



Grundlagen der Statistik: Faltung (von Funktionen) Data Science and Business Analytics

Dipl.-Ing. Hubert Schölnast, BSc.
Stand: 24. März 2020

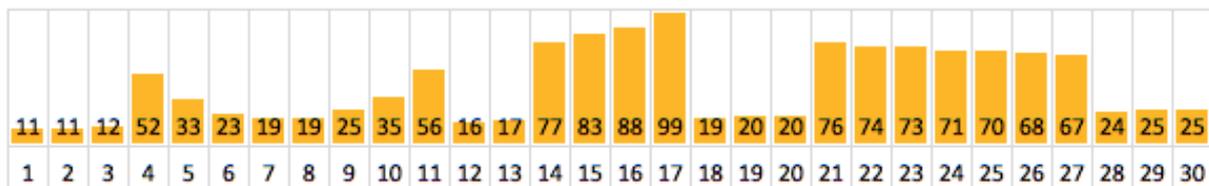
1	Welches Problem soll gelöst werden?	3
1.1	Rohdaten-Funktion	3
1.2	Geräte-Funktion	4
2	Faltung	5
2.1	Erklärung anhand eines Beispiels	5
2.2	Mathematische Beschreibung der Faltung	7
3	Anwendungen der Faltung in der Statistik	9
3.1	Kerndichteschätzer	9
3.2	Zentraler Grenzwertsatz	12
3.2.1	Selbstfaltung	12
3.2.2	Beispiel »Augensumme bei mehreren Würfeln eines Würfels«	13

1 Welches Problem soll gelöst werden?

Brillenträger kennen das Phänomen: Ohne Brille sieht man unscharf. Wer keine Brille trägt, kann dieses Erlebnis dadurch simulieren, dass man eine (starke) Brille aufsetzt. Der Grund dafür ist genau derselbe, der dafür sorgt, dass manche Fotografien unscharf sind: Die Brennweite der Linse hat nicht den optimalen Wert um einen Gegenstand, der sich in einer bestimmten Entfernung von der Linse befindet, auf die Sensor-Schicht abzubilden (Netzhaut beim Auge, CCD- oder CMOS-Chip bei einer Digitalkamera, Film mit lichtempfindlichen Chemikalien bei einer Kamera nach althergebrachter Technologie). Dadurch werden aus den erhofften scharfen Licht-Punkten verwaschene Licht-Flecken.

Sehen wir und das Anhand des Beispiels einer Digitalkamera an. Das hat nämlich den Vorteil, dass hier die einzelnen Sensoren (»Pixel«) in einem sehr regelmäßigen Gitter angeordnet sind, was bei Filmen und Netzhäuten nicht der Fall ist. Um nicht in einem zweidimensionalen Gitter arbeiten zu müssen, vereinfachen wir die Fragestellung noch weiter und betrachten nur eine einzelne Pixel-Zeile, und wir betrachten nur eine der drei Farben, die Kameras üblicherweise aufnehmen (rot, grün, blau), oder wird betrachten einfach nur die Gesamthelligkeit der Pixel, was denselben Effekt hat: Jetzt liefert jedes Pixel nur genau eine Zahl.

Die ersten 30 Pixel einer Pixel-Zeile könnten, wenn die Linse eine perfekt scharfe Abbildung liefern würde, beispielsweise diese Werte enthalten:



Unten steht die Pixel-Nummer, darüber die Helligkeit, die das jeweilige Pixel misst. Zur besseren Veranschaulichung wurden die Helligkeitswerte auch durch Säulen dargestellt.

1.1 Rohdaten-Funktion

Beachte, dass man den Zusammenhang zwischen Pixel-Nummer und dessen Helligkeit durch eine mathematische Funktion beschreiben kann. Es gibt zwar keinen analytischen Ausdruck (es gibt also keine Formel) in die man die Pixelnummer eingeben kann, um daraus die Helligkeit des Pixels zu berechnen, aber das ist auch gar nicht notwendig. Eine Werte-Liste reicht vollkommen.

Eine mathematische Funktion liegt bereits dann vor, wenn es zu jedem X-Wert aus der Grundmenge (das sind hier die ganzen Zahlen von 1 bis 30) genau einen Y-Wert gibt. Tatsächlich ist das sogar ganz genau die Definition des Begriffs Funktion: Zu jedem möglichen X-Wert gibt es genau einen Y-Wert. (Gleiche Y-Werte für verschiedene X-Werte sind erlaubt.) Eine Berechnungsformel hat natürlich viele praktische Vorteile, ist aber nicht Bestandteil der Definition des Begriffs »Funktion«.

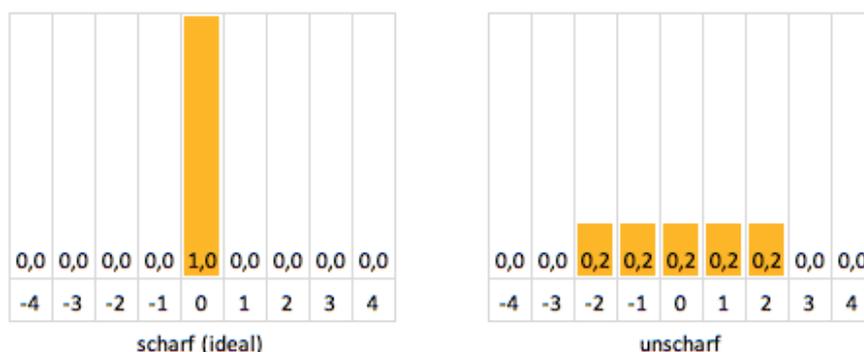
Zusatzwissen ohne direkten Zusammenhang zum Thema: Wenn es zu einem X-Wert mehrere Y-Werte geben kann (Beispiel: Die Quadratwurzel von 4 ist sowohl -2 als auch +2) dann liegt keine Funktion vor, sondern eine Relation. Funktionen sind Spezialfälle von Relationen.

