t) dt.$$

BEMERKUNG 12.5.15. Die Eigenschaft $\mathbb{P}[Z = a] = 0$ bedeutet, dass a kein Atom von X ist, bzw. dass die Verteilungsfunktion F_Z an der Stelle a stetig ist.

BEWEIS. SCHRITT 1. Wir verwenden das Fresnel-Integral

$$\lim_{c \to +\infty} S(c) = \frac{\pi}{2}, \text{ wobei } S(c) := \int_0^c \frac{\sin(x)}{x} dx.$$

Außerdem benötigen wir die Vorzeichenfunktion:

$$\operatorname{sgn}(\theta) = \begin{cases} 1, & \theta > 1, \\ 0, & \theta = 0, \\ -1, & \theta < 1. \end{cases}$$

Sei $\theta \in \mathbb{R}\backslash\{0\}$, dann gilt (wobei wir die Variable $s = |\theta|\, t$ einführen):

$$\int_0^c \frac{\sin(\theta t)}{t} dt = \operatorname{sgn}(\theta) \int_0^{|\theta| c} \frac{\sin(s)}{s/|\theta|} \frac{ds}{|\theta|} = \operatorname{sgn}(\theta) \int_0^{|\theta| c} \frac{\sin(s)}{s} ds = \operatorname{sgn}(\theta) \cdot S(|\theta| c).$$

Die Formel gilt auch für $\theta = 0$, denn beide Seiten sind dann gleich 0.

Schritt 2. Nun zum Beweis der Umkehrformel. Betrachte das Integral

$$
\begin{aligned}
I(c) &:= \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_z(t) dt \\
&= \int_{-c}^{c} \mathbb{E}\left[\frac{e^{it(Z-a)} - e^{it(Z-b)}}{it}\right] dt \\
&= \mathbb{E}\left[\int_{-c}^{c} \frac{e^{it(Z-a)} - e^{it(Z-b)}}{it} dt\right],
\end{aligned}
$$

wobei wir im letzten Schritt den Satz von Fubini benutzen durften, denn

$$
\left|\frac{e^{it(Z-a)} - e^{it(Z-b)}}{it}\right| = \left|\frac{e^{it(b-a)} - 1}{it}\right| \leq b - a.
$$

Da $e^{ix} = \cos x + i \sin x$, wobei cos eine gerade und sin eine ungerade Funktion ist, erhalten wir

$$
\begin{aligned}
I(c) &= 2 \cdot \mathbb{E}\left[\int_{0}^{c} \frac{\sin(t(Z-a)) - \sin(t(Z-b))}{t} dt\right] \\
&= 2 \cdot \mathbb{E}\left[\operatorname{sgn}(z-a) S(|Z-a|c) - \operatorname{sgn}(Z-b) S(|Z-b|c)\right].
\end{aligned}
$$

Für die Funktion unter dem Erwartungswert gilt (mit dem Fresnel-Integral aus Schritt 1)

$$
\lim_{c \to +\infty} (\operatorname{sgn}(z-a) S(|z-a|c) - \operatorname{sgn}(z-b) S(|z-b|c)) = \pi \psi_{a,b}(z),
$$

wobei

$$
\psi_{a,b}(z) = \frac{\pi}{2} (\operatorname{sgn}(z-a) - \operatorname{sgn}(z-b)) = \begin{cases} 0, & z < a, \\ \frac{1}{2}, & z = a, \\ 1, & a < z < b, \\ \frac{1}{2}, & z = b, \\ 0, & z > b. \end{cases}
$$

Aus der Existenz eines endlichen Grenzwerts $\lim_{c \to +\infty} S(c)$ folgt, dass $B := \sup_{c \geq 0} |S(c)| < \infty$. Außerdem gilt $|\operatorname{sgn} t| \leq 1$. Daher haben wir die obere Schranke

$$
|\operatorname{sgn}(z-a) S(|z-a|c) - \operatorname{sgn}(z-b) S(|z-b|c)| < 2B.
$$

Wir können also den Satz von der majorisierten Konvergenz benutzen und erhalten damit:

$$
\begin{aligned}
\frac{1}{2\pi} \lim_{c \to \infty} I(c) &= \frac{2}{2\pi} \lim_{c \to \infty} \mathbb{E}\left[\operatorname{sgn}(z-a) \cdot S(|z-a|c) - \operatorname{sgn}(z-b) \cdot S(|z-b|c)\right] \\
&= \frac{1}{\pi} \mathbb{E}\left[\lim_{c \to \infty} \operatorname{sgn}(z-a) \cdot S(|z-a|c) - \operatorname{sgn}(z-b) \cdot S(|z-b|c)\right] \\
&= \mathbb{E}[\psi_{a,b}(z)].
\end{aligned}
$$

