

Ausarbeitung der Vorlesung

Mathematik 3 für Maschinenbauer

Studiengang Bachelor Maschinenbau



Bergische Universität Wuppertal

Quellen

F. Beichelt: Stochastik für Ingenieure, Teubner-Verlag Stuttgart, 1995

J. Schwarze: Grundlagen der Statistik, NWB-Studienbücher, Verlag Neue Wirtschaftsbriefe, Herne/Berlin

K. Burg, H. Haf, F. Wille: Höhere Mathematik für Ingenieure, Band II, Teubner-Verlag

K. Burg, H. Haf, F. Wille: Höhere Mathematik für Ingenieure, Band IV, Teubner-Verlag

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Eigenwerttheorie für Matrizen | 5 |
| 1.1 | Determinanten in höherer Dimension | 5 |
| 1.1.1 | Eigenwerte von Matrizen | 8 |
| 1.1.2 | Anwendung: Gekoppelte Systeme linearer Differenzialgleichungen | 14 |
| 1.1.3 | Symmetrische und hermitesche Matrizen | 19 |
| 2 | Raumkurven | 23 |
| 2.1 | Allgemeine Kurventheorie | 23 |
| 2.1.1 | Die Weglänge | 23 |
| 2.1.2 | Ebene und räumliche Kurven | 27 |
| 2.2 | Wegintegrale | 37 |
| 2.2.1 | Allgemeine Definition | 37 |
| 2.2.2 | Stammfunktionen | 38 |
| 3 | Grundlagen der Statistik | 45 |
| 3.1 | Elementare Stochastik | 45 |
| 3.1.1 | Einführung: Zufallsexperimente und Wahrscheinlichkeiten | 45 |
| 3.1.2 | Diskrete Ergebnismengen | 48 |
| 3.1.3 | Einschub: Aus der Kombinatorik | 57 |
| 3.1.4 | Bedingte Wahrscheinlichkeiten | 59 |
| 3.1.5 | Zufallsvariablen und Verteilungsfunktionen | 63 |
| 3.1.6 | Erwartungswerte und Varianz | 67 |
| 3.1.7 | Kontinuierliche Zufallsvariable | 73 |
| 3.1.8 | Erwartungswerte u. Varianz | 81 |
| 3.1.9 | Zusammengesetzte Zufallsvariablen | 86 |
| 3.1.10 | Gesetze der großen Zahlen | 98 |
| 3.1.11 | Der zentrale Grenzwertsatz | 100 |
| 3.1.12 | Konfidenzschranken | 102 |

Kapitel 1

Eigenwerttheorie für Matrizen

1.1 Determinanten in höherer Dimension

Wir beginnen damit, die Determinante einer Matrix auch für quadratische Matrizen mit mehr als 3 Zeilen einzuführen.

Definition. Sei \mathcal{A} eine $n \times n$ -Matrix (reell oder komplex). Dann definieren wir

für $n = 4$:

$$\det \mathcal{A} = a_{11} \det \mathcal{A}_1 - a_{12} \det \mathcal{A}_2 + a_{13} \det \mathcal{A}_3 - a_{14} \det \mathcal{A}_4,$$

für $n = 5$:

$$\det \mathcal{A} = a_{11} \det \mathcal{A}_1 - a_{12} \det \mathcal{A}_2 + a_{13} \det \mathcal{A}_3 - a_{14} \det \mathcal{A}_4 + a_{15} \det \mathcal{A}_5,$$

wobei jedesmal die Matrix \mathcal{A}_i aus \mathcal{A} durch Weglassen der 1. Zeile und i .ten Spalte entsteht. Dabei ist a_{km} das Element von \mathcal{A} das in der k .ten Zeile und m .ten Spalte steht.

Allgemein: Haben wir schon die Determinante einer $(n - 1) \times (n - 1)$ -Matrix definiert, so setzen wir für eine $n \times n$ -Matrix fest:

$$\det \mathcal{A} := \sum_{i=1}^n (-1)^{i+1} \det \mathcal{A}_i,$$

wobei wieder die Matrix \mathcal{A}_i aus \mathcal{A} durch Weglassen der 1. Zeile und i .ten Spalte entsteht.

Hier sind

Beispiele. a) Sei $\mathcal{A} = \begin{pmatrix} 3 & 2 & -1 & -4 \\ -2 & 3 & 2 & -5 \\ 2 & 1 & 1 & 0 \\ 9 & 3 & 7 & 6 \end{pmatrix}$. Dann errechnen wir für ihre Determinante

$$\begin{aligned} \det \mathcal{A} &= 3 \underbrace{\begin{pmatrix} 3 & 2 & -5 \\ 1 & 1 & 0 \\ 3 & 7 & 6 \end{pmatrix}}_{=-14} - 2 \underbrace{\begin{pmatrix} -2 & 2 & -5 \\ 2 & 1 & 0 \\ 9 & 7 & 6 \end{pmatrix}}_{=-61} - \underbrace{\begin{pmatrix} -2 & 3 & -5 \\ 2 & 1 & 0 \\ 9 & 3 & 6 \end{pmatrix}}_{=-33} + 4 \underbrace{\begin{pmatrix} -2 & 3 & 2 \\ 2 & 1 & 1 \\ 9 & 3 & 7 \end{pmatrix}}_{=-29} \\ &= -42 + 122 + 33 - 116 \\ &= -3 \end{aligned}$$

b) Sei $\mathcal{A} = \begin{pmatrix} 8 & 2 & -4 & 7 & 9 \\ -12 & 2 & 7 & -4 & 3 \\ 2 & 2 & -1 & 0 & 8 \\ 6 & 2 & 11 & 0 & 5 \\ 6 & 2 & 1 & -1 & -2 \end{pmatrix}$.

Dann wird

$$\begin{aligned} \det \mathcal{A} &= 8 \begin{pmatrix} 2 & 7 & -4 & 3 \\ 2 & -1 & 0 & 8 \\ 2 & 11 & 0 & 5 \\ 2 & 1 & -1 & -2 \end{pmatrix} - 2 \begin{pmatrix} -12 & 7 & -4 & 3 \\ 2 & -1 & 0 & 8 \\ 6 & 11 & 0 & 5 \\ 6 & 1 & -1 & -2 \end{pmatrix} \\ &\quad - 4 \begin{pmatrix} -12 & 2 & -4 & 3 \\ 2 & 2 & 0 & 8 \\ 6 & 2 & 0 & 5 \\ 6 & 2 & -1 & -2 \end{pmatrix} - 7 \begin{pmatrix} -12 & 2 & 7 & 3 \\ 2 & 2 & -1 & 8 \\ 6 & 2 & 11 & 5 \\ 6 & 2 & 1 & -2 \end{pmatrix} \\ &\quad + 9 \begin{pmatrix} -12 & 2 & 7 & -4 \\ 2 & 2 & -1 & 0 \\ 6 & 2 & 11 & 0 \\ 6 & 2 & 1 & -1 \end{pmatrix} \\ &= 8 \cdot 840 - 2 \cdot 3770 - 4 \cdot (-100) - 7 \cdot 4040 + 9 \cdot 720 \\ &= -22220 \end{aligned}$$

Man kann den folgenden Satz zeigen

1.1.1 Satz. Die Determinante hat die folgenden Eigenschaften:

- Für die Einheitsmatrix \mathcal{E}_n gilt: $\det(\mathcal{E}_n) = 1$
- Stimmen 2 Spalten oder 2 Zeilen von \mathcal{A} überein, ist $\det(\mathcal{A}) = 0$.
- Sei $\lambda \neq 0$. Wenn dann die Matrix \mathcal{B} aus \mathcal{A} durch Multiplikation einer Spalte (oder Zeile) mit λ entsteht, so ist $\det \mathcal{B} = \lambda \det \mathcal{A}$.

d) Entsteht \mathcal{B} aus \mathcal{A} durch Vertauschen zweier Spalten (oder Zeilen), so ist $\det \mathcal{B} = -\det \mathcal{A}$.

e) Entsteht \mathcal{B} aus \mathcal{A} dadurch, dass man in \mathcal{A} zu einer Spalte ein Vielfaches einer anderen Spalte addiert, so ist $\det \mathcal{B} = \det \mathcal{A}$.

e') Entsteht \mathcal{B} aus \mathcal{A} dadurch, dass man in \mathcal{A} zu einer Zeile ein Vielfaches einer anderen Zeile addiert, so ist $\det \mathcal{B} = \det \mathcal{A}$.

f) Hat \mathcal{A} obere Dreiecksgestalt (Zeilenstufenform), so ist $\det \mathcal{A}$ das Produkt der Diagonalelemente von \mathcal{A} .

Damit erleichtert sich die Berechnung von Determinanten beträchtlich:

Beispiel. Sei wieder $\mathcal{A} = \begin{pmatrix} 3 & 2 & -1 & -4 \\ -2 & 3 & 2 & -5 \\ 2 & 1 & 1 & 0 \\ 9 & 3 & 7 & 6 \end{pmatrix}$. Dann nehmen wir Zeilenumformungen vor:

$$\begin{aligned} \det \mathcal{A} &= \det \begin{pmatrix} 3 & 2 & -1 & -4 \\ -2 & 3 & 2 & -5 \\ 0 & 4 & 3 & -5 \\ 9 & 3 & 7 & 6 \end{pmatrix} && (3.\text{Zeile} + 2.\text{Zeile}) \\ &= \det \begin{pmatrix} 3 & 2 & -1 & -4 \\ -2 & 3 & 2 & -5 \\ 0 & 4 & 3 & -5 \\ 0 & -3 & 10 & 18 \end{pmatrix} && (4.\text{Zeile} + 3 \times 2.\text{Zeile}) \\ &= \det \begin{pmatrix} 3 & 2 & -1 & -4 \\ 0 & \frac{13}{3} & \frac{4}{3} & -\frac{23}{3} \\ 0 & 4 & 3 & -5 \\ 0 & -3 & 10 & 18 \end{pmatrix} && (2.\text{Zeile} + \frac{2}{3} \times 1.\text{Zeile}) \\ &= 3 \cdot \frac{1}{3} \det \begin{pmatrix} 13 & 4 & -23 \\ 4 & 3 & -5 \\ -3 & 10 & 18 \end{pmatrix} = -3 \end{aligned}$$

Der Vollständigkeit halber geben wir noch eine Formel zur Determinantenberechnung an. Sie verallgemeinert unsere Definition der Determinante:

1.1.2 Satz (Entwicklungssatz). Ist $\mathcal{A} = (a_{jk})_{j,k=1}^n$ eine Matrix, und steht $\mathcal{A}(j, k)$ für diejenige Matrix, die aus \mathcal{A} durch Weglassen der j -ten Zeile und k -ten Spalte entsteht, so gilt

$$\det \mathcal{A} = \sum_{k=1}^n (-1)^{j+k} a_{jk} \det \mathcal{A}(j, k)$$

für jedes $j \in \{1, 2, \dots, n\}$ (Entwicklung nach der j -ten Zeile)
und

$$\det \mathcal{A} = \sum_{j=1}^n (-1)^{j+k} a_{jk} \det \mathcal{A}(j, k)$$

für jedes $j \in \{1, 2, \dots, n\}$ (Entwicklung nach der k .ten Spalte).

Analog zu früher gilt

1.1.3 Satz. Genau dann ist die Matrix \mathcal{A} invertierbar, wenn $\det \mathcal{A} \neq 0$ ist.

Weiter haben wir folgende wichtige Regel:

1.1.4 Satz (Multiplikationssatz). Sind \mathcal{A} und \mathcal{B} quadratische n -reihige Matrizen so ist

$$\det(\mathcal{A}\mathcal{B}) = \det \mathcal{A} \det \mathcal{B}$$

Vertauschen von Zeilen mit Spalten ändert an der Determinante nichts.

1.1.5 Satz. Für eine Matrix \mathcal{A} bezeichnen wir mit \mathcal{A}^t diejenige Matrix, die aus \mathcal{A} durch Vertauschen von Zeilen und Spalten entsteht. (Transponierte von \mathcal{A}). Dann ist $\det(\mathcal{A}^t) = \det \mathcal{A}$.

1.1.1 Eigenwerte von Matrizen

Beispiel. Gegeben sei ein gekoppeltes System linearer Differenzialgleichungen, also von der Form

$$\vec{u}' = \mathcal{A} \cdot \vec{u} + \vec{B},$$

mit einem Vektor \vec{B} , dessen Komponenten stetige Funktionen sein dürfen. Die Matrix $\mathcal{A} = (a_{jk})_{j,k=1}^n$ habe konstante Koeffizienten.

Eine Idee zur Lösung dieses DGL-systems für den Fall $\vec{B} = 0$ lautet so:

Man suche einen Vektor \vec{v} , so dass mit einer Zahl λ die Gleichung $\mathcal{A}\vec{v} = \lambda\vec{v}$ besteht.

Dann wird nämlich für die (vektorwertige) Funktion $u(t) = e^{\lambda t} \vec{v}$ gelten:

$$\vec{u}'(t) = \lambda e^{\lambda t} \vec{v} = e^{\lambda t} \mathcal{A} \vec{v} = \mathcal{A} \vec{u}$$

also wird \vec{u} eine Lösung.

Eine DGL-system wie das obige tritt bei gekoppelten Schwingungen oder beim Doppelpendel auf.

Definition. Ist \mathcal{A} eine quadratische n -reihige Matrix, so heißt eine Zahl $\lambda \in \mathbb{C}$ ein *Eigenwert* von \mathcal{A} , wenn ein Vektor $\vec{v} \neq \vec{0}$ mit $\mathcal{A}\vec{v} = \lambda\vec{v}$ existiert. Ein derartiger Vektor wird als *Eigenvektor*

zum Eigenwert λ bezeichnet. Die Menge aller Eigenvektoren zu \mathcal{A} zum Eigenwert λ zusammen mit dem Nullvektor bildet einen Unterraum des \mathbb{C}^n , den man als den *Eigenraum* von \mathcal{A} zum λ bezeichnet. Wir bezeichnen diesen Unterraum mit $E(\mathcal{A}, \lambda)$.

Das Polynom $\chi_{\mathcal{A}}(x) = \det(x\mathcal{E}_n - \mathcal{A})$ wird das *charakteristische Polynom* von \mathcal{A} genannt.

Wir gehen nun der Frage nach, wie man Eigenwerte und -vektoren berechnen kann. Wenn dann die im obigen DGL-System die Inhomogenität, also der Vektor \vec{B} nicht Null ist, lässt sich das DGL-System lösen, wenn man etwa eine Basis von Eigenvektoren für \mathcal{A} finden kann. Wann das möglich ist, wollen wir ebenfalls untersuchen.

Zur Berechnung der Eigenwerte dient uns

1.1.1.1 Hilfssatz. *Ist \mathcal{A} eine $n \times n$ -Matrix, so ist eine Zahl $\lambda \in \mathbb{C}$ genau dann ein Eigenwert von \mathcal{A} , wenn $\chi_{\mathcal{A}}(\lambda) = 0$ gilt.*

Beweis.

Denn folgende Aussagen sind äquivalent:

- i) $\lambda \in \mathbb{C}$ ist ein Eigenwert von \mathcal{A}
- ii) Das Gleichungssystem $\mathcal{A}\vec{x} = \lambda\vec{x}$ hat eine von Null verschiedene Lösung
- iii) Die Matrix $\lambda\mathcal{E}_n - \mathcal{A}$ ist nicht invertierbar
- iv) $\chi_{\mathcal{A}}(\lambda) = 0$.

□

Beispiel. Was sind die Eigenwerte von $\mathcal{A} := \begin{pmatrix} -78 & -4 & 22 \\ 36 & -32 & -55 \\ -72 & -2 & 77 \end{pmatrix}$, wenn \mathbb{R} der zugrunde liegende Zahlenbereich ist? Wir errechnen

$$\begin{aligned} \chi_{\mathcal{A}}(t) &= t^3 + 33t^2 - 4356t - 143748 \\ &= t^3 + 33t^2 - 4 \cdot 33^2t - 4 \cdot 33^3 \\ &= 33^3 \left(\left(\frac{t}{33}\right)^3 + \left(\frac{t}{33}\right)^2 - 4 \cdot \left(\frac{t}{33}\right) - 4 \right) \end{aligned}$$

Somit ist $\frac{t}{33}$ Nullstelle des Polynoms $g(s) = s^3 + s^2 - 4s - 4$. Da $g(-1) = g(2) = 0$, folgt mit dem Horner-Schema $g(s) = (s+1)(s-2)(s+2)$. Als Eigenwerte von \mathcal{A} finden wir also $-33, \pm 66$.

Die Eigenräume dazu sind die Lösungsräume zu den linearen Gleichungssystemen

$$(\mathcal{A} + 33\mathcal{E}_3)\vec{x} = 0, \quad (\mathcal{A} \pm 66\mathcal{E}_3)\vec{x} = 0$$

Nun ist $\mathcal{A} + 33\mathcal{E}_3 = \begin{pmatrix} -45 & -4 & 22 \\ 36 & 1 & -55 \\ -72 & -2 & 110 \end{pmatrix}$ und damit muss das äquivalente Gleichungssystem

mit Matrix $\begin{pmatrix} -45 & -4 & 22 \\ 0 & -11 & -187 \\ 0 & 0 & 0 \end{pmatrix}$ gelöst werden. So finden wir $\mathcal{A} \begin{pmatrix} 2 \\ -17 \\ 1 \end{pmatrix} = -33 \begin{pmatrix} 2 \\ -17 \\ 1 \end{pmatrix}$

und $E(\mathcal{A}, -33) = \mathbb{R} \begin{pmatrix} 2 \\ -17 \\ 1 \end{pmatrix}$.

In entsprechender Weise verfahren wir mit den Eigenräumen zu ± 66 . Wir finden

$$E(\mathcal{A}, 66) = \mathbb{R} \begin{pmatrix} 1 \\ -3 \\ 6 \end{pmatrix}, \quad E(\mathcal{A}, -66) = \mathbb{R} \begin{pmatrix} 4 \\ -1 \\ 2 \end{pmatrix}$$

Das nächste Beispiel lehrt, dass eine reelle Matrix nicht immer reelle Eigenwerte haben muss:

Beispiel. Sei $\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}$. Dann ist $\chi_{\mathcal{A}}(x) = x^4 + 1$. Die Eigenwerte zu \mathcal{A} sind $\lambda_1 = -\frac{1+j}{\sqrt{2}}, \lambda_2 = -\lambda_1, \lambda_3 = \bar{\lambda}_1$ und $\lambda_4 = \bar{\lambda}_2$, und

$$E(\mathcal{A}, \lambda_1) = \mathbb{C} \begin{pmatrix} -\lambda_1 \\ -j \\ j\lambda_1 \\ 1 \end{pmatrix}, \quad E(\mathcal{A}, \lambda_2) = \mathbb{C} \begin{pmatrix} \lambda_1 \\ -j \\ -j\lambda_1 \\ 1 \end{pmatrix}$$

$$E(\mathcal{A}, \lambda_3) = \mathbb{C} \begin{pmatrix} -\bar{\lambda}_1 \\ j \\ j\bar{\lambda}_1 \\ 1 \end{pmatrix}, \quad E(\mathcal{A}, \lambda_4) = \mathbb{C} \begin{pmatrix} \bar{\lambda}_1 \\ j \\ j\bar{\lambda}_1 \\ 1 \end{pmatrix}$$

Eigenwertberechnung bei 3×3 -Matrizen

Die Berechnung von Eigenwerten bei 3×3 -Matrizen läuft auf das Lösen einer algebraischen Gleichung 3. Grades hinaus. Ein Analog zur "p,q-Formel" ist in dieser Situation möglich. Sei also

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

ein Polynom 3. Grades. Wir suchen die Nullstellen von f (in \mathbb{C}).

1. *Schritt:* Normierung: Wir klammern a_3 aus und suchen die Nullstellen von

$$g(x) = x^3 + \frac{a_2}{a_3}x^2 + \frac{a_1}{a_3}x + \frac{a_0}{a_3} =: x^3 + b_2x^2 + b_1x + b_0$$

2. *Schritt:* Wir bringen den Term mit x^2 weg. Dazu gehen wir von g zum Polynom $h(x) = g(x - \frac{b_2}{3})$ über. Wir rechnen aus:

$$h(x) = x^3 + px + q, \quad \text{mit } p = b_1 - \frac{1}{3}b_2^2, \quad q = b_0 - \frac{1}{3}b_1b_2 + \frac{2}{27}b_2^3$$

3. Schritt. Wir suchen die Nullstelle von h . Wenn $q = 0$, so ist alles leicht, da dann $h(x) = x(x^2 + p)$.

Sei also $q \neq 0$. Wir unterscheiden 2 Fälle:

1. Fall: Es ist $\Delta := \left(\frac{q}{2}\right)^2 + \frac{p}{3} \geq 0$.

Dann schreiben wir $x := t - \frac{p}{3t}$ und setzen dies in $h(x) = 0$ ein. Wir erhalten eine neue Gleichung, und zwar

$$t^3 + q = \frac{p^3}{27t^3}$$

Damit löst t^3 die quadratische Gleichung

$$T^2 + qT = \frac{p^3}{27}$$

Diese hat (in \mathbb{C}) die Lösungen

$$T_{1,2} = -\frac{q}{2} \pm \sqrt{\Delta}.$$

Wir arbeiten mit $T_1 = -\frac{q}{2} + \sqrt{\Delta}$ weiter.

Dann wählen wir eine 3. Wurzel t_0 aus T_1 und bilden

$$z_1 = t_0 - \frac{p}{3t_0}$$

Das liefert uns eine Nullstelle für h .

Weitere Nullstellen finden wir so: Sei $\omega = \frac{-1+j\sqrt{3}}{2}$. Dann sind auch ωt_0 und $\bar{\omega} t_0$ wieder 3. Wurzeln aus T_1 und daher auch

$$z_2 = \omega t_0 - \frac{p}{3\omega t_0} = \omega t_0 - \frac{p}{3t_0} \bar{\omega}$$

und

$$z_3 = \bar{\omega} t_0 - \frac{p}{3\bar{\omega} t_0} = \bar{\omega} t_0 - \frac{p}{3t_0} \omega$$

Lösungen zur Gleichung $h(x) = 0$.

4. Schritt. Wir rechnen wieder "zurück", indem wir

$$x_j := z_j + \frac{1}{3} b_2, \quad j = 1, 2, 3$$

bilden. Das sind dann die gesuchten Nullstellen von g und damit von f .

Nun lässt sich die eine Nullstelle in einfacherer Form darstellen:

Es gilt nämlich

$$\begin{aligned}
z_1 &= t_0 - \frac{p}{3t_0} = t_0 - \frac{pt_0^2}{3T_1} \\
&= t_0 - \frac{p(\sqrt{\Delta} - \frac{q}{2})^{2/3}}{3(\sqrt{\Delta} - \frac{q}{2})} \\
&= t_0 - \frac{p(\sqrt{\Delta} - \frac{q}{2})^{2/3}(\sqrt{\Delta} + \frac{q}{2})}{3(\sqrt{\Delta}^2 - (\frac{q}{2})^2)} \\
&= t_0 - \frac{p(\sqrt{\Delta}^2 - (\frac{q}{2})^2)^{2/3}(\sqrt{\Delta} + \frac{q}{2})^{1/3}}{3(\sqrt{\Delta}^2 - (\frac{q}{2})^2)} \\
&= t_0 - \frac{p(\sqrt{\Delta} + \frac{q}{2})^{1/3}}{3(\sqrt{\Delta}^2 - (\frac{q}{2})^2)^{1/3}} \\
&= \sqrt[3]{\sqrt{\Delta} - \frac{q}{2}} - \sqrt[3]{\sqrt{\Delta} + \frac{q}{2}} \\
&= \sqrt[3]{\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} - \frac{q}{2}} - \sqrt[3]{\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} + \frac{q}{2}}
\end{aligned}$$

2. Fall: Es ist $\frac{q^2}{4} + \frac{p^3}{27} < 0$. Nun wird aber $p < 0$ und $\frac{|q|}{2} < (\frac{-p}{3})^{3/2}$. Dann wählen wir ein α mit

$$\cos(3\alpha) = -\frac{q}{2(\frac{-p}{3})^{3/2}}$$

Dann ist

$$x_1 = 2\sqrt{\frac{-p}{3}} \cos \alpha$$

Lösung zur kubischen Gleichung $h(x) = 0$. Denn es ist

$$\begin{aligned}
x_1^3 + px_1 + q &= (2\sqrt{\frac{-p}{3}})^3 \cos^3 \alpha + 2p\sqrt{\frac{-p}{3}} \cos \alpha + q \\
&= 2(\sqrt{\frac{-p}{3}})^3 (4 \cos^3 \alpha - 3 \cos \alpha - \cos(3\alpha)) \\
&= 0,
\end{aligned}$$

da

$$\begin{aligned}
\cos(3\alpha) &= \cos(2\alpha) \cos \alpha - \sin(2\alpha) \sin \alpha \\
&= (\cos(2\alpha) - 2 \sin^2 \alpha) \cos \alpha \\
&= (2 \cos^2 \alpha - 1 - 2 + 2 \cos^2 \alpha) \\
&= 4 \cos^3 \alpha - 3 \cos \alpha
\end{aligned}$$

Ein *Beispiel* dazu: Sei $\mathcal{A} = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 1 & 0 \\ -1 & 0 & 4 \end{pmatrix}$. Was sind die Eigenwerte von \mathcal{A} ?

Das charakteristische Polynom von \mathcal{A} ist

$$f(x) = x^3 - 8x^2 + 14x + 5$$

Das zugehörige Polynom g ist nun gleich f . Das Polynom h ist

$$h(x) = f\left(x + \frac{8}{3}\right) = x^3 - \frac{22}{3}x + \frac{119}{27}$$

Es wird also $p = -\frac{22}{3}$ und $q = \frac{119}{27}$. Nun ist $\Delta = -\frac{39}{4} < 0$ und

$$-\frac{q}{2(-p/3)^{3/2}} = -\frac{119}{44\sqrt{22}} = -0.576611 = \cos 2.18537$$

Wir berechnen nunmehr $x_1 = 2\sqrt{-p/3} \cos(2.18537/3)$ und finden $x_1 = 2.3333$ als Nullstelle von h . Somit finden wir $f(5) = 0$, und $f(x) = (x-5)(x^2 - 3x - 1)$. Die anderen Nullstellen von f sind nun $\frac{1}{2}(3 \pm \sqrt{13})$.

Definition. Wir nennen eine Matrix \mathcal{A} *diagonalisierbar* oder auch *diagonalähnlich*, wenn eine Basis $\vec{b}_1, \dots, \vec{b}_n \in \mathbb{C}^n$ existiert, die nur aus Eigenvektoren von \mathcal{A} besteht. Das heißt also:

Es muss gelten:

i) Jeder Vektor $\vec{v} \in \mathbb{C}^n$ ist Linearkombination von $\vec{b}_1, \dots, \vec{b}_n$ also $\vec{v} = \sum_{j=1}^n \alpha_j \vec{b}_j$ mit geeigneten Koeffizienten $\alpha_1, \dots, \alpha_n$,

ii) Mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von \mathcal{A} gilt $\mathcal{A}\vec{b}_j = \lambda_j \vec{b}_j$ für $j = 1, \dots, n$. (Dabei müssen die Eigenwerte nicht paarweise verschieden sein!)

Aus (i) folgt, dass die Matrix $\mathcal{S} = (\vec{b}_1, \dots, \vec{b}_n)$ invertierbar ist. Es gilt weiter

$$\mathcal{A}\mathcal{S} = (\lambda_1 \vec{b}_1, \dots, \lambda_n \vec{b}_n)$$

und damit

$$\mathcal{S}^{-1}\mathcal{A}\mathcal{S} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Das erklärt den Ausdruck "diagonalähnlich".

Wann ist nun eine Matrix diagonalisierbar? Sicher ist das nicht immer der Fall: Dazu betrachten wir die Matrix $\mathcal{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dann ist 0 einziger Eigenwert von \mathcal{A} . Wäre nun \mathcal{A} diagonalisierbar, so gäbe es eine invertierbare Matrix \mathcal{S} mit $\mathcal{S}^{-1}\mathcal{A}\mathcal{S} = 0$, was nicht geht, denn $\mathcal{A} \neq 0$.

Dagegen gilt aber:

1.1.1.2 Satz. Gilt für die Matrix \mathcal{A}

$$\chi_{\mathcal{A}}(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$, so ist \mathcal{A} diagonalisierbar.

Beweis. (Skizze). Wir wählen Eigenvektoren zu diesen Eigenwerten. Man kann zeigen, dass sie linear unabhängig sind. Daraus folgt alles. □

Folgender allgemeinere Satz gilt:

1.1.1.3 Satz. Angenommen, die Matrix \mathcal{A} habe das charakteristische Polynom

$$\chi_{\mathcal{A}}(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_r)^{d_r}$$

mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$ und positiven ganzen Zahlen d_1, \dots, d_r . Hat dann für jedes j der Eigenraum $E(\mathcal{A}, \lambda_j)$ die Dimension d_j , so ist \mathcal{A} diagonalisierbar. Es gilt auch der umgekehrte Schluss: Soll \mathcal{A} diagonalisierbar sein, so muss für jedes j der Eigenraum $E(\mathcal{A}, \lambda_j)$ die Dimension d_j haben.

1.1.2 Anwendung: Gekoppelte Systeme linearer Differenzialgleichungen

Beispiel. Gegeben sei ein lineares Differenzialgleichungssystem, also

$$\vec{u}' = \mathcal{A} \cdot \vec{u} + \vec{B},$$

mit einem Vektor \vec{B} , dessen Komponenten stetige Funktionen auf $[0, T)$ sein dürfen ($T > 0$ beliebig). Die Matrix $\mathcal{A} = (a_{jk})_{j,k=1}^n$ habe konstante Koeffizienten.

Wir nehmen an, dass \mathcal{A} diagonalisierbar sei. Wir finden dann eine invertierbare Matrix \mathcal{S} , so dass (mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von \mathcal{A}) gilt

$$\mathcal{S}^{-1} \mathcal{A} \mathcal{S} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Wir lösen zuerst das *homogene DGL-System*

$$\vec{v}' = \mathcal{A} \vec{v}$$

Dazu bilden wir die vektorwertigen Funktionen

$$\vec{u}_j(t) = e^{\lambda_j t} \mathcal{S} \vec{e}_j$$

mit $j = 1, \dots, n$. Sie lösen das DGL-System. Dann bilden wir die *Wronskimatrix*

$$\mathcal{W}(t) = (\vec{u}_1(t), \dots, \vec{u}_n(t))$$

Es gilt $\mathcal{W}'(t) = \mathcal{A}\mathcal{W}(t)$. Das DGL-System $\vec{u}' = \mathcal{A}\vec{u} + \vec{B}$ wird also durch

$$\vec{u}(t) = \mathcal{W}(t) \left(\mathcal{W}(0)^{-1} \vec{u}(0) + \int_0^t \mathcal{W}^{-1}(s) \vec{B}(s) ds \right)$$

gelöst, denn es gilt

$$\begin{aligned} \vec{u}'(t) &= \mathcal{W}'(t) \left(\mathcal{W}(0)^{-1} \vec{u}(0) + \int_0^t \mathcal{W}^{-1}(s) \vec{B}(s) ds \right) \\ &\quad + \mathcal{W}(t) \mathcal{W}^{-1}(t) \vec{B}(t) \\ &= \mathcal{A}\mathcal{W}(t) \left(\mathcal{W}(0)^{-1} \vec{u}(0) + \int_0^t \mathcal{W}^{-1}(s) \vec{B}(s) ds \right) + \vec{B}(t) \\ &= \mathcal{A}\vec{u}(t) + \vec{B} \end{aligned}$$

Gleichzeitig sehen wir, dass man einen Startwert $\vec{u}(0)$ vorgeben kann!

Beispiel: Drehgeschwindigkeit beim Gleichstrommotor. An einem Gleichstrommotor werde zur Zeit $t_0 = 0$ die Spannung U angelegt. Ist J der den Anker durchfließende Strom, M das Drehmoment des Ankers, Θ sein Trägheitsmoment, L die Induktivität der Ankerspule und R der Ohmsche Widerstand, so gilt für den Strom J und die Winkelgeschwindigkeit ω des Ankers das DGL-System

$$\begin{aligned} -\Theta\omega' &= -aJ + M \\ LJ' &= -RJ - a\omega + U \end{aligned}$$

Hierbei ist $a > 0$ eine Konstante.