Dafür ist man dann, wenn es zu bestimmten X-Werten keine Y-Werte gibt, etwas kulanter. Relationen, bei denen es zu bestimmten X-Werten keine Y-Werte gibt, werden manchmal trotzdem noch zu den Funktionen gezählt. Beispiel: Die Kehrwerte aller negativen und aller positiven reellen Zahlen existieren (Kehrwert von 5 ist $\frac{1}{5}$, Kehrwert von $-\frac{4}{3}$ ist $-\frac{3}{4}$), aber die Zahl 0 ist die einzige reelle Zahl ohne Kehrwert (der Ausdruck $\frac{1}{0}$ ist keine reelle Zahl). Trotzdem behandelt man den Kehrwert in vielen Fällen so als wäre er eine Funktion. Generell gilt aber trotzdem: Damit eine Relation als Funktion durchgeht, muss es zu jedem X-Wert genau einen Y-Wert geben.

Ich verwende für jene Funktion, die dem ideal scharfen Bild entspricht, den Begriff »Rohdaten-Funktion«. Das ist allerdings kein allgemein gängiger Ausdruck im Zusammenhang mit der Faltung. Allem Anschein nach gibt es für diese Funktion keinen allgemein etablierten Ausdruck. Innerhalb dieses Dokuments bleibe ich aber bei dem Begriff »Rohdaten-Funktion«. Merken muss man sich diesen Begriff aber nicht.

1.2 Geräte-Funktion

Im Gegensatz zu »Rohdaten-Funktion« ist der Begriff »Geräte-Funktion« aber sehr wohl allgemein etabliert und gängig. Das ist jene Funktion, die beschreibt, wie sich - im Fall unseres Beispiels - die Lichtmenge, die im Idealfall einem einzelnen Pixel »zusteht«, auf mehrere Pixel verteilt. Nehmen wir mal an, die Lichtmenge, die eigentlich für ein bestimmtes Pixel bestimmt ist, würde durch die falsche Brennweite der Linse gleichmäßig auf fünf Pixel verteilt werden:



Im linken Bild sieht man, dass die ganze Lichtmenge (100% = 1,0) auf ein Pixel abgebildet wird, und dieses Pixel hat die Nummer 0. Weder seine linken noch seine rechten Nachbarn bekommen etwas von dem Licht ab, das (bei idealer scharfer Abbildung) für das Pixel an Position 0 bestimmt ist.

Im rechten Bild sieht man, dass die 100%, die an Licht daherkommen, auf 5 Pixel aufgeteilt werden, wovon auf jedes dieser 5 Pixel je 20% entfallen (also das 0,2-fache des gesamten Lichts). Diese 5 Pixel sind symmetrisch um die Mitte angeordnet.

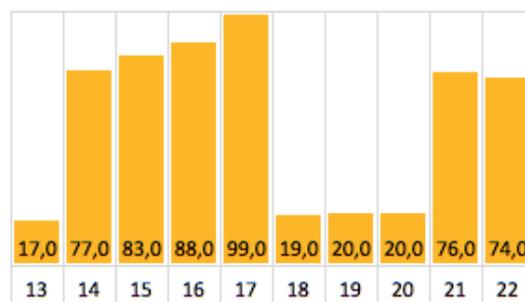
Wie kann man nun aus der Rohdaten-Funktion und der Geräte-Funktion berechnen, wieviel Licht tatsächlich auf die einzelnen Pixel entfällt? Genau das leistet die Operation, um die es in diesem Dokument geht: Die Faltung.

2 Faltung

2.1 Erklärung anhand eines Beispiels

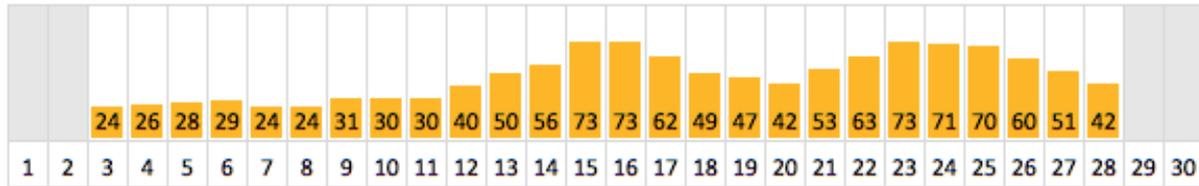
Wie man aus der Rohdaten-Funktion und der Geräte-Funktion die tatsächlichen Belichtungswerte berechnen kann, liegt eigentlich bereits auf der Hand:

Ich greife als Beispiele die beiden Pixel mit den Nummern 17 und 18 heraus. Hier sind nochmal die Werte der Rohdaten-Funktion im näheren Umfeld dieser beiden Pixel:

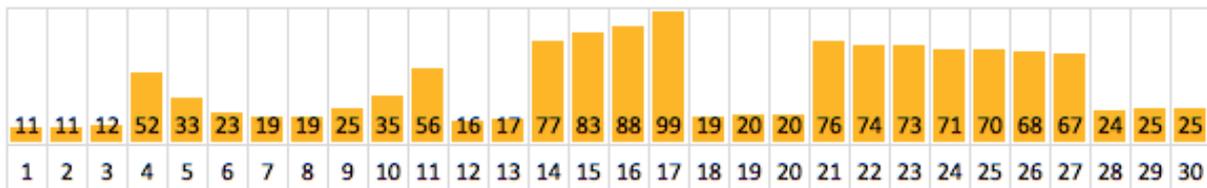


Von dem Licht, das für die Pixel 13 und 14 vorgesehen ist, landet auch bei unscharfer Abbildung nichts in die beiden Pixel 17 und 18, die wir hier betrachten wollen. Aber die 83 Licht-Einheiten, die bei scharfer Abbildung auf das Pixel 15 entfallen würden, verteilen sich wie folgt:

Rechnet man das für alle Pixel durch, erhält man diese Funktion:



Zum Vergleich nochmal die Rohdaten-Funktion:



Werte am Rand

Die Werte für die Pixel 1, 2, 29 und 30 konnten nicht berechnet werden, weil man dafür die Daten von Pixeln außerhalb des vorgegebenen Wertebereichs benötigt hätte, die aber nicht bekannt sind. Dieses Rand-Problem hat mit der eigentlichen Faltung nur wenig zu tun und wird in diesem Dokument daher nicht näher betrachtet. In der praktischen Anwendung (z.B. bei der Bildverarbeitung) muss man sich hier mit diversen Korrektur-Methoden behelfen, die im Wesentlichen darauf beruhen, dass man mehr oder weniger wohlbegründete Annahmen über die Werte der nicht bekannten Pixel macht.

