Kapitel IV
Regulär stetige Zufallsexperimente
und
geometrische Wahrscheinlichkeiten

Zufallsexperimente mit überabzählbarem Ergebnisraum lassen sich leider nicht so einfach formal sauber definieren. Man vergegenwärtige sich, daß es immerhin bis 1933 dauerte, ehe Kolmogorov, basierend auf den Fortschritten in der Maß- und Integrationstheorie, eine befriedigende Axiomatisierung gelang und so die Etablierung der W-Theorie als mathematische Teildisziplin einleitete. Wegen des schon erwähnten Verzichts auf jegliche Kenntnisse aus der Maß- und Integrationstheorie in diesem Text müssen wir uns in der folgenden Definition auf eine sehr spezielle Teilklasse von ZE mit überabzählbarem Ergebnisraum beschränken, für die lediglich auf den Begriff des Riemann-Integrals zurückgegriffen wird. Immerhin reicht dies aus, um neben einigen anderen nicht diskreten Verteilungen die ebenso wichtige wie prominente *Normalverteilung* einzuführen, die uns dann später beim wichtigen zentralen Grenzwertsatz wieder begegnen wird. Im Anschluß daran werden wir uns kurz mit einer interessanten Klasse von Modellen befassen, die man mit dem Namen "stetige Laplace-Experimente" versehen könnte und in denen die betrachteten Ereignisse als geometrische Objekte und ihre Wahrscheinlichkeiten durch deren Volumina (Länge, Fläche) gegeben sind. Man spricht deshalb in diesem Kontext von *geometrischen Wahrscheinlichkeiten*. Ein bekanntes Beispiel aus dieser Modellklasse bildet das *Buffonsche Nadelproblem*.

13. Regulär stetige Zufallsexperimente

Unter einem Intervall in $\mathbb{R}$ verstehen wir jede Teilmenge der Form

$$(a,b),\ (a,b],\ [a,b),\ [a,b],\ (-\infty,b),\ (-\infty,b],\ (a,\infty),\ [a,\infty)$$

mit $-\infty < a \leq b < \infty$ sowie $\mathbb{R}$ selbst.

13.1. Definition. Ein *regulär stetiges Zufallsexperiment* ist ein Paar (Ω, f), bestehend aus einem Ergebnisraum Ω, der ein Intervall in $\mathbb{R}$ mit inneren Punkten bildet, und einer stetigen (möglicherweise uneigentlich) Riemann-integrierbaren Funktion $f : \Omega \to [0,\infty)$, genannt *Dichte* oder *Dichtefunktion*, derart, daß $\int_\Omega f(x)\,dx = 1$.

Für jedes Teilintervall I von Ω setzen wir

$$P(I) \overset{\text{def}}{=} \int_I f(x)\, dx \tag{13.1}$$

und interpretieren $P(I)$ als Wahrscheinlichkeit für das Auftreten des Ereignisses I.

Die Forderung, daß Ω innere Punkte besitzt, schließt die Möglichkeit $\Omega = \{x\}$ und damit $P = \delta_x$ aus. Da f nichtnegativ ist, gibt das Riemann-Integral von f über ein Intervall $I \subset \Omega$ bekanntlich den Flächeninhalt unter der Kurve von f in den Intervallgrenzen von I an, d.h. den Flächeninhalt von

$$\{(x, y) : x \in I, 0 \le y \le f(x)\}.$$

Dieser ist also die Wahrscheinlichkeit von I, und die Bedingung $\int_\Omega f(x)\, dx = 1$ sichert, daß die Wahrscheinlichkeit des gesamten Ergebnisraums Ω wie im diskreten Fall 1 beträgt. Bezeichnen a, b, $a < b$, die Intervallgrenzen von I, so kann man statt $\int_I f(x)\, dx$ die vertrautere Schreibweise $\int_a^b f(x)\, dx$ benutzen. Da I jedoch weder abgeschlossen noch endlich zu sein braucht, ist das Integral i.a. als *uneigentliches Riemann-Integral* zu interpretieren, d.h.

$$\int_a^b f(x)\, dx \overset{\text{def}}{=} \lim_{c \downarrow a, d \uparrow b} \int_c^d f(x)\, dx.$$

Bild 13.1 zeigt die Wahrscheinlichkeit des Intervalls $I = [1, 2]$ als grau eingefärbte Fläche am Beispiel der Dichte einer *Standard-Normalverteilung.*

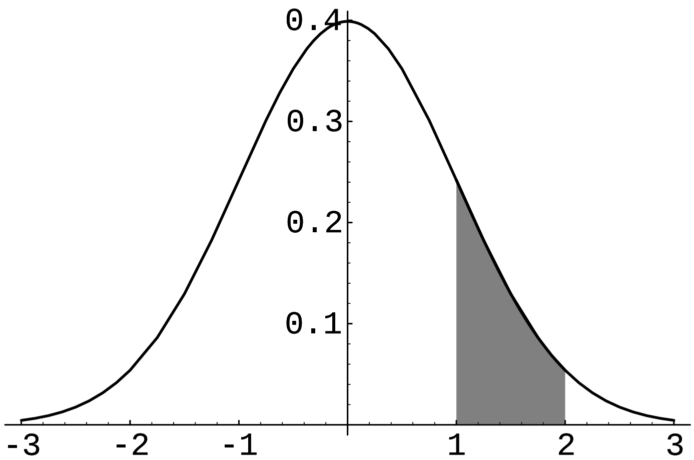


Bild 13.1. Wahrscheinlichkeit des Intervalls $[1, 2]$ unter einer Standard-Normalverteilung.

Die folgende Erweiterung von Definition 13.1 kann als Analogon zur Definition 1.11 für diskrete ZE gesehen werden.

13.2. Definition (Verallgemeinerung von 13.1). Ein *regulär stetiges Zufallsexperiment über* $\mathbb{R}$ ist ein Paar (Ω, f), bestehend aus einem Intervall $\Omega \subset \mathbb{R}$ mit nichtleerem Inneren und einer Funktion $f : \mathbb{R} \to [0, \infty)$, genannt *Dichte* oder *Dichtefunktion*, mit den folgenden Eigenschaften:

(1) f ist stetig auf Ω.
(2) f ist Riemann-integrierbar auf Ω mit $\int_\Omega f(x)\,dx = 1$.
(3) f verschwindet auf Ω^c, d.h. $f(x) = 0$ für alle $x \in \Omega^c$.

Für jedes Teilintervall I von $\mathbb{R}$ setzen wir

$$P(I) \stackrel{\text{def}}{=} \int_I f(x)\ dx \ = \int_{I\cap\Omega} f(x)\ dx \tag{13.2}$$

und interpretieren $P(I)$ als Wahrscheinlichkeit für das Auftreten des Ereignisses I.

Reguläre stetige ZE mit Ergebnisraum $\Omega \subset \mathbb{R}$ gemäß 13.1 lassen sich also stets als ZE über ganz $\mathbb{R}$ auffassen, indem man die Dichtefunktion f außerhalb von Ω gleich 0 setzt. Der Leser beachte, daß damit f auch auf ganz $\mathbb{R}$ (uneigentlich) Riemann-integrierbar und so das erste Integral auf der rechten Seite von (13.2) wohldefiniert bleibt.