Dann nähert sich mit wachsender Zeit der die Drehgeschwindigkeit $\omega(t)$ dem Grenzwert

$$\omega_\infty = \frac{U}{a} - \frac{RM}{a^2}.$$

Wir lösen dazu das DGL-System:

$$\frac{d}{dt} \begin{pmatrix} \omega \\ J \end{pmatrix} = \mathcal{A} \begin{pmatrix} \omega \\ J \end{pmatrix} + \begin{pmatrix} -M/\Theta \\ U/L \end{pmatrix}$$

mit $\mathcal{A} = \begin{pmatrix} 0 & a/\Theta \\ -a/L & -R/L \end{pmatrix}$. Wenn nur

$$\Delta = \left(\frac{R}{2L}\right)^2 - \frac{a^2}{L\Theta} > 0$$

ist, sind die Eigenwerte λ_+, λ_- von \mathcal{A} voneinander verschieden:

$$\lambda_{\pm} = -\frac{R}{2L} \pm \sqrt{\Delta}$$

Die Eigenvektoren dazu sind $\begin{pmatrix} a \\ \lambda_+\Theta \end{pmatrix}, \begin{pmatrix} a \\ \lambda_-\Theta \end{pmatrix}$. Mit

$$\mathcal{W}(t) = \begin{pmatrix} ae^{\lambda_+t} & ae^{\lambda_-t} \\ \lambda_+\Theta e^{\lambda_+t} & \lambda_-\Theta e^{\lambda_-t} \end{pmatrix}$$

ist eine Wronskimatrix zur homogenen Dgl gefunden. Als Lösung zur inhomogenen Dgl mit homogenen Startwerten dient dann

$$\begin{aligned} \begin{pmatrix} \omega \\ J \end{pmatrix} &= \mathcal{W}(t) \int_0^t \mathcal{W}(s)^{-1} \begin{pmatrix} -M/\Theta \\ U/L \end{pmatrix} ds \\ &= \frac{1}{2a\Theta\sqrt{\Delta}} \mathcal{W}(t) \begin{pmatrix} \frac{\lambda_-M + \frac{aU}{L}}{\lambda_- + \frac{R}{L}} \left(e^{(\lambda_- + \frac{R}{L})t} - 1 \right) \\ -\frac{\lambda_+M + \frac{aU}{L}}{\lambda_+ + \frac{R}{L}} \left(e^{(\lambda_+ + \frac{R}{L})t} - 1 \right) \end{pmatrix} \end{aligned}$$

Berechnen wir die erste Komponente dazu, so erhalten wir

$$\begin{aligned} \omega(t) &= \frac{1}{2\Theta\sqrt{\Delta}} \left(\frac{\lambda_-M + \frac{aU}{L}}{\lambda_- + \frac{R}{L}} \left(e^{(\lambda_- + \frac{R}{L})t + \lambda_+t} - e^{\lambda_+t} \right) \right. \\ &\quad \left. - \frac{\lambda_+M + \frac{aU}{L}}{\lambda_+ + \frac{R}{L}} \left(e^{(\lambda_+ + \frac{R}{L})t + \lambda_-t} - e^{\lambda_-t} \right) \right) \\ &= \frac{1}{2\Theta\sqrt{\Delta}} \left(\frac{\lambda_-M + \frac{aU}{L}}{\lambda_- + \frac{R}{L}} (1 - e^{\lambda_+t}) - \frac{\lambda_+M + \frac{aU}{L}}{\lambda_+ + \frac{R}{L}} (1 - e^{\lambda_-t}) \right) \end{aligned}$$

Unabhängig davon, ob Δ positiv oder negativ ist, gilt stets $e^{\lambda_{\pm}t} \rightarrow 0$, wenn $t \rightarrow \infty$. Wir

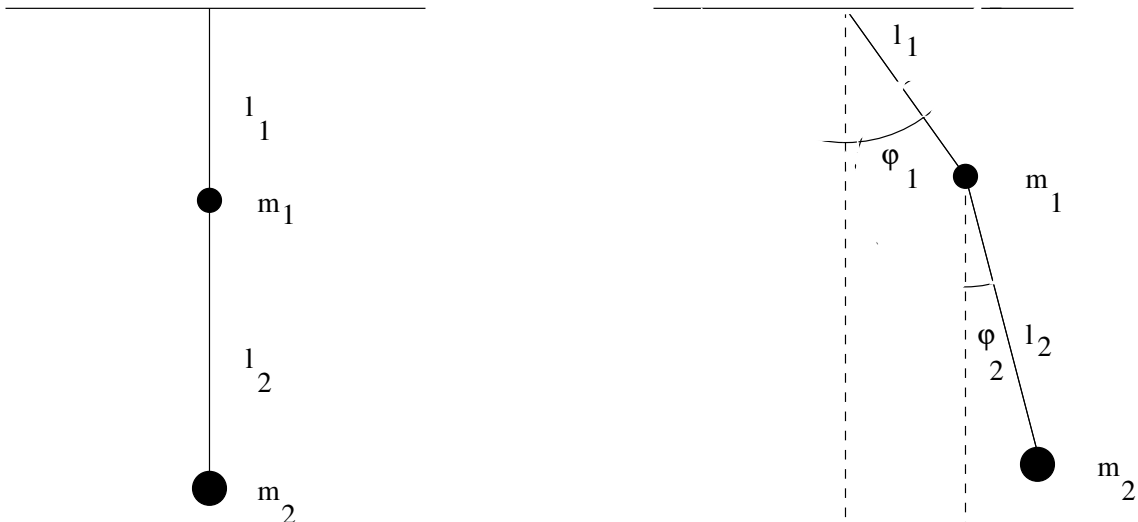
erhalten damit

$$\begin{aligned}
 \omega_\infty := \lim_{t \rightarrow \infty} \omega(t) &= \frac{1}{2\Theta\sqrt{\Delta}} \left(\frac{\lambda_- M + \frac{aU}{L}}{\lambda_- + \frac{R}{L}} - \frac{\lambda_+ M + \frac{aU}{L}}{\lambda_+ + \frac{R}{L}} \right) \\
 &= \frac{1}{2\Theta\sqrt{\Delta}} \left(\frac{(\lambda_- M + \frac{aU}{L})(\lambda_+ + \frac{R}{L}) - (\lambda_+ M + \frac{aU}{L})(\lambda_- + \frac{R}{L})}{(\lambda_+ + \frac{R}{L})(\lambda_- + \frac{R}{L})} \right) \\
 &= \frac{1}{2\Theta\sqrt{\Delta}} \left(\frac{\frac{aU-MR}{L}(\lambda_+ - \lambda_-)}{(\lambda_+ \lambda_- + \frac{R}{L}(\lambda_+ + \lambda_-) + (\frac{R}{L})^2)} \right) \\
 &= \frac{1}{\Theta\sqrt{\Delta}} \left(\frac{\frac{aU-MR}{L}\sqrt{\Delta}}{(a^2/L\Theta)} \right) = \frac{U}{a} - \frac{RM}{a^2}
 \end{aligned}$$

Dabei benutzen wir, dass $\lambda_+ \lambda_- = \det \mathcal{A} = \frac{a^2}{L\Theta}$ und $\lambda_+ + \lambda_- = -R/L$.

Das Doppelpendel

Zwei Teilchen mit den Massen m_1 und m_2 sind an 2 Fäden der Längen ℓ_1 und ℓ_2 aufgehängt, wobei der Faden mit Teilchen 2 am Teilchen 1 befestigt ist. Dann wird die Anordnung in Schwingung versetzt, siehe folgendes Bild:



Wir nehmen an, es sei $m_1 + m_2 = 1$. Die Bewegung des Doppelpendels wird nun durch das DGL-System

$$\mathcal{A}\Phi'' = -\mathcal{C}\Phi$$

gesteuert, sofern die Auslenkungen klein bleiben. Dabei ist

$$\Phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} \ell_1^2 & m_2\ell_1\ell_2 \\ m_2\ell_1\ell_2 & m_2\ell_2^2 \end{pmatrix}$$

und $\mathcal{C} = g \begin{pmatrix} \ell_1 & 0 \\ 0 & m_2\ell_2 \end{pmatrix}$, wobei g die Erdbeschleunigung bedeutet. Die Matrix \mathcal{A} ist invertierbar, und es gilt

$$\mathcal{A}^{-1} = \frac{1}{m_1m_2\ell_1^2\ell_2^2} \begin{pmatrix} m_2\ell_2^2 & -m_2\ell_1\ell_2 \\ -m_2\ell_1\ell_2 & \ell_1^2 \end{pmatrix}$$

Wir errechnen dann

$$\mathcal{B} := -\mathcal{A}^{-1}\mathcal{C} = \frac{g}{m_1} \begin{pmatrix} -\frac{1}{\ell_1} & \frac{m_2}{\ell_1} \\ \frac{1}{\ell_2} & -\frac{1}{\ell_2} \end{pmatrix}$$

Gesucht sind nun die Eigenwerte von \mathcal{B} mit den zugehörigen Eigenvektoren. Wir berechnen die Eigenwerte zur Matrix $\begin{pmatrix} -\frac{1}{\ell_1} & \frac{m_2}{\ell_1} \\ \frac{1}{\ell_2} & -\frac{1}{\ell_2} \end{pmatrix}$ und multiplizieren diese dann mit $\frac{g}{m_1}$. Die zugehörigen Eigenvektoren sind beidemal dieselben.

Die Eigenwerte sind Lösungen der Gleichung

$$\left(x + \frac{1}{\ell_1}\right)\left(x + \frac{1}{\ell_2}\right) = \frac{m_2}{\ell_1\ell_2}$$

Wir finden also als Eigenwerte

$$\lambda'_{1,2} = \frac{1}{2\ell_1\ell_2} \left(-\ell_1 - \ell_2 \pm \sqrt{-4m_1\ell_1\ell_2 + (\ell_1 + \ell_2)^2}\right)$$

und \mathcal{B} hat die Eigenwerte $\lambda_{1,2} = -\omega'_{1,2}$, mit

$$\omega'_1 = \frac{g}{2m_1\ell_1\ell_2} \left(\ell_1 + \ell_2 - \sqrt{-4m_1\ell_1\ell_2 + (\ell_1 + \ell_2)^2}\right)$$

$$\omega'_2 = \frac{g}{2m_1\ell_1\ell_2} \left(\ell_1 + \ell_2 + \sqrt{-4m_1\ell_1\ell_2 + (\ell_1 + \ell_2)^2}\right)$$

Beide ω'_j sind positiv.

Zu den Eigenvektoren: Wir finden als Eigenvektoren

$$\vec{v}_j = \begin{pmatrix} m_2 \\ 1 + \ell_1\lambda_j \end{pmatrix}$$

Das führt für die DGL des Doppelpendels auf die allgemeine Lösung:

$$\Phi(t) = A_1 \vec{v}_1 \sin(\omega_1 t) + A_2 \vec{v}_2 \sin(\omega_2 t)$$

Dabei ist $\omega_j = \sqrt{\omega_j'}$. Die Konstanten A_1 und A_2 hängen von den Startwerten ab.

Wir sehen uns den Fall an, dass $\ell_1 = \ell_2 =: \ell$. Dann haben wir

$$\omega_{1,2} = \frac{g}{2m_1 \ell} (1 \mp \sqrt{m_2})$$

Soll die Funktion Φ also die Bewegung des Doppelpendels periodisch sein, muss $q := \frac{\omega_2}{\omega_1}$ rational sein. In diesem Falle ist

$$m_1 = \frac{4q}{(1+q)^2}, \quad m_2 = \left(\frac{1-q}{1+q}\right)^2$$

1.1.3 Symmetrische und hermitesche Matrizen

Wir erinnern uns an das Skalarprodukt in \mathbb{R}^n :

Definition. Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ setzen wir

$$\langle \vec{x}, \vec{y} \rangle := \sum_{j=1}^n x_j y_j$$

und $\|\vec{x}\| := \sqrt{\langle \vec{x}, \vec{x} \rangle}$. Wir nennen 2 Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ *orthonormal*, wenn $\|\vec{x}\| = \|\vec{y}\| = 1$ und $\langle \vec{x}, \vec{y} \rangle = 0$ gilt.

1.1.3.1 Hilfssatz. Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ und jede reelle $n \times n$ -Matrix \mathcal{A} gilt

$$\langle \vec{x}, \mathcal{A} \vec{y} \rangle = \langle \mathcal{A}^t \vec{x}, \vec{y} \rangle$$

Ist also \mathcal{A} symmetrisch, d.h.: gilt $\mathcal{A}^t = \mathcal{A}$, so wird $\langle \vec{x}, \mathcal{A} \vec{y} \rangle = \langle \mathcal{A} \vec{x}, \vec{y} \rangle$.

1.1.3.2 Satz. Jede reelle symmetrische $n \times n$ -Matrix \mathcal{A} ist diagonalisierbar. Es gibt sogar eine Orthonormalbasis $\vec{b}_1, \dots, \vec{b}_n$, welche nur aus Eigenvektoren von \mathcal{A} besteht. Es gilt also $\langle \vec{b}_i, \vec{b}_j \rangle = 0$, wenn $i \neq j$ und $\|\vec{b}_i\| = 1$, für alle $i = 1, \dots, n$.

Beispiel. Die Matrix $\mathcal{A} = \frac{1}{81} \begin{pmatrix} 115 & -56 & 80 \\ -56 & 178 & 116 \\ 80 & 116 & 31 \end{pmatrix}$ ist symmetrisch und die Matrix $\mathcal{S} = \frac{1}{9} \begin{pmatrix} -4 & -4 & 7 \\ -8 & 1 & -4 \\ 1 & -8 & -4 \end{pmatrix}$ ist orthogonal. Es gilt

$$\mathcal{S}^t \mathcal{A} \mathcal{S} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Die Spalten von \mathcal{S} sind eine Orthonormalbasis des \mathbb{R}^3 und Eigenvektoren zu \mathcal{A} zu den Eigenwerten $-1, 2$ und 3 .

Man nennt eine symmetrische $n \times n$ -Matrix \mathcal{A} *positiv definit*, wenn gilt: $\langle \vec{x}, \mathcal{A}\vec{x} \rangle \geq 0$ und $\langle \vec{x}, \mathcal{A}\vec{x} \rangle > 0$, wenn $\vec{x} \neq \vec{0}$.

Solche Matrizen können an Hand ihrer Eigenwerte gekennzeichnet werden:

1.1.3.3 Satz. a) Genau dann ist die symmetrische $n \times n$ -Matrix \mathcal{A} positiv definit, wenn alle ihre Eigenwerte positiv sind.

b) Genau dann ist $\mathcal{A} = (a_{ij})_{i,j=1}^n$ positiv definit, wenn gilt $\det(a_{ij})_{i,j=1}^k > 0$, für alle $k = 1, \dots, n$.

Beweis. a) Wir wählen zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ zugehörige Eigenvektoren $\vec{b}_1, \dots, \vec{b}_n$ und beachten, dass für jeden Vektor \vec{x} die Darstellung $\vec{x} = \sum_{j=1}^n \langle x, \vec{b}_j \rangle \vec{b}_j$ gilt. Es folgt

$$\langle \vec{x}, \mathcal{A}\vec{x} \rangle = \sum_{j=1}^n \lambda_j \|\vec{b}_j\|^2 \langle x, \vec{b}_j \rangle^2$$

Wenn also alle Eigenwerte von \mathcal{A} positiv sind, so ist \mathcal{A} positiv definit.

Ist umgekehrt \mathcal{A} positiv definit, so gilt $\lambda_j \|\vec{b}_j\|^2 = \langle \vec{b}_j, \mathcal{A}\vec{b}_j \rangle > 0$.

b) Sei \mathcal{A} positiv-definit. Die Determinante jeder positiv-definiten Matrix ist positiv. Wenden wir dies an auf die positiv-definiten $k \times k$ -Matrizen $(a_{ij})_{i,j=1}^k$, so finden wir $\det(a_{ij})_{i,j=1}^k > 0$, für alle $k = 1, \dots, n$.

Umgekehrt angenommen, alle Zahlen $\delta_k := \det(a_{ij})_{i,j=1}^k$, seien positiv. Man kann zeigen (quadratische Ergänzung), dass mit einer geeigneten quadratischen Matrix \mathcal{W} die Gleichung

$$\mathcal{W}^t \mathcal{A} \mathcal{W} = \begin{pmatrix} \delta_1 & & & 0 \\ 0 & \delta_2/\delta_1 & & \\ \vdots & & \ddots & \\ 0 & & & \delta_n/\delta_{n-1} \end{pmatrix}$$

besteht. Die rechte und damit auch die linke Matrix sind positiv-definit.

□

Das Problem der strikten lokalen Minima einer Funktion in mehreren Variablen führt uns auf positiv-definite Matrizen (Hesse-Matrix).

Anwendung. Darstellung "gedrehter Ellipsen". Sei etwa (mit $b \neq 0$) $\mathcal{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ eine positiv-definite Matrix, also $a > 0, ac - b^2 > 0$. Frage: Welche Menge E wird durch die Gleichung

$$\langle \vec{x}, \mathcal{A}\vec{x} \rangle = r^2$$

beschrieben?

Dazu berechnen wir die Eigenwerte von \mathcal{A} : Es sind

$$\lambda_+ = \frac{1}{2}(a + c + \sqrt{(a - c)^2 + 4b^2}), \quad \lambda_- = \frac{1}{2}(a + c - \sqrt{(a - c)^2 + 4b^2})$$

mit Eigenvektoren

$$\vec{v}_+ = \frac{1}{\sqrt{b^2 + (a - \lambda_+)^2}} \begin{pmatrix} -b \\ a - \lambda_+ \end{pmatrix}, \quad \vec{v}_- = \frac{1}{\sqrt{b^2 + (a - \lambda_-)^2}} \begin{pmatrix} -b \\ a - \lambda_- \end{pmatrix}$$

Wir fassen \vec{v}_+ und \vec{v}_- zu einer Matrix \mathcal{S} zusammen. Dann gilt also

$$\mathcal{S}^t \mathcal{A} \mathcal{S} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}, \quad \mathcal{S} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \mathcal{S}^t = \mathcal{A}$$

Genau dann ist $\vec{x} \in E$, wenn

$$\langle \mathcal{S}^t \vec{x}, \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \mathcal{S}^t \vec{x} \rangle = \langle \vec{x}, \mathcal{S} \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \mathcal{S}^t \vec{x} \rangle = r^2$$

Das ist mit der Forderung

$$\mathcal{S}^t \vec{x} \in E' := \{\vec{x}' \mid \lambda_+ x_1'^2 + \lambda_- x_2'^2 = r^2\}$$

äquivalent. Die Menge E' ist aber eine Ellipse mit den Halbachsen $r/\sqrt{\lambda_{\pm}}$.

Die Matrix \mathcal{S} beschreibt eine Drehung um den Nullpunkt mit dem Drehwinkel

$$\alpha = \arccos \left(-\frac{b}{\sqrt{b^2 + (a - \lambda_-)^2}} \right)$$

Das zeigt: E entsteht aus E' durch Drehung um den Winkel α gegen den Uhrzeigersinn.

Sei etwa $E := \{4x_1^2 - 10x_1x_2 + 8x_2^2 = 4\}$. Dann ist also $\mathcal{A} = \begin{pmatrix} 4 & -5 \\ -5 & 8 \end{pmatrix}$. Die Eigenwerte von \mathcal{A} sind $\lambda_{\pm} = 6 \pm \sqrt{29}$, also $\lambda_- = 0.614835$, $\lambda_+ = 11.3852$.

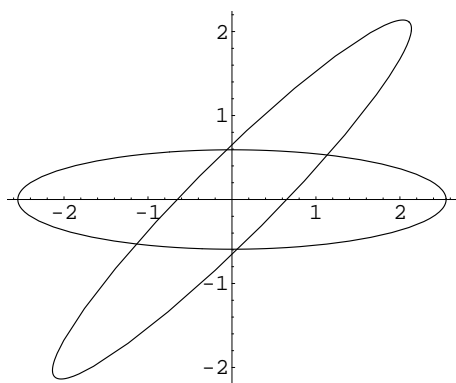
Die Matrix

$$\mathcal{S} = \begin{pmatrix} 0.828066 & -0.560629 \\ 0.560629 & 0.828068 \end{pmatrix}$$

bewirkt

$$\mathcal{S}^t \mathcal{A} \mathcal{S} = \begin{pmatrix} 0.614835 & 0 \\ 0 & 11.3852 \end{pmatrix}$$

Es gilt $\mathcal{S} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$, wobei $\alpha = 34^\circ$. Weiter ist $r = 2$ und $E' = \{\vec{x}' \mid \left(\frac{x_1'}{2.55065}\right)^2 + \left(\frac{x_2'}{0.59273}\right)^2 = 1\}$. Die Menge E ist also eine Ellipse, welche aus E' durch Drehen um 34° gegen den Uhrzeigersinn entsteht.



Kapitel 2

Raumkurven

2.1 Allgemeine Kurventheorie

2.1.1 Die Weglänge

Definition. Unter einer *Raumkurve* im \mathbb{R}^n verstehen wir eine stetige Abbildung $c : [a, b] \rightarrow \mathbb{R}^n$. Weiter sprechen wir von einem \mathcal{C}^k -Weg c , wenn jede Komponente von c schon k -mal stetig differenzierbar ist.

Wir diskutieren zunächst den Begriff der Länge eines Weges.

Dazu treffen wir weitere Definitionen

Definition. Sei $c : [a, b] \rightarrow \mathbb{R}^n$ ein Weg. Mit \underline{c} bezeichnen wir die Menge $c([a, b])$.

a) Unter einer *Zerlegung* \mathcal{Z} von c verstehen wir eine endliche Unterteilung $c(t_1), \dots, c(t_N)$ von c , wobei $t_1 = a, t_N = b$ und $t_1 < t_2 < \dots < t_N$ sei. Die Menge aller derartiger Zerlegungen von c bezeichnen wir mit $\mathcal{Z}(c)$.

b) Sind \mathcal{Z}_1 und \mathcal{Z}_2 zwei Zerlegungen von c , so sagen wir \mathcal{Z}_2 sei *feiner* als \mathcal{Z}_1 , wenn jeder Teilpunkt von \mathcal{Z}_1 auch in \mathcal{Z}_2 vorkommt.

c) Sei jetzt $\mathcal{Z} \in \mathcal{Z}(c)$ eine Zerlegung mit Teilpunkten $c(t_1), \dots, c(t_N)$. Dann setzen wir

$$L(c, \mathcal{Z}) := \sum_{j=1}^{N-1} \|c(t_{j+1}) - c(t_j)\|$$

Aus der Dreiecksungleichung folgt: $L(c, \mathcal{Z}_1) \leq L(c, \mathcal{Z}_2)$, wenn \mathcal{Z}_2 feiner als \mathcal{Z}_1 ist.

d) Wir nennen c *rektifizierbar*, wenn

$$L(c) := \sup\{L(c, \mathcal{Z}) \mid \mathcal{Z} \in \mathcal{Z}(c)\}$$

endlich ist.

Beispiel. Der folgende Weg c wird als Grenzwert einer Folge von Wegen $(c_k)_k$ definiert:

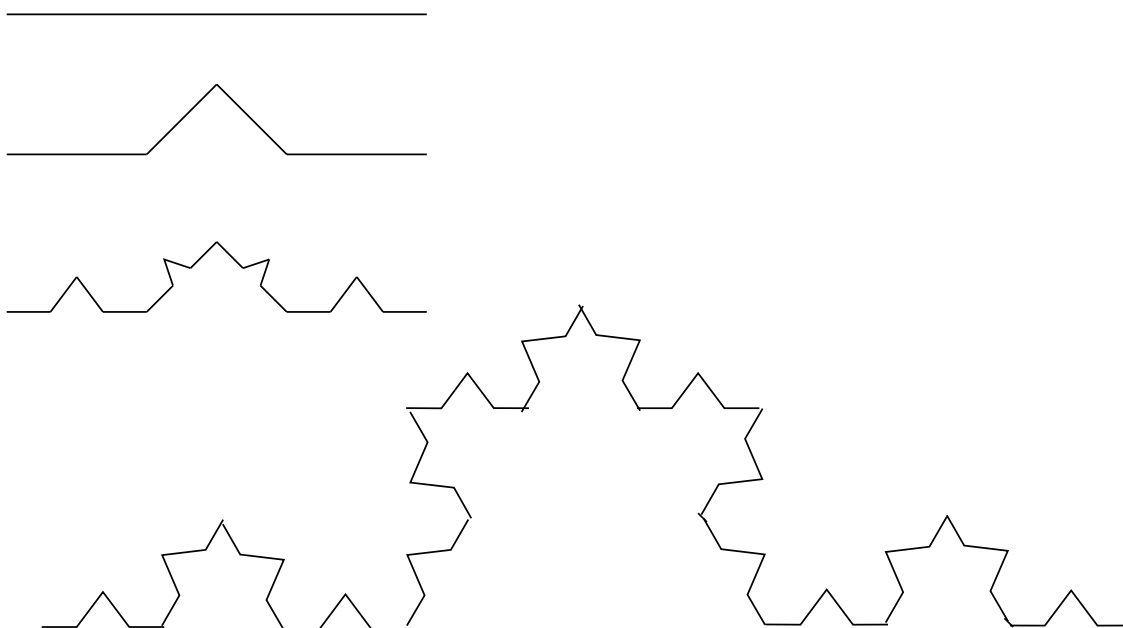
c_1 ist die Strecke von 0 bis 3 (in der Ebene \mathbb{R}^2)

Zu c_2 : Man ersetze in c_1 die Teilstrecke von 1 bis 2 durch die Schenkel des gleichseitigen Dreiecks mit Ecken bei 2 und 3 und $(1.5 | \frac{1}{2}\sqrt{3})$

Zu c_3 : Man teile jede Teilstrecke von c_2 in drei gleiche Teile und ersetze jeweils das mittlere Teilstück durch die Schenkel eines gleichseitigen Dreiecks.

So fahre man fort: Ist c_k konstruiert, so teile man jede Teilstrecke von c_k in drei gleiche Teile und ersetze jeweils das mittlere Teilstück durch die Schenkel eines gleichseitigen Dreiecks und erhält c_{k+1} .

Im Bild sehen wir untereinander c_1, c_2, c_3 und (in vergrößerter Form) c_4 :



Das so entstehende c ist nicht rektifizierbar, denn $L(c_k) = 4(\frac{4}{3})^{k-1}$.

Ein stetig differenzierbarer Weg ist immer rektifizierbar.

2.1.1.1 Satz. Ist $c : [a, b] \rightarrow \mathbb{R}^n$ auf (a, b) stetig differenzierbar, so ist c rektifizierbar, und es gilt

$$L(c) = \int_a^b \|\dot{c}(t)\| dt$$

(Hierbei bedeutet \dot{c} den Vektor der Ableitungen der Komponenten von c).

Beweis. Ist nämlich $\mathcal{Z} = \{c(t_1), \dots, c(t_N)\}$ eine Zerlegung von c , so ist (mit dem vorherigen

Hilfssatz)

$$L(c, \mathcal{Z}) = \sum_{j=1}^{N-1} \|c(t_{j+1}) - c(t_j)\| = \sum_{j=1}^{N-1} \left\| \int_{t_j}^{t_{j+1}} \dot{c}(t) dt \right\| \leq \sum_{j=1}^{N-1} \int_{t_j}^{t_{j+1}} \|\dot{c}(t)\| dt = \int_a^b \|\dot{c}(t)\| dt$$

Damit ist die Rektifizierbarkeit von c gezeigt, ebenso, dass $L(c) \leq \int_a^b \|\dot{c}(t)\| dt$.

Mit c_t bezeichnen wir den Weg $c|_{[a, t]}$, wenn $t \in (a, b)$. Auch c_t ist rektifizierbar, und $L(c_t) \leq L(c)$. Sei $t_0 \in (a, b)$ und \mathcal{Z} eine Zerlegung von c_{t_0} . Dann gilt für $t_0 < t < b$:

$$L(c_{t_0}, \mathcal{Z}) + \|c(t) - c(t_0)\| \leq L(c_t)$$

da durch Hinzunahme von $c(t)$ zu \mathcal{Z} eine Zerlegung für c_t entsteht. Da aber \mathcal{Z} beliebig ist, ist

$$L(c_{t_0}) + \|c(t) - c(t_0)\| \leq L(c_t)$$

Damit wird aber

$$\|c(t) - c(t_0)\| \leq L(c_t) - L(c_{t_0}) \leq \int_{t_0}^t \|\dot{c}(s)\| ds$$

Dasselbe gilt auch, wenn $t < t_0$. Teilen wir die Ungleichungskette durch $t - t_0$ und lassen dann t gegen t_0 gehen, so folgt, dass $L(c_t)$ eine differenzierbare Funktion ist und ihre Ableitung bei t_0 gerade

$$\left. \frac{d}{dt} L(c_t) \right|_{t=t_0} = \|\dot{c}(t_0)\|$$

Das ergibt aber

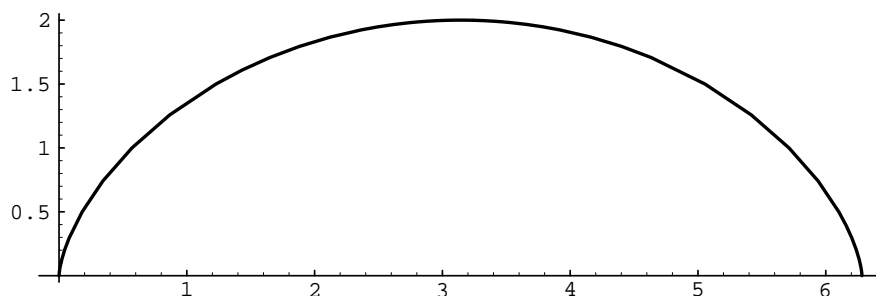
$$L(c) = L(c_b) = \int_a^b \|\dot{c}(t)\| dt$$

□

Beispiele. 1) Ein Kreis um \vec{M} mit Radius R . Wir setzen $c(t) = \vec{R} + R(\cos t, \sin t)$, für $t \in [0, 2\pi]$ und bilden \dot{c} . Es gilt $\dot{c}(t) = R(-\sin t, \cos t)$, also $\|\dot{c}(t)\| = R$. Das ergibt

$$L(c) = \int_0^{2\pi} \|\dot{c}(t)\| dt = 2\pi R$$

2) Die Zykloide $c(t) = (t - \sin t, 1 - \cos t)$



Dann ist

$$\dot{c}(t) = (1 - \cos t, \sin t), \quad \|\dot{c}(t)\|^2 = (1 - \cos t)^2 + \sin^2 t = 2(1 - \cos t) = 4 \sin^2(t/2)$$

Damit wird

$$L(c) = 2 \int_0^{2\pi} \sin(t/2) dt = 4 \int_0^\pi \sin s ds = 8$$

3) Die Parabel $c(t) = (t, t^2)$ mit $t \in [0, 4]$. Nun ist $\|\dot{c}(t)\| = \sqrt{1 + 4t^2}$, also

$$\begin{aligned} L(c) &= \int_0^4 \sqrt{1 + 4t^2} dt \\ &= \left(\frac{1}{2} t \sqrt{1 + 4t^2} + \frac{1}{4} \log(2t + \sqrt{1 + 4t^2}) \right) \Big|_0^4 \\ &= 2\sqrt{65} + \frac{1}{4} \log(8 + \sqrt{65}) = 16.8186 \end{aligned}$$

Wir können jede glatte Kurve in der Weise parametrisieren, dass ihr Geschwindigkeitsvektor die konstante Länge 1 bekommt.

2.1.1.2 Hilfssatz. Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine glatte Kurve, so sei $L(t) := \int_a^t \|\dot{\gamma}(\tau)\| d\tau$, für $t \in [a, b]$. Dann ist $L(a) = 0$, und s hat eine Umkehrfunktion $L^{-1} : [0, L(b)] \rightarrow [a, b]$. Die Kurve $c(s) = \gamma(L^{-1}(s))$ ist auf $[0, L(b)]$ definiert und der Vektor $\frac{d}{ds} c(s)$ hat die konstante Länge 1.

Beweis. Dies folgt aus der Kettenregel.

□

Man sagt, dass die Kurve γ aus c durch *Parametrisierung nach der Bogenlänge* hervorgehe.

Beispiel. Der Kreis um 0 mit Radius R wird durch

$$\gamma(s) := R \begin{pmatrix} \cos(s/R) \\ \sin(s/R) \end{pmatrix}, \quad 0 \leq s < 2\pi R$$

bogenlängenparametrisiert.

2.1.2 Ebene und räumliche Kurven

Das begleitende Frenetsche 2-Bein (3-Bein)

Der Fall ebener Kurven

Angenommen, wir haben eine nach der Bogenlänge parametrisierte glatte Kurve $c : [0, b] \rightarrow \mathbb{R}^2$, also $\|\dot{c}(s)\| = 1$ für alle $s \in [0, b]$. Dann bilden

$$\vec{e}_1(s) := \dot{c}(s), \quad \vec{n}(s) := \begin{pmatrix} -\dot{c}_2 \\ \dot{c}_1 \end{pmatrix}(s)$$

eine Orthonormalbasis im \mathbb{R}^2 für jede Wahl von s .

Das System $\{\vec{e}_1, \vec{e}_2\}$ wird als *begleitende Frenetsche 2-Bein* zu c bezeichnet.

Das begleitende Frenetsche 2-Bein ist in dem folgenden Sinne eindeutig bestimmt:

Ist $\{\vec{f}_1, \vec{f}_2\}$ ein System von Orthonormalbasen des \mathbb{R}^2 , so dass \vec{f}_1 in Richtung \dot{c} weist und $\det(\vec{f}_1, \vec{f}_2)$ stets positiv ist, so ist schon $\{\vec{f}_1, \vec{f}_2\} = \{\vec{e}_1, \vec{e}_2\}$.

Wir zweigen jetzt, dass jede ebene Kurve durch ihr begleitendes Frenetsches 2-Bein "im Wesentlichen" eindeutig festgelegt ist.

Dazu leiten wir für das Frenetsche 2-Bein eine Differenzialgleichung her:

Zunächst ist

$$\dot{\vec{e}}_1(s) = \ddot{c}$$

und weiter steht \ddot{c} auf \dot{c} senkrecht, denn $\langle \ddot{c}, \dot{c} \rangle(s) = \frac{1}{2} \frac{d}{ds} \|\dot{c}\|^2 = 0$. Das bedeutet aber

$$\ddot{c}(s) = \varkappa(s) \vec{e}_2(s),$$

wobei \varkappa die durch

$$\varkappa(s) := \langle \ddot{c}, \vec{e}_2 \rangle(s)$$

definierte Funktion sein soll. Also ist

$$\dot{\vec{e}}_1(s) = \varkappa(s) \cdot \vec{e}_2(s)$$

Ferner haben wir wegen $\langle \dot{\vec{e}}_2, \vec{e}_2 \rangle = 0$, dass

$$\dot{\vec{e}}_2(s) = \alpha(s) \vec{e}_1(s)$$

mit

$$\alpha(s) = \langle \dot{\vec{e}}_2, \vec{e}_1 \rangle = \frac{d}{ds} \langle \vec{e}_2, \vec{e}_1 \rangle - \langle \vec{e}_2, \dot{\vec{e}}_1 \rangle = -\kappa(s)$$

Das ergibt zusammen

$$\begin{aligned} \dot{\vec{e}}_1(s) &= \kappa(s) \vec{e}_2(s) \\ \dot{\vec{e}}_2(s) &= -\kappa(s) \vec{e}_1(s) \end{aligned} \quad \text{Frenetsche DGL}$$

Definition. Die Größe κ wird dabei *Krümmung* von c genannt.