Effekt der Faltung

Wie man an den beiden Bildern in der oberen Hälfte dieser Seite gut erkennen kann, entsteht eine Funktion, die aus abrupten Sprüngen sanftere Verläufe macht und die auch Extremwerte nivelliert. Das ist genau das, was auch tatsächlich passiert, wenn in einer Kamera oder in einem Auge das Licht nicht genau auf die Sensorfläche fokussiert wird sondern einen verwaschenen Fleck bildet.

2.2 Mathematische Beschreibung der Faltung

Man hat eine Rohdaten-Funktion, die hier $r(x)$ heißen soll, und eine Gerätefunktion $g(x)$. Die Faltung $f(x)$ der beiden Funktionen wird so geschrieben:

$$f(x) = (r * g)(x)$$

Berechnet wird sie, im Fall von Funktionen mit ganzzahligen Argumenten, so:

$$f(x) = (r * g)(x) = \sum_{t=-\infty}^{+\infty} r(t) \cdot g(x - t)$$

In Wahrheit ist diese Formel nichts weiter als eine sehr kompakte Beschreibung des Algorithmus, den ich davor schon anhand des Beispiels erklärt habe.

In dieser Formel sieht man auch, dass die Laufvariable t in der Gerätefunktion von x abgezogen wird. Das bedeutet: Während der Bildung der einzelnen Summanden werden von der Rohdaten-Funktion Werte in aufsteigender Reihenfolge verwendet:

$$t = \dots, -2, -1, 0, 1, 2, \dots$$

Zur gleichen Zeit läuft der Index der Gerätefunktion aber in absteigender Richtung:

$$x - t = \dots, 8, 7, 6, 5, 4, \dots \text{ (wenn man gerade die Summenbildung für } x = 6 \text{ durchläuft)}$$

In diesem Dokument ist die Gerätefunktion aber immer eine symmetrische Funktion, daher spielt diese Umkehrung der Reihenfolge weder für das Verständnis der Kerndichteschätzer noch für den zentralen Grenzwertsatz eine Rolle.

Möglicherweise irritieren vielleicht die Grenzen des Bereichs, den die Laufvariable der Summe durchlaufen soll. Laut Formel muss t ja alle Werte von $-\infty$ bis $+\infty$ durchlaufen, was nach einem großen Problem für die praktische Umsetzung klingt, denn es wird laut Formel verlangt unendlich viele Werte aufzusummieren, was auch beim schnellsten Supercomputer unendlich lang dauern würde. In Wahrheit ist das aber gar nicht so schlimm:

Weil nämlich die Gerätefunktion $g(x)$ in allen praktisch sinnvollen Fällen bei X -Werten, die weit weg vom Nullpunkt sind, den Funktionswert 0 hat (*mit anderen Worten: Weil $g(x)$ nur in einem kleinen Bereich links und rechts vom Nullpunkt interessante Werte hat und sonst überall 0 ist*), muss der Summierungs-Parameter t in Wahrheit nur jene Werte durchlaufen, bei denen das Argument von g (also $x - t$) in jenem Bereich liegt, in dem g Werte ungleich 0 hat. Wenn nämlich $g(x - t)$ den Wert 0 hat, hat auch der ganze Summand (also $r(t) \cdot g(x - t)$) den Wert 0 und trägt somit nichts zur Gesamtsumme bei. Daher kann man diese t -Werte komplett außer Acht lassen.

Viel häufiger braucht man die Faltung aber dann, wenn als X - und Y -Werte reelle Zahlen in Frage kommen, wenn es also für $f(x)$ und $g(x)$ auch dann Werte gibt, wenn x keine ganze Zahl ist. Dann muss man in der obigen Formel nur die Summe durch ein Integral ersetzen:

$$f(x) = (r * g)(x) = \int_{t=-\infty}^{+\infty} r(t) \cdot g(x - t) dt$$

In der numerischen Berechnung werden Integrale aber ohnehin durch Summen angenähert, womit also auch in diesem Fällen in der praktischen Anwendung die Summen-Formel verwendet wird.

3 Anwendungen der Faltung in der Statistik

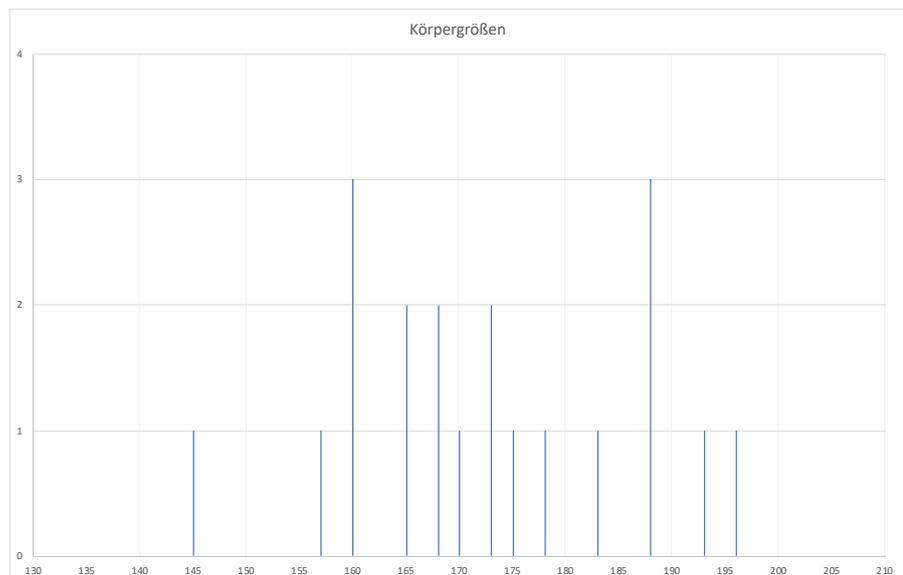
3.1 Kerndichteschätzer

Ein Beispiel:

An einer amerikanischen Uni wurden die Körpergrößen von 20 männlichen Studenten im Alter von 22 Jahren erhoben. Weil das in den USA geschah, wurden - wie dort üblich - alle Körpergrößen auf ganze Zoll gerundet. Diese Größen wurden später in Europa in Zentimeter umgerechnet und abermals gerundet. Diese Liste enthält nun folgende Werte (die Körpergrößen sind aufsteigend sortiert): 145, 157, 160, 160, 160, 165, 165, 168, 168, 170, 173, 173, 175, 178, 183, 188, 188, 188, 193, 196.