1.3 beinhaltet die Definition der zu (Ω, f) gehörenden W-Verteilung P für beliebige offene Teilintervalle I von Ω. Anders als im diskreten Fall liegt aber die Definition von $P(A)$ für beliebige Teilmengen A von Ω keineswegs auf der Hand, weil die kanonische Möglichkeit

$$P(A) \stackrel{\text{def}}{=} \int_A f(x)\ dx \tag{13.3}$$

nicht in Frage kommt. Das Riemann-Integral ist nämlich leider nur über Intervalle erklärt. In der Tat lassen sich W-Verteilungen, worunter jedes normierte und σ-additive P (☞ Korollar 1.3) verstanden wird, i.a. und für regulär stetige ZE im besonderen nicht auf der gesamten Potenzmenge des Ergebnisraums Ω definieren. Hierin besteht das fundamentale Problem für die Axiomatisierung der W-Theorie und ein wesentlicher Grund für die Entstehung der Maß- und Integrationstheorie. Basierend auf diesen Feststellungen, die wir hier im einzelnen nicht weiter erörtern wollen und können, stehen im Hinblick auf regulär stetige ZE zwei Aufgaben im Vordergrund:

- Finde ein möglichst großes Teilsystem $\mathfrak{A}$ der Potenzmenge von Ω, welches alle offenen Teilintervalle von Ω enthält und unter den wichtigsten Mengenoperationen abgeschlossen ist und auf das P gemäß (13.1) in eindeutiger Weise fortgesetzt werden kann.
- Finde ein Integral, das bei Integration über ein Intervall dem Riemann-Integral entspricht und zugleich die Definition von P durch (13.3) erlaubt.

Hinsichtlich der ersten Aufgabe definiert man zunächst:

13.3. Definition. Sei Ω eine nichtleere Menge. Ein System $\mathfrak{A}$ von Teilmengen von Ω heißt *σ-Algebra über* Ω, wenn es folgende Eigenschaften besitzt:

$$\Omega \in \mathfrak{A}, \tag{13.4}$$

$$A \in \mathfrak{A} \quad \Rightarrow \quad A^c \in \mathfrak{A}, \tag{13.5}$$

$$A_1, A_2, \ldots \in \mathfrak{A} \quad \Rightarrow \quad \bigcup_{j \geq 1} A_j \in \mathfrak{A}. \tag{13.6}$$

Wie man sofort nachweist, sind σ-Algebren ebenfalls abgeschlossen unter endlicher Vereinigungs- sowie endlicher und abzählbarer Durchschnittsbildung. Da *beliebige* Durchschnitte von σ-Algebren offenkundig wieder eine σ-Algebra ergeben, existiert zu jedem Teilsystem $\mathfrak{E}$ der Potenzmenge von Ω eine kleinste σ-Algebra (über Ω), die $\mathfrak{E}$ enthält. Sie wird mit $\sigma(\mathfrak{E})$ bezeichnet.

Kehren wir zurück zur Situation eines regulär stetigen ZE's (Ω, f): Sei $\mathfrak{I} \subset \mathfrak{P}(\Omega)$ das System aller offenen Teilintervalle von Ω. Die von $\mathfrak{I}$ erzeugte σ-Algebra $\mathfrak{B}(\Omega) \stackrel{\text{def}}{=} \sigma(\mathfrak{I})$ heißt *Borelsche σ-Algebra (über Ω)*, ihre Elemente *Borel-Mengen* oder auch *Borel-meßbare Mengen*. Man kann zeigen, daß

$$\mathfrak{B}(\Omega) \neq \mathfrak{P}(\Omega).$$

Dies ist die schlechte Nachricht. Es gibt aber auch eine gute: Um eine Menge zu konstruieren, die nicht Borelsch ist, bedarf es der Verwendung des Auswahlaxioms, was aus anschaulicher Sicht bedeutet, daß die nicht meßbaren Mengen in das Reich der Exotik gehören. In der Maß- und Integrationstheorie wird nun gezeigt, daß das gemäß (13.1) auf $\mathfrak{I}$ definierte P in *eindeutiger Weise* zu einer normierten, σ-additiven Mengenfunktion auf $\mathfrak{B}(\Omega)$ fortsetzbar ist, und zwar gemäß (13.3) mit einem verallgemeinerten Integral, dem sogenannten *Lebesgue-Integral*. Dieses stimmt bei Integration über Intervalle sowie abzählbare Vereinigungen von solchen mit dem bekannten Riemann-Integral überein. In Analogie zum diskreten Fall nennt man P die zu (Ω, f) gehörende *regulär stetige Wahrscheinlichkeitsverteilung* und umgekehrt f die zu P gehörende *Dichte* oder *Dichtefunktion*. Auch hier besitzt P die in Satz 2.1 formulierten Eigenschaften und ist wieder stetig von oben und unten (☞ Korollar 2.4). Den fundamentalen Unterschied zu diskreten W-Verteilungen gibt folgender Satz:

13.4. Satz. *Gegeben ein regulär stetiges ZE (Ω, f) mit zugehöriger W-Verteilung P gilt für alle $x \in \Omega$*

$$P(\{x\}) = 0. \tag{13.7}$$

In einem regulär stetigen ZE hat also jedes Ergebnis die Wahrscheinlichkeit 0.

Beweis: Sei x ein innerer Punkt von Ω und $\varepsilon > 0$ so klein, daß $I_\varepsilon \stackrel{\text{def}}{=} [x - \varepsilon, x + \varepsilon] \subset \Omega$. Dann folgt

$$P(\{x\}) \leq P(I_\varepsilon) = \int_{x-\varepsilon}^{x+\varepsilon} f(y)\, dy$$

und daraus die Behauptung bei Grenzübergang $\varepsilon \downarrow 0$. Ist x ein Randpunkt von Ω, etwa der linke, so wähle $I_\varepsilon = [x, x+\varepsilon]$ mit $\varepsilon > 0$ so klein und argumentiere analog. ♢

Unser letztes Ergebnis dieses Abschnitts hält nochmals in einem Satz fest, daß jede regulär stetige W-Verteilung bereits durch ihre Werte auf dem System der offenen Intervalle eindeutig festgelegt ist:

13.5. Eindeutigkeitssatz. *Es sei (Ω, f_1) ein regulär stetiges ZE mit zugehöriger W-Verteilung P_1 und $P_2 : \mathfrak{B}(\Omega) \to [0,1]$ irgendeine weitere W-Verteilung (= normierte, σ-additive Mengenfunktion) über Ω. Dann impliziert $P_1(I) = P_2(I)$ für alle Teilintervalle I von Ω bereits $P_1 = P_2$.*

Beweis: Es genügt der Hinweis auf das schon weiter oben genannte Ergebnis aus der Maßtheorie, daß es genau eine normierte, σ-additive Mengenfunktion von $\mathfrak{B}(\Omega)$ nach $[0,1]$ gibt, die für alle Intervalle $I \subset \Omega$ durch $P_1(I) = \int_I f_1(x)\,dx$ definiert ist. ♢

Wir beenden den Abschnitt mit einer Reihe prominenter regulär stetiger W-Verteilungen.

13.6. Beispiel. Sei $\Omega = (a,b)$ ein beliebiges Teilintervall der reellen Achse mit endlichen Randpunkten a, b, $a < b$. Die regulär stetige W-Verteilung P über Ω mit der Dichte

$$f(x) = \frac{1}{b-a}, \quad x \in \Omega, \tag{13.8}$$

heißt *Gleichverteilung* oder *Rechteckverteilung auf* Ω, kurz $R(a,b)$*-Verteilung.* Der Grund für die erste Bezeichnung liegt natürlich darin, daß Teilintervallen von Ω gleicher Länge dieselbe, zu ihrer Länge proportionale Wahrscheinlichkeit zugeordnet wird, während die zweite Bezeichnung mnemonischen Bezug auf die Gestalt der Dichte nimmt (☞ Bild 13.2). Die Gleichverteilung ist das Gegenstück zur Laplace-Verteilung im diskreten Fall. Sie läßt sich in offenkundiger Weise auch für Intervalle in $\mathbb{R}^d$, $d \geq 2$, und sogar noch allgemeinere Teilmengen definieren (☞ Definition 14.2).