Auch für nicht bogenlängenparametrisierte Kurven γ wird die Krümmung definiert.

2.1.2.1 Hilfssatz. Ist $\gamma : [a, b] \rightarrow \mathbb{R}^2$ eine glatte Kurve, so ist ihre Krümmung durch

$$\kappa_\gamma(t) := \frac{\det(\dot{\gamma}, \ddot{\gamma})}{\|\dot{\gamma}\|^3}$$

gegeben.

Beweis. Man schreibe $\gamma = c \circ h$ mit einer differenzierbaren Funktion h und einer nach der Bogenlänge parametrisierten Kurve c . Dann haben wir

$$\dot{\gamma} = c' \circ h \cdot h', \quad \ddot{\gamma} = c'' \circ h \cdot h'^2 + c' \circ h \cdot h''$$

also

$$\det(\dot{\gamma}, \ddot{\gamma}) = \det(\dot{c}, \ddot{c}) \circ h \cdot h'^3$$

und damit

$$\frac{\det(\dot{\gamma}, \ddot{\gamma})}{\|\dot{\gamma}\|^3} = \det(\dot{c}, \ddot{c}) \circ h$$

Zusammen mit $\kappa_c = \det(\dot{c}, \ddot{c})$ folgt die Behauptung.

Die Krümmung legt den Verlauf der Kurve fest:

2.1.2.2 Satz. Angenommen, es seien zwei Kurven $c, \gamma : [a, b] \rightarrow \mathbb{R}^2$ gegeben und beide haben dieselbe Krümmung. Dann gibt es eine orthogonale Matrix \mathcal{A} und ein $\vec{B} \in \mathbb{R}^2$ mit

$$c = \mathcal{A} \cdot \gamma + \vec{B}$$

Beweis. Sind $\{\vec{e}_1^c, \vec{e}_2^c\}$ und $\{\vec{e}_1^\gamma, \vec{e}_2^\gamma\}$ die begleitenden Frenetschen 2-Beine zu c und γ , so wählen wir \mathcal{A} als orthogonale Matrix mit

$$\mathcal{A} \cdot \vec{e}_1^\gamma(a) = \vec{e}_1^c(a), \quad \mathcal{A} \cdot \vec{e}_2^\gamma(a) = \vec{e}_2^c(a)$$

und dann $\vec{B} = c(a) - \mathcal{A} \cdot \gamma(a)$. Dann erfüllen $\{\vec{e}_1^c, \vec{e}_2^c\}$ und $\{\mathcal{A} \cdot \vec{e}_1^\gamma, \mathcal{A} \cdot \vec{e}_2^\gamma\}$ dieselbe Frenetsche Differenzialgleichung und stimmen bei a überein. Damit stimmen sie überall überein. Daraus erhalten wir die Behauptung.

□

Die Krümmung gibt uns bei ebenen Kurven ein geeignetes Maß für die Abweichung vom geraden Verlauf.

Ist $\gamma : (-a, a) \rightarrow \mathbb{R}^2$ eine glatte Kurve, also $\dot{\gamma}$ ohne Nullstellen, so wird die Tangenterichtungsvektor an $\gamma(t)$ durch $\dot{\gamma}(t)$ und der Richtungsvektor ihrer Normalen in diesem Punkt durch $\gamma^{nor}(t) = \begin{pmatrix} -\dot{\gamma}_2 \\ \dot{\gamma}_1 \end{pmatrix}$ gegeben.

Weiter erwarten wir, dass bei nicht gerade verlaufenden Kurven die Normalen nicht parallel verlaufen, sich also schneiden.

Es seien $t_0, t \in (-a, a)$. Dann schneiden sich die Normalen an γ durch die Punkte $\gamma(t_0)$ und $\gamma(t)$, wenn es Werte $\lambda_0, \lambda_t \in \mathbb{R}$ mit

$$\gamma(t_0) + \lambda_0 \gamma^{nor}(t_0) = \gamma(t) + \lambda_t \gamma^{nor}(t)$$

gibt. Es folgt durch Bilden des Skalarproduktes mit $\dot{\gamma}(t_0)$:

$$\langle \gamma(t_0) - \gamma(t), \dot{\gamma}(t_0) \rangle = \lambda_t \langle \gamma^{nor}(t_0), \dot{\gamma}(t_0) \rangle = \lambda_t \langle \gamma^{nor}(t) - \gamma^{nor}(t_0), \dot{\gamma}(t_0) \rangle,$$

wobei die letzte Gleichung aus $\langle \gamma^{nor}(t_0), \dot{\gamma}(t_0) \rangle = 0$ folgt. Taylorentwicklung von γ^{nor} um t_0 ergibt uns nun

$$\gamma^{nor}(t) - \gamma^{nor}(t_0) = \left(\frac{d}{dt} \gamma^{nor} \right) (t_0) (t - t_0) + \mathcal{O}_2(t - t_0)^2$$

in der Nähe von t_0 . Nun teilen wir für $t \neq t_0$ durch $t - t_0$ und finden

$$\frac{\gamma^{nor}(t) - \gamma^{nor}(t_0)}{t - t_0} = \left(\frac{d}{dt} \gamma^{nor} \right) (t_0) + \mathcal{O}_1(t - t_0) = \begin{pmatrix} -\ddot{\gamma}_2 \\ \ddot{\gamma}_1 \end{pmatrix} (t_0) + \mathcal{O}_1(t - t_0)$$

Einsetzen in die Gleichung für λ_t führt auf

$$\begin{aligned} \left\langle \frac{\gamma(t_0) - \gamma(t)}{t - t_0}, \dot{\gamma}(t_0) \right\rangle &= \lambda_t \left(\langle \gamma^{nor}(t_0), \dot{\gamma}(t_0) \rangle + \left\langle \begin{pmatrix} -\ddot{\gamma}_2 \\ \ddot{\gamma}_1 \end{pmatrix}, \dot{\gamma}(t_0) + \mathcal{O}_1(t - t_0) \right\rangle \right) \\ &= \lambda_t \left(\left\langle \begin{pmatrix} -\ddot{\gamma}_2 \\ \ddot{\gamma}_1 \end{pmatrix}, \dot{\gamma}(t_0) + \mathcal{O}_1(t - t_0) \right\rangle \right) \end{aligned}$$

Fordern wir also, dass $\left\langle \begin{pmatrix} -\ddot{\gamma}_2 \\ \ddot{\gamma}_1 \end{pmatrix}, \dot{\gamma}(t_0) \right\rangle \neq 0$ sein soll, so kann die Gleichung für einen Schnittpunkt zwischen beiden Normalen nach λ_t aufgelöst werden. Nun lassen wir t gegen t_0 streben und finden

$$\lim_{t \rightarrow t_0, t \neq t_0} \lambda_t = - \frac{\|\dot{\gamma}(t_0)\|^2}{\left\langle \begin{pmatrix} -\ddot{\gamma}_2 \\ \ddot{\gamma}_1 \end{pmatrix}, \dot{\gamma}(t_0) \right\rangle}$$

Der Schnittpunkt zwischen beiden Normalen strebt also mit $t \rightarrow t_0$ gegen den Punkt

$$M(t_0) := \gamma(t_0) + \frac{1}{\varkappa(t_0)} \frac{\gamma^{nor}}{\|\gamma^{nor}\|}(t_0).$$

Der Kreis um den Punkt $M(t_0)$ mit Radius $\frac{1}{\varkappa(t_0)}$ wird der *Schmiegekreis* an γ bei $\gamma(t_0)$ genannt und $M(t_0)$ der *Schmiegekreismittelpunkt*.

Beispiele. a) Der Kreis um a mit Radius R hat die Krümmung konstant gleich $1/R$.

b) Ist $f : [a, b] \rightarrow \mathbb{R}$ 2-mal stetig differenzierbar, so ist der Graph von f eine Kurve mit Parametrisierung $\alpha(t) = (t, f(t))$. Die Krümmung \varkappa ist nun

$$\varkappa(t) = \frac{f''(t)}{(1 + (f'(t))^2)^{3/2}}$$

c) Die Zykloide $\alpha(t) = r(t - \sin t, 1 - \cos t)$ ist regulär auf $(0, 2\pi)$ mit der Krümmung

$$\varkappa(t) = \frac{1}{4r} \frac{t \cos(t/2) - 2 \sin(t/2)}{\sin^2(t/2)}$$

Die Krümmung beschreibt, wie schnell sich die Richtung des Geschwindigkeitsvektors ändert:

2.1.2.3 Satz. Ist α eine glatte reguläre Kurve und $\dot{\alpha} = \|\dot{\alpha}\| \begin{pmatrix} \cos \psi(t) \\ \sin \psi(t) \end{pmatrix}$ mit der "Winkelgeschwindigkeit" $\dot{\psi}$, so gilt $\varkappa = \dot{\psi}$.

Beweis. Denn es gilt

$$\vec{e}_1^\alpha = \frac{\dot{\alpha}}{\|\dot{\alpha}\|} = \begin{pmatrix} \cos \psi(t) \\ \sin \psi(t) \end{pmatrix}, \quad \dot{\vec{e}}_1^\alpha = \dot{\psi} \begin{pmatrix} -\sin \psi(t) \\ \cos \psi(t) \end{pmatrix}$$

und

$$\vec{e}_2^\alpha = \frac{1}{\|\dot{\alpha}\|} \begin{pmatrix} -\dot{\alpha}_2 \\ \dot{\alpha}_1 \end{pmatrix} = \begin{pmatrix} -\sin \psi(t) \\ \cos \psi(t) \end{pmatrix}.$$

Die Behauptung folgt aus $\varkappa = \langle \dot{\vec{e}}_1^\alpha, \vec{e}_2^\alpha \rangle = \dot{\psi}$.

Evoluten und Evolventen

Definition. Sei γ eine glatte Kurve, deren Krümmung nirgends verschwindet. Dann laufen die Schmiegkreismittelpunkte zu γ auf einer Kurve γ^E , die man *Evolute* von γ nennt. Eine Kurve c heißt *Evolvente* zu γ , wenn $\gamma = c^E$ gilt.

Es ist also

$$\gamma^E(t) = \gamma(t) + \frac{1}{\varkappa(t)} \frac{\gamma^{nor}}{\|\gamma^{nor}\|}(t).$$

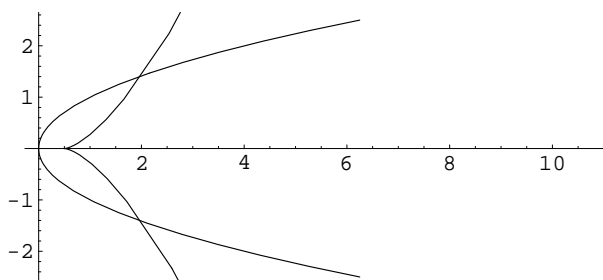
Die Berechnung von Evolventen spielt bei der Konstruktion von Zahnrädern eine große Rolle. Am bedeutendsten sind hierbei Zahnräder mit Evolventenverzahnung. Hier haben die Flanken der Zähne des Zahnrades die Form einer Evolvente einer geeigneten Kurve.

Beispiel: Die Parabel $\gamma(t) = (t^2, t)$. Nun ist

$$\dot{\gamma}(t) = \begin{pmatrix} 2t \\ 1 \end{pmatrix}, \quad \varkappa(t) = -\frac{2}{(1+4t^2)^{3/2}}$$

also

$$\gamma^E(t) = \begin{pmatrix} t^2 \\ t \end{pmatrix} + \frac{1}{2}(1+4t^2) \begin{pmatrix} 1 \\ -2t \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + 3t^2 \\ -4t^3 \end{pmatrix}$$



Hier haben wir eine Kennzeichnung der Evolventen einer glatten Kurve:

2.1.2.4 Satz. Angenommen, γ sei eine glatte nach der Bogenlänge parametrisierte Kurve.

a) Ist dann $t_0 > 0$, so ist $c(t) := \gamma(t) + (t_0 - t)\vec{e}_1^\gamma(t)$ eine auf $[0, t_0]$ definierte Evolvente zu γ .

b) Hat die Krümmung \varkappa und ihre Ableitung $\dot{\varkappa}$ nirgends eine Nullstelle, so hat jede Evolvente von γ die unter a) beschriebene Form.

Beispiel: Die Kettenlinie und die Ziehkurve. Hängt man eine Kette an ihren beiden Enden auf, so beschreibt die derart durchhängende Kette eine Kettenlinie genannte Kurve c . Parametrisiert man sie nach der Bogenlänge, so findet man

$$c(s) = a \left(\operatorname{Arsinh} \left(\frac{s}{a} \right), \sqrt{1 + (s/a)^2} \right)$$

Nun studieren wir die Evolvente

$$T(s) := c(s) - s c'(s)$$

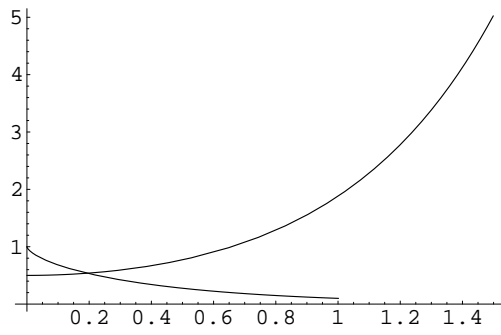
Dann ist

$$T'(s) = \frac{s}{a} \sqrt{1 + (s/a)^2}^{-3} \left(\frac{s}{a}, -1 \right)$$

Die Tangente an T durch $T(s)$ hat die Darstellung $\lambda \mapsto T(s) + \lambda T'(s)$ und trifft die x -Achse im Punkte $T(s) + \lambda_a T'(s)$, wobei $\lambda_a = \frac{a^2}{s} (1 + (s/a)^2)$. Der Abstand zwischen diesem Schnittpunkt und dem Punkt $T(s)$ ist also

$$\|T(s) - (T(s) + \lambda_a T'(s))\| = |\lambda_a| \|T'(s)\| = a,$$

also konstant. Die Kurve T wird *Ziehkurve (Traktrix)* genannt, denn sie entsteht auf folgende Weise: Man nehme ein Seil der Länge a und lege es längs der x -Achse hin, so dass ein Ende im Nullpunkt liegt. Dann bewege man dieses Ende längs der y -Achse. Das andere Ende beschreibt dann eine Kurve vom Typ T .



Die aufstrebende Kurve ist die Kettenlinie, die andere (eine Evolvente dazu) die Traktrix.

Der Fall räumlicher Kurven im \mathbb{R}^3

Angenommen, es sei c eine nach der Bogenlänge parametrisierte Kurve und weiter, es sei auch \ddot{c} nullstellenfrei.

Dann gilt

$$\langle \ddot{c}, \dot{c} \rangle(s) = \frac{d}{ds} \|c(s)\|^2 = 0$$

also steht \ddot{c} auf dem Geschwindigkeitseinheitsvektor \dot{c} senkrecht. Wir nennen daher $\vec{n}_c = \ddot{c}/\|\ddot{c}\|$ den *Normalenvektor* an c . Als den *Binormalenvektor* bezeichnen wir das Vektorprodukt $\vec{b}_c = \dot{c} \times \vec{n}_c$.

Durch $\{\dot{c}(s), \vec{n}_c(s), \vec{b}_c(s)\}$ wird dann an jeder Stelle $c(s)$ der Kurve eine Orthonormalbasis gegeben. Dies nennt man das *begleitende (Frenetsche) 3-Bein* zu c .

Wir stellen jetzt ein DGL-System für das Frenet-3-Bein auf: Offenbar gilt

$$\ddot{c} = \|\ddot{c}\| \vec{n}_c$$

Weiter haben wir $\dot{\vec{n}}_c(s) = \alpha \dot{c} + \beta \vec{n}_c(s) + \gamma \vec{b}_c(s)$, wobei

$$\alpha = \langle \dot{\vec{n}}_c(s), \dot{c} \rangle = -\langle \vec{n}_c(s), \ddot{c} \rangle = -\|\ddot{c}\|$$

$$\beta = \langle \dot{\vec{n}}_c(s), \vec{n}_c(s) \rangle = 0$$

$$\gamma = \langle \dot{\vec{n}}_c(s), \vec{b}_c \rangle = -\langle \vec{n}_c(s), \dot{\vec{b}}_c \rangle = -\langle \vec{n}_c(s), \dot{c} \times \vec{n}_c \rangle$$

Definieren wir also

$$\varkappa(s) := \|\ddot{c}(s)\| \quad \text{und} \quad \tau := \langle \dot{\vec{n}}_c(s), \vec{b}_c \rangle,$$

so wird

$$\dot{\vec{n}}_c(s) = -\varkappa \dot{c} + \tau \vec{b}_c(s)$$

Schließlich haben wir noch

$$\dot{\vec{b}}_c(s) = \tilde{\alpha} \dot{c} + \tilde{\beta} \vec{n}_c(s) + \tilde{\gamma} \vec{b}_c(s)$$

Dabei ist $\tilde{\alpha} = \langle \dot{\vec{b}}_c(s), \dot{c} \rangle = -\langle \vec{b}_c(s), \ddot{c} \rangle = -\varkappa \langle \vec{b}_c(s), \vec{n}_c \rangle = 0$. Ebenso ist $\tilde{\gamma} = \langle \dot{\vec{b}}_c(s), \vec{b}_c \rangle = \frac{1}{2} \frac{d}{ds} \|\vec{n}_c\|^2 = 0$. Nun gilt noch

$$\tilde{\beta} = \langle \dot{\vec{b}}_c(s), \vec{n}_c \rangle = -\langle \vec{b}_c(s), \dot{\vec{n}}_c \rangle = -\tau \vec{b}_c$$

Also insgesamt:

$$\frac{d}{ds} \begin{pmatrix} \dot{c} \\ \vec{n}_c \\ \vec{b}_c \end{pmatrix} = \begin{pmatrix} \varkappa \vec{n}_c \\ -\varkappa \dot{c} + \tau \vec{b}_c(s) \\ -\tau \vec{n}_c \end{pmatrix}$$

Nun berechnen wir die Torsion der Kurve: Aus

$$\dot{\vec{n}}_c = \frac{\ddot{\vec{c}}}{\|\dot{\vec{c}}\|} - \frac{\langle \ddot{\vec{c}}, \dot{\vec{c}} \rangle \dot{\vec{c}}}{\|\dot{\vec{c}}\|^3}$$

$$\tau = \langle \dot{\vec{n}}_c, \vec{b}_c \rangle = \left\langle \frac{\ddot{\vec{c}}}{\|\dot{\vec{c}}\|} - \frac{\langle \ddot{\vec{c}}, \dot{\vec{c}} \rangle \dot{\vec{c}}}{\|\dot{\vec{c}}\|^3}, \vec{b}_c \right\rangle = \left\langle \frac{\ddot{\vec{c}}}{\|\dot{\vec{c}}\|}, \vec{b}_c \right\rangle = \frac{1}{\|\dot{\vec{c}}\|^2} \langle \ddot{\vec{c}}, \dot{\vec{c}} \times \dot{\vec{c}} \rangle = \frac{\det(\dot{\vec{c}}, \ddot{\vec{c}}, \dot{\vec{c}})}{\|\dot{\vec{c}}\|^2}$$

Wir berechnen anschließend Krümmung und Torsion einer Kurve, die nicht notwendig nach der Bogenlänge parametrisiert sein muss.

Sei also nun γ eine glatte Kurve, für die $\dot{\gamma}$ und $\ddot{\gamma}$ stets linear unabhängig sein sollen. Ist dann h eine geeignet gewählte Umparametrisierung, so können wir γ in der Form $\gamma = \vec{c} \circ h$ darstellen, wobei nun c nach der Bogenlänge parametrisiert ist.

Dann gilt

$$\gamma' = \dot{c} \circ h h'$$

$$\gamma'' = \ddot{c} \circ h h'^2 + \dot{c} \circ h h''$$

$$\gamma''' = \ddot{\ddot{c}} \circ h h'^3 + 3\ddot{c} \circ h h' h'' + \dot{c} \circ h h'''$$

Es folgt leicht, dass

$$\det(\gamma', \gamma'', \gamma''') = \det(\dot{c}, \ddot{c}, \ddot{\ddot{c}}) \circ h h'^6$$

Weiter haben wir

$$\gamma' \times \gamma'' = (\dot{c} \times \ddot{c}) \circ h h'^3$$

Da \dot{c} auf \ddot{c} senkrecht steht, haben wir

$$\varkappa \circ h = \|\ddot{c}\| \circ h = \frac{\|\gamma' \times \gamma''\|}{h'^3}$$

Also ist

$$\tau \circ h = \frac{\det(\dot{c}, \ddot{c}, \ddot{\ddot{c}})}{\|\dot{\vec{c}}\|^2} = \frac{\det(\gamma', \gamma'', \gamma''')}{\|\gamma' \times \gamma''\|^2}$$

Wegen $h' = \|\gamma'\|$ folgt gleichzeitig

$$\varkappa = \frac{\|\gamma' \times \gamma''\|}{\|\gamma'\|^3}$$

Die Torsion misst, ob die Kurve eben ist:

2.1.2.5 Satz. Genau dann verläuft eine Kurve $c : [0, L] \rightarrow \mathbb{R}^3$ in einer Ebene, wenn $\tau = 0$ ist.

Beweis. Ist die Kurve c eben, so sind \dot{c}, \ddot{c} und $\ddot{\ddot{c}}$ immer linear abhängig, also $\tau = 0$, da $\det(\dot{c}, \ddot{c}, \ddot{\ddot{c}}) = 0$.

Umgekehrt angenommen, es sei $\tau = 0$. Dann ist $\dot{\vec{b}}_c = 0$, also \vec{b}_c konstant gleich $\vec{b}_c(0)$, also für alle t :

$$\frac{d}{dt}\langle c, \vec{b}_c(0) \rangle = \frac{d}{dt}\langle c, \vec{b}_c \rangle(t) = \langle \dot{c}, \vec{b}_c \rangle(t) + \langle c, \dot{\vec{b}}_c \rangle(t) = \langle \dot{c}, \vec{b}_c \rangle(t) = \langle \dot{c}(t), \dot{c}(t) \times \vec{n}_c(t) \rangle = 0.$$

Es folgt

$$\langle c, \vec{b}_c(0) \rangle = c_0 := \langle c(0), \vec{b}_c(0) \rangle$$

Damit verläuft c in der Ebene $E := \{\vec{x} \mid \langle \vec{x}, \vec{b}_c(0) \rangle = c_0\}$. □

Beispiel: Für die Schraubenlinie $\gamma(t) = \begin{pmatrix} R \cos(\omega t) \\ R \sin(\omega t) \\ kt \end{pmatrix}$ haben wir

$$\gamma'(t) = \begin{pmatrix} -R\omega \sin(\omega t) \\ R\omega \cos(\omega t) \\ k \end{pmatrix}, \quad \gamma''(t) = -R\omega^2 \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \\ 0 \end{pmatrix}, \quad \gamma''' = R\omega^3 \begin{pmatrix} \sin(\omega t) \\ -\cos(\omega t) \\ 0 \end{pmatrix}$$

Also

$$\gamma' \times \gamma'' = R\omega^2 \begin{pmatrix} k \sin(\omega t) \\ -k \cos(\omega t) \\ R\omega \end{pmatrix}$$

Das liefert uns

$$\varkappa = \frac{\|\gamma' \times \gamma''\|}{\|\gamma'\|^3} = R\omega^2 \frac{\sqrt{R^2\omega^2 + k^2}}{\sqrt{R^2\omega^2 + k^2}^3} = \frac{R\omega^2}{R^2\omega^2 + k^2}$$

und

$$\tau = \frac{\det(\gamma', \gamma'', \gamma''')}{\|\gamma' \times \gamma''\|^2} = \frac{R^2\omega^5 k}{(R^2\omega^2 + k^2)R^2\omega^4} = \frac{\omega k}{R^2\omega^2 + k^2}$$

Zum Schluss berechnen wir das begleitende Frenetsche 3-Bein bei allgemein parametrisierten Kurven. Sei also α eine solche glatte Kurve und c bogenlängenparametrisiert, so dass $\alpha = c \circ h$ mit einer glatten Funktion h , so dass $h' > 0$. Dann gilt

$$\alpha' = \dot{c} \circ h h', \quad \alpha'' = \ddot{c} \circ h h'^2 + \dot{c} \circ h h''$$

Es folgt

$$\alpha' \times \alpha'' = (\dot{c} \times \ddot{c}) \circ h h'^3$$

und mit der Rechenregel für das 3-fache Vektorprodukt

$$(\alpha' \times \alpha'') \times \alpha' = (\|\dot{c} \circ h\|^2 \ddot{c} \circ h - \langle \dot{c}, \ddot{c} \rangle \circ h) h'^4 = \ddot{c} \circ h h'^4$$

Das zeigt, dass $(\alpha' \times \alpha'') \times \alpha'$ in Richtung \vec{n}_c weist.

So finden wir, dass

$$\left\{ \frac{\alpha'}{\|\alpha'\|}, \frac{(\alpha' \times \alpha'') \times \alpha'}{\|(\alpha' \times \alpha'') \times \alpha'\|}, \frac{\alpha' \times \alpha''}{\|\alpha' \times \alpha''\|} \right\}$$

als begleitendes Frenet-3-Bein gewählt werden kann.

2.2 Wegintegrale

2.2.1 Allgemeine Definition

Wir können nun Wegintegrale stetig differenzierbarer Funktionen wie folgt einführen:

Definition. Ist $c : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg, so soll für jede (vektorwertige) stetige Funktion $f : \underline{c} \rightarrow \mathbb{R}^n$ unter

$$\int_c \langle f(\vec{s}), d\vec{s} \rangle$$

das Integral

$$\int_c \langle f(\vec{s}), d\vec{s} \rangle := \int_a^b \langle f(c(t)), \dot{c}(t) \rangle dt$$

verstanden werden.

Das folgende zeigt, dass diese Definition von der gewählten Parametrisierung nicht abhängt:

2.2.1.1 Hilfssatz. Ist $h : [a', b'] \rightarrow [a, b]$ umkehrbar und stetig mit $h(a') = a$, $h(b') = b$ und ist weiter h auf (a', b') stetig differenzierbar, dann ist für jeden glatten Weg $c : [a, b] \rightarrow \mathbb{R}^n$

$$\int_a^b \langle f(c(t)), \dot{c}(t) \rangle dt = \int_{a'}^{b'} \langle f(c^*(t)), \dot{c}^*(t) \rangle dt,$$

wobei $c^* := c \circ h$.

Beispiel. Ist etwa $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definiert als $f(\vec{x}) := (x_1, x_2 - x_1, x_3 + x_1)$ und $c : [0, 1] \rightarrow \mathbb{R}^3$ die "Schraubenlinie"

$$c(t) := (R \cos(2\pi t), R \sin(2\pi t), 3t)$$

so wird

$$\begin{aligned}
 \int_c \langle f(\vec{s}), d\vec{s} \rangle &= \int_0^1 f(R \cos(2\pi t), R \sin(2\pi t), 3t) \cdot \begin{pmatrix} -2\pi R \sin(2\pi t) \\ 2\pi R \cos(2\pi t) \\ 3 \end{pmatrix} dt \\
 &= \int_0^1 (-2\pi R^2 \cos(2\pi t) \sin(2\pi t) + 2\pi R^2 (\sin(2\pi t) - \cos(2\pi t)) \cos(2\pi t)) dt \\
 &\quad + 3 \int_0^1 (R \cos(2\pi t) + 3t) dt \\
 &= -2\pi R^2 \int_0^1 \cos^2(2\pi t) dt + 3 \int_0^1 (R \cos(2\pi t) + 3t) dt \\
 &= -\pi R^2 + \frac{9}{2}
 \end{aligned}$$

2.2.2 Stammfunktionen

Definition. Ist $f : B \rightarrow \mathbb{R}^n$ stetig auf einer Kugel B , so nennen wir eine skalare Funktion $U : B \rightarrow \mathbb{R}$, welche stetig differenzierbar ist, auch *Stammfunktion* zu f , wenn $\nabla U = f$ gilt.

Wann hat eine vektorwertige Funktion $f : B \rightarrow \mathbb{R}^n$ eine Stammfunktion? Dazu muss eine Bedingung erfüllt sein, ohne die eine Stammfunktion nicht gefunden werden kann.

2.2.2.1 Hilfssatz. *Hat $f : B \rightarrow \mathbb{R}^n$ auf B eine Stammfunktion U , so gilt:*

Ist $c : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg, so gilt

$$\int_c \langle f(\vec{s}), d\vec{s} \rangle = U(c(b)) - U(c(a))$$

Ist also c sogar ein geschlossener Weg ist, also $c(a) = c(b)$, so ist

$$\int_c \langle f(\vec{s}), d\vec{s} \rangle = 0$$

Manchmal schreibt man in diesem Fall auch

$$\oint_c \langle f(\vec{s}), d\vec{s} \rangle = 0$$

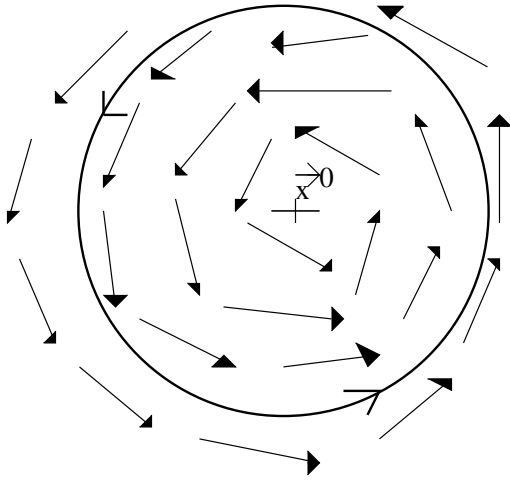
Beweis. Die Kettenregel lehrt, dass

$$U(c(b)) - U(c(a)) = \int_a^b \frac{d(U \circ c)}{dt} dt = \int_a^b \nabla U(c(t)) \cdot \dot{c} dt = \int_c \langle f(\vec{s}), d\vec{s} \rangle$$

□

Wir streben nun an, eine Differentialbedingung an eine vektorwertige Funktion f zu finden, die für Existenz einer Stammfunktion notwendig ist.

Dazu sehen wir uns nun das folgende \mathcal{C}^2 -Richtungsfeld $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ an und wollen f längs eines Kreises c mit Radius R integrieren, der in einer Ebene liegt, die durch die (orthogonalen) Einheitsvektoren \vec{u} und \vec{v} aufgespannt wird und durch einen Punkt \vec{x}^0 geht.



Dazu schreiben wir

$$c(t) = \vec{x}^0 + R(\vec{u} \cos t + \vec{v} \sin t), \quad \dot{c}(t) = R(-\vec{u} \sin t + \vec{v} \cos t)$$

und rechnen aus

$$\begin{aligned} \int_c \langle f(\vec{s}), d\vec{s} \rangle &= \int_0^1 f(c(t)) \cdot \dot{c}(t) dt \\ &= -R \int_0^1 \vec{u} \cdot f(\vec{x}^0 + (R\vec{u} \cos t + R\vec{v} \sin t)) \sin t dt \\ &\quad + R \int_0^1 \vec{v} \cdot f(\vec{x}^0 + (R\vec{u} \cos t + R\vec{v} \sin t)) \cos t dt \end{aligned}$$

Wir nehmen weiter an, dass R klein sei. Dann können wir die Taylorentwicklung von f um \vec{x}^0 verwenden. Es gilt

$$f(\vec{x}^0 + (R\vec{u} \cos t + R\vec{v} \sin t)) = f(\vec{x}^0) + J_f(\vec{x}^0) \cdot (R\vec{u} \cos t + R\vec{v} \sin t)$$

$$+ \sum_{j=1}^n \vec{e}_j \left((R\vec{u} \cos t + R\vec{v} \sin t) \cdot \text{Hes}_{f_j}((1 - \tau_{j,t})\vec{x}^0 + \tau_{j,t}(R\vec{u} \cos t + R\vec{v} \sin t)) (R\vec{u} \cos t + R\vec{v} \sin t) \right)$$

Mit einer Zahlen $\tau_{j,t} \in (0, 1)$. Nun kürzen wir für den Restterm mit $\Phi_2(t)$ ab und sehen, dass

$$\|\Phi_2(t)\| \leq MR^2$$

mit irgendeiner geeigneten Schranke M . Nun setzen wir dies alles in das Wegintegral ein und erhalten

$$\begin{aligned} \int_c \langle f(\vec{s}), d\vec{s} \rangle &= -R^2 \vec{u} \cdot J_f(\vec{x}^0) \vec{v} \int_0^{2\pi} \sin^2 t dt + R^2 \vec{v} \cdot J_f(\vec{x}^0) \vec{u} \int_0^{2\pi} \cos^2 t dt \\ &\quad + \int_0^{2\pi} \Phi_2(t) \cdot (-R\vec{u} \sin t + R\vec{v} \cos t) dt \\ &= \pi R^2 (\vec{v} \cdot J_f(\vec{x}^0) \vec{u} - \vec{u} \cdot J_f(\vec{x}^0) \vec{v}) + \mathcal{E}(\vec{x}^0, R) \end{aligned}$$

mit einem Fehlerterm $E(\vec{x}^0, R)$, der durch

$$|E(\vec{x}^0, R)| \leq M'R^3$$

abzuschätzen ist.