Tabelle mit den absoluten Häufigkeiten und Histogramm:

Größe	Anzahl
145	1
157	1
160	3
165	2
168	2
170	1
173	2
175	1
178	1
183	1
188	3
193	1
196	1



Wie man sicherlich ahnt, ist diese Verteilung nicht die Verteilung aller 22-jährigen Männer in den USA. Es gibt sicherlich auch junge Männer, die 176 cm groß sind, oder auch welche, die 176,237 cm groß sind, denn, auch wenn die Körpergröße im Lauf eines Tages ein wenig schwankt (Bandscheiben werden bei aufrechter Haltung zusammengedrückt und dehnen sich

nachts wieder aus, was bis zu 2 cm Unterschied ausmachen kann), so hat doch jeder Mensch zu einem bestimmten Zeitpunkt eine ganz bestimmte tatsächliche Größe, und die richtet sich nicht nach den Einheiten unseres Maßsystems. Daher kommen für die tatsächliche Körpergröße auch sämtliche denkbaren Zwischenwerte zwischen ganzen Zentimetern in Betracht.

Wir suchen nun also eine Möglichkeit, aus der Verteilung der Stichprobe, die in der Grafik aus einigen wenigen unendlich dünnen Peaks besteht und dazwischen überall 0 ist, eine stetige Verteilung zu machen, die der wahrscheinlichsten Verteilung der Werte in der Grundgesamtheit (hier: 22-jährige männliche US-Bürger) entspricht.

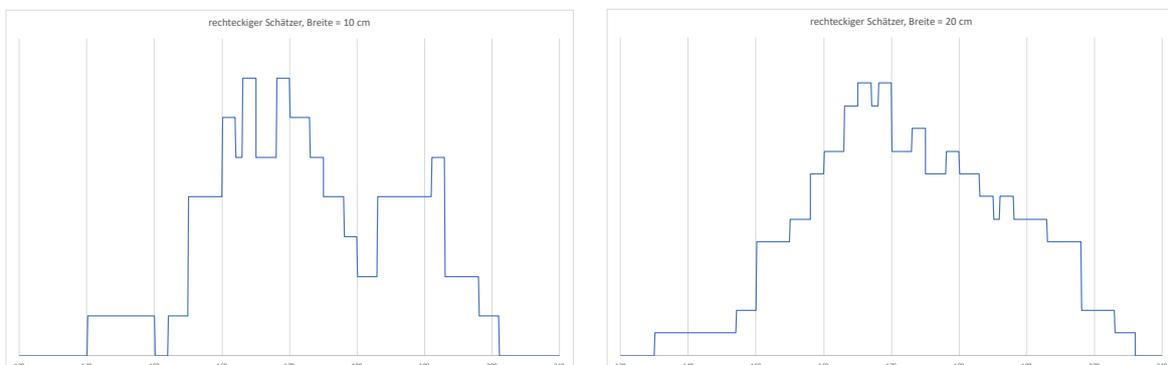
Die Lösung dieses Problems heißt: **Faltung**.

Die Liste mit den absoluten Häufigkeiten ist - im Jargon dieses Dokuments - die Rohdaten-Funktion, mit der Erweiterung, dass die Rohdaten-Funktion für jede theoretisch denkbare Körpergröße einen Wert hat. Dieser Default-Wert ist 0. Nur wenn in der Liste mit den Häufigkeiten ein expliziter Wert steht, gilt dieser Wert. Ansonsten ist der Funktionswert 0.

Die Werte für die Körpergrößen -80, 0, 52,987, 169, 200, 2000, 18.000.000 usw. sind also definiert, aber allesamt genau 0.

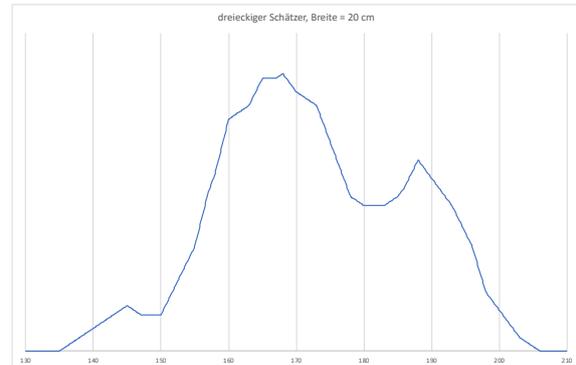
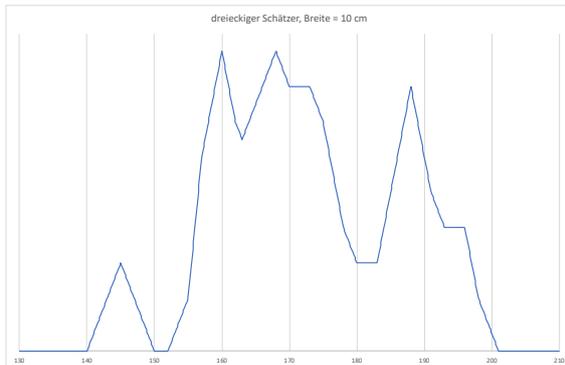
Nun benötigen wir noch eine »Gerätefunktion«, die in diesem Fall aber einen anderen Namen hat. Diese Funktion heißt »Kerndichteschätzer«.