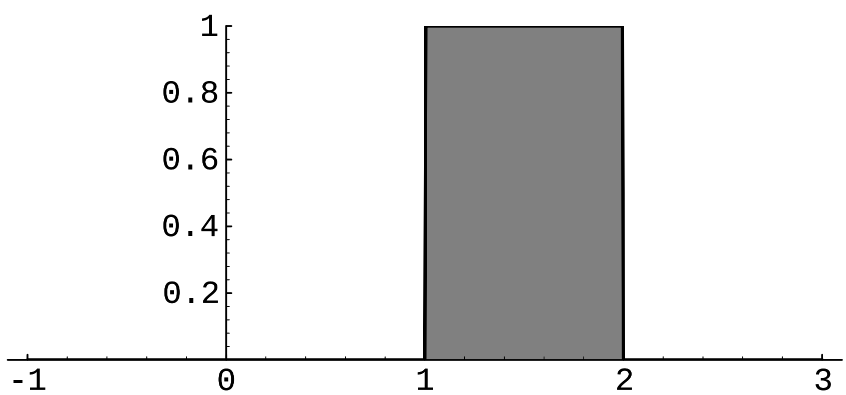


Bild 13.2. Die Dichte der $R(1,2)$-Verteilung.

13.7. Beispiel. Die regulär stetige W-Verteilung P über $\Omega = (0, \infty)$ mit der Dichte

$$f(x) \;=\; \theta e^{-\theta x}, \quad x, \theta > 0, \tag{13.9}$$

heißt *Exponentialverteilung mit Parameter* θ, kurz *Exp*(θ)*-Verteilung.* Sie wird häufig zur Modellierung von zufälligen Lebensdauern oder Wartezeiten benutzt.

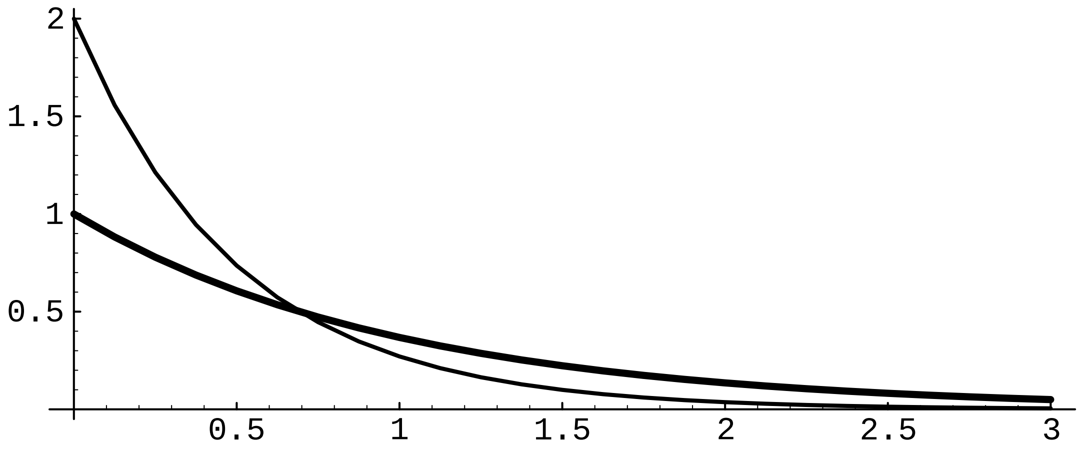


BILD 13.3. Die Dichten der Exp(1)- (fett) und Exp(2)-Verteilung.

13.8. Beispiel. Die regulär stetige W-Verteilung P auf $\Omega = \mathbb{R}$ mit der Dichte

$$f(x) \;=\; \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), \quad \mu \in \mathbb{R}, \sigma > 0, \tag{13.10}$$

heißt *Normal-* oder *Gauß-Verteilung mit Parametern* μ *und* σ^2, kurz $N(\mu, \sigma^2)$-Verteilung. Im Fall $\mu = 0$ und $\sigma^2 = 1$ spricht man auch von einer *Standard-Normalverteilung.* f ist die bekannte *Gaußsche Glockenkurve*, die man auf jedem Zehnmarkschein findet. Die Standard-Normalverteilung wird uns in Abschnitt 1? im zentralen Grenzwertsatz wiederbegegnen. Sie spielt in der W-Theorie eine herausragende Rolle.

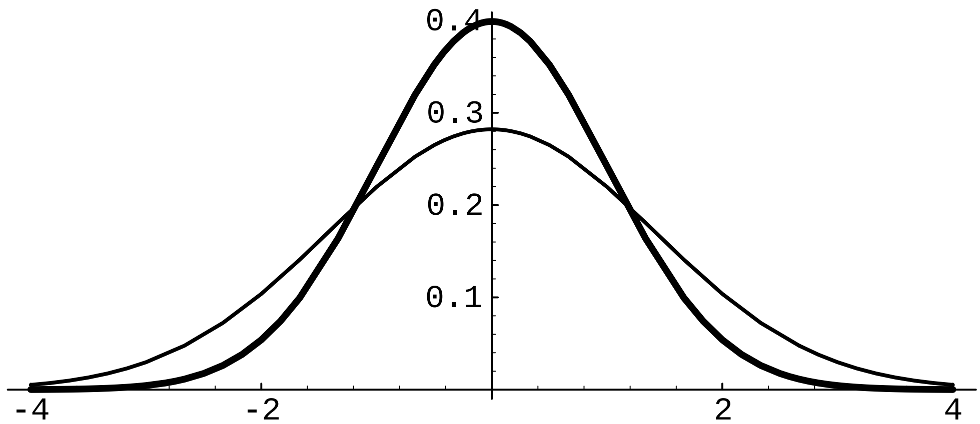


BILD 13.4. Die Dichten der $N(0,1)$- (fett) und $N(0,2)$-Verteilung.

14. Geometrische Wahrscheinlichkeiten

In der Literatur finden sich eine ganze Reihe interessanter Beispiele stochastischer Problemstellungen im Kontext von Zufallsexperimenten, denen eine Gleichverteilung auf einer geometrischen Menge, z.B. ein Kreis, ein Dreieck oder eine Kugel, zugrundeliegt. Um dies zu präzisieren, beginnen wir mit einer Definition, wobei im folgenden $\lambda\!\!\lambda^d(A)$ stets das Volumen einer Menge $A \subset \mathbb{R}^d$ bezeichnet (= Fläche im Fall $d = 2$ und = Länge im Fall $d = 1$).