So können wir sehen, dass der Quotient

$$\frac{\int_c \langle f(\vec{s}), d\vec{s} \rangle}{\pi R^2}$$

mit $R \rightarrow 0$ einem Grenzwert zustrebt, den man als lokale "Wirbeldichte" des Feldes bei \vec{x}^0 ansehen kann.

Definition. a) Wir nennen eine Menge $W \subset \mathbb{R}^n$ *offen*, wenn zu jedem Punkt $\vec{x} \in W$ eine Kugel mit Mittelpunkt \vec{x} gefunden werden kann, die noch ganz in W enthalten ist.

b) Ist $f : W \rightarrow \mathbb{R}^n$ ein \mathcal{C}^2 -Richtungsfeld auf einer offenen Menge, dann bezeichnet man als *Rotation* von f in $\vec{x}^0 \in W$ die durch

$$\text{Rot} f(\vec{x}^0)(\vec{u}, \vec{v}) := \vec{v} \cdot J_f(\vec{x}^0) \vec{u} - \vec{u} \cdot J_f(\vec{x}^0) \vec{v}$$

definierte alternierende 2-Form auf \mathbb{R}^n .

Beispiele. a) Wenn $n = 2$, ist

$$\text{Rot} f(\vec{x}^0)(\vec{u}, \vec{v}) = \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \det(\vec{u}, \vec{v})$$

Im Spezialfall $n = 3$ gilt in Koordinaten

$$\text{Rot} f(\vec{x}^0)(\vec{u}, \vec{v}) = \nabla \times f(\vec{x}^0) \cdot (\vec{u} \times \vec{v})$$

wie nun ohne Mühe nachgerechnet werden kann. Dabei bedeutet $\nabla \times f(\vec{x}^0)$ den Vektor

$$\nabla \times f(\vec{x}^0) := \begin{pmatrix} \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \\ \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \end{pmatrix} (\vec{x}^0)$$

Mitunter wird $\nabla \times f(\vec{x}^0)$ mit dem Symbol $\text{rot } f(\vec{x}^0)$ bezeichnet.

Wir können jetzt Richtungsfelder kennzeichnen, die sich als $f = \nabla U$ mit einer Funktion (Potential) U darstellen lassen.

Folgendes wissen wir schon:

2.2.2.2 Hilfssatz. *Ist $f : W \rightarrow \mathbb{R}^n$ ein \mathcal{C}^2 -Richtungsfeld mit Potential U , also $f = \nabla U$ mit einer \mathcal{C}^3 -Funktion, so ist $\int_c \langle f(\vec{s}), d\vec{s} \rangle = 0$ für jeden geschlossenen Weg c in W , ebenso ist $\text{Rot } f(\vec{x}) = 0$ für jedes $\vec{x} \in W$.*

Sind diese Kriterien auch hinreichend für die Existenz eines Potentials?

Dazu sehen wir uns ein Beispiel an. Es sei W ein Kreisring in der Ebene, etwa $W = \{\vec{x} \mid r_1 < \|\vec{x}\| < r_2\}$ mit $r_1 < 1 < r_2$, und

$$f(\vec{x}) = \frac{1}{\|\vec{x}\|^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

Dann ist

$$\frac{\partial f_1}{\partial x_2} = -\frac{1}{\|\vec{x}\|^2} + \frac{2x_2^2}{\|\vec{x}\|^4}, \quad \frac{\partial f_2}{\partial x_1} = \frac{1}{\|\vec{x}\|^2} - \frac{2x_1^2}{\|\vec{x}\|^4}$$

also

$$\text{Rot } f(\vec{x}) = -\frac{2}{\|\vec{x}\|^2} + \frac{2x_2^2}{\|\vec{x}\|^4} + \frac{2x_1^2}{\|\vec{x}\|^4} = 0$$

Aber ist nun $c(t) = (\cos t, \sin t)$, so folgt

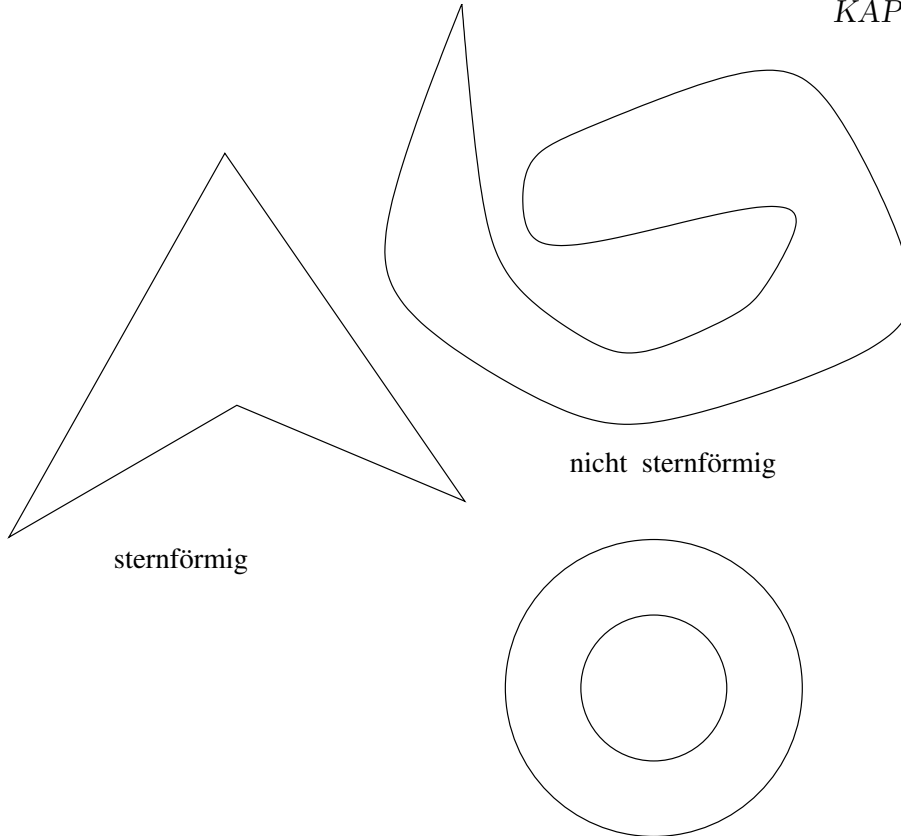
$$\int_c \langle f(\vec{s}), d\vec{s} \rangle = 2\pi$$

Also kann f nicht über ganz W als ∇U mit irgendeiner differenzierbaren Funktion U dargestellt werden.

Ob ein Richtungsfeld f Gradient einer Funktion U ist oder nicht, hängt also nicht nur von f ab, sondern auch vom Definitionsbereich von f .

Definition. Wir nennen eine offene Menge $W \subset \mathbb{R}^n$ sternförmig, wenn ein Punkt $\vec{x}^0 \in W$ derart gefunden werden kann, dass für jeden Punkt $\vec{P} \in W$ die Verbindungsstrecke von \vec{P} nach \vec{x}^0 ganz innerhalb W verläuft. Man nennt dann \vec{x}^0 ein Zentrum für U .

Hier sind Beispiele:



Über einem sternförmigen Gebiet ist jedes Rot-freie Feld Gradient einer Potentialfunktion. Um das zu zeigen benötigen wir ein Hilfsmittel über die Vertauschung von Limesbildung und Integral.

2.2.2.3 Hilfssatz. Angenommen $u : [a, b] \times V \rightarrow \mathbb{R}$ sei eine stetige Funktion.

a) Dann ist auch

$$v(\vec{x}) := \int_a^b u(t, \vec{x}) dt$$

eine auf V stetige Funktion.

b) Sind ferner die 1. partiellen Ableitungen von $u(t, \vec{x})$ bezgl. \vec{x} stetig, so ist auch v stetig differenzierbar, und

$$\frac{\partial v}{\partial x_k}(\vec{x}) = \int_a^b \frac{\partial u}{\partial x_k}(t, \vec{x}) dt$$

für alle $1 \leq k \leq n$.

Nun sind wir bereit für den Beweis des Satzes über die Existenz eines Potentials:

2.2.2.4 Satz. Ist $W \subset \mathbb{R}^n$ ein sternförmiges Gebiet und $f : W \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Richtungsfeld, so gibt es genau dann eine stetig differenzierbare Funktion $U : W \rightarrow \mathbb{R}$ mit $\nabla U = f$, wenn $\text{Rot} f = 0$. Wir erhalten dann eine Stammfunktion U auf folgende Weise: Ist $\vec{x}^0 \in U$ ein Zentrum für U , so setze man

$$U(\vec{x}) := \int_{c_{\vec{x}}} \langle f(\vec{s}), d\vec{s} \rangle,$$

wobei $c_{\vec{x}}$ die Verbindungsstrecke $c_{\vec{x}}(t) := \vec{x}^0 + t(\vec{x} - \vec{x}^0)$, $0 \leq t \leq 1$ von \vec{x}^0 nach \vec{x} bedeutet.

Beweis. Dazu schreiben wir U aus:

$$U(\vec{x}) = \int_0^1 u(t, \vec{x}) dt,$$

mit

$$u(t, \vec{x}) = f(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) \cdot (\vec{x} - \vec{x}^0)$$

Die Voraussetzung $\text{Rot} f = 0$ bedeutet, dass

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}, \quad 1 \leq i, j \leq n$$

Der vorherige Hilfssatz ist auf dieses u anwendbar und ergibt

$$\frac{\partial U}{\partial x_k}(\vec{x}) = \int_0^1 \frac{\partial u}{\partial x_k}(t, \vec{x}) dt$$

Nun errechnen wir aber

$$\begin{aligned} \frac{\partial u}{\partial x_k}(t, \vec{x}) &= \frac{\partial}{\partial x_k} \sum_{j=1}^n f_j(\vec{x}^0 + t(\vec{x} - \vec{x}^0))(x_j - x_j^0) \\ &= \sum_{j=1}^n t \cdot \frac{\partial f_j}{\partial x_k}(\vec{x}^0 + t(\vec{x} - \vec{x}^0))(x_j - x_j^0) + f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) \\ &= t \cdot \sum_{j=1}^n \frac{\partial f_k}{\partial x_j}(\vec{x}^0 + t(\vec{x} - \vec{x}^0))(x_j - x_j^0) + f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) \\ &= t \cdot \frac{d}{dt} f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) + f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) \\ &= \frac{d}{dt} (t \cdot f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0))) \end{aligned}$$

Das führt auf

$$\frac{\partial U}{\partial x_k}(\vec{x}) = \int_0^1 \frac{\partial u}{\partial x_k}(t, \vec{x}) dt = \int_0^1 \frac{d}{dt} (t \cdot f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0))) dt = t \cdot f_k(\vec{x}^0 + t(\vec{x} - \vec{x}^0)) \Big|_0^1 = f_k(\vec{x})$$

wie gewünscht. □

Kapitel 3

Grundlagen der Statistik

3.1 Elementare Stochastik

3.1.1 Einführung: Zufallsexperimente und Wahrscheinlichkeiten

Ziel der Vorlesung soll es sein, eine Einführung in die Methoden der *Mathematischen Statistik* zu geben. Das ist die Sparte der Angewandten Mathematik, welche Hilfsmittel entwickelt, die es ermöglichen, auf Grund einer Stichprobe Aussagen über die Parameter zu machen, welche eine Gesamtheit von Dingen kennzeichnen.

Beispiele: 1) *Qualitätskontrolle.* Angenommen, man will eine Lieferung von 10000 Transistoren auf ihre Qualität hin überprüfen und mit dem Hersteller ist eine gewisse Quote von defekten Transistoren vertraglich vereinbart. Da man sicher nicht alle Transistoren überprüfen kann, wird man eine Stichprobe von einem gewissen Umfang, etwa 50 Stück, entnehmen und testen.

Man wird dann irgendeine Ausschussquote in der Stichprobe beobachten und sie mit der vom Hersteller angegebenen vergleichen. Ist sie höher als diese, wird man erwägen, diese Lieferung abzulehnen.

2) *Benzinverbrauch eines Autos.* Angenommen, ein Autotyp A wird angeboten, von dem der Hersteller versichert, dass der Benzinverbrauch nicht höher als 6.4 Liter auf 100 km sei. Man ermittelt den Benzinverbrauch bei einer Zahl von Wagen und berechnet den mittleren Benzinverbrauch. Ist der Wert höher als 6.4, so stellt man die Herstellerangabe in Frage. Inwieweit ist das gerechtfertigt?

3) Bei der Produktion von Zylinderscheiben soll überprüft werden, ob ein Sollwert für deren Durchmesser, etwa 20.2mm, eingehalten wird. Angenommen, man entnimmt aus der laufenden Produktion 16 Stück und stellt fest, dass der mittlere Durchmesser 20.6 mm beträgt, wobei ein Toleranzspanne von 0.5 mm eingeräumt wird. Kann man jetzt mit genügender Sicherheit sagen, dass der Sollwert eingehalten wurde?

Diese Beispiele haben eines gemeinsam: Stets wird zu einer Hypothese (Vermutung oder Behauptung über die Eingrenzung eines technischen Parameters) diskutiert, ob man sie auf der Grundlage einer Stichprobe annehmen oder verwerfen will. Abweichungen des untersuchten Parameters von einem "Sollwert" können verschiedene Ursachen haben, etwa

- "Grobe" Fehler, das sind Fehler bei der Auswertung der Stichproben die durch menschliche Unzulänglichkeiten entstehen, wie etwa falsche Messmethoden, defekte Messgeräte, und ähnliches,
- Abhängigkeit des Prüfergebnisses von der jeweiligen Stichprobe: Die Überprüfung einer anderen Stichprobe hätte vielleicht günstiger ausfallen können

Man ist daher an einer Entscheidungshilfe interessiert, also an einem Kriterium dafür, wann man eine Hypothese annehmen soll oder nicht, wobei zusätzlich die Wahrscheinlichkeit einer Fehlentscheidung (Risiko) unterhalb einer sinnvollen Grenze (in der Regel 5%) bleiben sollte.

Ähnliche Überlegungen werden angestellt, wenn man Prognosen über die Lebensdauer einer Autokarosserie, die Anzahl der bei einer Tankstelle ankommenden Autos oder die Anzahl der weltweit pro Monat stattfindenden Havarien von Hochseeschiffen abgeben will.

Die Mathematische Statistik bedient sich der mathematischen Hilfsmittel aus der *Wahrscheinlichkeitstheorie*, welche nach Gesetzmäßigkeiten sucht, denen das Eintreten von solchen Ereignissen unterliegt, die vom Zufall regiert werden und nicht von einem Naturgesetz, wie etwa in der Physik oder Chemie.

Definition. Unter einem *Zufallsexperiment* verstehen wir jedes Experiment, dessen Ausgang nicht durch Naturgesetze vorherbestimmt ist, sondern nur vom Zufall abhängt.

Beispiele hierfür sind:

- Das Werfen eines Würfels oder einer Münze
- Eine Ziehung der Lottozahlen
- Die Entnahme einer Stichprobe aus einer Lieferung von Werkstücken (Glühbirnen und ähnliches)
- Testen der Reißlast bei Drahtseilen
- Registrieren der Lebensdauern von Leuchtstoffröhren
- Umfragen unter den Wählern innerhalb eines Wahlkreises

Kein Zufallsexperiment liegt vor, wenn etwa Eisensulfat mit Salzsäure versetzt wird, denn was herauskommt, weiß man vorher: Es wird H_2S (Schwefelwasserstoff) freigesetzt.

Ebenso ist vorherbestimmt, was geschieht, wenn man ein Pendel anstößt: Es beginnt zu schwingen, und man kann seine Bewegung vorausberechnen.

Man möchte quantitativ erfassen, was man für den Ausgang eines Zufallsexperimentes zu erwarten hat. Dazu ordnet man dem Zufallsexperiment einen Ergebnisraum zu:

Definition. Der *Ergebnisraum* eines Zufallsexperimentes ist definiert als die Menge Ω aller möglichen Ausgänge, die das Zufallsexperiment nehmen kann. Die Elemente von Ω heißen auch *Elementarereignisse*.

Beispiele: a) Würfeln wir mit 2 Würfeln, so ist der Ergebnisraum gegeben durch

$$\Omega := \{(a, b) \mid a, b \in \{1, 2, 3, 4, 5, 6\}\}$$

b) Angenommen, ein Kaufhaus führe 3 Typen von Taschenrechnern und man beobachte bei 4 Kunden, welchen der Typen sie kaufen. Dann ist der Ergebnisraum hierzu

$$\Omega := \{(a, b, c, d) \mid a, b, c \in T\},$$

wobei $T = \{t_1, t_2, t_3\}$ ist und t_1, t_2, t_3 für die 3 Typen von Taschenrechnern steht.

c) Registrieren wir die weltweit pro Monat stattgefundenen Havarien von Hochseeschiffen, so ist \mathbb{N}_0 der geeignete Ergebnisraum.

d) Messen wir die Lebensdauer einer Leuchtstoffröhre einer bestimmten Marke, so wird \mathbb{R}^+ der Ergebnisraum werden.

Bei einem Zufallsexperiment interessieren mehr als die Elementarereignisse die Teilmengen des Ergebnisraumes. Beim Würfeln mit 2 Würfeln etwa interessiert ja vor allem die Summe der Augenzahlen und weniger, welche Augenzahl auf den einzelnen Würfeln steht. So beschreibt die Teilmenge

$$A := \{(a, b) \in \Omega \mid a + b = 10\}$$

das Ereignis, dass man eine 10 geworfen hat.

Definition. Ist Ω der Ergebnisraum eines Zufallsexperimentes, so ist bezeichnen wir eine Teilmenge von Ω als *Ereignis*.

Mit Hilfe des Teilmengenkalküls können wir Ereignisse miteinander kombinieren:

- Sind A und B Ereignisse, so bedeutet $A \cap B$ das Ereignis, dass das Ergebnis des Zufallsexperimentes sowohl in A als auch in B liegt, oder einfacher: Die Menge $A \cap B$ steht für das gleichzeitige Eintreten der Ereignisse A und B .
- Sind A und B Ereignisse, so steht $A \cup B$ für das Ereignis, dass das Ergebnis des Zufallsexperimentes in einer der Mengen A oder B liegt. Wenn $A \cap B = \emptyset$, nennen wir A und B *unvereinbar miteinander*. Das gleichzeitige Eintreten von A und B ist nicht möglich.
- Ω repräsentiert das sichere Ereignis, \emptyset das nie auftretende Ereignis.
- Ist $A \subset \Omega$, so steht $A^c := \Omega \setminus A$ dafür, dass A nicht eingetreten ist.

Hierzu sehen wir uns *Beispiele* an:

(i) Das Kaufhaus mit den 3 Typen von Taschenrechnern. Das Ereignis, dass ein Taschenrechner unter den 4 Kunden keinen Käufer gefunden hat, ist

$$A := A_1 \cup A_2 \cup A_3$$

wobei

$$A_1 := \{(a, b, c, d) \in \Omega \mid a, b, c, d \in \{t_2, t_3\}\}$$

$$A_2 := \{(a, b, c, d) \in \Omega \mid a, b, c, d \in \{t_1, t_3\}\}$$

$$A_3 := \{(a, b, c, d) \in \Omega \mid a, b, c, d \in \{t_1, t_2\}\}$$

(ii) Lebensdauer von Leuchtstoffröhren. Beobachtet man, dass eine Leuchtstoffröhre zwischen 8 und 10 Monaten funktioniert hat, so ist dies durch das Intervall $[8, 10]$ als Ereignis zu beschreiben.

(iii) Werfen mit 2 Würfeln. Das Ereignis

$$A := \{(a, b) \mid a, b \leq 5\}$$

beschreibt das Resultat, dass die geworfene Augenzahl höchstens 10 war.

3.1.2 Wahrscheinlichkeitsverteilungen auf diskreten Ergebnismengen

Gegeben sei ein Zufallsexperiment mit Ergebnisraum Ω . Mit $\mathfrak{P}(\Omega)$ bezeichnen wir die Menge aller Teilmengen von Ω .

Wir nehmen jetzt für das Folgende an, dass Ω abzählbar sei:

$$\Omega = \{x_1, x_2, x_3, x_4, \dots\}$$

(Zum Beispiel: $\Omega = \{x_1, \dots, x_N\}$: endliche Menge, oder $\Omega = \mathbb{Z}, \mathbb{N}$ oder \mathbb{Q} . Nicht zugelassen werden etwa: $\Omega = \mathbb{R}$ oder $\Omega = (0, \infty)$, $\Omega = (0, 1]$).

Definition. Unter einem *Wahrscheinlichkeitsmaß* (auch kurz *W-Maß*) auf Ω verstehen wir eine Funktion $P : \mathfrak{P}(\Omega) \rightarrow [0, 1]$, so dass folgende Regeln erfüllt werden:

- $P(\emptyset) = 0$, $P(\Omega) = 1$,
- Sind A_1, A_2, A_3, \dots abzählbar viel Ereignisse, die paarweise unvereinbar sind, so ist

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

Es gilt für nicht notwendig unvereinbare Ereignisse:

3.1.2.1 Hilfssatz. Sei $P : \mathfrak{P}(\Omega) \rightarrow [0, 1]$ ein W-Maß. Dann gilt für zwei Mengen $A, B \subset \Omega$:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ist $A \subset \Omega$, so gilt $P(A^c) = 1 - P(A)$.

Beweis. Denn sei $C := A \cup B \setminus (A \cap B) = A_1 \cup A_2$, wobei $A_1 = A \setminus (A \cap B)$ und $A_2 = B \setminus (A \cap B)$. Dann ist $A \cup B = C \cup (A \cap B)$ und C ist mit $(A \cap B)$ unvereinbar. Also

$$P(A \cup B) = P(C) + P(A \cap B)$$

Aber auch A_1 und A_2 sind miteinander unvereinbar, also

$$P(C) = P(A_1) + P(A_2)$$

Da $A = A_1 \cup (A \cap B)$ und $B = A_2 \cup (A \cap B)$, wobei dies Zerlegungen in unvereinbare Ereignisse sind, haben wir weiter

$$P(A) = P(A_1) + P(A \cap B), \quad P(B) = P(A_2) + P(A \cap B)$$

Also

$$\begin{aligned} P(A \cup B) &= P(C) + P(A \cap B) \\ &= P(A_1) + P(A_2) + P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Die zweite Behauptung folgt aus $A \cup A^c = \Omega$, wobei A und A^c natürlich miteinander nicht vereinbar sind. □

Auf $\mathfrak{P}(\Omega)$ können wir auf folgende Weise ein W-Maß erklären:

- Bei endlichem Ω (mit N Elementen) wähle man Zahlen $p_1, \dots, p_N \geq 0$, so dass $p_1 + p_2 + \dots + p_N = 1$,
- Bei abzählbar unendlichem Ω wähle man Zahlen $p_j \geq 0$, so dass die Reihe $\sum_{j=1}^{\infty} p_j$ konvergiert und

$$\sum_{j=1}^{\infty} p_j = 1$$

- Für $A \subset \Omega$ setze man

$$P(A) := \sum_{j: x_j \in A} p_j$$

Insbesondere ist dann $P(\{x_j\}) = p_j$ für alle $j \geq 1$.

So erhalten wir sogar alle Wahrscheinlichkeitsmaße auf $\mathfrak{P}(\Omega)$.

A) Die Laplaceverteilung

Nun ist Ω endlich (etwa mit N Elementen), und alle p_j haben den Wert $p_j = 1/N$. Dann ergibt

$$P(A) = \frac{|A|}{N}$$

die relative Anzahl der Elemente von A im Vergleich zu Ω an. (Hier steht $|A|$ für die Anzahl der Elemente von A).

Zufallsexperimente, zu denen ein W -raum dieses Typs passt, heißen auch *Laplace-Experimente*.

Beispiele. (1) *Werfen mit einem Würfel.* Nun ist $\Omega = \{1, 2, 3, 4, 5, 6\}$ der adäquate Ergebnisraum und $P(A) = |A|/6$ für jedes Ereignis A . Die W 'keit, eine gerade Zahl zu werfen ist also $1/2$.

(2) *Werfen mit 3 Würfeln.* Jetzt wird $\Omega = \{(a, b, c) \mid a, b, c \in \{1, 2, 3, 4, 5, 6\}\}$. Das Ereignis, dass man eine 10 geworfen hat, wird nun beschrieben durch

$$A = \{(a, b, c) \in \Omega \mid a + b + c = 10\}$$

Dann wird

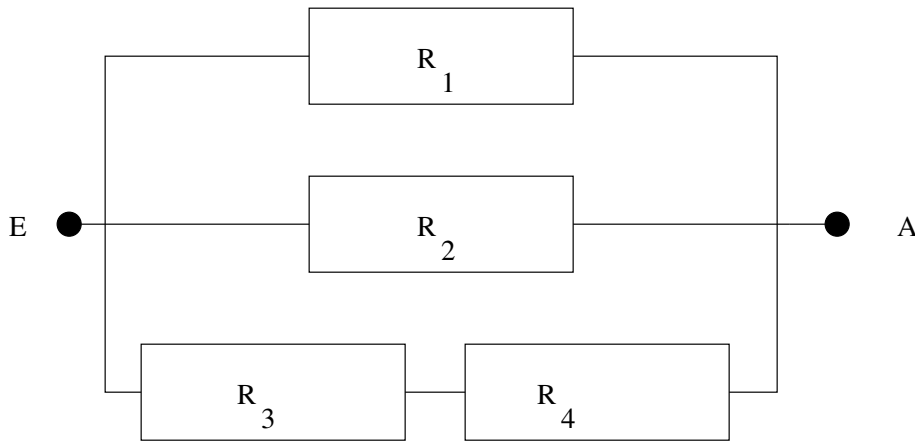
$$\begin{aligned} A = & \{(1, b, c) \mid (b, c) \in \{(3, 6), (4, 5), (6, 3), (5, 4)\}\} \\ & \cup \{(2, b, c) \mid (b, c) \in \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}\} \\ & \cup \{(3, b, c) \mid (b, c) \in \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}\} \\ & \cup \{(4, b, c) \mid (b, c) \in \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}\} \\ & \cup \{(5, b, c) \mid (b, c) \in \{(1, 4), (2, 3), (3, 2), (4, 1)\}\} \\ & \cup \{(6, b, c) \mid (b, c) \in \{(1, 3), (2, 2), (3, 1)\}\} \end{aligned}$$

Das Ereignis, eine 9 zu werfen, ist zu beschreiben durch die Menge

$$\begin{aligned} B = & \{(a, b, c) \in \Omega \mid a + b + c = 9\} \\ & \cup \{(1, b, c) \mid (b, c) \in \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}\} \\ & \cup \{(2, b, c) \mid (b, c) \in \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}\} \\ & \cup \{(3, b, c) \mid (b, c) \in \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}\} \\ & \cup \{(4, b, c) \mid (b, c) \in \{(1, 4), (2, 3), (3, 2), (4, 1)\}\} \\ & \cup \{(5, b, c) \mid (b, c) \in \{(1, 3), (2, 2), (3, 1)\}\} \\ & \cup \{(6, b, c) \mid (b, c) \in \{(1, 2), (2, 1)\}\} \end{aligned}$$

Wir sehen, dass $|A| = 27$, $|B| = 25$, somit $P(A) = 0.125$, $P(B) = 0.12$.

3) *Geschlossene Relais.* Angenommen, ein Schaltkreis (siehe Bild) bestehe aus 4 Relais R_1, \dots, R_4 und jedes sei mit W'keit 0.5 geschlossen, bzw. offen. Mit welcher W'keit führt ein Weg vom Eingang E des Schaltkreises bis zu seinem Ausgang A nur über geschlossene Relais ?



Der Ergebnisraum ist nun

$$\Omega = \{(a_1, a_2, a_3, a_4) \mid a_1, \dots, a_4 \in \{0, 1\}\}$$

Das Elementarereignis $(1, 0, 1, 0)$ bedeutet zum Beispiel: Die Relais R_1 und R_3 sind offen, und R_2 und R_4 sind geschlossen. Sei jetzt S_j das Ereignis, dass R_j geschlossen ist. Das Ereignis S , dass ein Weg vom Eingang E des Schaltkreises bis zu seinem Ausgang A nur über geschlossene Relais führt, wird dann durch die Menge

$$S = S_1 \cup S_2 \cup (S_3 \cap S_4)$$

gegeben. Wir errechnen nun

$$\begin{aligned} P(S) &= P(S_1 \cup S_2) + P(S_3 \cap S_4) - P((S_1 \cup S_2) \cap (S_3 \cap S_4)) \\ &= P(S_1) + P(S_2) - P(S_1 \cap S_2) + P(S_3 \cap S_4) - P((S_1 \cap S_3 \cap S_4) \cup (S_2 \cap S_3 \cap S_4)) \\ &= P(S_1) + P(S_2) - P(S_1 \cap S_2) + P(S_3 \cap S_4) \\ &\quad - (P(S_1 \cap S_3 \cap S_4) + P(S_2 \cap S_3 \cap S_4) - P(S_1 \cap S_2 \cap S_3 \cap S_4)) \end{aligned}$$

Jedes Elementarereignis hat W'keit $1/16$. Nun ist aber $P(S_j) = 1/2$ für $j = 1, 2, 3, 4$, ferner $P(S_1 \cap S_2) = P(S_3 \cap S_4) = 1/4$ und $P(S_1 \cap S_3 \cap S_4) = P(S_2 \cap S_3 \cap S_4) = 1/8$. Schließlich ist $P(S_1 \cap S_2 \cap S_3 \cap S_4) = 1/16$. Das ergibt zusammen

$$P(S) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} + \frac{1}{4} - \left(\frac{1}{8} + \frac{1}{8} - \frac{1}{16}\right) = \frac{13}{16} = 0.81$$

2) Binomialverteilung

Ist $n \in \mathbb{Z}^+$ und $0 < p < 1$, so definieren wir für die Menge $\Omega = \{0, 1, 2, 3, \dots, n\}$ die W'erteilung

$$P(\{j\}) = \binom{n}{j} p^j (1-p)^{n-j}$$

für $j = 0, 1, 2, \dots, n$. Dann ist

$$P(\Omega) = \sum_{j=0}^n \binom{n}{j} p^j (1-p)^{n-j} = (p + 1 - p)^n = 1$$

Man nennt diese W'erteilung auch die Binomialverteilung mit Parametern n und p , oder auch $\mathcal{B}_{n,p}$ -Verteilung.

Wenn man ein Zufallsexperiment mit 2 möglichen Ergebnissen e_1, e_2 eine Anzahl n -mal durchführt (wobei die Wiederholungen unabhängig voneinander sein sollen), und tritt das Einzelergebnis e_1 mit W'keit p auf, so gibt die $\mathcal{B}_{n,p}$ -Verteilung an, mit welcher W'keit es bei den Wiederholungen k -mal auftritt.

Beispiele. 1) *n*-maliges Werfen einer Münze. Wenn man eine Münze wirft, bei der mit W'keit p Kopf (K) herauskommt und mit W'keit $1 - p$ das Resultat Zahl (Z), so ist bei n -maligem Werfen der Münze die W'keit dafür, dass j -mal das Resultat K auftritt, durch

$$\binom{n}{j} p^j (1-p)^{n-j}$$

zu berechnen.

2) *Stichprobe.* Angenommen, man führt bei einer große Lieferung von Transistoren eine Qualitätskontrolle durch. Der Hersteller gibt die Ausschussquote mit 2% an. Um dies zu überprüfen nimmt man eine Stichprobe, etwa vom Umfang $n = 100$. Die Lieferung wird abgelehnt, wenn mehr als 4 defekte Transistoren in der Stichprobe entdeckt werden. Die Entnahme der Stichprobe ist formal ein Zufallsexperiment mit einer $\mathcal{B}_{n,p}$ -verteilten Zahl von defekten Transistoren, wobei $n = 100, p = 0.02$ ist. Die W'keit, j defekte Transistoren zu finden, ist dann

$$p_j = \binom{100}{j} 0.02^j (0.98)^{n-j}$$

Wir finden: ($p_j := P(\{j\})$):

$$p_0 = 0.98^{100} = 0.1326$$

$$p_1 = \binom{100}{1} 0.02(0.98)^{99} = 0.2706$$

$$p_2 = \binom{100}{2} 0.02^2(0.98)^{98} = 0.2734$$

$$p_3 = \binom{100}{3} 0.02^3(0.98)^{97} = 0.1822$$

$$p_4 = \binom{100}{4} 0.02^4(0.98)^{96} = 0.0902$$

Die W'keit, dass nicht mehr als 4 Transistoren defekt sind, ist also

$$w = p_0 + p_1 + p_2 + p_3 + p_4 = 0.9492$$

Das Produzentenrisiko, also die W'keit, dass die Lieferung abgelehnt wird, ist $1 - w = 0.051 = 5.1\%$.

3) Hypergeometrische Verteilung

Bei Stichproben im Rahmen einer Qualitätskontrolle gewinnt die folgende Wahrscheinlichkeitsverteilung Bedeutung:

Angenommen, eine Lieferung von N Bauteilen enthalte $d \leq N$ defekte Bauteile. Der Empfänger der Lieferung nimmt eine Stichprobe von n Bauteilen. Mit welcher Wahrscheinlichkeit enthält die Stichprobe k defekte Bauteile?