Im Beispiel mit dem Fotosensoren war die Gerätefunktion der Einfachheit halber eine Rechteckfunktion. Von einer Rechteckfunktion spricht man, wenn alle Funktionswerte außerhalb eines bestimmten Intervalls 0 sind, und innerhalb des Intervalls einen konstanten Wert ungleich 0 haben. Die Fläche, die unter diesen Bedingungen zwischen der Funktion und der X-Achse liegt, ist dann nämlich ein Rechteck. Wenn man so eine Rechteckfunktion als Schätzer verwendet, erhält man so etwas. (Links ist das Rechteck 10 cm breit, rechtes 20):



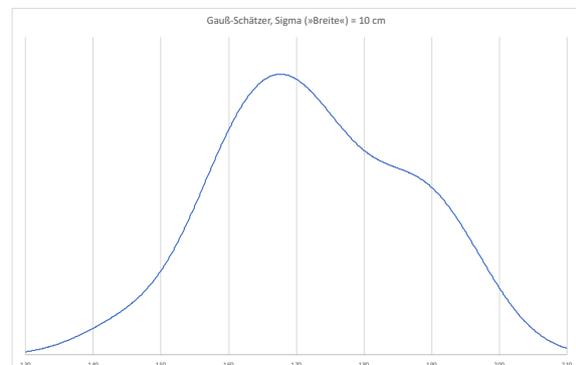
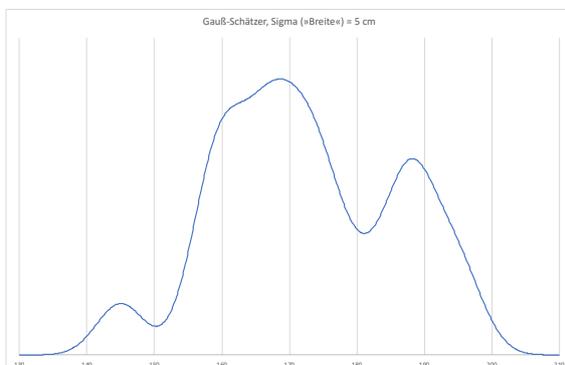
Die kantige Beschaffenheit des verwendeten Kerndichteschätzers spiegelt sich deutlich im Ergebnis der Faltung wider. Die gefaltete Funktion hat vielen Stufen.

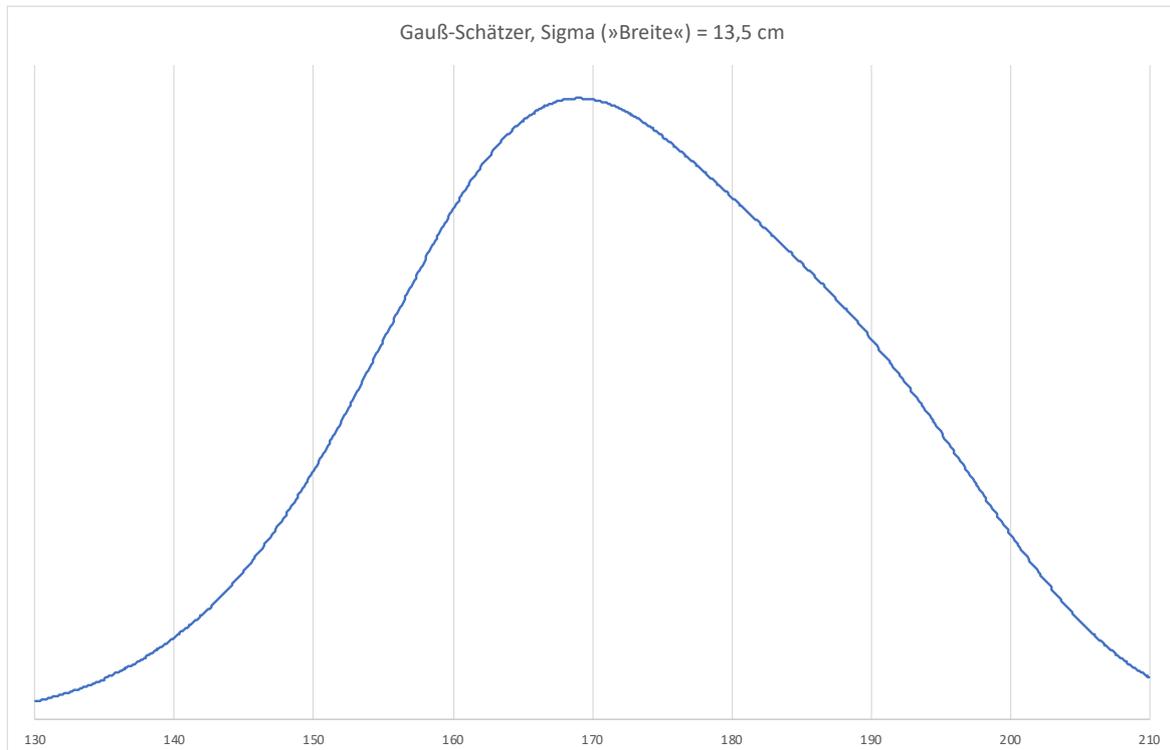
Wenn der Kerndichteschätzer eine Dreiecksfunktion ist, erhält man so etwas. (Länge der Basis des Dreiecks: links 10 cm, rechts 20 cm):



Das ist schon etwas besser, aber die gefaltete Funktion ist noch immer von geraden Teilstücken begrenzt, die scharfe Knicke bilden.

Eine »Gerätefunktion« (also ein Kerndichteschätzer) in Form einer Glockenkurve erzeugt bei der Faltung mit unserer Rohdaten-Funktion eine Funktion, die viel eher aussieht wie eine echte Körpergrößenverteilung. Verwendet wird eine Gaußsche Glockenkurve mit Sigma = 5 cm, 10 cm und 13,5 cm. 13,5 cm ist die Standardabweichung der Stichprobe):





Die Methoden zur Auswahl des richtigen Kerndichteschätzers (welche Funktion? wie breit?) ist ein eigener Bereich der Statistik, der hier, in diesem Dokument über Faltung, nicht vertieft werden soll. Wie die obigen Beispiele zeigen, sind Glockenkurven gut geeignet, aber auch davon gibt es viele verschiedene Arten. Bei der Breite des Schätzfunktion kann man sich an der Standardabweichung der Stichprobe orientieren.