14.1. Definition. Eine beschränkte Teilmenge $\Omega \subset \mathbb{R}^d$, $d \geq 1$, heißt *geometrisch regulär*, wenn sie von innen und außen durch endliche Vereinigungen p.d. d-dimensionale Intervalle beliebig genau approximierbar ist, d.h. für jedes $\varepsilon > 0$ existieren endliche Folgen $(I_j)_{1\leq j\geq m}$ und $(J_k)_{1\leq k\leq n}$ jeweils p.d. Intervalle, so daß

$$\sum_{j=1}^{m} I_j \;\subset\; \Omega \;\subset\; \sum_{k=1}^{n} J_k$$

und

$$0 \;\leq\; \sum_{k=1}^{n} \lambda\!\!\lambda^d(J_k) \;-\; \sum_{j=1}^{m} \lambda\!\!\lambda^d(I_j) \;<\; \varepsilon,$$

wobei $\lambda\!\!\lambda^d(I)$ das Volumen (Länge, Fläche) von I bezeichnet.

Bild 14.1 veranschaulicht in Dimension 2, wie man eine innere und äußere Approximation einer gegebenen Menge Ω erhalten kann, indem man über Ω ein Raster gleichgroßer Quadrate legt. Die innere Approximation besteht dann aus allen Quadraten, die ganz in Ω enthalten sind, während die äußere Approximation zusätzlich auch diejenigen Quadrate einschließt, die mit Ω einen nichtleeren Schnitt besitzen. Mit zunehmender Verfeinerung des Rasters strebt dann das Gesamtvolumen der nur in der äußeren Approximation auftretenden Quadrate im Fall einer geometrisch regulären Menge gegen 0.

Jede geometrisch reguläre Menge besitzt somit ein Volumen, und es ist offensichtlich, daß alle wohlbekannten geometrischen Objekte wie etwa der Kreis, das Dreieck, die Kugel, der Quader oder die Pyramide geometrisch reguläre Mengen bilden.

Auf geometrisch regulären Ergebnisräumen können wir nun als Verallgemeinerung von Beispiel 13.6 wiederum die Gleichverteilung in kanonischer Weise einführen.

14.2. Definition und Satz. Sei Ω eine nichtleere geometrisch reguläre Teilmenge des $\mathbb{R}^d$ mit $\lambda\!\!\lambda^d(\Omega) > 0$, und sei $\mathcal{G}$ die Klasse aller geometrisch regulären Teilmengen von Ω. Dann wird durch

$$P(A) \stackrel{\text{def}}{=} \frac{\lambda\!\!\lambda^d(A)}{\lambda\!\!\lambda^d(\Omega)} \;=\; \frac{\text{Volumen von } A}{\text{Volumen von } \Omega}, \quad A \in \mathcal{G}, \tag{14.1}$$

eine W-Verteilung auf Ω definiert, genannt *Gleichverteilung auf* Ω.

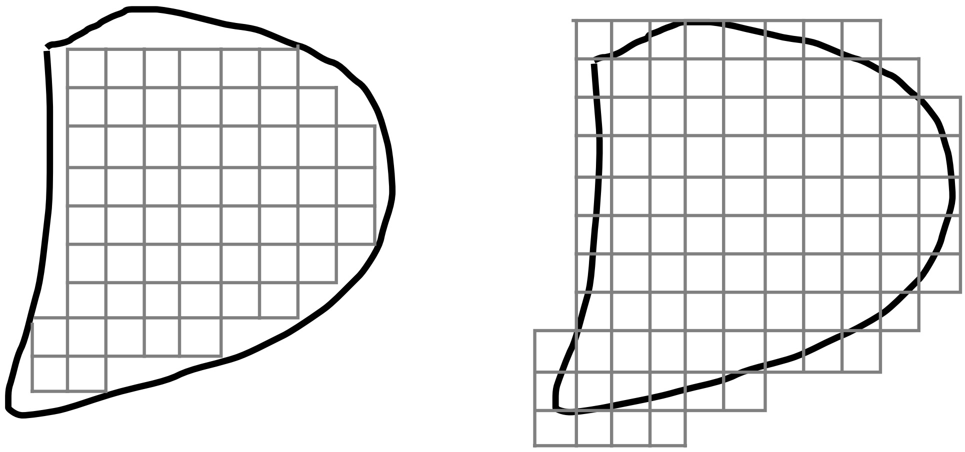


Bild 14.1. Eine Menge Ω mit innerer (links) und äußerer (rechts) Approximation durch Intervalle

Beweisen können wir diesen Satz nicht, bemerken aber noch, daß P definiert ist auf der σ-Algebra $\mathfrak{B}(\Omega)$, der kleinsten σ-Algebra, die alle Mengen $\Omega \cap I$, $I \subset \mathbb{R}^d$ ein Intervall, und damit auch $\mathcal{G}$ enthält. Für die Betrachtung der nachfolgenden Beispiele reicht es aber aus zu wissen, daß sich die Wahrscheinlichkeit eines geometrisch regulären Ereignisses als ihr relatives Volumen bezogen auf das Volumen von Ω ergibt. Man spricht dann von *geometrischen Wahrscheinlichkeiten.*

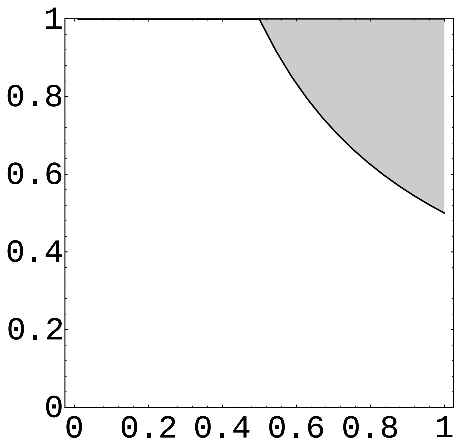


Bild 14.2. Das Ereignis E in Beispiel 14.3 (graue Fläche)

14.3. Beispiel. Wähle zwei Punkte x, y zufällig im Einheitsintervall (0,1) und betrachte das möglicherweise degenerierte Rechteck in $(0,1)^2$ mit den Eckpunkten $(0,0), (x,0), (0,y)$ und (x,y). Wie groß ist die Wahrscheinlichkeit, daß dieses Rechteck eine Fläche von mehr als $\frac{1}{2}$ besitzt? Als Ergebnisraum für die Wahl von (x,y) wählen wir $\Omega \stackrel{\text{def}}{=} (0,1)^2$, und die zufällige Auswahl von x und y rechtfertigt, auf Ω als W-Verteilung P die Gleichverteilung zugrundezulegen. Da das o.a. Rechteck offenkundig die Fläche xy besitzt, lautet das gesuchte Ereignis

$$E \stackrel{\text{def}}{=} \left\{(x,y) \in (0,1)^2 : xy > \frac{1}{2}\right\},$$

das sich auch in der Form

$$E = \left\{(x,y) \in (0,1)^2 : x > \frac{1}{2}, y > \frac{1}{2x}\right\}$$

schreiben läßt und in Bild 14.2 als graue Fläche ausgewiesen ist. Wegen $\lambda\!\!\lambda^2(\Omega) = 1$ gilt $P(E) = \lambda\!\!\lambda^2(E)$, so daß wir nur noch den Flächeninhalt von E berechnen müssen, der sich aber offenkundig als $\frac{1}{2}$ minus der Fläche unter der Kurve $x \mapsto \frac{1}{2x}$ für $x \in (\frac{1}{2}, 1)$ ergibt, also

$$P(E) = \frac{1}{2} - \int_{1/2}^{1} \frac{1}{2x}\, dx = \frac{1}{2}(1 - \log 2).$$

14.4. Beispiel. (Das Buffonsche Nadelproblem) Eine Ebene sei mit einem Raster paralleler Geraden vom Abstand 1 versehen. Auf dieses Raster wird eine Nadel der Länge $a < 1$ geworfen. Wir fragen nach der Wahrscheinlichkeit, daß sie eine der Geraden schneidet.