Man arbeitet nun mit dem Ergebnisraum

$$\Omega = \{A \subset \{1, \dots, N\} : |A| = n\}$$

Wir nummerieren die Bauteile so, dass die ersten d gerade die defekten sind und setzen $D := \{1, \dots, d\}$. Das Ereignis, k defekte Bauteile in der Stichprobe zu finden, wird durch die Menge

$$K := \{A_1 \cup A_2 \mid A_1 \subset D, A_2 \subset \{1, 2, \dots, n\} \setminus D, |A_1| = k, |A_2| = n - k\}$$

beschrieben. Wir zählen ab, wieviele Elemente K hat und erhalten

$$|K| = \binom{d}{k} \binom{N-d}{n-k}.$$

So führt die Gleichverteilung auf Ω auf eine neue Verteilung auf $\mathcal{X} := \{0, 1, 2, \dots, N\}$, nämlich

$$\mathcal{H}_{N,d,n}(\{k\}) := \frac{\binom{d}{k} \binom{N-d}{n-k}}{\binom{N}{n}}$$

und

$$\mathcal{H}_{N,d,n}(A) = \sum_{k:k \in A} \mathcal{H}_{N,d,n}(\{k\}),$$

für alle $A \in \mathfrak{P}(\mathcal{X})$.

(Hypergeometrische Verteilung mit Parameter N, d und n , wobei $n \leq N$)

Dabei ist $\mathcal{H}_{N,d,n}(\{k\}) = 0$, wenn $k > \min\{d, n\}$.

Man beachte, dass

$$\sum_{k=0}^{\min\{d,n\}} \binom{d}{k} \binom{N-d}{n-k} = \binom{N}{n}$$

Denn setzen wir $f(x) = (1+x)^d$ und $g(x) = (1+x)^{N-d}$, so liefert der Binomialsatz, dass

$$f(x) = \sum_{i=0}^d \binom{d}{i} x^i, \quad g(x) = \sum_{j=0}^{N-d} \binom{N-d}{j} x^j$$

Ausmultiplizieren ergibt dann

$$\sum_{s=0}^N \binom{N}{n} x^s = (1+x)^N = f(x)g(x) = \sum_{s=0}^N \left(\sum_{k=0}^{\min\{d,n\}} \binom{d}{k} \binom{N-d}{n-k} \right) x^s$$

Die Behauptung ergibt sich durch Koeffizientenvergleich.

Ein kleines Anwendungsbeispiel: Ein Unternehmer bezieht von einer Firma gewisse Bauteile. Angenommen, man habe vertraglich vereinbart, dass eine Ausschussquote von 10% akzeptiert wird. Zum Zweck der Überprüfung einer Lieferung von 50 Stück entnimmt der Unternehmer eine Stichprobe von 5 Teilen in der Absicht, die Lieferung abzulehnen, wenn mehr als 1 defektes Teil in der Stichprobe vorkommt. Mit welcher Wahrscheinlichkeit würde eine Lieferung mit 8 defekten Bauteilen abgelehnt?

Die Anzahl der defekten Bauteile ist verteilt nach $\mathcal{H}_{50,8,5}$. Das zur Ablehnung führende Ereignis ist also hier die Menge $A := \{0, 1\}^c$. Nun gilt

$$\mathcal{H}_{50,8,5}(\{0, 1\}) = \frac{\binom{8}{0} \binom{42}{5} + \binom{8}{1} \binom{42}{4}}{\binom{50}{5}} = \frac{850668 + 8 \cdot 111930}{2118760} = \frac{1746108}{2118760} = 0.82$$

Also

$$P(A) = 1 - 0.82 = 0.18$$

Die Wahrscheinlichkeit, aufgrund der Stichprobe die Lieferung abzulehnen, ist nur 0.18.

Die folgende Frage stellt sich: Formal ist dieses Beispiel genauso gebaut wie das vorherige Beispiel der Qualitätskontrolle bei Transistoren. Dennoch wird jetzt die hypergeometrische Verteilung angewendet. Warum arbeiten wir nicht mit der Binomialverteilung mit Parametern $n = 5, p = 0.1$?

Der Unterschied zum Beispiel mit den Transistoren besteht in dem Umfang der Gesamtheit. Beim Beispiel mit den Transistoren ist $n = 100$ und die Gesamtheit sollte einen Umfang $N \gg n$ haben. Im zweiten Beispiel sind die Zahlen viel kleiner: $N = 50, n = 5$. Erst, wenn $N \gg n$, kann man zu Recht davon ausgehen, dass bei einer Stichprobe sinnvollen Umfangs der Anteil an defekten Teilen der Relation

$$\frac{k}{n} \approx \frac{d}{N}$$

genügt.

In diesem Fall hat man folgende Approximation

3.1.2.2 Hilfssatz. Wenn man annimmt, es sei $N \gg n$, so wird die hypergeometrische Verteilung $\mathcal{H}_{N,d,n}$ durch eine Binomialverteilung mit Parametern n und $p = d/N$ approximiert:

$$\mathcal{H}_{N,d,n}(\{k\}) \approx \binom{n}{k} \left(\frac{d}{N}\right)^k \left(1 - \frac{d}{N}\right)^{n-k}$$

Beweis. Das rechnen wir einfach nach:

$$\begin{aligned} \mathcal{H}_{N,d,n}(\{k\}) &= \frac{\binom{d}{k} \binom{N-d}{n-k}}{\binom{N}{n}} \\ &= \frac{d!(N-d)!(N-n)!n!}{k!(d-k)!(n-k)!(N-d-(n-k))!N!} \\ &= \binom{n}{k} \frac{d!(N-d)!(N-n)!}{(d-k)!(N-d-(n-k))!N!} \\ &= \binom{n}{k} \frac{d(d-1)\cdots(d-k+1)}{N(N-1)\cdots(N-n+1)} \frac{(N-d)!}{(N-d-(n-k))!} \\ &= \binom{n}{k} \frac{d(d-1)\cdots(d-k+1)}{N(N-1)\cdots(N-n+1)} ((N-d)\cdots(N-d-(n-k)+1)) \\ &= \binom{n}{k} \left(\frac{d}{N}\right)^k \frac{(1-\frac{1}{N})\cdots(1-\frac{k-1}{N})}{(1-\frac{1}{N})\cdots(1-\frac{n-1}{N})} ((N-d)\cdots(N-d-(n-k)+1)) \\ &= \binom{n}{k} \left(\frac{d}{N}\right)^k \frac{(1-\frac{1}{N})\cdots(1-\frac{k-1}{N})}{(1-\frac{1}{N})\cdots(1-\frac{n-1}{N})} \left(1-\frac{d}{N}\right)\cdots\left(1-\frac{d}{N}-\frac{n-k-1}{N}\right) \end{aligned}$$

In diesem Ausdruck dürfen wir die Approximation

$$\frac{(1-\frac{1}{N})\cdots(1-\frac{k-1}{N})}{(1-\frac{1}{N})\cdots(1-\frac{n-1}{N})} \approx 1$$

vornehmen, denn $\frac{n}{N} \approx \frac{k}{d}$ ist, wie die Erfahrung lehrt, gerechtfertigt, also wird $\frac{k}{d} \approx \frac{n}{N} \ll 1$. Das führt auf

$$\mathcal{H}_{N,d,n}(\{k\}) \approx \binom{n}{k} \left(\frac{d}{N}\right)^k \left(1-\frac{d}{N}\right)\cdots\left(1-\frac{d}{N}-\frac{n-k-1}{N}\right)$$

Aus einem ähnlichen Grunde haben wir

$$\left(1-\frac{d}{N}\right)\cdots\left(1-\frac{d}{N}-\frac{n-k-1}{N}\right) \approx \left(1-\frac{d}{N}\right)^{n-k}$$

Das beweist die Behauptung. □

4) Poissonverteilung

Nun nehmen wir an, dem Zufallsexperiment ist die Menge IN_0 angemessen. Dann sagen wir, die Ergebnisse des Zufallsexperimentes seien *poissonverteilt* mit Parameter $\lambda > 0$, wenn die W'keit, dass die Zahl k als Ereignis auftritt gegeben ist durch

$$P(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Beispiele: a) Die Anzahl der

- weltweit pro Monat stattfindenden Havarien von Hochseeschiffen,
- bei einer Tankstelle ankommenden Autos,
- in einer Telefonzentrale eingehenden Anrufe,
- der in der preußischen Armee an den Folgen eines Pferdehuftrittes gestorbenen Soldaten

ist poissonverteilt.

Die Poissonverteilung ist unter gewissen Voraussetzungen geeignet, die Binomial- und auch die hypergeometrische Verteilung zu approximieren.

3.1.2.3 Hilfssatz. Die Binomialverteilung mit Parametern n, p lässt sich näherungsweise durch eine Poissonverteilung mit Parameter $\lambda = np$ beschreiben, wenn nur $n \gg 1$ und $p \ll 1$:

$$P(\{k\}) \approx \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{wenn } n - k \gg 1$$

Beweis. Dazu benutzen wir die *Stirlingsche Formel*:

$$n! \approx \sqrt{2\pi n} (n/e)^n$$

Dann wird für $n - k \gg 1$:

$$\begin{aligned} \binom{n}{k} p^k (1-p)^{n-k} &= \frac{n!}{(n-k)! k!} p^k (1-p)^{n-k} \\ &\approx \frac{\sqrt{2\pi n}}{\sqrt{2\pi(n-k)}} e^{-k} \left(\frac{n}{n-k}\right)^{n-k} \frac{(np)^k}{k!} \left(1 - \frac{np}{n}\right)^{n-k} \\ &\approx \frac{\sqrt{n}}{\sqrt{n-k}} e^{-k} \left(1 + \frac{k}{n-k}\right)^{n-k} \frac{(np)^k}{k!} \left(1 - \frac{np}{n}\right)^{n-k} \\ &\approx \frac{\sqrt{n}}{\sqrt{n-k}} e^{-k} \left(1 + \frac{k}{n-k}\right)^{n-k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &\approx \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

denn unter den obigen Voraussetzungen wird

$$\frac{\sqrt{n}}{\sqrt{n-k}} e^{-k} \left(1 + \frac{k}{n-k}\right)^{n-k} \approx 1, \quad \left(1 - \frac{\lambda}{n}\right)^{n-k} \approx e^{-\lambda}.$$

□

Vergleich beider Verteilungen

Angenommen, $p = 0.01$ und $n = 1000$. Wir schreiben: $p_k = \binom{n}{k} p^k (1-p)^{n-k}$ und $q_k = \frac{(1000p)^k}{k!} e^{-1000p}$.

Dann wird

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|-----------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| p_k/q_k | 0.96 | 0.96 | 0.97 | 0.98 | 0.98 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

5) Geometrische Verteilung

Wieder ist der Ergebnisraum \mathbb{Z}^+ . Für eine Zahl $0 < p < 1$ ist nun

$$P(\{k\}) = p(1-p)^{k-1}$$

Diese Verteilung tritt in folgender Situation auf:

Angenommen, man testet wieder eine Lieferung von Bauteilen und man weiß dass ein einzelnes Bauteil mit W'keit p defekt ist. Dann gibt $p(1-p)^{k-1}$ die W'keit dafür an, dass das k -te Bauteil bei der Überprüfung das erste defekte ist.

3.1.3 Einschub: Aus der Kombinatorik

Um nachvollziehen zu können, warum die Ergebnisse beim wiederholten Werfen einer Münze einer Binomialverteilung unterliegen, erinnern wir uns an ein paar Ergebnisse aus der Kombinatorik.

Vertauschungen endlich vieler Elemente

3.1.3.1 Hilfssatz. Angenommen, eine Menge S_n habe n -Elemente, also

$$S_n = \{x_1, \dots, x_n\}$$

Dann lassen sich diese Elemente auf $n!$ Weisen anordnen.

Beispiel. Wenn etwa $n = 3$, so haben wir folgende Möglichkeiten:

$$(x_1, x_2, x_3), (x_1, x_3, x_2), (x_2, x_1, x_3) \\ (x_2, x_3, x_1), (x_3, x_1, x_2), (x_3, x_2, x_1)$$

Das sind $6 = 3!$ an der Zahl.

Beweis. Es bietet sich ein Induktionsargument an. Für $n = 1$ ist nichts zu tun. Gilt die Behauptung für $n - 1$, so auch für n . Denn ist M_j die Menge der Vertauschungen der Elemente von S_n , bei denen x_j an die n .te Stelle kommt, so sind diese Vertauschungen nur an einer Stelle festgelegt; um die anderen Elemente, also $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n$ auf die ersten $n - 1$ Stellen zu verteilen, bestehen nach Induktionsannahme $(n - 1)!$ Möglichkeiten. Die Mengen M_j sind paarweise disjunkt, und es gilt

$$S_n = M_1 \cup M_2 \cup \dots \cup M_n$$

Da jede einzelne Menge $(n - 1)!$ Elemente hat, finden wir insgesamt genau $n \cdot (n - 1)! = n!$ Elemente in S_n .

3.1.3.2 Hilfssatz. *Wenn sich in einer Urne n unterscheidbare Kugeln befinden und man zieht nacheinander $k \leq n$ von ihnen, so ergeben sich eine Anzahl von möglichen Resultaten für diese Ziehung:*

- a) *Wenn wir jede gezogene Kugel wieder zurücklegen, ergeben sich n^k mögliche Ergebnisse,*
- b) *wenn wir die gezogenen Kugeln nicht zurücklegen, haben wir $n(n-1) \cdot \dots \cdot (n-k+1) = \frac{n!}{(n-k)!}$ mögliche Ergebnisse.*

Dabei ist die Reihenfolge, in der die Kugeln gezogen werden, signifikant.

Beweis. Wir numerieren die Kugeln mit $1, 2, 3, \dots, n$ (vergleiche Lotto). Im Falle (a) bestehen für jede Kugel n Möglichkeiten, da man jede Kugel zurücklegt. Außerdem hat es keinen Einfluss auf die Ziehung jeder Kugel, welche Kugeln vorher gezogen worden sind. Jedes Ziehungsergebnis ist dann zu beschreiben durch einen k -komponentigen Vektor mit Einträgen in der Menge $\{1, 2, \dots, n\}$. So finden wir insgesamt n^k Möglichkeiten.

Im Falle (b) gibt es n Möglichkeiten die erste Kugel zu ziehen, $(n - 1)$ Möglichkeiten, die zweite zu ziehen, da ja die erstgezogene Kugel nicht wieder zurückgelegt wird. Allgemein bleiben für die ℓ -te Kugel $n - (\ell - 1)$ Möglichkeiten. All diese Möglichkeiten bestehen unabhängig voneinander, so dass man zur Auffindung der Gesamtzahl aller möglichen Ergebnisse nur noch all diese Zahlen miteinander multiplizieren muss. Das ergibt $\frac{n!}{(n-k)!}$ mögliche Ziehungen. □

Beim Lotto (also einer speziellen Ziehung ohne Zurücklegen der Kugeln) interessieren aber nur die 6 Zahlen, also nicht die Reihenfolge, in der sie gezogen wurden. Das Resultat einer Lottoziehung ist also eine 6-elementige Teilmenge von $\{1, 2, 3, \dots, 49\}$. Die Anzahl der Teilmengen einer solchen Menge kann ebenfalls ausgerechnet werden.

3.1.3.3 Hilfssatz. Eine n -elementige Menge hat genau $\binom{n}{k}$ Teilmengen mit je k Elementen.

Beweis. Die oben unter Fall (b) gefundene Zahl muss daher noch geteilt werden durch die Anzahl der Vertauschungsmöglichkeiten von k Elementen, also durch $k!$. Das liefert den Wert

$$\frac{n!}{(n-k)!k!} = \binom{n}{k}$$

□

Nun lässt sich das bei der Definition der Binomialverteilung angegebene Beispiel verstehen:

Angenommen, man habe eine Münze mit Kopf (K) und Zahl (Z). Mit W'keit p komme K, mit W'keit $1-p$ komme Z. Werfen wir die Münze n -mal, so ist das Resultat zu beschreiben durch ein n -Tupel (x_1, \dots, x_n) , wobei jedes der x_j entweder 1 (für K) oder 0 für Z sein kann. Unter allen n -Tupeln zählen wir also diejenigen ab, bei denen an k Stellen 1 steht. Die Menge dieser n -Tupel entspricht der Menge aller k -elementigen Teilmengen von $\{1, \dots, n\}$. Das sind $\binom{n}{k}$ viele. Jedes Resultat, bei dem k -mal K gekommen ist, erhält die W'keit $p^k(1-p)^{n-k}$ (denn k -mal ist K gekommen, und $(n-k)$ -mal nicht!). Das führt auf die W'keit

$$\binom{n}{k} p^k (1-p)^{n-k},$$

dass beim n -maligen Werfen einer Münze k -mal K geworfen wird.

All dies überträgt sich leicht auf das n -malige unabhängige Wiederholen eines Experimentes, bei dem nur 2 Ausgänge (Erfolg bzw. Misserfolg) relevant sind.

3.1.4 Bedingte Wahrscheinlichkeiten, stochastische Unabhängigkeit

Angenommen, ein Zufallsexperiment werde im mathematischen Modell durch den diskreten Ergebnisraum Ω und das W'maß P beschrieben. Manchmal kommt es vor, dass das Eintreten eines Ereignisses B die W'keit eines anderen Ereignisses A beeinflusst. Dazu sehen wir uns Beispiele an:

Beispiele: 1) Angenommen, eine Fabrik fertigt gewisse Schrauben mit Hilfe von 3 Maschinen, die wir M_1, M_2, M_3 nennen. Dabei ist bekannt, dass Maschine M_j einen Produktionsanteil p_j und eine Ausschussquote q_j hat.

1. Frage: Mit welcher W'keit ist dann eine zufällig der Produktion entnommene Schraube defekt?

Man wird nun mit folgenden Ereignissen arbeiten wollen: B_j bedeutet das Ereignis, dass die entnommene Schraube von M_j hergestellt ist, und A bedeute das Ereignis, dass die Schraube

defekt ist. Hier wird A durch B_j beeinflusst. Man wird eine W'keit $P(A|B_j)$ benutzen, dass eine von M_j hergestellte Schraube defekt ist. Sie wird q_j betragen. Ferner kennt man die W'keit $P(B_j)$, dass die Schraube von M_j stammt: $P(B_j) = p_j$. Da nun die Schraube in jedem Fall von einer der 3 Maschinen M_1, M_2, M_3 stammen muss, wird man die W'keit, dass sie defekt ist, mit

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = p_1q_1 + p_2q_2 + p_3q_3$$

bezeichnen.

(Beispiel: Wenn M_1 den Produktionsanteil 25% und die Fehlerquote 5% und entsprechend M_2 die Werte 40% und 2% aufweist und für M_3 die Werte 35% und 4% gelten, so wird

$$P(A) = 0.05 \cdot 0.25 + 0.02 \cdot 0.4 + 0.04 \cdot 0.35 = 0.0345$$

also im Schnitt 3.5% Fehlerquote)

2) Angenommen, man wendet zur Prüfung von Schaltkreisen ein bestimmtes Verfahren an. Die Erfahrung lehrt, dass dieses mit W'keit 0.9 keinen Fehler anzeigt, wenn der Schaltkreis fehlerfrei ist und mit W'keit 0.95 einen Fehler anzeigt, wenn der Schaltkreis fehlerhaft ist. Die W'keit, dass ein Schaltkreis defekt ist, sei 0.04.

1. Frage *Mit welcher W'keit zeigt das Verfahren einen defekten Schaltkreis an?*

2. Frage: *Mit welcher W'keit ist ein Schaltkreis wirklich defekt, wenn das Prüfverfahren einen Fehler behauptet hat?*

Wieder arbeitet man mit 2 Ereignissen die einander bedingen: A bedeute: Das Prüfverfahren zeigt einen Fehler an. Das Ereignis B bestehe darin, dass der Schaltkreis einen Fehler hat.

Man kennt die "bedingte W'keit" $P(A|B)$, dass berechtigterweise der Schaltkreis als fehlerhaft eingestuft wird: $P(A|B) = 0.95$. Ebenso hat man $P(A|B^c) = 0.1$ als die W'keit, dass fälschlicherweise ein Fehler angezeigt wird. Man wird dem Ereignis, dass ein Fehler angezeigt wird, wieder das gewichtete Mittel der W'keiten, also

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) = 0.95 \cdot 0.04 + 0.1 \cdot 0.96 = 0.134$$

als Eintretensw'keit zuweisen.

Die 2. Frage wird später behandelt werden.

Definition. Sei Ω der zu einem Zufallsexperiment gehörige diskrete Ergebnisraum, auf dem das W'maß P definiert sei. Sind $A, B \in \mathcal{A}$ Ereignisse, so verstehen wir, wenn $P(B) > 0$, unter der *bedingten* W'keit von A unter der Bedingung B den Wert

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Das bedeutet: Man nimmt an, man weiß, dass B eingetreten ist. Dann misst $P(A|B)$ die W'keit, dass nun noch A eintreten kann: Der Ergebnisraum Ω wird also ersetzt durch das nunmehr

sichere Ereignis B . Das W'keitsmaß P wird durch Division durch $P(B)$ normiert und so zu einem W'keitsmaß auf diesem neuen Ergebnisraum B .

Nun können wir die 2. Frage nach der W'keit $P(B|A)$ beantworten, mit der ein als fehlerhaft gemeldeter Schaltkreis auch wirklich defekt ist: Es ist

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{0.95 \cdot 0.04}{0.134} = 0.283$$

Das bedeutet: Nur etwa jeder 4. als defekt angezeigte Schaltkreis ist auch wirklich defekt. Das Ergebnis kann auf die kleine Zahl der defekten Schaltkreise zurückgeführt werden und zeigt an, dass das Prüfverfahren "überevorsichtig" genannt werden kann. Auf der anderen Seite ist die W'keit, dass ein fehlerfrei gemeldeter Schaltkreis auch wirklich fehlerfrei ist:

$$P(B^c|A^c) = \frac{P(B^c \cap A^c)}{1 - P(A)} = \frac{P(A^c|B^c)P(B^c)}{1 - P(A)} = \frac{0.9 \cdot 0.96}{1 - 0.134} = 0.997$$

was für die Verlässlichkeit des Verfahrens spricht.

Vom folgenden Satz haben wir bei den Beispielen schon Gebrauch gemacht:

3.1.4.1 Hilfssatz. (Formel von Bayes) Sind B_1, \dots, B_n Ereignisse in dem diskreten Ergebnisraum Ω , welche sich paarweise ausschließen, also $B_i \cap B_j = \emptyset$ für $i \neq j$, so dass aber $\Omega = B_1 \cup \dots \cup B_n$, so gilt für jedes weitere Ereignis A :

$$P(A) = P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)$$

Ferner ist

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)}$$

Beweis. Denn es ist

$$P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n) = P(A \cap B_1) + \dots + P(A \cap B_n) = P(A \cap (B_1 \cup \dots \cup B_n)) = P(A)$$

Die 2. Formel folgt aus dieser und der Definition von $P(B_j|A)$:

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A)}$$

□

Stochastisch unabhängige Ereignisse

Definition. Wir nennen 2 Ereignisse A, B *stochastisch unabhängig*, wenn $P(A|B) = P(A)$ (oder äquivalent: $P(B|A) = P(B)$) also

$$P(A \cap B) = P(A)P(B)$$

gilt.

3.1.4.2 Hilfssatz. Sind A, B stochastisch unabhängig, so sind es auch A^c, B und A, B^c und A^c, B^c .

Beweis. Denn $P(B) = P(A^c \cap B) + P(A \cap B) = P(A^c \cap B) + P(A)P(B)$, also

$$P(A^c \cap B) = P(B) - P(A)P(B) = (1 - P(A))P(B) = P(A^c)P(B)$$

Ferner ist (wegen $A^c \cap B^c = (A \cup B)^c$)

$$\begin{aligned} P(A^c \cap B^c) &= 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A)P(B)) \\ &= 1 - P(A) - P(B) + P(A)P(B) = (1 - P(A))(1 - P(B)) = P(A^c)P(B^c) \end{aligned}$$

Das beweist die letzte Behauptung. □

Rekursiv treffen wir die folgende

Definition. Wir nennen mehrere Ereignisse A_1, \dots, A_r , mit $r \geq 3$, stochastisch unabhängig, wenn jeweils $k < r$ von ihnen es sind und weiter

$$P(A_1 \cap \dots \cap A_r) = P(A_1) \cdot \dots \cdot P(A_k)$$

gilt.

Für 3 Ereignisse A, B, C bedeutet das: Es muss

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C), \quad P(B \cap C) = P(B)P(C),$$

und

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

sein.

Beispiele. a) Würfeln wir 2-mal mit einem Würfel, so hängt das Ergebnis des 2. Wurfes sicher nicht von dem des 1. Wurfes ab. Die beiden Ereignisse $A :=$ "beim 1. Wurf eine 6 geworfen" und $B :=$ "beim zweiten Wurf eine gerade Zahl geworfen" sind stochastisch unabhängig:

$$P(A \cap B) = P(\{(6, 2), (6, 4), (6, 6)\}) = 1/12$$

$$P(A) = P(\{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}) = 1/6$$

$$\begin{aligned} P(B) &= P(\{(1, 2), (1, 4), (1, 6), (2, 2), (2, 4), (2, 6), (3, 2), (3, 4), (3, 6), (4, 2), (4, 4), (4, 6), (5, 2), \\ &\quad (5, 4), (5, 6), (6, 2), (6, 4), (6, 6)\}) = 1/2 \end{aligned}$$

Somit $P(A) \cdot P(B) = \frac{1}{12}$.

b) Werfen wir den Würfel n -mal und kommt mit W'keit p Kopf und mit W'keit $1 - p$ das Ergebnis Zahl zum Vorschein, so können wir die W'keit, dass k -mal Kopf geworfen wird, darstellen

als $P(K_k)$, wobei K_k die Menge aller n -Tupel ist, bei denen an k Stellen eine 1 steht und an den anderen Stellen eine 0. Ist nun $\omega \in K_k$ ein Elementarereignis, bei dem beim Wurf Nr. i_1 , sowie beim Wurf i_2, \dots , bis zum Wurf i_k "Kopf" geworfen wurde, ($1 \leq i_1 < i_2 < \dots < i_k \leq n$), so können wir $\{\omega\}$ als

$$\{\omega\} = A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k} \cap \bigcap_{j \notin \{i_1, \dots, i_k\}} A_j^c$$

darstellen, wobei A_j das Ereignis bedeutet, dass beim j .ten Wurf "Kopf" erscheint. Offenbar haben wir $P(A_j) = p$ für alle $j = 1, \dots, n$. Da alle Ereignisse paarweise stochastisch unabhängig sind, erhalten wir

$$P(\{\omega\}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k}) \cdot \prod_{j \notin \{i_1, \dots, i_k\}} P(A_j^c) = p^k (1-p)^{n-k}$$

Da es aber $\binom{n}{k}$ derartige Elementarereignisse gibt, finden wir

$$P(k \text{ - mal Kopf}) = \binom{n}{k} p^k (1-p)^{n-k}$$

c) Angenommen, wir werfen mit 2 Würfeln. Dann sind die Ereignisse $A =$ "der erste Würfel zeigt eine 1", $B :=$ "der zweite Würfel zeigt eine 1", und $C :=$ "beide Würfel zeigen dieselbe Zahl" nicht stochastisch unabhängig:

$$P(A) = P(B) = P(C) = 1/6, \quad P(A \cap B) = P(A \cap C) = P(B \cap C) = 1/36$$

aber

$$P(A \cap B \cap C) = P(A \cap B) = \frac{1}{36} > \frac{1}{216} = P(A)P(B)P(C)$$

3.1.5 Zufallsvariablen und Verteilungsfunktionen

Diskrete Zufallsvariablen

Definition. Eine diskrete Zufallsvariable X ist eine Messgröße mit Werten in einer abzählbaren Menge \mathcal{X} , also einer Menge von der Form $\mathcal{X} = \{x_1, x_2, x_3, \dots\}$ (endlich oder unendlich). (Beispiele hierfür sind etwa: $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}$, nicht jedoch \mathbb{R}).

Beispiele. Diskrete Zufallsvariablen sind etwa:

- Die Energie eines harmonischen Oszillators,
- die Ladung eines Teilchens,

- die Anzahl der bei einer Telefonzentrale eingehenden Anrufe
- die Anzahl der Autos, die am Tag an einer Tankstelle haltmachen

Definition. Sei X eine diskrete Zufallsvariable mit Werten in der Menge $\mathcal{X} := \{x_1, x_2, \dots\}$ und P ein W-Maß auf \mathcal{X} . Dann sagen wir, X unterliege der Wahrscheinlichkeitsverteilung P , wenn $P(A)$ für alle $A \subset \mathcal{X}$ die Wahrscheinlichkeit angibt, dass X einen Wert aus A annimmt. Wir schreiben auch $P(X \in A)$ statt $P(A)$.

Weiter bezeichnen wir als *Verteilungsfunktion* von X die auf \mathbb{R} definierte Funktion

$$F^X(t) := \sum_{k: x_k \leq t} P(\{x_k\})$$

Das ist eine monoton wachsenden Funktion, die bei den Werten x_l Sprünge haben kann.

3.1.5.1 Hilfssatz. Die Verteilungsfunktion F^X einer diskreten Zufallsvariablen X hat folgende Eigenschaften:

- $\lim_{t \rightarrow -\infty} F^X(t) = 0$
- $\lim_{t \rightarrow \infty} F^X(t) = 1$
- $P(\{a < X \leq b\}) = F^X(b) - F^X(a)$
- $P(\{X > a\}) = 1 - F^X(a)$, $P(\{X \geq a\}) = P(\{X = a\}) + 1 - F^X(a)$

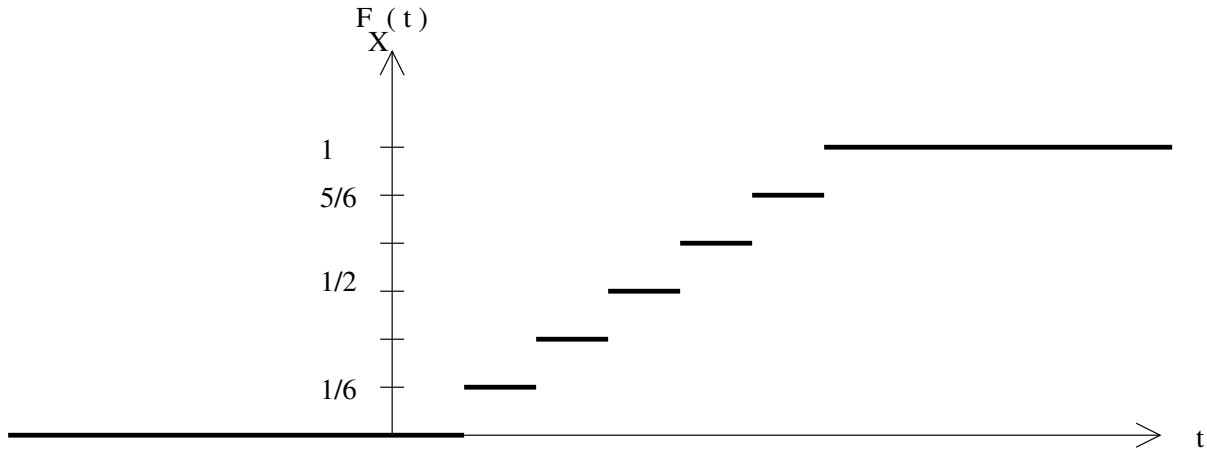
Hier sind Beispiele für einige W'keitsverteilungen:

i) Laplaceverteilung: Ist etwa X Laplaceverteilt, also $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$, so gilt

$$F^X(t) = \frac{|\{k : x_k \leq t\}|}{n}$$

also etwa beim Werfen mit einem Würfel: $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ hat man

$$F^X(t) = \frac{[t]}{6}$$



Beispiel. Werfen mit 2 Würfeln. Mit welcher W'keit wird eine Augenzahl ≥ 10 geworfen? Jetzt ist $\mathcal{X} = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ und wir haben für die Augenzahl X die W'keiten:

| k | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------------|------|------|------|-----|------|-----|------|-----|------|------|------|
| $P(X = k)$ | 1/36 | 1/18 | 1/12 | 1/9 | 5/36 | 1/6 | 5/36 | 1/9 | 1/12 | 1/18 | 1/36 |

Das liefert uns: Die gesuchte W'keit ist

$$p = \frac{6}{36} = 1/6$$

(ii) Binomialverteilung mit Parametern $n = 10, p = 0.2$.

$$F^X(t) = \sum_{0 \leq k \leq t} \binom{10}{k} 0.2^k \cdot 0.8^{10-k}$$

Jetzt ist $\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 10\}$. Weiter haben wir

$$\binom{10}{0} 0.2^0 \cdot 0.8^{10} = 0.11$$

$$\binom{10}{1} 0.2^1 \cdot 0.8^9 = 0.27$$

$$\binom{10}{2} 0.2^2 \cdot 0.8^8 = 0.31$$

$$\binom{10}{3} 0.2^3 \cdot 0.8^7 = 0.20$$

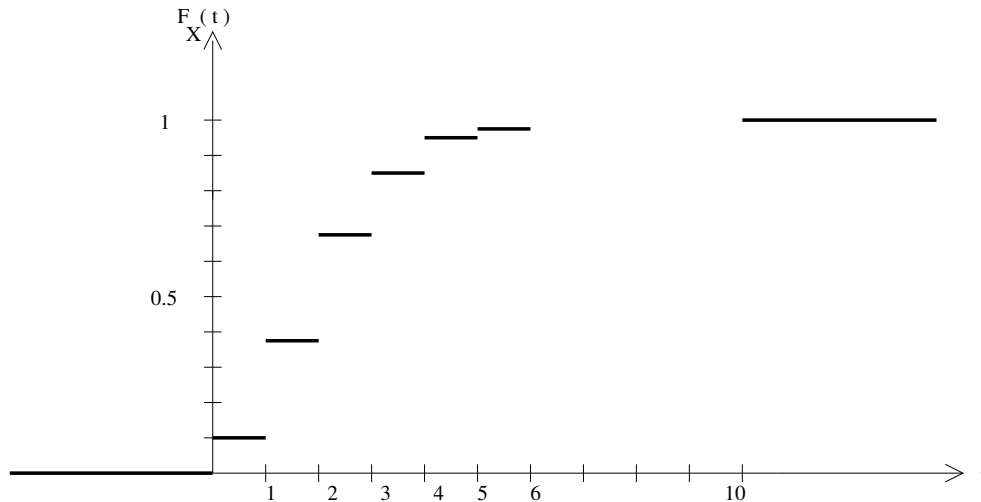
$$\binom{10}{4} 0.2^4 \cdot 0.8^6 = 0.09$$

$$\binom{10}{5} 0.2^5 \cdot 0.8^5 = 0.026$$

$$\binom{10}{6} 0.2^6 \cdot 0.8^4 = 0.005$$

$$\binom{10}{7} 0.2^7 \cdot 0.8^3 = 0.0007.$$

Man findet:



Beispiel: Eine "gezinkte" Münze (mit $p = 0.4$) werde 10-mal geworfen. Mit welcher W'keit kommt mehr als 4-mal Kopf heraus?