Da nun a und b keine Atome von X sind, ist die rechte Seite gleich $\mathbb{P}[a < z < b]$. □

Satz 12.5.16 (Eindeutigkeitssatz). *Die charakteristische Funktion bestimmt die Verteilungsfunktion vollständig, das heißt: Sind X, Y Zufallsvariablen mit $\varphi_X(t) = \varphi_Y(t)$ für alle $t \in \mathbb{R}$, dann gilt:*

$$F_X(t) = F_Y(t) \quad \textit{für alle } t \in \mathbb{R}.$$

Beweis. Schritt 1. Die Menge aller Punkte, wo beide Verteilungsfunktionen F_X und F_Y stetig sind, bezeichnen wir mit

$$S = \{t \in \mathbb{R} : F_X \text{ und } F_Y \text{ sind stetig an der Stelle } t\}.$$

Die Menge $\mathbb{R} \backslash S$ ist höchstens abzählbar. Nun benutzen wir die Umkehrformel, die besagt, dass für alle $a, b \in S$ mit $a < b$ gilt:

$$\begin{aligned} F_X(b) - F_X(a) &= \frac{1}{2\pi} \lim_{c \to \infty} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt \\ &= \frac{1}{2\pi} \lim_{c \to \infty} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_Y(t) dt \\ &= F_Y(b) - F_Y(a). \end{aligned}$$

Dabei haben wir benutzt, dass $\varphi_X = \varphi_Y$ laut Voraussetzung. Für alle $a, b \in S$ mit $a < b$ gilt also $F_X(b) - F_X(a) = F_Y(b) - F_Y(a)$.

Schritt 2. Da das Komplement von S höchstens abzählbar ist, können wir für jedes $n \in \mathbb{N}$ ein $a_n < -n$ mit $a_n \in S$ finden. Daraus folgt, dass $\lim_{n\to\infty} a_n = -\infty$. Für alle $b \in S$ gilt nun laut Schritt 1:

$$\begin{aligned} F_X(b) &= F_X(b) - \lim_{n \to \infty} F_X(a_n) \\ &= \lim_{n \to \infty} (F_X(b) - F_X(a_n)) \\ &= \lim_{n \to \infty} (F_Y(b) - F_Y(a_n)) \\ &= F_Y(b) - \lim_{n \to \infty} F_Y(a_n) \\ &= F_Y(b). \end{aligned}$$

Für alle $b \in S$ gilt also $F_X(b) = F_Y(b)$.

Schritt 3. Sei nun $b \in \mathbb{R}$ beliebig. Da das Komplement von S höchstens abzählbar ist, kann man für jedes $n \in \mathbb{N}$ ein $b_n \in (b, b + 1/n) \cap S$ finden. Insbesondere gilt $\lim_{n\to\infty} b_n = b$ und dabei ist $b_n > b$. Dann gilt:

$$\begin{aligned} F_X(b) &= \lim_{n \to \infty} F_X(b_n) \quad (\text{da } F_X \text{ rechtsstetig}) \\ &= \lim_{n \to \infty} F_Y(b_n) \quad (\text{mit Schritt 2}) \\ &= F_Y(b) \quad (\text{da } F_Y \text{ rechtsstetig}). \end{aligned}$$

Für alle $b \in \mathbb{R}$ gilt also $F_X(b) = F_Y(b)$. □

Die Umkehrformel gilt für beliebige Zufallsvariablen. Sie ist aber nicht sehr schön, weswegen wir noch eine einfachere Darstellung zeigen, die aber nur für absolut stetige Zufallsvariablen gilt.