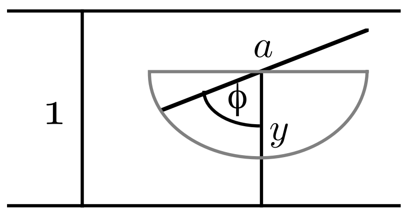


Bild 14.3. Illustration zum Buffonschen Nadelproblem

Um die Frage beantworten zu können, müssen wir zunächst ein Modell finden, welches das Ergebnis des Nadelwurfs plausibel beschreibt. Da die Nadel kürzer als der Abstand zwischen zwei benachbarten Rasterlinien ist, können wir uns auf die Beschreibung ihrer Lage zwischen zwei Linien beschränken. Als Koordinaten wählen wir die Ordinate des Mittelpunkts, y, und den Winkel mit der Ordinate, ϕ. Zur Illustration ☞ Bild 14.3. Wirft man die Nadel aufs Geratewohl auf das Raster, wird (ϕ, y) im Rechteck $[-\frac{\pi}{2}, \frac{\pi}{2}) \times [0,1)$ gleichverteilt sein, was insbesondere die Vorstellung berücksichtigt, daß sich Ordinate y und Winkel ϕ unabhängig voneinander ergeben. Der zugrundegelegte W-Raum lautet also

$$([-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \times [0,1), \mathfrak{B}^2_{[-\frac{\pi}{2},\frac{\pi}{2})\times[0,1)}, P), \tag{14.2}$$

wobei P die Gleichverteilung auf $[-\frac{\pi}{2}, \frac{\pi}{2}) \times [0,1)$ angibt.

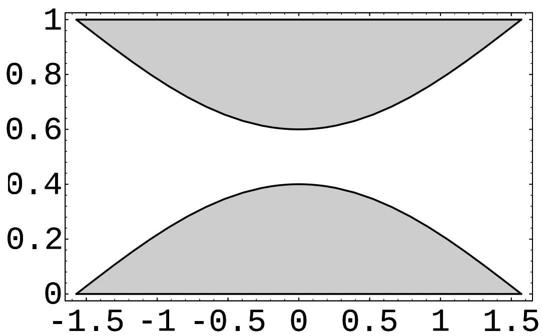


Bild 14.4. Das Ereignis "Nadel schneidet eine der Geraden" (schattierte Fläche)

Eine Nadel mit den Koordinaten (ϕ, y) schneidet die untere Gerade genau dann, wenn $y < \frac{a}{2} \cos \phi$, und die obere Gerade, wenn $1 - y < \frac{a}{2} \cos \phi$. Das gesuchte Ereignis E lautet also

$$E \stackrel{\text{def}}{=} \left\{ (\phi, y) \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \times [0, 1) : y < \frac{a}{2} \cos \phi \text{ oder } 1 - y < \frac{a}{2} \cos \phi \right\},$$

das in Bild 14.4 grau schattiert dargestellt ist ($a = 0.8$). Die Wahrscheinlichkeit von E unter der Gleichverteilung ergibt sich demnach als

$$P(E) \;=\; \frac{1}{\pi} \lambda\!\!\lambda^2(E) \;=\; \frac{a}{\pi} \int_{-\pi/2}^{\pi/2} \cos \phi \, d\phi \;=\; \frac{2a}{\pi}. \tag{14.3}$$

Bild 14.5. Georges-Louis Leclerc de Buffon (1707-1788)

Das beschriebene Problem eröffnet aufgrund seiner Verbindung zur Kreiszahl π eine auf Wahrscheinlichkeiten beruhende, ungewöhnliche Methode zur approximativen Berechnung von π. Diese Methode stammt von Georges-Louis Leclerc de Buffon (☞ Bild 14.5), der sie im Alter von 20 Jahren erfand. Buffon warf Stöcke über die Schulter auf einen gekachelten Fußboden. Anschließend zählte er, wie oft sie die Fugen trafen. Eine praktikablere Variante beschrieb Jakow I. Perelman im Buch "Unterhaltsame Geometrie". Man nehme eine kurze, ca. 2 cm lange Nadel oder einen anderen Metallstift mit ähnlicher Länge und Durchmesser, am besten ohne Spitze, und zeichne auf ein Blatt Papier eine Reihe dünner paralleler Striche, die um die doppelte Länge der Nadel voneinander entfernt sind. Dann läßt man die Nadel sehr häufig (mehrere hundert- oder tausendmal) aus einer beliebigen Höhe auf das Blatt fallen und notiert, ob die Nadel eine Linie schneidet oder nicht. Das Berühren eines Striches durch ein Nadelende zählt dabei als Schnittpunkt. Die Division der Gesamtzahl der Nadelwürfe durch die Zahl der Fälle, in denen die Nadel eine Linie geschnitten hat, ergibt im Ergebnis eine Näherung von π (Gesetz der großen Zahlen). Die Nadel kann dabei auch gebogen oder mehrfach geknickt sein, wobei in diesem Fall auch mehr als ein Schnittpunkt pro Wurf möglich ist und entsprechend gezählt werden muß. In der Mitte des 19. Jahrhunderts kam der Schweizer Astronom Johann Rudolf Wolf durch 5.000 Nadelwürfe auf einen Wert von $\pi = 3.159$. Ähnliche Experimente wurden später auch von anderen durchgeführt. Die Ergebnisse (nach Umrechnung auf Rasterabstand 1) sind in folgender Tabelle zusammengefaßt und [?, S. 97] entnommen:

Autor	Nadellänge a	Zahl der Würfe n	Zahl der Überschneidungen	Schätzwert für π
Wolf (1850)	0.8	5000	2531	3.159
Smith (1855)	0.6	3204	1218	3.155
De Morgan (1860)	1.0	600	382	3.137
Fox (1864)	0.75	1030	489	3.160
Lazzerini (1901)	0.83	3408	1808	3.142
Reina (1925)	0.5419	2520	859	3.180

14.5. Beispiel. (Das Bertrandsche Paradoxon) Man wähle in einem Kreis eine zufällige Sehne. Wie groß ist die Wahrscheinlichkeit dafür, daß die Sehne länger ist als die Seiten eines in denselben Kreis eingeschriebenen regelmäßigen Dreiecks? War beim Buffonschen Nadelproblem die Präzisierung eines mathematischen Modells für das Experiment des Nadelwurfs relativ problemlos, so bieten sich für die Modellierung der zufälligen Auswahl einer Sehne mehrere Möglichkeiten. Auf diesen Umstand aufmerksam zu machen, war das Ziel der von Joseph Louis François Bertrand (1822-1900) aufgeworfenen Frage [2]. Wir stellen im folgenden drei Modelle vor, die alle zu verschiedenen Wahrscheinlichkeiten für das gesuchte Ereignis führen, ein Umstand, der früher paradox erschien, weil man nicht verstand, daß die jeweiligen Berechnungen tatsächlich auf verschiedenen Modellannahmen basieren.