Wir bilden dazu

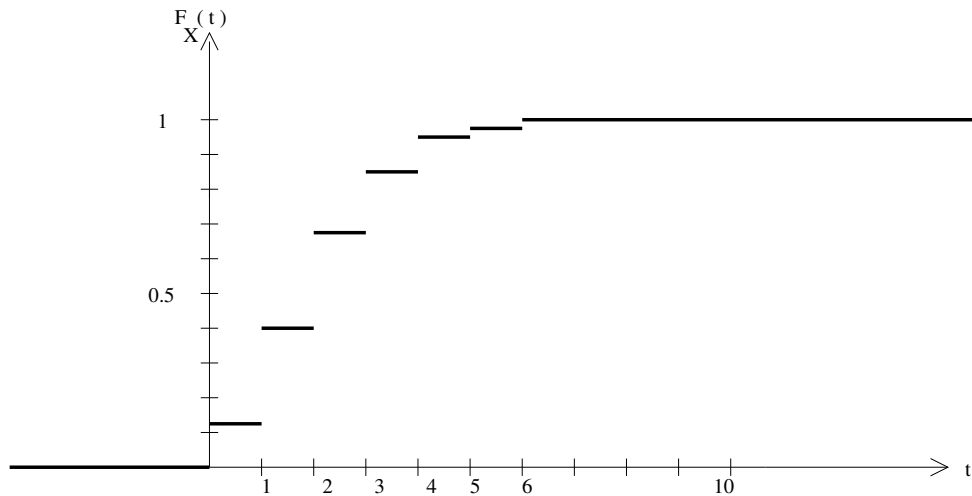
$$F^X(4) = \binom{10}{0} 0.4^0 \cdot 0.6^{10} + \binom{10}{1} 0.4 \cdot 0.6^9 + \binom{10}{2} 0.4^2 \cdot 0.6^8 + \binom{10}{3} 0.4^3 \cdot 0.6^7 + \binom{10}{4} 0.4^4 \cdot 0.6^6 = 0.633$$

Die gesuchte W'keit ist daher $1 - F^X(4) = 0.367$.

(iii) Poissonverteilung mit Parameter λ : Nun wird

$$F^X(t) = \sum_{k \leq t} \frac{\lambda^k}{k!} e^{-\lambda}$$

Für $\lambda = 2$ erhalten wir:



Beispiel: Angenommen, man weiß dass die Anzahl der in einer Telefonzentrale innerhalb 5 Min. eingehenden Telefonate poissonverteilt mit Parameter $\lambda = 10$ sei. Dann ist

$$1 - e^{-10} \sum_{k=0}^5 \frac{10^k}{k!} = 1 - e^{-10}(1 + 10 + 50 + 100/6 + 1000/24 + 1000/12) = 1 - 0.067 = 0.933$$

die W'keit dafür, dass in der nächsten 5 Min. mindestens 6 Anrufe eingehen werden.

Ebenso ist

$$1 - e^{-10} \sum_{k=0}^{12} \frac{10^k}{k!} = 1 - 0.79 = 0.21$$

die W'keit dafür, dass es mehr als 12 und

$$1 - e^{-10} \sum_{k=0}^{15} \frac{10^k}{k!} = 1 - 0.95126 = 0.048$$

die W'keit dafür, dass es mehr als 15 sein werden.

3.1.6 Erwartungswerte und Varianz

Waren wir davon ausgegangen, dass eine Zufallsvariablen jeden Wert mit einer gewissen W'keit annimmt, so liegt es nahe, ihr einen geeigneten gewichteten Mittelwert zuzuweisen.

Definition. Angenommen, es sei X eine diskrete Zufallsgröße mit Werten in $\mathcal{X} := \{x_1, x_2, x_3, \dots\}$, welche Der Wahrscheinlichkeitsverteilung P unterliege. Dann definieren wir den *Erwartungswert* $\mathcal{E}(X)$ wie folgt:

Wenn die Reihe $\sum_{j=0}^{\infty} |x_j| P(\{x_j\}) < \infty$, so sei

$$\mathcal{E}(X) = \sum_{j=0}^{\infty} x_j P(\{x_j\})$$

Ist für X auch $\mathcal{E}(X^2)$ endlich, so setzen wir

$$\text{Var}(X) := \mathcal{E}((X - \mathcal{E}(X))^2)$$

Diese Größe wird *Varianz* von X genannt. Die *Standardabweichung* von X ist definiert als

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

Dann haben wir den folgenden kleinen

3.1.6.1 Hilfssatz. *Ist die Varianz der Zufallsvariablen X definiert, dann gilt*

$$\text{Var}(X) = \mathcal{E}(X^2) - \mathcal{E}(X)^2$$

Beweis. Schreiben wir $p_j = P(\{x_j\})$, so wird

$$\begin{aligned} \text{Var}(X) &= \sum_{j=1}^{\infty} p_j (x_j - \mathcal{E}(X))^2 \\ &= \sum_{j=1}^{\infty} p_j x_j^2 - 2\mathcal{E}(X) \sum_{j=1}^{\infty} p_j x_j + (\mathcal{E}(X))^2 \\ &= \sum_{j=1}^{\infty} p_j x_j^2 - 2\mathcal{E}(X)^2 + (\mathcal{E}(X))^2 \\ &= \mathcal{E}(X^2) - \mathcal{E}(X)^2 \end{aligned}$$

Für die wichtigen Verteilungen können wir Erwartungswert und Varianz ausrechnen.

Beispiele. a) Laplaceverteilung auf $\mathcal{X} = \{a_1, \dots, a_n\}$. Nun ist

$$\mathcal{E}(X) = \sum_{j=1}^n a_j P(\{a_j\}) = \frac{1}{n} \sum_{j=1}^n a_j$$

und

$$\text{Var}(X) = \frac{1}{n} \sum_{j=1}^n a_j^2 - \frac{1}{n^2} \left(\sum_{j=1}^n a_j \right)^2$$

b) Bernoulliverteilung auf $\mathcal{X} = \{0, \dots, n\}$ mit Parametern $p \in (0, 1)$ und $n \in \mathbb{Z}^+$. Nun ist

$$\begin{aligned} \mathcal{E}(X) &= \sum_{k=1}^n k P(\{k\}) = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{l=0}^{n-1} \binom{n-1}{l} p^l (1-p)^{n-1-l} \\ &= np \end{aligned}$$

und zur Berechnung der Varianz:

$$\begin{aligned} \mathcal{E}(X^2) &= \sum_{k=1}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\ &= np \sum_{l=0}^{n-1} (l+1) \binom{n-1}{l} p^l (1-p)^{n-1-l} \\ &= np \underbrace{\sum_{l=1}^{n-1} l \binom{n-1}{l} p^l (1-p)^{n-1-l}}_{=(n-1)p} + np \underbrace{\sum_{l=0}^{n-1} \binom{n-1}{l} p^l (1-p)^{n-1-l}}_{=1} \\ &= np(n-1)p + np \\ &= (np)^2 - np^2 + np \end{aligned}$$

Es folgt

$$\text{Var}(X) = (np)^2 - np^2 + np - (np)^2 = np(1-p)$$

c) Poissonverteilung. Jetzt wird $\mathcal{X} = \mathbb{N}_0$ und daher

$$\mathcal{E}(X) = \sum_{k=1}^{\infty} e^{-\lambda} k \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

Ähnlich errechnet man

$$\mathcal{E}(X^2) = \lambda^2 + \lambda, \quad \text{also } \text{Var}(X) = \lambda$$

d) Hypergeometrische Verteilung mit Parametern N, d und n . Kürzen wir ab: $\nu = \min\{n, d\}$,

so gilt wegen $\frac{n}{N} \binom{N}{n} = \binom{N-1}{n-1}$:

$$\begin{aligned}
 \binom{N}{n} \mathcal{E}(X) &= \sum_{k=1}^{\nu} k \binom{d}{k} \binom{N-d}{n-k} \\
 &= d \sum_{k=1}^{\nu} \binom{d-1}{k-1} \binom{N-1-(d-1)}{(n-1)-(k-1)} \\
 &= d \sum_{l=0}^{\nu-1} \binom{d-1}{l} \binom{N-1-(d-1)}{(n-1)-l} \\
 &= d \binom{N-1}{n-1} = \frac{dn}{N} \binom{N}{n}
 \end{aligned}$$

also $\mathcal{E}(X) = dn/N$.

Der Erwartungswert für X^2 wird ähnlich berechnet

$$\begin{aligned}
 \binom{N}{n} \mathcal{E}(X^2) &= \sum_{k=1}^{\nu} k^2 \binom{d}{k} \binom{N-d}{n-k} \\
 &= d \sum_{k=1}^{\nu} k \binom{d-1}{k-1} \binom{N-1-(d-1)}{(n-1)-(k-1)} \\
 &= d \sum_{l=1}^{\nu-1} l \binom{d-1}{l} \binom{N-1-(d-1)}{(n-1)-l} + d \underbrace{\sum_{l=0}^{\nu-1} \binom{d-1}{l} \binom{N-1-(d-1)}{(N-1)-l}}_{= \frac{n}{N} \binom{N}{n}} \\
 &= d(d-1) \underbrace{\sum_{l=1}^{\nu-1} l \binom{d-2}{l-1} \binom{N-2-(d-2)}{(n-2)-l}}_{= \binom{N-2}{n-2} = \frac{n(n-1)}{N(N-1)} \binom{N}{n}} + d \frac{s}{N} \binom{N}{n} \\
 &= \left(d(d-1) \frac{n(n-1)}{N(N-1)} + \frac{dn}{N} \right) \binom{N}{n}
 \end{aligned}$$

Es folgt

$$\text{Var}(X) = \mathcal{E}(X^2) - (\mathcal{E}(X))^2 = \frac{dn}{N^2(N-1)}(N-d)(N-n)$$

e) Geometrische Verteilung. Wir erhalten leicht:

$$\begin{aligned}
\mathcal{E}(X) &= p \sum_{k=1}^{\infty} k(1-p)^{k-1} \\
&= p \sum_{k=0}^{\infty} (k+1)(1-p)^k \\
&= p \frac{1}{(1-(1-p))^2} = \frac{1}{p}
\end{aligned}$$

und

$$\begin{aligned}
\text{Var}(X) &= p \sum_{k=1}^{\infty} k^2(1-p)^{k-1} - \frac{1}{p^2} \\
&= p \sum_{k=0}^{\infty} (k+1)^2(1-p)^k - \frac{1}{p^2} \\
&= p \sum_{k=0}^{\infty} (k+1)(k+2)(1-p)^k - p \sum_{k=0}^{\infty} (k+1)(1-p)^k - \frac{1}{p^2} \\
&= p \frac{2}{(1-(1-p))^3} - p \frac{1}{(1-(1-p))^2} - \frac{1}{p^2} \\
&= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}
\end{aligned}$$

Folgendes ist klar

3.1.6.2 Hilfssatz. Ist $X : \Omega \rightarrow \mathcal{X}$ eine Zufallsvariable, so gilt

- a) $\mathcal{E}(aX + b) = a\mathcal{E}(X) + b$
- b) $\text{Var}(aX + b) = a^2\text{Var}(X)$, wenn $a \neq 0, b \in \mathbb{R}$.

Zwischen der Varianz und der Abweichung einer Zufallsvariablen von ihrem Erwartungswert hat man eine Abschätzung

3.1.6.3 Satz (Tschebycheff-Ungleichung). Ist $X : \Omega \rightarrow \mathcal{X}$ eine Zufallsvariable, für welche $\mathcal{E}(X)$ und $\text{Var}(X)$ endlich sind, so gilt für jedes $\varepsilon > 0$:

$$P(\{\mathcal{E}(X) - \varepsilon \leq X \leq \mathcal{E}(X) + \varepsilon\}) \geq 1 - \frac{\text{Var}(X)}{\varepsilon^2}$$

Beweis. Es gilt mit $\mu = \mathcal{E}(X)$:

$$\begin{aligned} \text{Var}(X) &= \mathcal{E}((X - \mu)^2) \\ &= \sum_{k:|x_k-\mu|>\varepsilon} p_k(x_k - \mu)^2 + \sum_{k:|x_k-\mu|\leq\varepsilon} p_k(x_k - \mu)^2 \\ &\geq \varepsilon^2 \sum_{k:|x_k-\mu|>\varepsilon} p_k = \varepsilon^2 P(|X - \mu| > \varepsilon) \\ &= \varepsilon^2(1 - P(|X - \mu| \leq \varepsilon)) \end{aligned}$$

woraus alles folgt. □

Stochastische Unabhängigkeit

Definition. Gegeben seien 2 diskrete Zufallsvariable X und Y mit Werten in den abzählbaren Mengen $\mathcal{X} \subset \mathbb{R}$ bzw. $\mathcal{Y} \subset \mathbb{R}$. Angenommen, X unterliege der W'verteilung P^X und Y einer W-Verteilung P^Y . Dann nennen wir X und Y *stochastisch unabhängig*, wenn $P^X(A)P^Y(B)$ für $A \subset \mathcal{X}$ und $B \subset \mathcal{Y}$ die W'keit dafür angibt, dass X einen Wert in A und Y einen Wert in B annimmt.

ii) Angenommen, für beide Zufallsvariablen X, Y seien die Erwartungswerte definiert. Dann bezeichnen wir als *Covarianz* von X und Y die Zahl

$$\text{Cov}(X, Y) = \mathcal{E}((X - \mathcal{E}(X))(Y - \mathcal{E}(Y)))$$

(iii) Ferner heißen X und Y *unkorreliert*, wenn

$$\text{Cov}(X, Y) = 0$$

Wir stellen fest:

3.1.6.4 Hilfssatz. a) Für 2 diskrete reelle Zufallsvariable X, Y gilt

$$\text{Cov}(X, Y) = \mathcal{E}(X \cdot Y) - \mathcal{E}(X)\mathcal{E}(Y)$$

b) Insbesondere sind stochastische unabhängige Zufallsvariable X und Y auch unkorreliert.

Beweis. Wir schreiben $X' = X - \mathcal{E}(X), Y' = Y - \mathcal{E}(Y)$. Dann wird $\mathcal{E}(X') = \mathcal{E}(Y') = 0$ und damit

$$\begin{aligned} \mathcal{E}(XY) &= \mathcal{E}((X' + \mathcal{E}(X))(Y' + \mathcal{E}(Y))) \\ &= \mathcal{E}(X'Y' + Y'\mathcal{E}(X) + X'\mathcal{E}(Y)) + \mathcal{E}(X)\mathcal{E}(Y) \\ &= \mathcal{E}(X'Y') + \mathcal{E}(Y')\mathcal{E}(X) + \mathcal{E}(X')\mathcal{E}(Y) + \mathcal{E}(X)\mathcal{E}(Y) \\ &= \mathcal{E}(X'Y') + \mathcal{E}(X)\mathcal{E}(Y) \\ &= \text{Cov}(X, Y) + \mathcal{E}(X)\mathcal{E}(Y) \end{aligned}$$

Daraus folgt die erste Behauptung.

b) Angenommen, $\mathcal{X} = \{x_1, x_2, x_3, \dots\}$ und $\mathcal{Y} = \{y_1, y_2, y_3, \dots\}$. Ferner sehen wir, dass

$$\mathcal{E}(X \cdot Y) = \sum_{j,k=1}^{\infty} x_j y_k P(X = x_j, Y = y_k) = \sum_{j,k=1}^{\infty} x_j y_k P(X = x_j) P(Y = y_k) = \mathcal{E}(X) \mathcal{E}(Y)$$

Das beweist b). □

3.1.6.5 Hilfssatz. Sind $X : \Omega \rightarrow \mathcal{X}, Y : \Omega \rightarrow \mathcal{Y}$ diskrete Zufallsvariablen, so wird

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

Bei unkorrelierten X, Y gilt also

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

Beweis. Wir setzen wieder $X' = X - \mathcal{E}(X), Y' = Y - \mathcal{E}(Y)$. Dann ist $\text{Var}(X') = \text{Var}(X), \text{Var}(Y') = \text{Var}(Y)$. Es folgt:

$$\begin{aligned} \text{Var}(X + Y) &= \text{Var}(X' + Y') = \mathcal{E}((X' + Y')^2) \\ &= \mathcal{E}(X'^2) + \mathcal{E}(Y'^2) + 2\mathcal{E}(X'Y') \\ &= \text{Var}(X') + \text{Var}(Y') + 2\text{Cov}(X, Y) \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \end{aligned}$$

3.1.7 Kontinuierliche Zufallsvariable

Definition. Zufallsvariablen, deren Wertebereich ein Kontinuum, wie etwa \mathbb{R} oder Intervalle in \mathbb{R} umfasst, nennt man auch *kontinuierlich*.

Eine Möglichkeit, W'keitsverteilungen für die Werte kontinuierlicher Zufallsvariablen sinnvoll zu beschreiben, basiert auf W'keitsdichten:

Definition. Sei X eine kontinuierliche Zufallsvariable mit Werten in der Menge $\mathcal{X} \subset \mathbb{R}$. Ist $f : \mathcal{X} \rightarrow \mathbb{R}_0^+$ stückweise stetig und außerhalb \mathcal{X} gleich 0, so dass $\int_{-\infty}^{\infty} f(x) dx = 1$, so definiert

$$F(t) = \int_{-\infty}^t f(x) dx$$

für $t \in \mathbb{R}$ eine stetige Funktion. Man nennt f die Verteilungsdichte zu X , wenn F die W'keit dafür angibt, dass X einen Wert kleiner oder gleich t annimmt.

Das folgende ist ein Satz aus der W'keitstheorie:

3.1.7.1 Satz. Ist $f : \mathcal{X} \rightarrow \mathbb{R}_0^+$ eine stückweise stetige Verteilungsdichtefunktion und $F(t) = \int_{-\infty}^t f(x)dx$ für $t \in \mathbb{R}$, so wird durch

$$P([a, b]) = F(b) - F(a) = \int_a^b f(x)dx$$

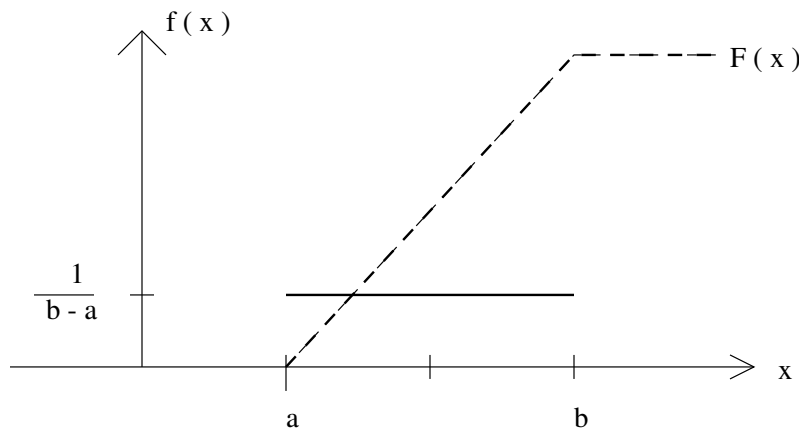
ein Wahrscheinlichkeitsmaß definiert.

Wir sehen uns nun einige wichtige Beispiele an:

Beispiele. a) Die *Rechteckverteilung*: Hier ist $f(x) = \frac{1}{b-a}$, wenn $a \leq x \leq b$ und $f(x) = 0$ sonst. Im Bild:

Die Verteilungsfunktion dazu ist

$$F(t) = \int_{-\infty}^t f(x)dx = \begin{cases} 0 & \text{wenn } t \leq a \\ \frac{t-a}{b-a} & \text{wenn } a \leq t \leq b \\ 1 & \text{wenn } t \geq b \end{cases}$$



Beispiel. Angenommen, man lasse durch einen dünnen Kupferdraht Strom fließen. Die Stromstärke (in mA) sei rechteckverteilt mit $a = 0, b = 10$. Dann ist die W'keit, dass ein Strom von weniger als 4mA fließt, durch

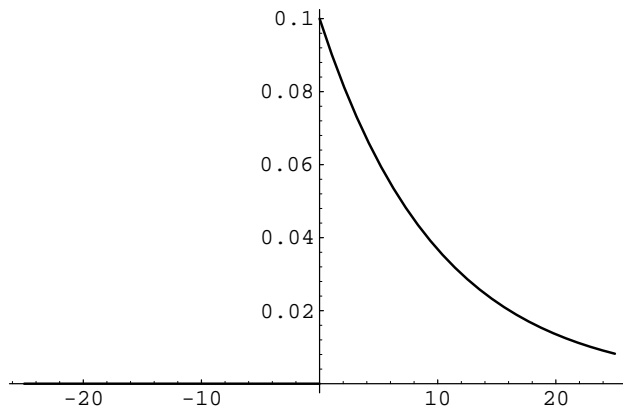
$$P(X \leq 4) = \frac{4}{10} = 0.4$$

gegeben.

b) Die *Exponentialverteilung*. Ist $\lambda > 0$, so sei

$$f_\lambda(x) = \begin{cases} 0 & , \text{ wenn } x \leq 0 \\ \lambda e^{-\lambda x} & , \text{ wenn } x > 0 \end{cases}$$

Graph von f_λ für $\lambda = 0.1$:



Beispiel. Die Lebensdauer X (in Std) einer Leuchtstoffröhre sei exponentialverteilt mit $\lambda = 10^{-4}/\text{Std}$. Die W'keit, dass die Leuchtstoffröhre mehr als 12000 Std hält, ist dann

$$P(X > 12000) = \lambda \int_{12000}^{\infty} e^{-\lambda x} dx = -e^{-\lambda x} \Big|_{12000}^{\infty} = e^{-12000\lambda} = e^{-1.2} = 0.301$$

Mit der W'keit

$$\begin{aligned} P(9000 < X < 12000) &= \lambda \int_{9000}^{12000} e^{-\lambda x} dx = -e^{-\lambda x} \Big|_{9000}^{12000} \\ &= e^{-9000\lambda} - e^{-12000\lambda} = e^{-0.9} - e^{-1.2} = 0.406 - 0.301 = 0.105 \end{aligned}$$

liegt ihre Lebensdauer zwischen 9000 und 12000 Std.

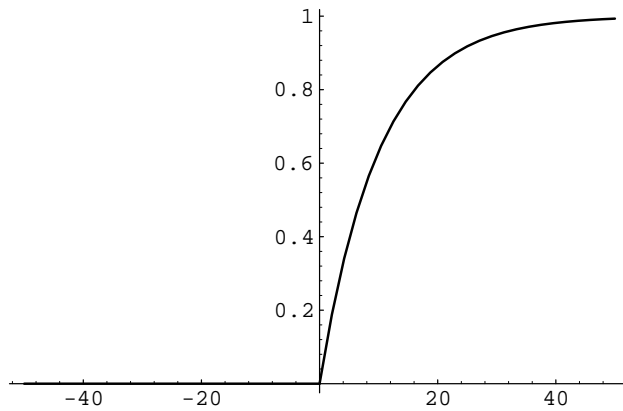
Weiß man, dass sie schon mindestens 8000 Std. gehalten hat, so ist die W'keit dafür, dass sie sogar mehr als 12000 Std halten wird, gerade

$$P(X > 12000 | X \geq 8000) = \frac{P(X > 12000)}{P(X \geq 8000)} = \frac{e^{-1.2}}{e^{-0.8}} = e^{-0.4} = 0.67$$

Die zugehörige Verteilungsfunktion der Exponentialverteilung

$$F_{\lambda}(t) = \begin{cases} 0 & , \text{ wenn } t < 0 \\ \int_0^t f_{\lambda}(x) dx = 1 - e^{-\lambda t} & , \text{ wenn } t \geq 0 \end{cases}$$

hat den folgenden Graphen:



Die Gammaverteilung

Für ein $\alpha > 0$ setzen wir

$$f_{\alpha}(x) = \begin{cases} 0 & , \text{ wenn } x \leq 0 \\ \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} & , \text{ wenn } x > 0 \end{cases}$$

Dabei ist

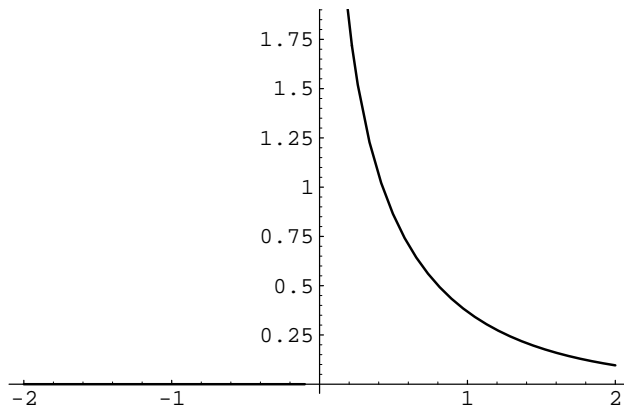
$$\Gamma(s) := \int_0^{\infty} x^{s-1} e^{-x} dx$$

die Eulersche *Gammafunktion*. Sie ist beliebig oft differenzierbar auf \mathbb{R}^+ und erfüllt

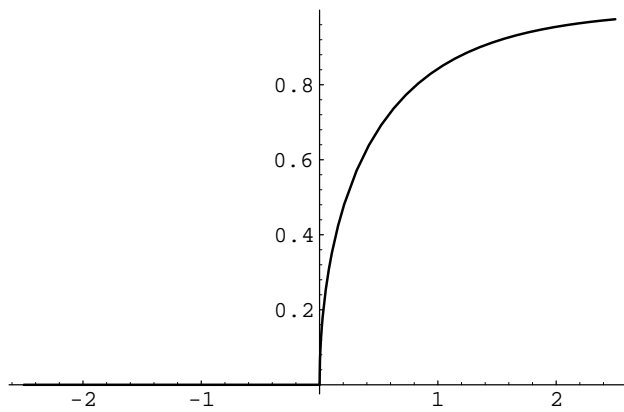
- $\Gamma(s+1) = s\Gamma(s)$, $\Gamma(n+1) = n!$ für ganze $n \geq 0$.
- Man hat die Funktionalgleichung $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$

Das liefert etwa $\Gamma(1/2) = \sqrt{\pi}$.

Der Graph von f_{α} sieht für $\alpha = 0.5$ so aus:



und ihre Verteilungsfunktion F_α so:



Die Weibull-Verteilung

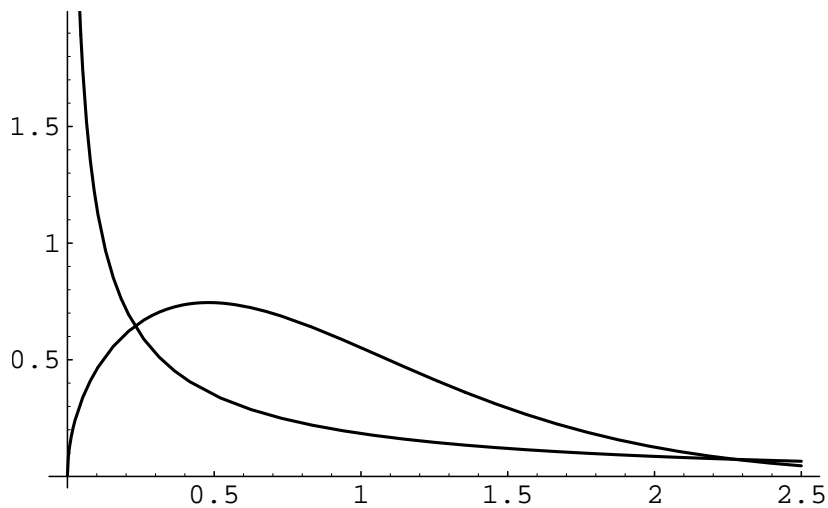
Sie ist nach dem schwedischen Ingenieur benannt, der Ende der 40iger Jahre herausfand, dass die Lebensdauer vieler Typen von Verschleißteilen dieser Verteilung unterliegt. Ihre Verteilungsdichte ist

$$f_{\beta,\theta}(x) = \begin{cases} 0 & , \text{ wenn } x \leq 0 \\ \frac{\beta}{\theta} \left(\frac{x}{\theta}\right)^{\beta-1} e^{-\left(\frac{x}{\theta}\right)^\beta} & , \text{ wenn } x > 0 \end{cases}$$

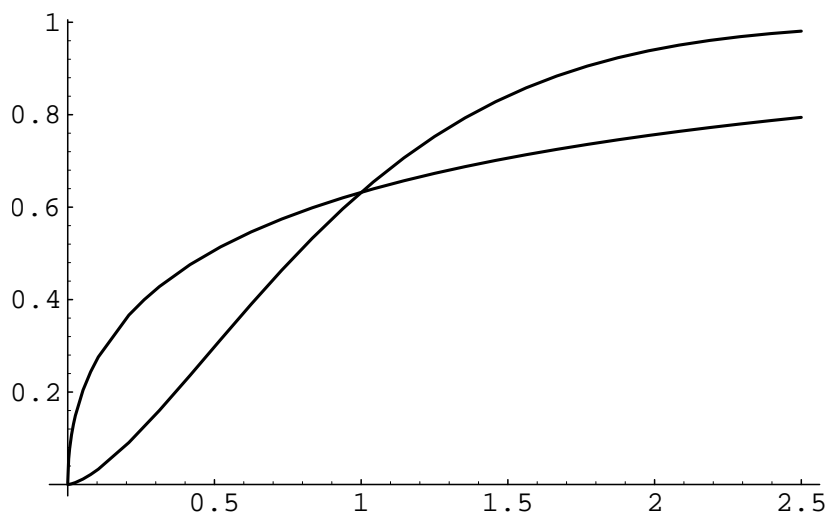
Das ergibt als Verteilungsfunktion

$$F_{\beta,\theta}(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^\beta}$$

Hier sind die Graphen von $f_{\beta,\theta}$ mit $\theta = 1$, einmal für ein $\beta > 1$ und einmal für den Fall $\beta < 1$.



und hier für die zugehörigen Verteilungsfunktionen:



Die *Normalverteilung* oder *Gaußverteilung*

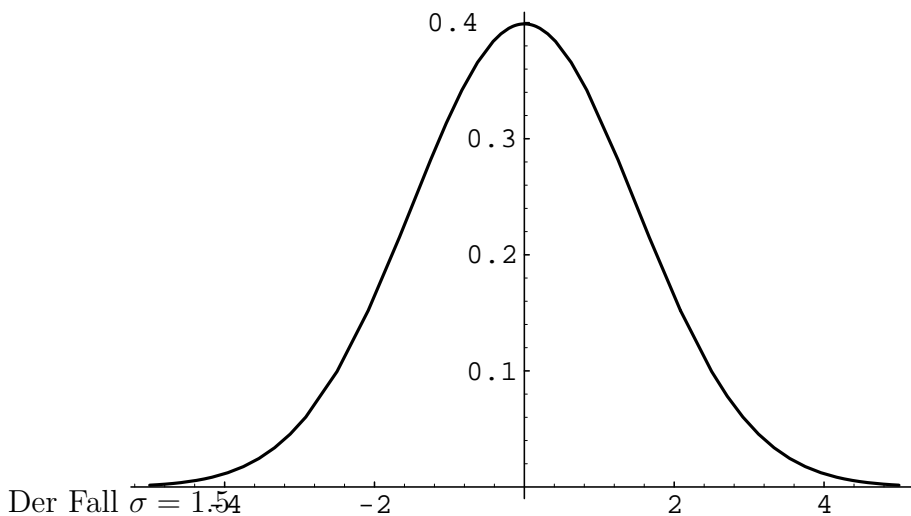
Ihre Verteilungsfunktion ist

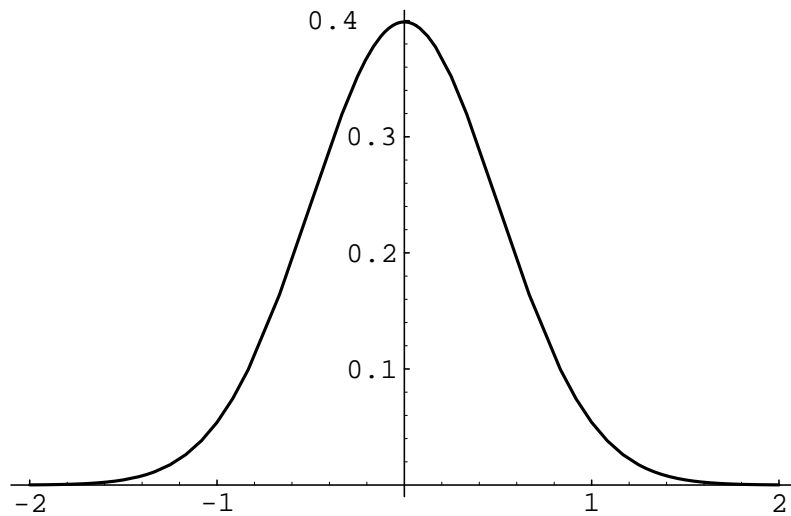
$$f_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

für $x \in \mathbb{R}$.