Satz 12.5.17 (Umkehrformel für die Dichte). *Sei Z eine Zufallsvariable, deren charakteristische Funktion φ_Z integrierbar ist, d.h. $\int_{-\infty}^{\infty} |\varphi_Z(t)| dt < \infty$. Dann gilt: Z ist absolut stetig mit Dichte*

$$f_Z(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \varphi_Z(t) dt.$$

Beweis. Wir bezeichnen die rechte Seite mit

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \varphi_Z(t) dt.$$

Die Funktion g ist wohldefiniert, denn $|e^{-ity}\varphi_Z(t)| = |\varphi_Z(t)|$ und $|\varphi_Z|$ ist integrierbar (laut Voraussetzung). Außerdem ist g stetig. (Beweis identisch mit dem Beweis für Stetigkeit der charakteristischen Funktion).
Nun wollen wir zeigen, dass die Zufallsvariable Z eine Dichte besitzt und dass g diese Dichte ist. Dazu betrachten wir das Integral

$$\begin{aligned}
\int_a^b g(y) dy &= \int_a^b \frac{1}{2\pi} \left(\int_{-\infty}^{\infty} e^{-ity} \varphi_Z(t) dt \right) dy \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_a^b e^{-ity} \varphi_Z(t) dy \right) dt \quad \text{(Fubini)} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\varphi_Z(t) \cdot \int_a^b e^{-ity} dy \right) dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_Z(t) \cdot \frac{e^{-ita} - e^{-itb}}{it} dt \\
&= \mathbb{P}[a < Z < b] \quad \text{(mit Umkehrformel).}
\end{aligned}$$

Wir haben gezeigt, dass für alle $a, b \in \mathbb{R}$ mit $a < b$, die Stetigkeitspunkte von F_Z sind,

$$\int_a^b g(y) dy = \mathbb{P}[a \le Z \le b]$$

Daraus folgt (genauso wie im Beweis von Satz 12.5.16), dass g die Dichte von Z ist. $\square$

Beispiel 12.5.18. In diesem Beispiel berechnen wir die charakteristische Funktion einer Cauchy-verteilten Zufallsvariable. Sei zuerst X eine Zufallsvariable mit der Dichte $f_X(y) = \frac{1}{2} e^{-|y|}$ (zweiseitige Exponentialverteilung). Die charakteristische Funktion von X ist dann gegeben durch:

$$\begin{aligned}
\varphi_X(t) &= \int_{-\infty}^{\infty} \frac{1}{2} e^{-|y|} e^{ity} dy = \frac{1}{2} \left(\int_0^{\infty} e^{-|y|} e^{ity} dy + \int_{-\infty}^{0} e^{-|y|} e^{ity} dy \right) \\
&= \frac{1}{2} \left[\frac{1}{1 - it} + \frac{1}{1 + it} \right] = \frac{1}{1 + t^2}.
\end{aligned}$$

Nun kann man die Umkehrformel für die Dichte anwenden:

$$\frac{1}{2} e^{-|y|} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ity} \frac{1}{1 + t^2} dt.$$

Nun ersetzen wir t durch $-t$ und multiplizieren beide Seiten mit 2:

$$e^{-|y|} = \int_{-\infty}^{\infty} e^{ity} \left(\frac{1}{\pi} \cdot \frac{1}{1+t^2} \right) dt = \int_{-\infty}^{\infty} e^{ity} f_Z(t) dt = \varphi_Z(y),$$

wobei Z eine Cauchy-verteilte Zufallsvariable ist. Wir haben somit gezeigt, dass die charakteristische Funktion einer Cauchy-verteilten Zufallsvariable gleich $\varphi_Z(y) = e^{-|y|}$ ist.

Beispiel 12.5.19. Seien $X_1 \sim N(\mu_1, \sigma_1^2)$ und $X_2 \sim N(\mu_2, \sigma_2^2)$ unabhängige, normalverteilte Zufallsvariablen. Dann gilt:

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Beweis. Wir haben das bereits mithilfe der Faltungsformel gezeigt. Hier führen wir den Beweis mit charakteristischen Funktionen. Die charakteristischen Funktionen von X_1 und X_2 sind gegeben durch

$$\varphi_{X_k}(t) = e^{i\mu_k t - \frac{1}{2}\sigma_k^2 t^2}, \quad k = 1, 2.$$

Nun berechnen wir die charakteristische Funktion von $X_1 + X_2$:

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) = e^{i\mu_1 t - \frac{1}{2}\sigma_1^2 t^2} \cdot e^{i\mu_2 t - \frac{1}{2}\sigma_2^2 t^2} = e^{i(\mu_1+\mu_2)t - \frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}.$$

Auf der rechten Seite erkennen wir die charakteristische Funktion einer $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$-Verteilung. Mit dem Eindeutigkeitssatz folgt dann, dass

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$