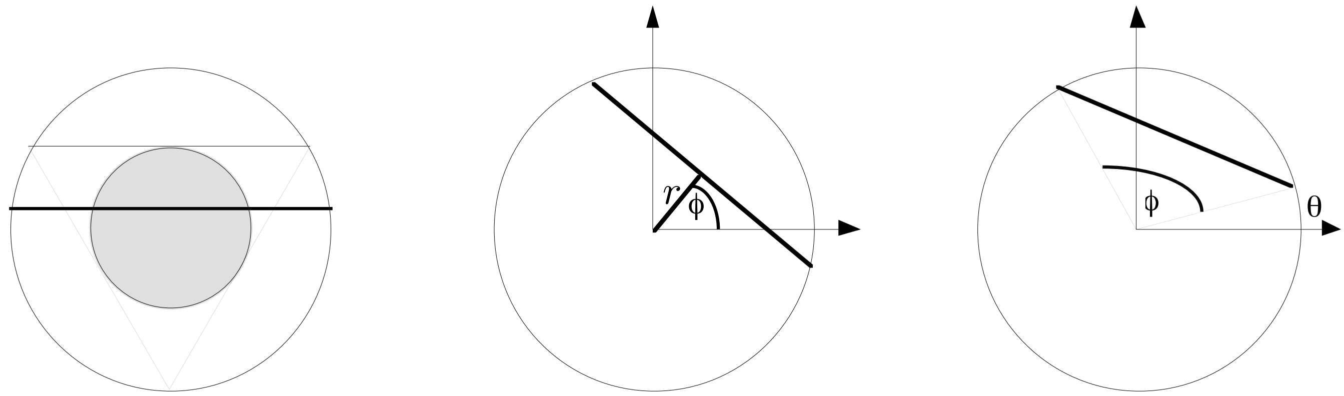
Ohne Beschränkung der Allgemeinheit dürfen wir in allen drei nachfolgenden Modellen voraussetzen, daß der gegebene Kreis den Mittelpunkt 0 und den Radius 1 besitzt. Außerdem bezeichne P stets die Gleichverteilung auf dem jeweiligen Ergebnisraum.

Modell 1: Jede Sehne ist durch ihren Mittelpunkt festgelegt. Wir können daher die zufällige Auswahl einer Sehne als Auswahl ihres Mittelpunkts gemäß der Gleichverteilung auf der Einheitskreisscheibe $K(1) \stackrel{\text{def}}{=} \{(x,y) : 0 \leq (x^2+y^2)^{1/2} < 1\}$ auffassen. Dies führt zum W-Raum $(K(1), \mathfrak{B}^2_{K(1)}, P)$. Da jedes eingeschriebene Dreieck die senkrecht auf den Seiten stehenden Radien halbiert (☞ Bild 14.6), erhalten wir als gesuchtes Ereignis diejenigen Sehnenmittelpunkte, die vom Kreismittelpunkt 0 nicht weiter als $\frac{1}{2}$ entfernt sind, d.h.

$$E \stackrel{\text{def}}{=} K(1/2) \;=\; \{(x,y) : 0 \leq (x^2+y^2)^{1/2} < 1/2\}.$$

Die gesuchte Wahrscheinlichkeit lautet

$$P(E) \;=\; \frac{1}{\pi}\lambda\!\!\lambda^2(K(1/2)) \;=\; \frac{1}{\pi}\cdot\pi\cdot\left(\frac{1}{2}\right)^2 \;=\; \frac{1}{4}. \tag{14.4}$$



Modell 1 Modell 2 Modell 3

Bild 14.6. Illustrationen zum Betrandschen Paradoxon

Modell 2: Die Länge der Sehne ist durch den Abstand $r \in [0,1)$ ihres Mittelpunkts vom Kreismittelpunkt festgelegt, die Sehne selbst, wenn wir zusätzlich den Winkel $\phi \in [0,2\pi)$ zwischen dem Radius durch den Sehnenmittelpunkt und der Abszisse angeben. Verstehen wir unter zufälliger Auswahl einer Sehne die Auswahl von (r,ϕ) gemäß der Gleichverteilung auf $[0,1)\times[0,2\pi)$, erhalten wir als W-Raum $([0,1)\times[0,2\pi), \mathfrak{B}^2_{[0,1)\times[0,2\pi)}, P)$ und als gesuchtes Ereignis

$$E \;\stackrel{\text{def}}{=}\; [0,1/2)\times[0,2\pi)$$

mit der Wahrscheinlichkeit

$$P(E) \;=\; \frac{1}{2\pi}\lambda\!\!\lambda^2([0,1/2)\times[0,2\pi)) \;=\; \frac{1}{2}. \tag{14.5}$$

Modell 3: Die Sehne ist festgelegt durch einen Punkt θ auf der Kreislinie, die wir durch $[0,2\pi)$ parametrisieren, und dem Winkel ϕ zwischen den Radien der beiden Sehnenendpunkte. Als zufällige Auswahl verstehen wir diesmal die Auswahl von (θ,ϕ) gemäß einer Gleichverteilung auf $[0,2\pi)\times[0,\pi)$, was zum W-Raum $([0,2\pi)\times[0,\pi), \mathfrak{B}^2_{[0,2\pi)\times[0,\pi)}, P)$ führt. Die Sehne ist offensichtlich länger als die Seiten eines eingeschriebenen gleichseitigen Dreiecks, wenn $\phi > \frac{2\pi}{3}$ gilt (☞ Bild 14.6), so daß sich

$$E \;\stackrel{\text{def}}{=}\; [0,2\pi)\times(2\pi/3,\pi)$$

als gesuchtes Ereignis mit Wahrscheinlichkeit

$$P(E) \;=\; \frac{1}{2\pi^2}\lambda\!\!\lambda^2([0,2\pi)\times(2\pi/3,\pi)) \;=\; \frac{1}{3} \tag{14.6}$$

ergibt.

Der Leser wird sich vermutlich verwirrt fragen, welche der drei Antworten auf die zu Beginn gestellte Frage denn nun richtig ist. Bertrand bemerkt hierzu [2]: "Keine von den dreien ist falsch, keine ganz richtig; die Frage ist schlecht gestellt." Das Problem besteht darin, daß nicht vollkommen klar ist, was unter der zufälligen Auswahl einer Sehne zu verstehen ist. Jedes vorgeschlagene Modell wählt zunächst eine Codierung (Parametrisierung) der Sehne innerhalb einer Grundmenge und dann die Gleichverteilung auf dieser Grundmenge. Die Konsequenz ist dann offenbar, daß verschiedene (plausible) Codierungen zu verschiedenen Ergebnissen führen. Mittels Argumenten, zu deren Begründung und präzisen Formulierung uns hier die Mittel fehlen, kann man aber zeigen, daß dem 2. Modell gegenüber den anderen beiden der Vorzug zu geben ist (☞ [1] oder [3]).

15. Regulär stetige Zufallsgrößen

In diesem Abschnitt werden wir den Begriff der Zufallsvariable auf die allgemeine in der W-Theorie zugrundegelegte Situation erweitern. Dem weitestgehenden Verzicht auf maß- und integrationstheoretische Hilfsmittel tragen wir auch weiterhin Rechnung, indem wir anschließend nur solche Zufallsvariablen betrachten, deren Verteilungen regulär stetig sind.

15.1. Definition. Ein *Wahrscheinlichkeitsraum* (kurz W-Raum) ist ein Tripel $(\Omega, \mathfrak{A}, P)$, bestehend aus

- einer nichtleeren Menge Ω, genannt *Ergebnisraum*,
- einer σ-Algebra $\mathfrak{A}$ über Ω, deren Elemente die *meßbaren Mengen* (Ereignisse) bilden, denen eine Wahrscheinlichkeit zugeordnet wird

sowie

- einer normierten, σ-additiven Mengenfunktion $P : \mathfrak{A} \to [0,1]$, genannt *W-Verteilung* oder *W-Maß über* Ω.