Sie ist eine besonders wichtige Verteilung, und ihr Name rührt von der Tatsache her, dass Folgen von W'keitsverteilungen unter vernünftigen Voraussetzungen in einem später zu präzisierenden Sinne gegen die Normalverteilung streben. (So kann man etwa die Binomialverteilung mit Parametern n und p durch die Normalverteilung mit Parametern $\mu = np$ und $\sigma = \sqrt{np(1-p)}$ approximieren, wenn nur n groß wird).

Hier ist der Graph für $\mu = 0$ und verschiedene σ -Werte.





Der Fall $\sigma = 0.5$

3.1.8 Erwartungswerte und Varianz bei kontinuierlichen Zufallsvariablen

Den Erwartungswert definiert man für kontinuierliche Zufallsvariable mit Verteilungsdichte wie folgt:

Definition. Ist X eine kontinuierliche Zufallsvariable mit Verteilungsdichtefunktion f , so bezeichnen wir als *Erwartungswert* das Integral

$$\mathcal{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Entsprechend definieren wir die *Varianz* und die *Streuung* wieder durch

$$\text{Var}(X) = \mathcal{E}(X^2) - \mathcal{E}(X)^2, \quad \sigma(X) = \sqrt{\text{Var}(X)}$$

ausführlicher:

$$\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - \mathcal{E}(X)^2$$

Die Kovarianz von 2 Zufallsvariablen X, Y ist definiert wie früher als

$$\text{Cov}(X, Y) = \mathcal{E}((X - \mathcal{E}(X))(Y - \mathcal{E}(Y)))$$

Der Satz über die Varianz einer Summe von Zufallsvariablen bleibt wörtlich gültig:

3.1.8.1 Hilfssatz. *Ist X eine Zufallsvariable, so gilt*

a) $\mathcal{E}(aX + b) = a\mathcal{E}(X) + b$

b) $\text{Var}(aX + b) = a^2\text{Var}(X)$, wenn $a \neq 0, b \in \mathbb{R}$.

c) Sind X, Y Zufallsvariablen, so ist

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

Ebenso überträgt sich die Tschebyscheff-Ungleichung

3.1.8.2 Satz (Tschebyscheff-Ungleichung). *Ist X eine Zufallsvariable, für welche $\mathcal{E}(X)$ und $\text{Var}(X)$ endlich sind, so gilt für jedes $\varepsilon > 0$:*

$$P(\{\mathcal{E}(X) - \varepsilon \leq X \leq \mathcal{E}(X) + \varepsilon\}) \geq 1 - \frac{\text{Var}(X)}{\varepsilon^2}$$

Wir sehen uns einige Beispiele an:

Rechteck-Verteilung: Hier ist

$$f(x) = \begin{cases} 0 & \text{wenn } x \leq a \\ \frac{1}{b-a} & \text{wenn } a \leq x \leq b \\ 0 & \text{wenn } x \geq b \end{cases}$$

und daher

$$\begin{aligned} \mathcal{E}(X) &= \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2} \\ \text{Var}(X) &= \frac{1}{b-a} \int_a^b x^2 dx - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{1}{3} \frac{b^3 - a^3}{b-a} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{a^2 - 2ab + b^2}{12} \\ &= \frac{(a-b)^2}{12} \end{aligned}$$

Exponentialverteilung. Hier ist

$$f(x) = \begin{cases} 0 & \text{wenn } x \leq 0 \\ \lambda e^{-\lambda x} & \text{wenn } x > 0 \end{cases}$$

und daher

$$\begin{aligned} \mathcal{E}(X) &= \lambda \int_0^{\infty} x e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^{\infty} t e^{-t} dt = \frac{1}{\lambda} \Gamma(2) = \frac{1}{\lambda} \\ \mathcal{E}(X^2) &= \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = \frac{1}{\lambda^2} \int_0^{\infty} t^2 e^{-t} dt = \frac{1}{\lambda^2} \Gamma(3) = \frac{2}{\lambda^2} \end{aligned}$$

Das ergibt

$$\text{Var}(X) = \frac{1}{\lambda^2}$$

Bei der

Gammaverteilung

ist die Dichte

$$f_{\alpha}(x) = \begin{cases} 0 & , \text{ wenn } x \leq 0 \\ \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} & , \text{ wenn } x > 0 \end{cases}$$

mit einem Parameter $\alpha > 0$. Wir erhalten

$$\mathcal{E}(X) = \int_0^{\infty} x f_{\alpha}(x) dx = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha} e^{-x} dx = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha$$

Die Weibullverteilung

Ihre Verteilungsdichtefunktion ist

$$f_{\beta,\theta}(x) = \begin{cases} 0 & , \text{ wenn } x \leq 0 \\ \frac{\beta}{\theta} \left(\frac{x}{\theta}\right)^{\beta-1} e^{-\left(\frac{x}{\theta}\right)^{\beta}} & , \text{ wenn } x > 0 \end{cases}$$

So finden wir für den Erwartungswert einer Weibull-verteilten Zufallsvariablen:

$$\begin{aligned} \mathcal{E}(X) &= \frac{\beta}{\theta} \int_0^{\infty} x \left(\frac{x}{\theta}\right)^{\beta-1} e^{-\left(\frac{x}{\theta}\right)^{\beta}} dx \\ &= \theta \int_0^{\infty} t^{1/\beta} e^{-t} dt, \quad (\text{Substitution } t := \left(\frac{x}{\theta}\right)^{\beta}) \\ &= \theta \Gamma\left(\frac{1}{\beta} + 1\right) \end{aligned}$$

Für die Varianz finden wir auf entsprechende Weise

$$\text{Var}(X) = \theta^2 \Gamma\left(\frac{2}{\beta} + 1\right) - \mathcal{E}(X)^2 = \theta^2 \left(\Gamma\left(\frac{2}{\beta} + 1\right) - \Gamma\left(\frac{1}{\beta} + 1\right)^2 \right)$$

Wenn der Erwartungswert bekannt ist und etwa den Wert E hat ebenso die Varianz (mit Wert V), kann man damit die Parameter β und θ bestimmen. Es gilt

$$1 + \frac{V}{E^2} = \frac{\Gamma(\frac{2}{\beta} + 1)}{\Gamma(\frac{1}{\beta} + 1)^2}$$

Dazu sucht man den richtigen β -Wert. Der Wert für θ ergibt sich dann aus

$$\theta = \frac{E}{\Gamma(\frac{1}{\beta} + 1)}$$

Beispiel: Ein Motor habe eine Lebensdauer von durchschnittlich 8 Jahren mit einer Streuung von 4 Jahren. Die Erfahrung lehrt, dass seine Lebensdauer Weibull-verteilt ist. Was ist dann θ , was β ?

Nun muss man ($V = 16$):

$$\frac{\Gamma(\frac{2}{\beta} + 1)}{\Gamma(\frac{1}{\beta} + 1)^2} = 1.25$$

lösen. Die Lösung dazu ist $\beta = 2.1013$. Dann folgt

$$\theta = \frac{8}{\Gamma(1.4759)} = 9.0325$$

Die W'keit, dass der Motor im 10. Jahr ausfällt, ist nun

$$P(X < 11) = F(11) = 1 - e^{-\left(\frac{11}{9.0325}\right)^{2.1013}} = 0.779$$

Wenn man schon weiß, dass der Motor 9 Jahre alt ist, dann ist die W'keit, dass er noch mindestens 1 weiteres Jahr hält:

$$P(X \geq 10 \mid X \geq 9) = \frac{P(X \geq 10)}{P(X \geq 9)} = \frac{1 - F(10)}{1 - F(9)} = e^{\left(\frac{9}{9.0325}\right)^{2.1013} - \left(\frac{10}{9.0325}\right)^{2.1013}} = e^{-0.246} = 0.782$$

Die Normalverteilung

Angenommen, eine Zufallsvariable unterliege der Normalverteilung mit Parametern μ, σ (wir sagen auch: X sei nach $\mathcal{N}(\mu, \sigma)$ verteilt). Dann ist μ der Erwartungswert und σ^2 die Varianz von X , denn

$$\mathcal{E}(X) = \int_{-\infty}^{\infty} x f_{\mu, \sigma}(x) dx = \mu + \int_{-\infty}^{\infty} (x - \mu) f_{\mu, \sigma}(x) dx = \mu$$

denn das 2. Integral verschwindet, weil der Integrand eine ungerade Funktion ist.

Nun zur Varianz

$$\begin{aligned}
 \mathcal{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f_{\mu,\sigma}(x) dx \\
 &= \int_{-\infty}^{\infty} (x - \mu)^2 f_{\mu,\sigma}(x) dx + 2 \int_{-\infty}^{\infty} (x - \mu) f_{\mu,\sigma}(x) dx + \mu^2 \\
 &= \int_{-\infty}^{\infty} (x - \mu)^2 f_{\mu,\sigma}(x) dx + \mu^2 \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2} dt + \mu^2 \\
 &= \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s^2 e^{-\frac{1}{2}s^2} ds + \mu^2 \\
 &= \sigma^2 + \mu^2
 \end{aligned}$$

denn

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s^2 e^{-\frac{1}{2}s^2} ds = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} s^2 e^{-\frac{1}{2}s^2} ds = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \sqrt{2t} e^{-t} dt = \frac{2}{\sqrt{\pi}} \Gamma(3/2) = \frac{1}{\sqrt{\pi}} \Gamma(1/2) = 1$$

Daraus folgt

$$\text{Var}(X) = \sigma^2$$

Um Werte der Verteilungsfunktion bequem berechnen zu können, schreiben wir

$$\text{Erf}(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx$$

Dann ist für $t > 0$:

$$\text{Erf}(t) = 1 - \text{Erf}(-t)$$

und damit

$$\begin{aligned}
 P(X - \mu < -t) &= P(X - \mu \leq -t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu-t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{-t} e^{-\frac{x^2}{2\sigma^2}} dx \\
 &= \text{Erf}(-t/\sigma) = 1 - \text{Erf}(t/\sigma)
 \end{aligned}$$

Wegen der Symmetrie der Verteilungsdichtefunktion $f_{\mu,\sigma}$ bezügl. der Stelle $x = \mu$ gilt ebenso

$$P(X - \mu > t) = P(X - \mu \geq t) = 1 - \text{Erf}(t/\sigma)$$

also

$$P(X - \mu < t) = P(X - \mu \leq t) = 1 - P(X - \mu > t) = \text{Erf}(t/\sigma)$$

und

$$P(-t < X - \mu < t) = P(X - \mu < t) - P(X - \mu < -t) = 2\text{Erf}(t/\sigma) - 1$$

Beispiel. Angenommen, in 2 Kiesgruben K_1 und K_2 werde Sand verarbeitet. Mit diesem Sand wird eine Glasschmelze hergestellt. Der (in Einheiten von 0.01% gemessene) Verunreinigungsgrad X sei bei K_1 nach $\mathcal{N}(60, 0.04)$ und bei K_2 nach $\mathcal{N}(62, 0.03)$ verteilt. Ein Abnehmer will aber nur Sand mit Verunreinigungsgrad maximal gleich 0.64% haben. Aus welcher Kiesgrube wird er den Sand dann bestellen?

Für K_1 gilt:

$$P(X < 0.64) = P(X - 0.60 < 0.04) = \text{Erf}(0.04/0.04) = \text{Erf}(1) = 0.841$$

und für K_2 :

$$P(X < 0.64) = P(X - 0.62 < 0.02) = \text{Erf}(2/3) = 0.747$$

Die 1. Kiesgrube erhält also den Vorzug.

3.1.9 Zusammengesetzte Zufallsvariablen

Wir untersuchen nun, wie sich die Erwartungswerte von Zufallsvariablen verhalten, die dadurch entstehen, dass man eine Zufallsvariable in eine stetige Funktion einsetzt.

3.1.9.1 Hilfssatz. Angenommen, X sei eine kontinuierliche stetige auf \mathbb{R} definierte Zufallsvariable mit einer Verteilungsfunktion F^X und $h : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige, streng monotone Funktion, die eine Umkehrfunktion besitzt. Dann besitzt $h \circ X$ die Verteilungsfunktion

$$F^{h \circ X} = F^X \circ h^{-1}$$

wenn h streng monoton wächst. Wenn F^X stetig und h streng monoton fallend ist, so ist

$$F^{h \circ X} = 1 - F^X \circ h^{-1}$$

Sind schließlich F^X und h sogar differenzierbar, so wird durch

$$f^{h \circ X}(t) = f^X \circ h^{-1} \cdot |(h^{-1})'|$$

eine W'keitsdichtefunktion definiert.

Beweis. Denn es gilt für streng monoton wachsendes h :

$$F^{h \circ X}(t) = P(h \circ X \leq t) = P(X \leq h^{-1}(t)) = F^X(h^{-1}(t))$$

und für streng monoton fallendes h :

$$F^{h \circ X}(t) = P(h \circ X \leq t) = P(X \geq h^{-1}(t)) = 1 - P(X < h^{-1}(t)) = 1 - F(h^{-1}(t))$$

denn $P(X = s) = 0$ für alle s , wenn F^X stetig ist. Die letzte Aussage folgt durch Differenziation. \square

Beispiele. Sei X eine kontinuierliche stetige Zufallsvariable mit stetiger Verteilungsfunktion.

(i) $h(x) = \alpha x + \beta$. Dann gilt für jede Zufallsvariable X mit Verteilungsfunktion F^X :

$$f_{\alpha X + \beta}(t) = |\alpha| f^X\left(\frac{t - \beta}{\alpha}\right)$$

und

$$F^{\alpha X + \beta}(t) = \begin{cases} \int_{-\infty}^{\frac{t - \beta}{\alpha}} f^X(s) ds & , \alpha > 0 \\ \int_{\frac{t - \beta}{\alpha}}^{\infty} f^X(s) ds & , \alpha < 0 \end{cases}$$

(ii) Sei $h(s) := e^{|s|}$. Dann ist h auf \mathbb{R}^+ monoton wachsend und auf \mathbb{R}^- monoton fallend. Wir schreiben für $t > 1$:

$$P(h \circ X \leq t) = P(-\ln t \leq X \leq \ln t) = F^X(\ln t) - F^X(-\ln t)$$

Ist zusätzlich F^X differenzierbar, so erhalten wir in

$$f^{h \circ X}(t) = \frac{f^X(\ln t) + f^X(-\ln t)}{t}$$

eine W'keitsdichtefunktion.

Für den Erwartungswert einer zusammengesetzten Zufallsvariablen gilt

3.1.9.2 Hilfssatz. Ist X eine Zufallsvariable mit differenzierbarer Verteilungsfunktion und $h : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare streng monoton wachsende Funktion, so ist

$$\mathcal{E}(h \circ X) = \int_{-\infty}^{\infty} h(x) f^X(x) dx$$

Beweis. Das ergibt sich aus der Substitutionsregel

$$\int_{-\infty}^{\infty} t f^{h \circ X}(t) dt = \int_{-\infty}^{\infty} h(h^{-1}(t)) f^X \circ h^{-1}(t) (h^{-1})'(t) dt = \int_{-\infty}^{\infty} h(x) f(x) dx$$

□

Aus zwei Zufallsvariablen zusammengesetzte Zufallsvariablen

Definition. Sind $X_1, X_2 : \mathbb{R} \rightarrow \mathbb{R}$ Zufallsvariable, so verstehen wir unter der *gemeinsamen Verteilung* von X_1 und X_2 die (nun auf \mathbb{R}^2 erklärte) Funktion

$$F^{X_1, X_2}(t_1, t_2) = P(\{X_1 \leq t_1\} \cap \{X_2 \leq t_2\})$$

Ist diese Funktion 2-mal stetig differenzierbar, so gilt

$$F(t_1, t_2) = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \frac{\partial^2 F}{\partial s_1 \partial s_2}(s_1, s_2) ds_1 ds_2$$

Wir bezeichnen daher die Funktion

$$f_{X_1, X_2}(s_1, s_2) = \frac{\partial^2 F}{\partial s_1 \partial s_2}(s_1, s_2)$$

als *gemeinsame W'keitsdichtefunktion* für X_1 und X_2 .

Für stochastisch unabhängige Zufallsvariablen gilt der

3.1.9.3 Hilfssatz. Angenommen, X_1 und X_2 besitzen stetige Verteilungsfunktionen F_1 und F_2 und W'dichten f_1 und f_2 . Dann sind folgende Aussagen äquivalent:

- $F^{X_1, X_2}(t_1, t_2) = F_1(t_1)F_2(t_2)$
- $f_{X_1, X_2}(s_1, s_2) = f_1(s_1)f_2(s_2)$
- X_1 und X_2 sind stochastisch unabhängig.

Beweis. Folgt leicht aus

$$P(X_1 \leq t_1, X_2 \leq t_2) = P(X_1 \leq t_1)P(X_2 \leq t_2)$$

und der Definition von f_{X_1, X_2} .

□

Eine Anwendung für spätere Zwecke ist

3.1.9.4 Hilfssatz. Angenommen, X_1 und X_2 sind stochastisch unabhängig und $\mathcal{N}(0, 1)$ -verteilt. Sind dann Y_1 und Y_2 die durch

$$Y_1 = aX_1 + bX_2, \quad Y_2 = cX_1 + dX_2$$

definierte Zufallsvariablen, so gilt:

Sind die Vektoren $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ Einheitsvektoren, die aufeinander senkrecht stehen, so sind auch Y_1 und Y_2 wieder stochastisch unabhängig.

Beweis. Denn man kann zeigen, dass ihre gemeinsame W'keitsdichtefunktion wieder Produktstruktur hat. □

Mit Hilfe der gemeinsamen W'keitsdichtefunktion lassen sich Verteilungsfunktionen für Summen, Produkte und Quotienten von 2 Zufallsvariablen X_1 und X_2 bestimmen.

Summe zweier Zufallsvariablen

Angenommen, es sind X_1 und X_2 Zufallsvariablen mit einer 2-mal stetig differenzierbaren gemeinsamen Verteilungsfunktion F und W'keitsdichtefunktion f .

Wir setzen $X := X_1 + X_2$.

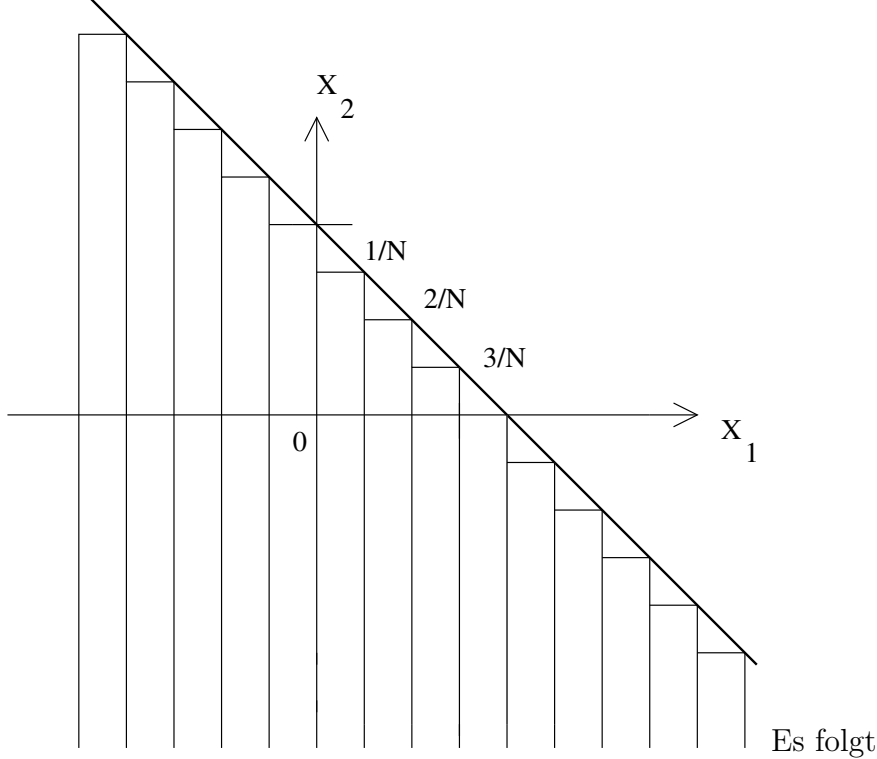
Dann sei für jedes $N \in \mathbb{Z}^+$:

$$A_N^L(s) := \bigcup_{k=-L \cdot N}^{L \cdot N} \left\{ \frac{k-1}{N} < X_1 \leq \frac{k}{N}, X_2 \leq s - \frac{k}{N} \right\} \subset \{X \leq s\}$$

und

$$A_N \subset A_{N+1} \nearrow \{X \leq s\} \cap \{-L < X_1 \leq L\}.$$

Im Bild:



$$P(\{X \leq s\} \cap \{-L < X_1 \leq L\}) = \lim_{N \rightarrow \infty} P(A_N^L)$$

Aber

$$\begin{aligned} P(A_N^L) &= \sum_{k=-L \cdot N}^{L \cdot N} P\left(\left\{\frac{k-1}{N} < X_1 \leq \frac{k}{N}, X_2 \leq s - \frac{k}{N}\right\}\right) \\ &= \sum_{k=-L \cdot N}^{L \cdot N} \left(F\left(\frac{k}{N}, s - \frac{k}{N}\right) - F\left(\frac{k-1}{N}, s - \frac{k}{N}\right) \right) \\ &= \sum_{k=-L \cdot N}^{L \cdot N} \int_{\frac{k-1}{N}}^{\frac{k}{N}} \frac{\partial F}{\partial t_1}\left(t_1, s - \frac{k}{N}\right) dt_1 \\ &= \sum_{k=-L \cdot N}^{L \cdot N} \int_{\frac{k-1}{N}}^{\frac{k}{N}} \left(\int_{-\infty}^{s - \frac{k}{N}} f(t_1, t_2) dt_2 \right) dt_1 \end{aligned}$$

für jede Wahl von N und L .

Nun ist aber

$$\int_{\frac{k-1}{N}}^{\frac{k}{N}} \left(\int_{-\infty}^{s - t_1 - \frac{1}{N}} f(t_1, t_2) dt_2 \right) dt_1 \leq \int_{\frac{k-1}{N}}^{\frac{k}{N}} \left(\int_{-\infty}^{s - \frac{k}{N}} f(t_1, t_2) dt_2 \right) dt_1 \leq \int_{\frac{k-1}{N}}^{\frac{k}{N}} \left(\int_{-\infty}^{s - t_1} f(t_1, t_2) dt_2 \right) dt_1$$

Summieren wir das über alle $k = -L \cdot N, \dots, L \cdot N$, so folgt

$$\begin{aligned} \int_{-L-1}^L \left(\int_{-\infty}^{s-t_1-\frac{1}{N}} f(t_1, t_2) dt_2 \right) dt_1 &\leq P(A_N^L) \\ &\leq \int_{-L-1}^L \left(\int_{-\infty}^{s-t_1} f(t_1, t_2) dt_2 \right) dt_1 \end{aligned}$$

Nun lassen wir erst $N \rightarrow \infty$ und dann $L \rightarrow \infty$ streben und finden

$$F^X(s) = P(X \leq s) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{s-t_1} f(t_1, t_2) dt_2 \right) dt_1 = \int_{-\infty}^{\infty} \left(\int_{-\infty}^s f(t_1, t_2 - t_1) dt_2 \right) dt_1$$

Durch Differenzieren nach s finden wir als W'keitsdichtefunktion

$$f^X(s) = \int_{-\infty}^{\infty} f(t_1, s - t_1) dt_1$$

Sollten nun die beiden Zufallsvariablen stochastisch unabhängig sein, so folgt, wenn f_j die W'keitsdichte von X_j bedeutet:

$$f^X(s) = \int_{-\infty}^{\infty} f_1(t) f_2(s - t) dt$$

Beispiele. (i) *Exponentialverteilungen.* X_1 und X_2 seien 2 exponentialverteilte Zufallsvariablen mit W'keitsdichten $f_\lambda(x) = \lambda e^{-\lambda x} h_0(x)$ und f_μ , (wobei $h_0(x) = 1$ für $x \geq 0$ und $h_0(x) = 0$ für $x < 0$ sein soll). Angenommen, es sei $\lambda \neq \mu$. Dann ist

$$\begin{aligned} f^X(s) &= \int_{-\infty}^{\infty} f_\lambda(t) f_\mu(s - t) dt = \lambda \mu \int_0^s e^{-\lambda t} e^{-\mu(s-t)} dt \\ &= \lambda \mu e^{-\mu s} \int_0^s e^{-(\lambda-\mu)t} dt = -\frac{\lambda \mu}{\lambda - \mu} e^{-\mu s} (e^{-(\lambda-\mu)s} - 1) \\ &= -\frac{\lambda \mu}{\lambda - \mu} (e^{-\lambda s} - e^{-\mu s}) \end{aligned}$$

Wenn $\lambda = \mu$ ist, so wird

$$f^X(s) = \lambda^2 s e^{-\lambda s}$$

(ii) *Normalverteilte Zufallsvariablen.* Ist X_1 nach $\mathcal{N}(\mu_1, \sigma_1)$ verteilt und X_2 nach $\mathcal{N}(\mu_2, \sigma_2)$, so errechnen wir für die Dichte der Verteilung von $X_1 + X_2$:

$$\frac{1}{2\pi} \frac{1}{\sigma_1 \sigma_2} \int_{-\infty}^{\infty} e^{-\frac{(s-\mu_1)^2}{2\sigma_1^2} - \frac{(t-\mu_2-s)^2}{2\sigma_2^2}} ds$$

Aber es ist

$$\begin{aligned} -\frac{(s-\mu_1)^2}{2\sigma_1^2} - \frac{(t-\mu_2-s)^2}{2\sigma_2^2} &= \frac{x^2}{2\sigma_1^2} + \frac{(t-(\mu_1+\mu_2)-x)^2}{2\sigma_2^2} \\ &= \frac{(t-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)} + \frac{1}{2} \frac{\sigma_1^2+\sigma_2^2}{(\sigma_1\sigma_2)^2} \left(x - \frac{\sigma_1^2(t-(\mu_1+\mu_2))}{\sigma_1^2+\sigma_2^2}\right)^2 \end{aligned}$$

wobei $x = s - \mu_1$ sein soll. Setzen wir das ein und benutzen die Substitutionsregel, folgt

$$\begin{aligned} \frac{1}{2\pi} \frac{1}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} e^{-\frac{(s-\mu_1)^2}{2\sigma_1^2} - \frac{(t-\mu_2-s)^2}{2\sigma_2^2}} ds &= \frac{1}{2\pi} \frac{1}{\sigma_1\sigma_2} e^{-\frac{(t-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \frac{\sigma_1^2+\sigma_2^2}{(\sigma_1\sigma_2)^2} x^2} dx \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{(t-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{-\infty}^{\infty} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{(t-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}} \end{aligned}$$

Damit können wir sehen, dass $X_1 + X_2$ wieder normalverteilt ist, und zwar nach

$$\mathcal{N}(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}).$$

In entsprechender Weise zeigt man:

3.1.9.5 Hilfssatz. Sind X_1, \dots, X_k endlich viele stochastisch unabhängige Zufallsvariable, welche nach $\mathcal{N}(\mu_j, \sigma_j)$ verteilt sind, so ist ihre Summe nach

$$\mathcal{N}(\mu_1 + \dots + \mu_k, \sqrt{\sigma_1^2 + \dots + \sigma_k^2})$$

normalverteilt.

Beispiel. Angenommen, ein Schüttgutfrachter soll mit mindestens 2000t Zinnerz beladen werden. Dieses soll in Eisenbahnwaggons angeliefert werden. Man weiß dass deren Lademengen X_1, \dots, X_k alle normalverteilt nach $\mathcal{N}(50, 8)$ sind. (Angabe in t). Wieviele Waggons müssen bereitgestellt werden, damit der Frachter mit einer W'keit von ≥ 0.99 mit der gewünschten Menge beladen werden kann?

Die Gesamtlademenge von k Waggons ist verteilt nach $\mathcal{N}(50k, 8\sqrt{k})$. Gesucht ist ein k , so dass

$$P(X_1 + \dots + X_k < 2000) \leq 0.01$$

Schreiben wir

$$Y_k := \frac{X_1 + \dots + X_k - 50k}{8\sqrt{k}}$$

so ist Y_k normalverteilt nach $\mathcal{N}(0, 1)$, und es muss gelten

$$P(Y_k < \frac{2000 - 50k}{8\sqrt{k}}) < 0.01$$

also

$$\text{Erf}(\frac{2000 - 50k}{8\sqrt{k}}) < 0.01$$

Nun ist aber (Tabelle) $\text{Erf}(-2.32) = 0.01$, also muss die Ungleichung

$$\frac{2000 - 50k}{8\sqrt{k}} < -2.32$$

gelöst werden. Die Zahl k erfüllt die (quadratische) Ungleichung

$$(k - 40.0689)^2 > 5.5, \text{ also } |k - 40.0689| > 2.345$$

Wir erhalten also $k \geq 43$. D. h.:

Es müssen mindestens 43 Waggons bereitgestellt werden, damit der Frachter mit der gewünschten Menge Erz beladen werden kann.

(iii) *Summen von Quadraten normalverteilter Zufallsvariablen.*

Definition. Wir sagen, eine nichtnegative Zufallsvariable X sei χ_n^2 -verteilt, wenn

$$P(X \leq t) = c_n \int_0^t x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx$$

Dabei ist c_n der Normierungsfaktor:

$$c_n = 2^{-\frac{n}{2}} \cdot \frac{1}{\Gamma(\frac{n}{2})}$$

Angenommen, X_1 und X_2 seien $\chi_{n_1}^2$ bzw. $\chi_{n_2}^2$ -verteilt, dann gilt

$$\begin{aligned} f^X(s) &= c_1 c_2 \int_0^s t^{\frac{n_1}{2}-1} e^{-t} (s-t)^{\frac{n_2}{2}-1} e^{-(s-t)} ds \\ &= c_1 c_2 e^{-s} \int_0^s t^{\frac{n_1}{2}-1} (s-t)^{\frac{n_2}{2}-1} ds \\ &= c_1 c_2 s^{\frac{n_1+n_2}{2}-1} e^{-s} \int_0^s \left(\frac{t}{s}\right)^{\frac{n_1}{2}-1} \left(1 - \frac{t}{s}\right)^{\frac{n_2}{2}-1} \frac{dt}{s} \\ &= c_1 c_2 s^{\frac{n_1+n_2}{2}-1} e^{-s} \int_0^1 x^{\frac{n_1}{2}-1} (1-x)^{\frac{n_2}{2}-1} dx \\ &= c_{12} s^{\frac{n_1+n_2}{2}-1} e^{-s} \end{aligned}$$

Dabei ist

$$c_{12} = c_1 c_2 \int_0^1 x^{\frac{n_1}{2}-1} (1-x)^{\frac{n_2}{2}-1} dx$$

Benutzen wir nun die Formel

$$\int_0^1 x^{k-1} (1-x)^{\ell-1} dx = \frac{\Gamma(k)\Gamma(\ell)}{\Gamma(k+\ell)}$$

finden wir mit der Formel für die Werte von c_1 und c_2 , dass

$$c_{12} = 2^{-\frac{n_1+n_2}{2}} \cdot \frac{1}{\Gamma(\frac{n_1+n_2}{2})}$$

wird.

Damit ist auch X wieder chi-quadrat-verteilt, und zwar mit $n_1 + n_2$ Freiheitsgraden.

Wo tritt die chi-quadrat-Verteilung auf?

3.1.9.6 Hilfssatz. Sind X_1, \dots, X_n stochastisch unabhängige $\mathcal{N}(0, 1)$ -verteilte Zufallsvariablen, so ist

$$X_1^2 + X_2^2 + \dots + X_n^2$$

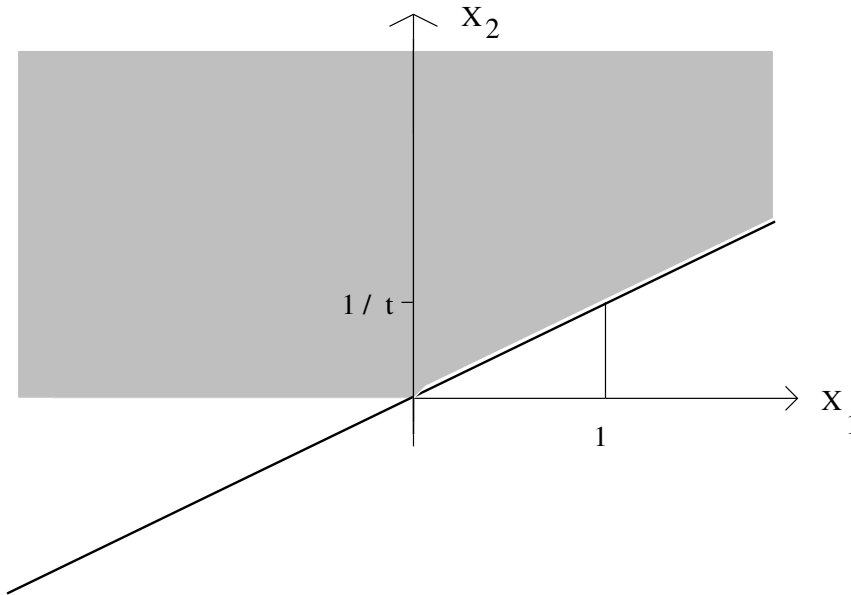
eine mit n Freiheitsgraden chi-quadrat-verteilte Zufallsvariable.