3.2 Zentraler Grenzwertsatz

Vereinfachter Wortlaut des zentralen Grenzwertsatzes:

»Die Überlagerung vieler kleiner voneinander unabhängiger Zufallseffekte kann in ihrer Gesamtheit durch eine Normalverteilung angenähert werden. Dieser Effekt ist unabhängig von der statistischen Verteilung, welche die einzelnen Zufallseffekte tatsächlich haben.«

3.2.1 Selbstfaltung

Bisher waren die Rohdatenfunktion und die Gerätefunktion immer zwei verschiedene Funktionen. Aber es gibt keinen mathematischen Grund zu verbieten, dass diese beiden Funktionen identisch sind. Man kann also eine Funktion durchaus auch mit sich selbst falten.

Zur Erinnerung nochmals die Formel (der Einfachheit halber nur die Summen-Formel):

$$f(x) = (r * g)(x) = \sum_{t=-\infty}^{+\infty} r(t) \cdot g(x - t)$$

Wenn r und g identisch sind, kann man r durch g ersetzen:

$$f(x) = (g * g)(x) = \sum_{t=-\infty}^{+\infty} g(t) \cdot g(x - t)$$

Statt $(g * g)(x)$ schreibt man der Einfachheit halber auch $g^2(x)$. Bei dieser Schreibweise muss man aber vorsichtig sein, weil man mit $g^2(x)$ in anderen Kontexten manchmal etwas anderes meint, nämlich $g(x)^2$, also $g(x) \cdot g(x)$ und das ist etwas ganz anderes als $(g * g)(x)$. (Der Ausdruck $g(x)^2$ bedeutet: Nimm den Wert x , speise ihn in die Funktion g ein, und berechne dann von dem Ergebnis das Quadrat. Da kommen gar keine Summen vor. Hingegen bedeutet $(g * g)(x)$: Falte die Funktion $g(x)$ mit sich selbst.)

Nachdem wir uns diese Mehrdeutigkeit der Schreibweise bewusst gemacht haben, legen wir hiermit fest, dass in diesem Dokument folgendes gemeint sein:

$$g^2(x) = (g * g)(x)$$

und analog

$$\begin{aligned} g^3(x) &= (g * g * g)(x) \\ g^4(x) &= (g * g * g * g)(x) \\ &\text{usw.} \end{aligned}$$

Natürlich gilt auch das:

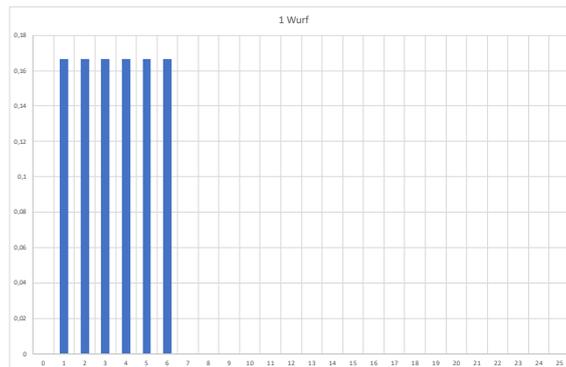
$$\begin{aligned} g^4(x) &= (g^2 * g^2)(x) \\ g^8(x) &= (g^4 * g^4)(x) \\ &\text{usw.} \end{aligned}$$

*(Für die, die es interessiert: Diese vereinfachte Schreibweise ist erlaubt, weil die Faltung assoziativ ist: $((g * g) * g)(x) = (g * (g * g))(x)$. Sie gehorcht übrigens auch dem Kommutativgesetz, das heißt Rohdatenfunktion und Gerätefunktion kann man gegeneinander austauschen, ohne dass das am Resultat etwas verändert.)*

3.2.2 Beispiel »Augensumme bei mehreren Würfeln eines Würfels«

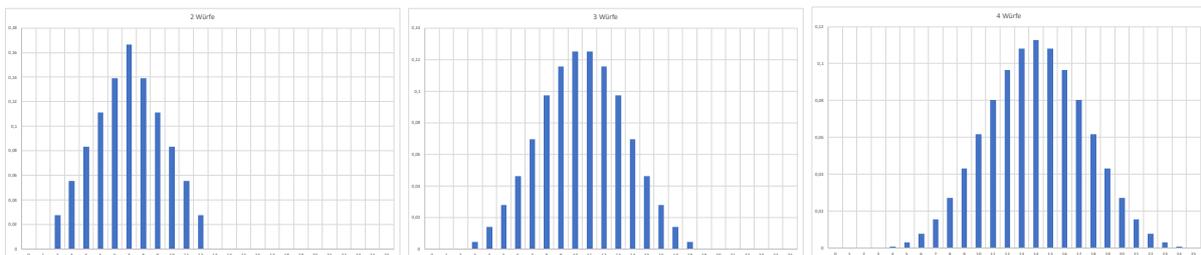
Wir gehen von einem fairen Würfel aus, das heißt, wir nehmen an, dass beim Würfeln keine Augenzahl bevorzugt oder benachteiligt ist. Jede Augenzahl kann mit derselben

Wahrscheinlichkeit auftreten. Weil sicher ist, dass eine der sechs möglichen Augenzahlen kommen muss, und weil es nicht sein kann, dass bei einem Wurf zugleich zwei verschiedene Augenzahlen oben liegen, hat jede Augenzahl die Wahrscheinlichkeit $\frac{1}{6}$. Die Wahrscheinlichkeitsdichtefunktion hat also für die Zahlen 1 bis 6 jeweils den Wert $\frac{1}{6}$. Für alle anderen Zahlen ist die Wahrscheinlichkeit, sie beim Wurf mit einem normalen fairen Würfel zu erhalten, gleich 0.



Um die Wahrscheinlichkeitsdichtefunktion für die Augensumme von 2 Würfeln zu erhalten, muss man jedes mögliche Ergebnis aus dem ersten Wurf mit jedem möglichen Ergebnis aus dem zweiten kombinieren, und muss dann die Wahrscheinlichkeiten aller Möglichkeiten, welche die gleiche Augensumme ergeben, aufsummieren.