In Übereinstimmung mit früheren Definitionen heißt $(\Omega, \mathfrak{A}, P)$ außerdem

- *diskret*, falls P eine diskrete W-Verteilung bildet, also eine abzählbare Menge $\Omega_0 \in \mathfrak{A}$ mit $P(\Omega_0) = 1$ existiert (in diesem Fall kann man immer $\mathfrak{A} = \mathfrak{P}(\Omega)$ wählen),
- *regulär stetig*, falls Ω ein Intervall in $\mathbb{R}$ mit inneren Punkten, $\mathfrak{A} = \mathfrak{B}(\Omega)$ die Borelsche σ-Algebra über Ω und P eine regulär stetige W-Verteilung über Ω ist.

Bei der anschließenden Einführung von Zufallsgrößen (= reellwertigen Zufallsvariablen) auf allgemeinen W-Räumen können wir im wesentlichen die im vorherigen Abschnitt gemachten Definitionen übernehmen. Lediglich der Tatsache, daß die auf dem Urbildraum gegebene σ-Algebra i.a. nicht der Potenzmenge entspricht, muß in geeigneter Weise Rechnung getragen werden (☞ (14.1)). Wir weisen darauf hin, daß die in Lemma 4.2 nachgewiesenen Eigenschaften der Urbildfunktion X^{-1} einer Abbildung $X : \Omega \to \mathbb{R}$ ohne jede Voraussetzung an die Menge Ω gilt.

15.2. Definition und Satz. Sei $(\Omega, \mathfrak{A}, P)$ ein W-Raum. Eine Abbildung $X : \Omega \to \mathbb{R}$, die der sogenannten *Meßbarkeitsbedingung*

$$X^{-1}(C) \in \mathfrak{A} \quad \text{für alle } C \in \mathfrak{B}(\mathbb{R}) \tag{15.1}$$

genügt, heißt *Zufallsgröße.* Die durch

$$P^X(C) \stackrel{\text{def}}{=} P(X^{-1}(C)) \tag{15.2}$$

definierte Mengenfunktion $P^X : \mathfrak{B}(\mathbb{R}) \to [0,1]$ ist normiert und σ-additiv, also eine W-Verteilung über $\mathbb{R}$, und wird als die *durch X induzierte Verteilung* oder auch *Verteilung von X (unter P)* bezeichnet. Wir nennen X *diskret* bzw. *regulär stetig*, wenn P^X diese Eigenschaft besitzt.

Der Beweis der Normiertheit und σ-Additivität von P^X entspricht vollkommen dem Beweis im diskreten Fall (☞ Satz 4.3) und braucht deshalb nicht nochmals gegeben zu werden. Wenngleich wir hiernach auf den Nachweis der Meßbarkeitsbedingung (15.1) für speziell gegebene Zufallsgr¨ßen stets verzichten werden, ist es wichtig, daß der Leser zumindest ihre Notwendigkeit versteht: Erst sie garantiert, daß die Festlegung (15.2) für P^X überhaupt Sinn macht, denn die über dem Urbildraum Ω gegebene W-Verteilung P ist nur auf $\mathfrak{A}$ definiert. Entspricht $\mathfrak{A}$ der Potenzmenge von Ω, ist die Bedingung natürlich automatisch erfüllt. Sie spielt deshalb bei Zufallsvariablen auf diskreten W-Räumen keine Rolle.

Wir erwähnen an dieser Stelle, daß die in Abschnitt 4 eingeführten Schreibweisen ebenso wie Definition 4.11 und Satz 4.12 über die *Randverteilungen von Zufallsvektoren* in der nunmehr vorliegenden Situation unverändert übernommen werden können. Auch heißt, gegeben eine Zufallsgröße X, die durch

$$F^X(x) \stackrel{\text{def}}{=} P(X \leq x)$$

definierte Funktion $F^X : \mathbb{R} \to [0,1]$ weiterhin *Verteilungsfunktion von X* oder P^X (☞ Definition 5.1). Der anschließende Satz verdeutlicht im Vergleich mit Satz 5.2 den fundamentalen Unterschied zwischen diskreten und regulär stetigen Zufallsgrößen.

15.3. Satz. *Die Verteilungsfunktion F^X einer regulär stetigen Zufallsgröße X besitzt folgende Eigenschaften:*

(a) F^X ist monoton wachsend, d.h. $x_1 \leq x_2$ impliziert $F^X(x_1) \leq F^X(x_2)$.

(b) $\lim_{x \to -\infty} F^X(x) = 0$ und $\lim_{x \to \infty} F^X(x) = 1$.

(c) F^X ist stetig und mit Ausnahme von höchstens zwei Punkten stetig differenzierbar.

Beweis: (a) und (b) ergeben sich vollkommen analog zum diskreten Fall, so daß wir uns gleich Teil (c) zuwenden können. Da X regulär stetig ist, existiert ein Intervall $I = (a, b)$, $-\infty \le a < b \le \infty$, sowie eine auf I stetige und Riemann-integrierbare Funktion f^X mit $f^X_{|[a,b]^c} \equiv 0$ und $\int_a^b f^X(x)\, dx = 1$, so daß

$$F^X(t) \;=\; P^X((-\infty, t]) \;=\; \int_{I\cap(-\infty,t]} f^X(x)\, dx$$
$$= \begin{cases} 0, & \text{falls } t < a, \\ \displaystyle\int_a^t f^X(x)\, dx, & \text{falls } a \le t < b, \\ \displaystyle\int_a^b f^X(x)\, dx \;=\; 1, & \text{falls } t \ge b. \end{cases}$$

Dies zeigt die Stetigkeit der Funktion F^X und ferner auch ihre Differenzierbarkeit für alle $x \in \mathbb{R}\backslash\{a, b\}$. $\diamondsuit$

Wir erwähnen an dieser Stelle, daß über die Differenzierbarkeit von F^X in den Randpunkten a, b (sofern diese endlich sind) keine allgemeine Aussage gemacht werden kann. Wählt man beispielsweise eine auf $(0, 1)$ gleichverteilte Zufallsgröße X mit Dichte $f^X = \mathbf{1}_{(0,1)}$, so hat X die Verteilungsfunktion

$$F^X(x) \;=\; \begin{cases} 0, & \text{falls } x \le 0, \\ x, & \text{falls } 0 < x < 1, \\ 1, & \text{falls } x \ge 1, \end{cases}$$

und diese ist in den Punkten 0 und 1 offenbar nicht differenzierbar.