Beweis. Denn die Betrachtungen aus dem ersten Teil dieses Abschnittes ergeben, dass X_j mit einem Freiheitsgrad chi-quadrat-verteilt ist. Dann ist aber $X_1^2 + X_2^2$ mit 2 Freiheitsgraden chi-quadrat-verteilt, und, (wie man etwa durch ein Induktionsargument begründen kann) die Quadratsumme $X_1^2 + X_2^2 + \dots + X_n^2$ mit n Freiheitsgraden chi-quadrat-verteilt.

□

Quotient zweier Zufallsvariablen.

Angenommen, X_1, X_2 seien zwei Zufallsvariablen mit der gemeinsamen Verteilungsfunktion F und Verteilungsdichte f . Wir setzen weiter voraus, dass $X_2 > 0$ gelte. Dann hat für $t \in \mathbb{R}$ die Menge $\{X_1/X_2 \leq t\}$ die Gestalt:



Ähnlich wie bei der Verteilungsfunktion für die Summe von X_1 und X_2 finden wir für die Verteilungsfunktion von $X = \frac{X_1}{X_2}$:

$$P(X \leq t) = \int_{-\infty}^0 \left(\int_{ts_2}^{\infty} f(s_1, s_2) ds_1 \right) ds_2 + \int_0^{\infty} \left(\int_{-\infty}^{ts_2} f(s_1, s_2) ds_1 \right) ds_2$$

Durch Ableiten finden wir, dass

$$f^X(t) = \int_{-\infty}^{\infty} |s| f(ts, s) ds$$

die zugehörige W'keitsdichtefunktion ist.

Wenn auch X_1 nichtnegativ ist, folgt

$$F^X(t) = \int_0^{\infty} \left(\int_0^{ts_2} f(s_1, s_2) ds_1 \right) ds_2, \quad f^X(t) = \int_0^{\infty} s f(ts, s) ds$$

Setzen wir voraus, dass X_1 und X_2 sogar stochastisch unabhängig seien und Verteilungsdichtefunktionen f_1 bzw. f_2 besitzen, so folgt

$$P(X \leq t) = \int_{-\infty}^0 \left(\int_{ts_2}^{\infty} f_1(s_1) f_2(s_2) ds_1 \right) ds_2 + \int_0^{\infty} \left(\int_{-\infty}^{ts_2} f_1(s_1) f_2(s_2) ds_1 \right) ds_2$$

und

$$f^X(t) = \int_{-\infty}^{\infty} |s| f_1(ts) f_2(s) ds$$

Beispiel: Die Studentsche t-Verteilung

Der Name "Student" wurde von W. Gosset als Pseudonym verwendet. Nach ihm ist die als t_n -Verteilung bezeichnete W'keitsverteilung benannt.

Ist Y_1 eine $\mathcal{N}(0, 1)$ -verteilte Zufallsvariable und Y_2 mit Y_1 stochastisch unabhängig, nicht-negativ und χ_n^2 -verteilt, so ist

$$T := \sqrt{n} \frac{Y_1}{\sqrt{Y_2}}$$

eine t_n -verteilte Zufallsvariable.

Wir berechnen ihre Dichtefunktion f_{t_n} .

Die W'keitsdichte zu Y_1 ist

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

und die W'dichte zu Y_2 ist gegeben durch

$$f_2(x) = c_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad c_n = 2^{-\frac{n}{2}} \cdot \frac{1}{\Gamma(\frac{n}{2})}$$

Daraus bestimmen wir die Dichtefunktion zu $\sqrt{Y_2}$:

$$P(\sqrt{Y_2} \leq t) = P(Y_2 \leq t^2) = \int_0^{t^2} f_2(x) dx = 2 \int_0^t f_2(y^2) y dy$$

Also

$$f^{\sqrt{Y_2}}(y) = 2f_2(y^2)y = 2c_n y^{n-1} e^{-\frac{y^2}{2}}$$

Die Formel für die W'keitsdichte zu $\frac{1}{\sqrt{n}}T$ liefert jetzt

$$\begin{aligned} f^{\frac{1}{\sqrt{n}}T}(t) &= \int_0^{\infty} s f_1(ts) f_2(s) ds \\ &= \frac{2}{\sqrt{2\pi}} c_n \int_0^{\infty} s^n e^{-\frac{(ts)^2}{2}} e^{-\frac{s^2}{2}} ds \\ &= \frac{2}{\sqrt{2\pi}} c_n \int_0^{\infty} s^n e^{-\frac{\sqrt{1+t^2}s^2}{2}} ds = \sqrt{\frac{2}{\pi}} c_n \frac{1}{(1+t^2)^{\frac{n+1}{2}}} \int_0^{\infty} x^n e^{-x^2/2} dx \\ &= \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}) \sqrt{\pi}} \frac{1}{(1+t^2)^{\frac{n+1}{2}}} \end{aligned}$$

Hieraus folgt leicht

$$f^T(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \frac{1}{(1 + \frac{t^2}{n})^{\frac{n+1}{2}}}$$

Anwendungsbeispiel: Schätzung von Erwartungswerten normalverteilter Zufallsvariablen auf der Basis von Stichproben vom Umfang n

Angenommen, es seien n normalverteilte stochastisch unabhängige Zufallsvariablen X_1, \dots, X_n gegeben mit Erwartungswert μ und Varianz σ . Mit $\bar{X}_{(n)}$ bezeichnen wir das arithmetische Mittel

$$\bar{X}_{(n)} = \frac{1}{n}(X_1 + \dots + X_n)$$

Die folgende Funktion ist eine geeignet zur Schätzung der Streuung (mehr dazu im nächsten Kapitel)

$$\overline{S}_{(n)} := \sqrt{\frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X}_{(n)})^2}$$

Frage: Was ist die Verteilungsdichte von

$$T_n := \sqrt{n} \frac{(\bar{X}_{(n)} - \mu)}{S_{(n)}} ?$$

Wir benötigen einen

3.1.9.7 Hilfssatz. Die Funktion $\frac{n-1}{\sigma^2} \overline{S}_{(n)}^2$ ist χ_{n-1} -verteilt.

Beweis. Es gilt

$$\frac{n-1}{\sigma^2} \overline{S}_{(n)}^2 = \sum_{j=1}^n \left(\frac{X_j - \bar{X}_{(n)}}{\sigma} \right)^2 = \sum_{j=1}^n (U_j - \bar{U}_{(n)})^2$$

wobei

$$U_j = \frac{X_j - \mu}{\sigma}, \quad \bar{U}_{(n)} = \frac{1}{n} \sum_{j=1}^n U_j$$

Jedes U_j ist $\mathcal{N}(0,1)$ -verteilt. Weiter sind auch die U_j paarweise stochastisch unabhängig. Die Funktionen $U_j - \bar{U}_{(n)}$ in der Summe auf der rechten Seite sind es aber nicht. Jedoch kann man zeigen, dass sich diese Summe zu einer Summe von Quadraten von $n-1$ stochastisch unabhängigen $\mathcal{N}(0,1)$ -verteilten Zufallsvariablen umformen lässt.

Zum Beispiel ist für $n=2$ diese Summe gerade $\left(\frac{U_1 - U_2}{\sqrt{2}}\right)^2$ und für $n=3$ hat sie die Gestalt $Z_1^2 + Z_2^2$, wobei

$$Z_1 = \frac{1}{\sqrt{2}}(U_1 - U_2), \quad Z_2 = \frac{1}{\sqrt{6}}(U_1 + U_2 - 2U_3)$$

Damit wird also $\sum_{j=1}^n (U_j - \bar{U}_{(n)})^2$ eine χ_{n-1}^2 verteilte Zufallsvariable.

□

Auf diese Weise sehen wir, dass

$$T_n = \sqrt{n-1} \frac{\frac{\bar{X}_{(n)} - \mu}{\sigma/\sqrt{n}}}{\frac{\sqrt{n-1}}{\sigma} \bar{S}_{(n)}}$$

dargestellt als $\sqrt{n-1}$ mal dem Quotienten einer $\mathcal{N}(0,1)$ -verteilten und der Wurzel einer χ_{n-1}^2 -verteilten Zufallsgröße, also t_{n-1} -verteilt ist.

3.1.10 Gesetze der großen Zahlen

Definition. Angenommen, wir haben eine Folge $(X_n)_n$ von Zufallsvariablen, für die der Erwartungswert $\mu_n := \mathcal{E}(X_n)$ definiert ist. Für $m \geq 1$ setzen wir

$$\bar{X}_{(m)} := \frac{1}{m} \sum_{j=1}^m X_j$$

Dann sagen wir, es gelte für $(X_n)_n$ das (*schwache*) *Gesetz der großen Zahlen*, wenn für jedes positive ε

$$\lim_{m \rightarrow \infty} P \left(\left| \bar{X}_{(m)} - \frac{1}{m} \sum_{j=1}^m \mu_j \right| > \varepsilon \right) = 0$$

Wir erinnern uns an die Tschebyscheff-Ungleichung. Wenn wir annehmen, dass die Zufallsvariablen X_n sogar eine endliche Varianz besitzen, so gilt

$$\begin{aligned} P \left(\left| \bar{X}_{(m)} - \frac{1}{m} \sum_{j=1}^m \mu_j \right| > \varepsilon \right) &\leq \frac{1}{\varepsilon^2} \text{Var} \left(\bar{X}_{(m)} - \frac{1}{m} \sum_{j=1}^m \mu_j \right) \\ &= \frac{1}{m^2 \varepsilon^2} \text{Var} \left(\sum_{j=1}^m (X_j - \mu_j) \right) \\ &= \frac{1}{m^2 \varepsilon^2} \text{Var} \left(\sum_{j=1}^m X_j \right) \end{aligned}$$

Hieraus können wir zwei Sätze ableiten:

3.1.10.1 Satz (*Tschebyscheff*) Angenommen, die Zufallsvariablen $(X_n)_n$ haben eine Varianz und sind paarweise unkorreliert, also

$$\text{Cov}((X_n - \mu_n)(X_k - \mu_k)) = 0, \quad n \neq k$$

Gibt es dann eine Schranke S für die Varianzen von X_n , also

$$\text{Var}(X_n) \leq S, \text{ für alle } n,$$

so gilt für die Folge $(X_n)_n$ das Gesetz der großen Zahlen.

Beweis. Denn ist $\varepsilon > 0$ beliebig vorgegeben, so haben wir

$$\begin{aligned} P\left(\left|\bar{X}_{(m)} - \frac{1}{m} \sum_{j=1}^m \mu_j\right| > \varepsilon\right) &\leq \frac{1}{m^2 \varepsilon^2} \text{Var}\left(\sum_{j=1}^m X_j\right) \\ &= \frac{1}{m^2 \varepsilon^2} \sum_{j=1}^m \text{Var}(X_j) \\ &\leq \frac{S}{\varepsilon^2} \frac{1}{m} \longrightarrow 0 \end{aligned}$$

wenn $m \longrightarrow \infty$.

□

Hieraus folgt der Satz von Poisson:

3.1.10.2 Satz (Poisson). Angenommen, jede der Zufallsvariablen X_n nehme nur die Werte 0 oder 1 an und sei binomialverteilt mit Parameter 1 und p_n , also $P(X_n = 1) = p_n$. Sind dann die X_n paarweise stochastisch unabhängig, so gilt das Gesetz der großen Zahlen für $(X_n)_n$.

Beweis. Denn $\text{Var}(X_n) = p_n(1 - p_n) \leq 1$ für alle n . Der Satz von Tschebyscheff ist anwendbar.

□

Ein Spezialfall dieses Satzes ist das Gesetz von Bernoulli:

3.1.10.3 Satz (Bernoulli). Angenommen, es sei $m \geq 1$ und man führe ein Zufallsexperiment m -mal unter gleichen Bedingungen durch, wobei die Wiederholungen unabhängig voneinander sein sollen. Es sei A eines der möglichen Ereignisse, die bei jeder Wiederholung eintreten können. Dann konvergiert die relative Häufigkeit mit der dabei A eingetreten, ist nach W 'keit gegen die W 'keit $P(A)$ für das Ereignis A . Das heißt: Ist

$$h_m(A) := \frac{1}{m} \text{Anzahl der Male, bei denen } A \text{ eingetreten ist}$$

so folgt für jedes beliebige positive ε

$$P(|h_m(A) - P(A)| > \varepsilon) \longrightarrow 0, \text{ wenn } m \longrightarrow \infty$$

Beweis. Wir überlegen, was die Situation dieses Satzes mit dem Poissonschen Gesetz zu tun hat. Dazu wählen wir im Poisson-Satz X_n als die Zufallsvariable, die die Werte 0 oder 1 annehmen kann, und zwar hat X_n den Wert 1, wenn A beim n -ten Versuch eintritt und den Wert 0 sonst.

Wir setzen $p_n = P(A)$ für jedes n . Dann ist für alle n der Erwartungswert von X_n gerade $P(A)$, und weiter

$$X_{(m)} = h_m(A)$$

Der Satz von Poisson ergibt jetzt die Behauptung. □

Beispiel. Dieser Satz rechtfertigt es zum Beispiel, beim Münzwurf die W'keit, dass "Zahl" geworfen wird, dadurch zu approximieren, dass man sehr oft wirft und dann die relative Häufigkeit für "Zahl" verwendet.

3.1.11 Der zentrale Grenzwertsatz

Angenommen, wir haben eine Folge $(X_n)_n$ von Zufallsvariablen gegeben. Lässt sich dann etwas aussagen über das Verhalten der Verteilungen der X_n , wenn $n \rightarrow \infty$?

Die Antwort darauf ist der "zentrale Grenzwertsatz". Zu seiner Formulierung vereinbaren wir eine Sprechweise:

Definition. a) Eine Funktion $F : \mathbb{R} \rightarrow [0, 1]$ heißt eine Verteilungsfunktion, wenn sie die folgenden Eigenschaften hat

- F ist rechtsseitig stetig, also $\lim_{h \rightarrow 0+} F(x+h) = F(x)$ für alle x ,
- $\lim_{t \rightarrow -\infty} F(t) = 0$
- $\lim_{t \rightarrow \infty} F(t) = 1$

b) Gegeben sei eine Verteilungsfunktion F auf \mathbb{R} . Ist dann $(X_n)_n$ eine Folge von Zufallsvariablen und F_n die Verteilungsfunktion von X_n , so sagen wir F_n strebe gegen F , wenn in allen Stetigkeitsstellen x von F gilt

$$\lim_{n \rightarrow \infty} F_n(x) = F(x).$$

Dann gilt:

3.1.11.1 Satz. Gegeben sei eine Verteilungsfunktion F auf \mathbb{R} und eine Folge $(X_n)_n$ von Zufallsvariablen mit Verteilungsfunktion F_n , die gegen eine Verteilungsfunktion F strebe. Dann gilt für jedes Intervall $I = (a, b)$, dass

$$\lim_{n \rightarrow \infty} P(X_n \in I) = \begin{cases} F(b) - F(a) & , \text{ wenn } -\infty < a < b < \infty \\ F(b) & , \text{ wenn } a = -\infty \\ 1 - F(a) & , \text{ wenn } b = \infty \end{cases} ,$$

Der folgende Satz von Glivenko-Cantelli stellt eine Verbindung zwischen relativen Häufigkeiten, dem Gesetz der großen Zahlen und Verteilungsfunktionen her.

3.1.11.2 Satz (Glivenko-Cantelli). Angenommen, es sei $(X_n)_n$ eine Folge von stochastisch unabhängigen und identisch verteilten Zufallsgrößen, d.h.: Alle X_n haben dieselben Verteilungsfunktion F . Für $t \in \mathbb{R}$ sei dann $h_n(t, X_1, \dots, X_n) := \frac{k}{n}$, wenn k der Zufallsvariablen X_1, \dots, X_n einen Wert kleiner oder gleich t annehmen. Dann konvergiert $h_n(t, X_1, \dots, X_n)$ gegen $F(t)$, wenn $n \rightarrow \infty$.

Die Voraussetzungen dieses Satzes sind etwa bei unabhängigen Wiederholungen einer Messung irgendeiner physikalischen Größe erfüllt. Wenn wir also genügend oft messen, so liefert uns die Kurve der relativen Häufigkeiten auf empirischem Wege die Verteilungsfunktion der in Rede stehenden Messgröße.

Unter vergleichsweise milden Voraussetzungen nun kann man zeigen, dass bei Folgen stochastisch unabhängiger Zufallsvariablen die Verteilungsfunktionen näherungsweise durch die Normalverteilung darstellbar ist.

3.1.11.3 Satz (Zentraler Grenzwertsatz). Angenommen, es sei $(X_n)_n$ eine Folge von stochastisch unabhängigen Zufallsvariablen mit Erwartungswerten $\mathcal{E}(X_n) = \mu_n$ und Varianz $\text{Var}(X_n) = \sigma_n^2$. Ferner nehmen wir an, es gebe ein $\delta > 0$, so dass der Erwartungswert von $|X_n - \mu_n|^{2+2\delta}$ definiert sei, und dass

$$\frac{\sum_{j=1}^n \mathcal{E}(|X_j - \mu_n|^{2+2\delta})}{\left(\sum_{j=1}^n \sigma_j^2\right)^{1+\delta}} \rightarrow 0, \quad \text{wenn } n \rightarrow \infty \quad (\text{Ljapunoff - Bedingung})$$

Dann gilt

$$P(\{U_n \leq t\}) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx, \quad \text{wenn } n \rightarrow \infty$$

wobei

$$U_n = \frac{\sum_{j=1}^n (X_j - \mu_j)}{\sqrt{\sum_{j=1}^n \sigma_j^2}}$$

Der Satz greift in fast allen praxisrelevanten Situationen.

Beispiel. a) Sind alle X_n identisch verteilt (etwa bei unabhängigen Wiederholung einer Messung), so hängen $\mathcal{E}(X_n)$, $\text{Var}(X_n)$ und $\mathcal{E}(|X_n - \mu_n|^{2+2\delta})$ nicht von n ab, haben also feste Werte μ, σ^2 bzw. E_δ . Ferner ist

$$\left(\sum_{j=1}^n \sigma_j^2\right)^{1+\delta} = n^{1+\delta} \sigma^2$$

Somit wird die Ljapunoffbedingung erfüllt:

$$\frac{\sum_{j=1}^n \mathcal{E}(|X_j - \mu_n|^{2+2\delta})}{\left(\sum_{j=1}^n \sigma_j^2\right)^{1+\delta}} = \frac{nE_\delta}{n^{1+\delta} \sigma^{2+2\delta}} = \frac{E_\delta}{n^\delta \sigma^{2+2\delta}} \rightarrow 0, \quad \text{wenn } n \rightarrow \infty$$

Somit streben die Verteilungen von

$$U_n = \frac{\sum_{j=1}^n (X_j - \mu)}{\sigma\sqrt{n}}$$

gegen die $\mathcal{N}(0, 1)$ -Verteilung.

b) Angenommen, $p \in (0, 1)$ und alle X_n haben nur die Werte 0 und 1, so dass $P(\{X = 1\}) = p$, $P(\{X = 0\}) = 1 - p$, so gilt, wenn wieder alle X_n als stochastisch unabhängig vorausgesetzt werden:

$$P\left(\frac{\sum_{j=1}^n X_j - np}{\sqrt{np(1-p)}} \leq t\right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx, \quad \text{wenn } n \rightarrow \infty$$

Das ist die Aussage des Grenzwertsatzes von Moivre-Laplace.

Man kann sagen:

Regel: Man kann diesen Satz anwenden, wenn $n > \frac{4}{p}$ und $n > \frac{4}{1-p}$.

3.1.12 Konfidenzschranken

Wir befassen uns nun mit Problemen vom folgenden Typ:

Angenommen, es sei eine Familie $(P_\vartheta)_{\vartheta \in T}$ von W'keitsverteilungen vorgegeben, wobei T eine Teilmenge von \mathbb{R} oder \mathbb{R}^d für irgendein $d \in \mathbb{Z}^+$ sein soll. Angenommen weiter, eine Zufallsvariable X unterliege einer W'keitsverteilung aus dieser Familie $(P_\vartheta)_{\vartheta \in T}$. Wie kann man einen möglichst sicheren Schätzwert des Parameters ϑ für die W'keitsverteilung unserer Zufallsvariablen ermitteln?

Dies soll an Beispielen erläutert werden:

1) *Qualitätskontrolle.* Eine Lieferung von elektronischen Bauteilen wird auf Qualität überprüft. Angenommen, der Hersteller garantiere eine Ausschussquote von maximal 4%. Man entnimmt eine Stichprobe von 300 Stück und findet darunter 15 defekte Bauteile. Die relevante Familie von W'keitsverteilungen ist jetzt, was wegen des ausreichend großen Umfanges der Stichprobe sinnvoll ist, die Familie der Normalverteilungen $\mathcal{N}(300p, \sqrt{300p(1-p)})$. Zu schätzen ist p (die Ausschussquote).

2) *Sollwertkontrolle.* Bei der Produktion von Zylinderscheiben soll überprüft werden, ob ein Sollwert für deren Durchmesser, etwa 20.2mm, eingehalten wird. Angenommen, man entnimmt aus der laufenden Produktion 16 Stück und stellt fest, dass der mittlere Durchmesser 20.6 mm beträgt, wobei ein Toleranzspanne von 0.5 mm eingeräumt wird. Kann man jetzt mit genügender Sicherheit sagen, dass der Sollwert eingehalten wurde?

Die zu testende Familie von Verteilungen ist $(\mathcal{N}(\mu, \sigma_0^2))_\mu$, wobei $\sigma_0 = 0.05$ ist. Zur Diskussion steht die Hypothese $\mu = 20.2$.

3) *Benzinverbrauch eines Autos.* Angenommen, ein Autotyp A wird angeboten, von dem der Hersteller versichert, dass der Benzinverbrauch nicht höher als $6.4 \ell/100 \text{ km}$ sei.

Es kann davon ausgegangen werden, dass unsere Zufallsvariable X , das ist hier der Benzinverbrauch, normalverteilt sei. Die zur Diskussion stehende Familie ist also die Familie der Normalverteilungen $\mathcal{N}(\mu, \sigma^2)$. Man will μ schätzen, genauer: Man will wissen, ob mit annehmbarer W'keit der Wert von μ unterhalb 6.4 liegt.

Man muss sagen:

Grundsätzlich gibt es kein Rechenverfahren, mit dem man ausrechnen kann, ob eine Behauptung über einen Parameter einer W'keitsverteilung wahr ist.

Wohl aber kann man zu einer vorgegebenen Zahl $\alpha \in (0, 1)$ eine obere (untere) KONFIDENZ-SCHRANKE für den fraglichen Parameter berechnen. Dies ist eine Schranke s_o (bezw. s_u) mit der Eigenschaft, dass mit einer W'keit von $1 - \alpha$ der Parameter unterhalb s_o (bezw. oberhalb s_u) gelegen ist.

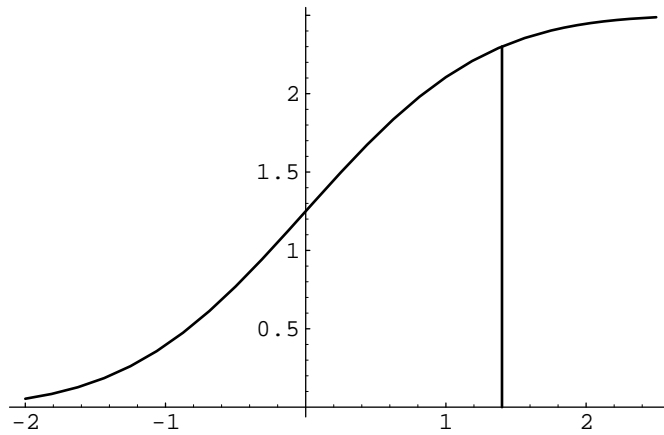
Wir wenden uns der Berechnung dieser Konfidenzschranken zu.

Definition. Ist F eine stetige streng monotone Verteilungsfunktion und ε eine Zahl aus $(0, 1)$, so nennen wir die Lösung q_ε der Gleichung

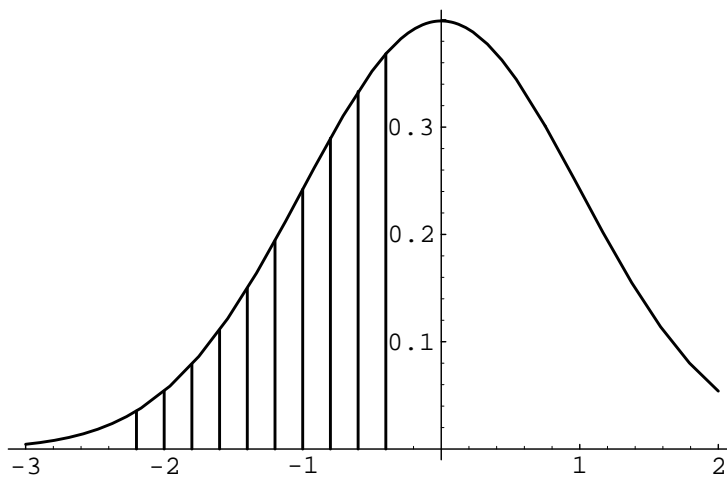
$$F(q_\varepsilon) = \varepsilon$$

das ε -Quantil oder ε -Fraktil der Verteilung F .

Im Bild:



In einer Graphik mit der Wahrscheinlichkeitsdichtefunktion würde das ε -Fraktile folgende Bedeutung haben:



Der Flächeninhalt der schraffierten Fläche ist ε .

Für die einzelnen Verteilungen (Normalverteilung, chi-quadrat-Verteilungen, oder Student-Verteilungen) ist das ε -Fraktile tabelliert.

Hier ist ein kleiner Auszug:

Normalverteilung $\mathcal{N}(0, 1)$

| | | | | | | |
|-----------------|--------|--------|--------|--------|--------|--------|
| ε | 0.001 | 0.005 | 0.01 | 0.025 | 0.05 | 0.1 |
| q_ε | -3.090 | -2.576 | -2.326 | -1.960 | -1.645 | -1.282 |

Wegen der Symmetrie der Dichtefunktion haben wir

$$\text{Erf}(t) = 1 - \text{Erf}(-t)$$

woraus für die Fraktile folgt:

$$q_{1-\varepsilon} = -q_\varepsilon, \quad q_\varepsilon = -q_{1-\varepsilon}$$

Damit ergibt sich die folgende Tabelle:

| | | | | | | |
|-----------------|-------|-------|-------|-------|-------|-------|
| ε | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
| q_ε | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 3.090 |

Für die Chi-quadrat-Verteilung erhalten wir:

| k | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\varepsilon = 0.025$ | 3.25 | 3.82 | 4.4 | 5.01 | 5.63 | 6.26 | 6.91 | 7.56 | 8.23 | 8.91 | 9.59 |
| $\varepsilon = 0.975$ | 20.48 | 21.92 | 23.34 | 24.74 | 26.12 | 27.49 | 28.85 | 30.19 | 31.53 | 32.85 | 34.17 |
| $\varepsilon = 0.005$ | 2.16 | 2.60 | 3.07 | 3.57 | 4.07 | 4.60 | 5.14 | 5.70 | 6.26 | 6.84 | 7.43 |
| $\varepsilon = 0.995$ | 25.19 | 26.76 | 28.30 | 29.82 | 31.32 | 32.80 | 34.27 | 35.72 | 37.16 | 38.58 | 40.00 |

Für die Student-Verteilung erhalten wir:

| k | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\varepsilon = 0.90$ | 1.372 | 1.363 | 1.356 | 1.35 | 1.345 | 1.341 | 1.337 | 1.333 | 1.330 | 1.328 | 1.325 |
| $\varepsilon = 0.95$ | 1.812 | 1.796 | 1.782 | 1.771 | 1.761 | 1.753 | 1.746 | 1.740 | 1.734 | 1.729 | 1.725 |
| $\varepsilon = 0.975$ | 2.228 | 2.201 | 2.179 | 2.160 | 2.145 | 2.131 | 2.120 | 2.220 | 2.101 | 2.093 | 2.086 |
| $\varepsilon = 0.99$ | 2.764 | 2.718 | 2.681 | 2.650 | 2.624 | 2.602 | 2.583 | 2.567 | 2.552 | 2.539 | 2.528 |
| 0.995 | 3.169 | 3.106 | 3.055 | 3.012 | 2.977 | 2.947 | 2.921 | 2.898 | 2.878 | 2.861 | 2.845 |

Die Dichtefunktion zur Studentverteilung ist symmetrisch (zur y -Achse), so dass für die Fraktile wieder gilt

$$q_{1-\varepsilon}^{t_k} = -q_\varepsilon^{t_k}, \quad q_\varepsilon^{t_k} = -q_{1-\varepsilon}^{t_k}$$

Konfidenzschranken und -intervall für den Mittelwert einer normalverteilten Zufallsvariablen

1. Fall: Die Varianz σ^2 sei bekannt

Angenommen, wir haben n unabhängige nach $\mathcal{N}(\mu, \sigma^2)$ verteilte Zufallsvariablen X_1, \dots, X_n gegeben. Dann setzen wir

$$U_n := \frac{\bar{X}_{(n)} - \mu}{\sigma/\sqrt{n}},$$

wobei wieder

$$\bar{X}_{(n)} = \frac{1}{n} \sum_{j=1}^n X_j$$

Ist dann q_ε das ε -Fraktile der Normalverteilung, so gilt

$$P(U_n \geq q_\alpha) = 1 - \alpha$$

Das bedeutet:

Mit W'keit $1 - \alpha$ besteht die Beziehung $U_n \geq q_\alpha$.

Formen wir das um, folgt:

$$\mu \leq \bar{X}_{(n)} - \frac{\sigma}{\sqrt{n}} q_\alpha = \bar{X}_{(n)} + \frac{\sigma}{\sqrt{n}} q_{1-\alpha}$$

Wissen wir also, dass die X_j gewisse Werte x_j annehmen, so können wir schließen, dass mit der W'keit $1 - \alpha$ der Mittelwert μ durch

$$\mu \leq \frac{1}{n} \sum_{j=1}^n x_j + \frac{\sigma}{\sqrt{n}} q_{1-\alpha}$$

abgeschätzt werden kann.

Die rechte Seite nennt man *obere Konfidenzschranke* g_o für μ zum Niveau $1 - \alpha$.

Genauso findet man wegen

$$P(U_n \leq q_{1-\alpha}) = 1 - \alpha$$

dass mit W'keit $1 - \alpha$ die Abschätzung $U_n \leq q_{1-\alpha}$, also

$$\mu \geq \bar{X}_{(n)} - \frac{\sigma}{\sqrt{n}} q_{1-\alpha}$$

besteht. Haben die X_j die Werte x_j angenommen, so führt dies auf die *untere Konfidenzschranke*

$$g_u(x_1, \dots, x_n) := \frac{1}{n} \sum_{j=1}^n x_j - \frac{\sigma}{\sqrt{n}} q_{1-\alpha}$$

zum Niveau $1 - \alpha$.

Für das Intervall

$$I_\alpha := [-q_{1-\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}]$$

gilt die Beziehung

$$P(U_n \in I_\alpha) = 1 - \alpha$$

also kann man sagen, dass mit W'keit $1 - \alpha$ die Ungleichung

$$\bar{X}_{(n)} - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \leq \bar{X}_{(n)} + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}$$

besteht.

Damit finden wir, dass mit W'keit $1 - \alpha$ der wahre Wert von μ im Intervall

$$I = \left[\bar{X}_{(n)} - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}, \bar{X}_{(n)} + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \right]$$

liegt. Liegen konkrete Werte x_j für die X_j vor, so finden wir mit

$$I = \left[\frac{1}{n} \sum_{j=1}^n x_j - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}, \frac{1}{n} \sum_{j=1}^n x_j + \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} \right]$$

ein *Konfidenzintervall* zum Niveau $1 - \alpha$ für μ .

Dazu ein

Beispiel. Sollwertkontrolle. Bei der Produktion von Zylinderscheiben soll überprüft werden, ob ein Sollwert für deren Durchmesser, etwa 20.2mm, eingehalten wird. Angenommen, man entnimmt aus der laufenden Produktion 16 Stück und stellt fest, dass der mittlere Durchmesser 20.6 mm beträgt, wobei ein Toleranzspanne von 0.5 mm eingeräumt wird. Kann man jetzt mit genügender Sicherheit sagen, dass der Sollwert eingehalten wurde?

Nun ist $n = 16$ und X_j der Durchmesser der j -ten Zylinderscheibe. Eine Schätzung des Parameters μ ist gesucht. Die Varianz ist $\sigma^2 = 0.25$. Wir wählen etwa $\alpha = 0.05$. Dann erhalten wir als Konfidenzschranken für μ :

$$g_0 = 20.6 + \frac{0.5}{4} q_{0.95} = 20.6 + \frac{0.5}{4} \cdot 1.645 = 20.806$$

und

$$g_u = 20.6 - \frac{0.5}{4} \cdot 1.645 = 20.39$$

Als Konfidenzintervall zum Niveau 0.95 ergibt sich

$$I = \left[20.6 - \frac{0.5}{4} \cdot q_{0.975}, 20.6 + \frac{0.5}{4} \cdot q_{0.975} \right] = [20.335, 20.845]$$

Auf der Grundlage der Stichprobe muss man daher urteilen, dass der wahre Mittelwert innerhalb I , bzw. oberhalb g_u liegt, also der Sollwert von 20.2 mm nicht eingehalten wird. Dieses Urteil ist mit Sicherheitsw'keit 95% abgesichert!