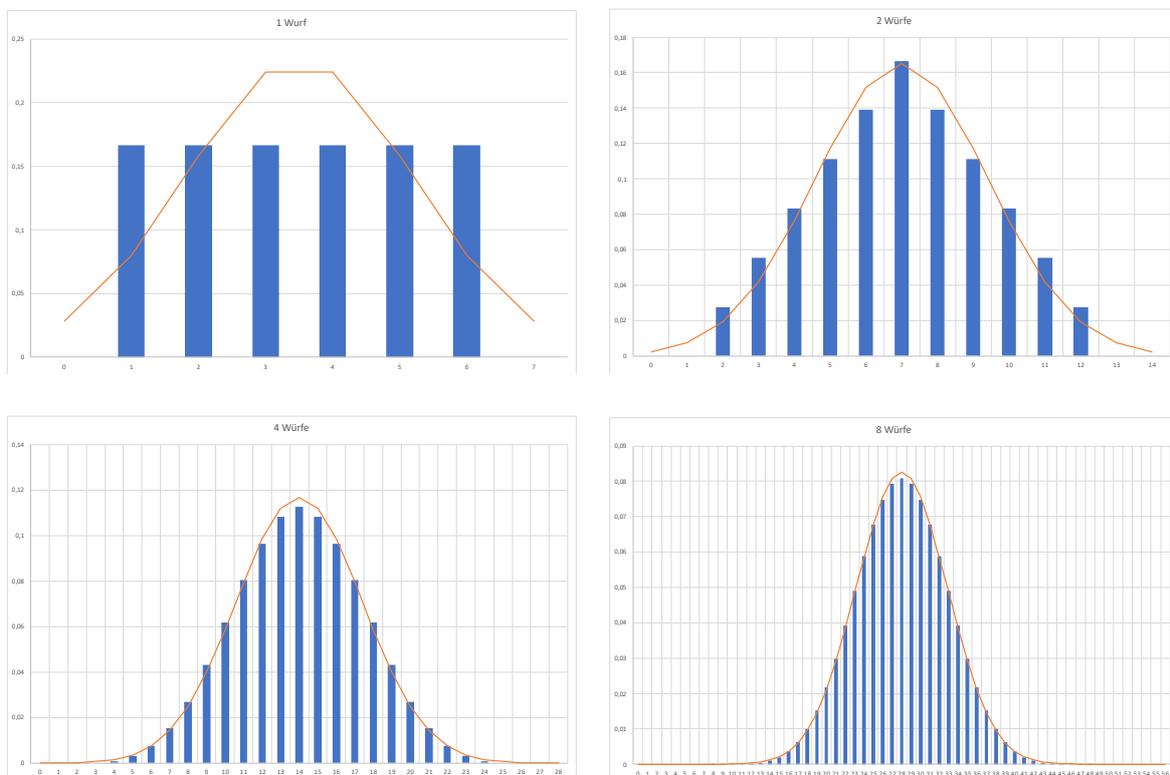
Das ist genau das, was man bei einer Faltung macht. Wenn $w(x)$ die Wahrscheinlichkeitsdichtefunktion eines einzelnen Wurfes ist, ist $w^2(x)$ (in unserer Schreibweise) die Dichtefunktion für zwei Würfe, $w^3(x)$ für 3 Würfe und $w^4(x)$ wenn man viermal würfelt. Die Funktionen sehen so aus:



Die Gerätefunktionen aus den vorherigen Beispielen hatten alle ihren Schwerpunkt bei genau $x = 0$. Daher haben sie aus der Rohdaten-Funktion zwar etwas gemacht, das in einem gewissen Sinn »breiter« war, aber insgesamt wurde nichts nach links oder rechts verschoben. Die Lage der Funktion wurde nicht verändert. Jetzt hat die Gerätefunktion ihren Mittelwert bei $x = 3,5$ (das ist der Erwartungswert für einen Wurf). Und daher verschiebt sich bei jeder Faltung der Mittelwert des Ergebnisses um genau diesen Wert.

Viel interessanter (in Bezug auf den zentralen Grenzwertsatz) ist aber die Form der Kurve. Sie nähert sich immer mehr einer Gaußschen Glockenkurve an, also genau jener Form, die eine Normalverteilung hat. In den folgenden Abbildungen gibt das blaue Säulendiagramm die durch Faltung ermittelte Wahrscheinlichkeitsdichtefunktion wieder. Zusätzlich wird die orange Kurve für eine Normalverteilung eingezeichnet, die denselben Mittelwert (μ) und dieselbe Standardabweichung (σ) hat.

Von Bild zu Bild wird die Anzahl der Würfel verdoppelt (1, 2, 4, 8).



Schon bei 8 Würfeln fällt es schwer, einen Unterschied zwischen der mehrfach gefalteten Gleichverteilungs-Funktion und der prognostizierten Normalverteilung zu unterscheiden. Bei noch mehr Würfeln nähert sich die Verteilung der Augensummen immer stärker an die Normalverteilung an.

Das gilt sogar für Anfangs-Verteilungen, die besonders widerspenstig sind, wie z.B. für die U-Verteilungen. Unter diesem Namen findet man in der Fachliteratur zwei verschiedene Verteilungen:

- **U-Verteilung 1:** »U« steht für »uniform«, und »u-distribution« oder »U-Verteilung« daher für »Gleichverteilung«. Daher ist damit eine bereits bestens bekannte Verteilung gemeint, nämlich jene, die durch eine Rechteckfunktion beschrieben wird. Diese Verteilung ist hier nicht gemeint.