Wenden wir uns schließlich der Frage zu, wie *Erwartungswert* und *Varianz* einer regulär stetigen Zufallsgröße definiert werden. Dabei müssen wir uns mit ad hoc Definitionen begnügen, da uns der Begriff des allgemeinen Maßintegrals nicht zur Verfügung steht. Damit unsere Definitionen jedoch nicht vollkommen "vom Himmel fallen", geben wir im folgenden ein Approximationsargument, welches diese plausibel macht. Zu diesem Zweck sei X eine regulär stetige Zufallsgröße mit Werten in einem beschränkten Intervall $I = (a, b)$ der Länge $l \overset{\text{def}}{=} b - a$, also $X : \Omega \to (a, b)$. Die Dichte f^X verschwinde also außerhalb von I und sei auf I stetig und Riemann-integrierbar. Wir approximieren X von oben und unten durch diskrete Zufallsgrößen $\underline{X}_n : \Omega \to [a, b]$ und $\overline{X}_n : \Omega \to [a, b]$, $n \ge 1$, indem wir

$$\underline{X}_n(\omega) \;\overset{\text{def}}{=}\; \sum_{k=1}^{n} \left(a + \frac{(k-1)l}{n}\right) \mathbf{1}_{(a+(k-1)l/n,\, a+kl/n]}(X(\omega))$$

und

$$\overline{X}_n(\omega) \;\overset{\text{def}}{=}\; \sum_{k=0}^{n-1} \left(a + \frac{kl}{n}\right) \mathbf{1}_{(a+(k-1)l/n,\, a+kl/n]}(X(\omega))$$

für $\omega \in \Omega$ und $n \geq 1$ setzen. Dann gilt offenbar $\underline{X}_n \leq X \leq \overline{X}_n$,

$$\begin{aligned}
E\underline{X}_n &= \sum_{k=1}^{n} \Big(a + \frac{(k-1)l}{n}\Big) P\Big(\underline{X}_n = a + \frac{(k-1)l}{n}\Big) \\
&= \sum_{k=1}^{n} \Big(a + \frac{(k-1)l}{n}\Big) P\Big(a + \frac{(k-1)l}{n} < X \leq a + \frac{kl}{n}\Big) \\
&= \sum_{k=1}^{n} \Big(a + \frac{(k-1)l}{n}\Big) \int_{a+(k-1)l/n}^{a+kl/n} f^X(x)\ dx \\
&\leq \sum_{k=1}^{n} \int_{a+(k-1)l/n}^{a+kl/n} x f^X(x)\ dx \\
&= \int_a^b x f^X(x)\ dx
\end{aligned}$$

sowie per analoger Rechnung

$$\begin{aligned}
E\overline{X}_n &= \sum_{k=1}^{n} \Big(a + \frac{kl}{n}\Big) P\Big(\overline{X}_n = a + \frac{kl}{n}\Big) \\
&= \sum_{k=1}^{n} \Big(a + \frac{kl}{n}\Big) P\Big(a + \frac{(k-1)l}{n} < X \leq a + \frac{kl}{n}\Big) \\
&= \sum_{k=1}^{n} \Big(a + \frac{kl}{n}\Big) \int_{a+(k-1)l/n}^{a+kl/n} f^X(x)\ dx \\
&\geq \sum_{k=1}^{n} \int_{a+(k-1)l/n}^{a+kl/n} x f^X(x)\ dx \\
&= \int_a^b x f^X(x)\ dx,
\end{aligned}$$

was insgesamt

$$E\underline{X}_n \;\leq\; \int_a^b x f^X(x)\ dx \;\leq\; E\overline{X}_n \tag{15.3}$$

für alle $n \geq 1$ zeigt. Wegen der Monotonie des Erwartungswertes impliziert $\underline{X}_n \leq X \leq \overline{X}_n$ außerdem

$$E\underline{X}_n \;\leq\; EX \;\leq\; E\overline{X}_n, \tag{15.4}$$

und da

$$E\overline{X}_n - E\underline{X}_n \;=\; \frac{l}{n} \sum_{k=1}^{n} P\Big(a + \frac{(k-1)l}{n} < X \leq a + \frac{kl}{n}\Big) \;=\; \frac{l}{n}$$

für $n \to \infty$ gegen 0 strebt, erhalten wir schließlich per Kombination von (15.3) und (15.4)

$$EX \;=\; \int_a^b x f^X(x)\ dx.$$

Die Verallgemeinerung dieser Formel auf allgemeine regulär stetige Zufallsgrößen bildet den Inhalt der nun folgenden Definition (vgl. 6.2):

15.4. Definition. Sei X eine regulär stetige Zufallsgröße mit der Dichte f^X, für die

$$\int_{-\infty}^{\infty} |x| f^X(x)\, dx \ < \ \infty \tag{15.5}$$

gilt. Dann ist auch

$$EX \ \stackrel{\text{def}}{=} \ \int_{-\infty}^{\infty} x f^X(x)\, dx \tag{15.6}$$

endlich und heißt *Erwartungswert* oder auch *Mittelwert von X* (unter P).

Mittels eines ähnlichen Approximationsarguments durch diskrete Zufallsgrößen läßt sich auch die folgende Definition der Varianz einer regulär stetigen Zufallsgröße rechtfertigen.

15.5. Definition. Sei X eine regulär stetige Zufallsgröße mit der Dichte f^X, für die

$$\int_{-\infty}^{\infty} x^2 f^X(x)\, dx \ < \ \infty \tag{15.7}$$

gilt. Dann ist auch

$$\mathrm{Var}X \ \stackrel{\text{def}}{=} \ \int_{-\infty}^{\infty} (x - EX)^2 f^X(x)\, dx \tag{15.8}$$

endlich und heißt *Varianz von X* (unter P).

Wir möchten an dieser nochmals Stelle betonen, daß sich die hier lediglich durch ein Approximationsargument gerechtfertigten Definitionen von EX und $\mathrm{Var}X$ für eine regulär stetige Zufallsgröße X im Rahmen der Maß- und Integrationstheorie selbstverständlich in derselben Form, aber auf natürlichere Weise ergeben. Alle interessierten Leser seien auf den anschließenden Kurs "Wahrscheinlichkeitstheorie" verwiesen.

15.6. Beispiel. Sei X eine $N(\mu, \sigma^2)$-verteilte und somit regulär stetige Zufallsgröße, wobei $\mu \in \mathbb{R}$ und $\sigma^2 \in (0, \infty)$. Ihre Dichte f^X ist gemäß (13.10) definiert. Der Leser beweise als Übung, daß μ den Erwartungswert und σ^2 die Varianz von X angibt, d.h.

$$EX \ = \ \int_{-\infty}^{\infty} x f^X(x)\, dx \ = \ \mu$$

und

$$\mathrm{Var} X \;=\; \int_{-\infty}^{\infty} (x-\mu)^2 f^X(x)\, dx \;=\; \sigma^2.$$

Zum Ende dieses Kapitels notieren wir noch, wiederum ohne Beweis, die folgende Erweiterung des Transformationssatzes. Eine Funktion heiße stückweise stetig, wenn sie außerhalb einer abzählbaren Teilmenge ihres Definitionsbereichs ohne Häufungspunkte stetig ist.

15.7. Transformationssatz (Verallgemeinerung). *Sei X eine regulär stetige Zufallsgröße mit Dichte f^X und $G : \mathbb{R} \to \mathbb{R}$ eine stückweise stetige Funktion derart, daß*

$$\int_{-\infty}^{\infty} |G(x)| f^X(x)\, dx \;<\; \infty. \tag{15.9}$$

Dann ist $G(X)$ eine integrierbare Zufallsgröße, und es gilt

$$EG(X) \;=\; \int_{-\infty}^{\infty} G(x) f^X(x)\, dx. \tag{15.10}$$

Literatur

[1] ALSMEYER, G. *Wahrscheinlichkeitstheorie (4. Auflage).* Skripten zur Mathematischen Statistik, Nr. 30, Universität Münster (2005).

[2] BERTRAND, J.L.F. *Calcul des probabilités.* Paris (1888/89), S. 4f (aus Kapitel I 'Bestimmung der Chancen'), in: *"Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933":* Einf. U.Texte/Hrsg. Ivo Schneider. Darmstadt: Wiss. Buchges. (1988).

[3] PFANZAGL, J. *Elementare Wahrscheinlichkeitsrechnung.* De Gruyter, Berlin (1988).