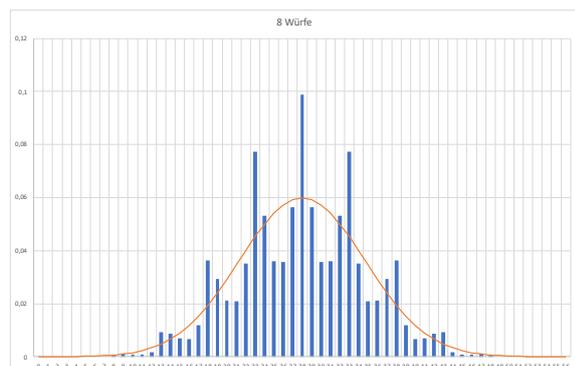
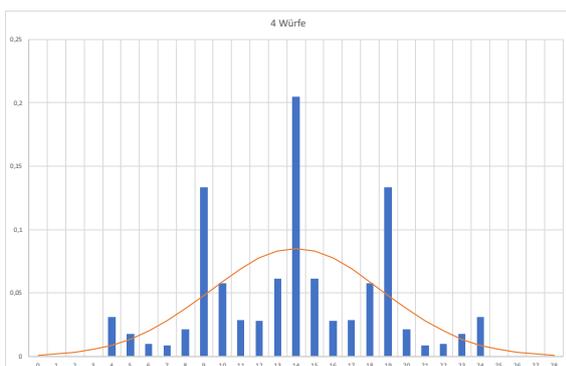
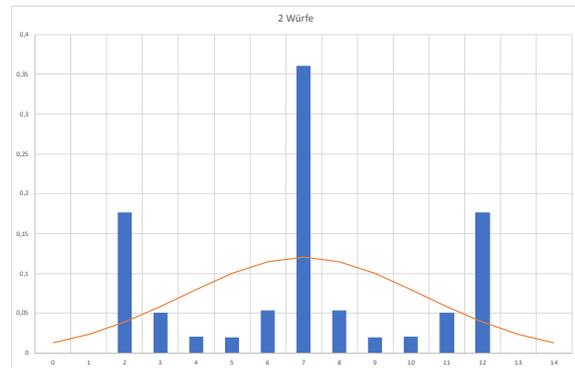
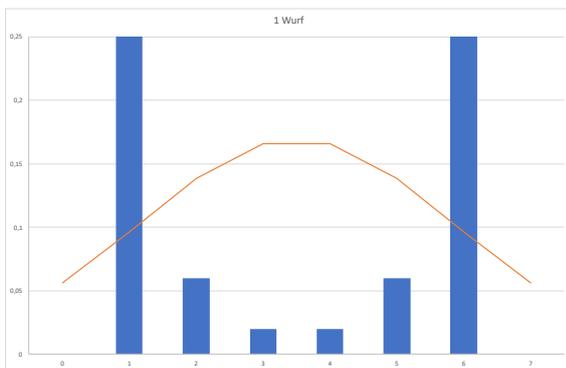
- U-Verteilung 2:** Hier soll der Buchstabe »U« die Form der Kurve symbolisieren. Wenn man zu einem zufälligen Zeitpunkt ein Foto von einer Pendeluhr macht, befindet sich das Pendel auf diesem Foto irgendwo zwischen seinen beiden Wendepunkten. Die U-Verteilung beschreibt genau die Wahrscheinlichkeitsdichte dieser harmonischen Schwingung. Wenn man hinter dem Pendel eine Skala anbringt, die beim linken Wendepunkt den Nullpunkt hat, und wo die Zahl 1 genau beim rechten Wendepunkt steht, dann kann man die Wahrscheinlichkeitsdichte im Intervall von 0 bis 1 durch diese Funktion beschreiben:

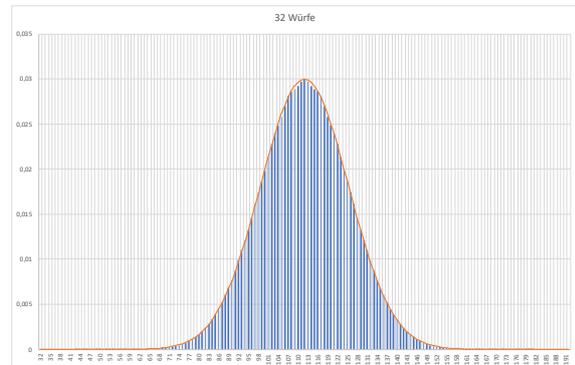
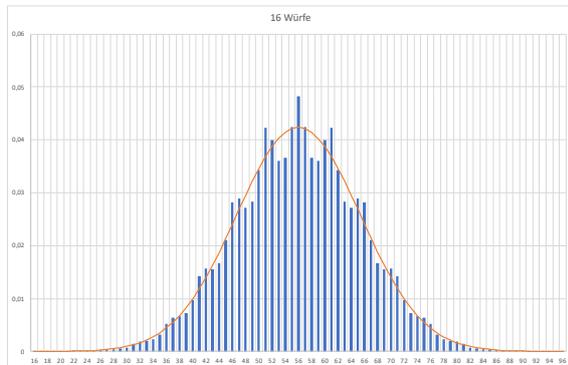
$$U(x) = \frac{1}{\pi} \cdot \sqrt{x \cdot (1 - x)}$$

Diese Funktion gibt also an, mit welcher Wahrscheinlichkeit man auf einem Foto von einem Pendel das Pendel an der Position x sehen kann.

Ganz allgemein bezeichnet man aber auch jede symmetrische Verteilung, die in der Mitte ein Minimum hat und umso größer wird, je weiter man sich von der Mitte entfernt, als U-Verteilung. Eine mögliche Ausprägung für einen unfairen Würfel ist im nächsten Bild dargestellt, die anderen Bilder zeigen wieder die wiederholte Selbstfaltung dieser Funktion.

Man kann sich vorstellen, dass der »Würfel« nun ein Quader ist, bei dem die Flächen für \square und \square besonders groß sind, und die Flächen für \square und \square dafür etwas kleiner.





Wie man sieht, benötigt man deutlich mehr Wiederholungen (32 statt 8), bis sich die Kurve so an die Form der Normalverteilung anpasst, dass man kaum noch Unterschiede sieht. Aber selbst der Würfel mit den U-verteiltern Wahrscheinlichkeiten liefert, wenn man ihn nur oft genug wirft und dabei die Augensummen addiert, eine Verteilung, die von der Normalverteilung nicht mehr zu unterscheiden ist.