Skript zur Vorlesung

Analysis IV

für Physiker

SS 2007

Peter Junghanns

Hinweis: Das vorliegende Skript stellt nur ein Gerüst zu den Inhalten der Vorlesung dar. Die Vorlesung selbst bietet weiterführende Erläuterungen, Beweise und die ausführliche Behandlung der Beispiele.

Inhaltsverzeichnis

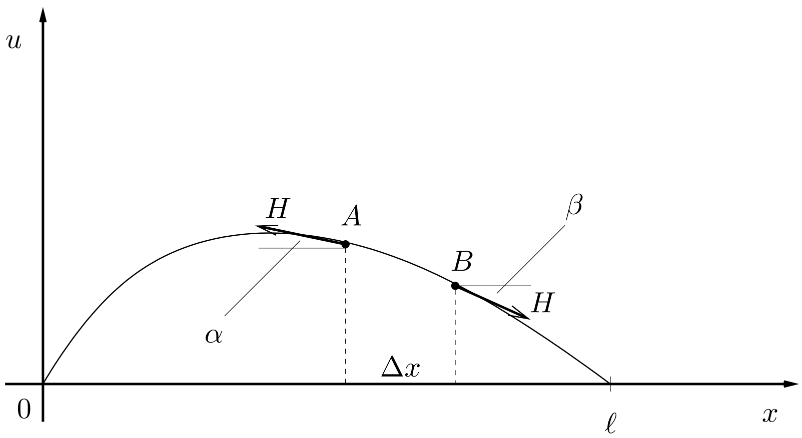
Kapitel 5

Gewöhnliche Differentialgleichungen. Dynamische Systeme (Fortsetzung)

5.11 Randwertaufgaben

5.11.1 Das Problem der schwingenden Saite

Wir setzen voraus, dass das Gewicht der Saite vernachlässigt werden kann und dass nur kleine Auslenkungen zu betrachten sind (d.h. Längen- und Spannungsänderungen sind vernachlässigbar).



Die gesuchte Auslenkung der Saite im Punkt x zum Zeitpunkt t bezeichnen wir mit $u(x,t)$. Die Summe der vertikalen Komponenten der Kräfte, die in den Punkten A und B wirken, ist gleich $F = H(\sin\alpha - \sin\beta)$. Da α und β nach Voraussetzung klein sind, gilt $\sin\alpha \approx \tan\alpha$ und $\sin\beta \approx \tan\beta$ und somit nach dem Mittelwertsatz der Differentialrechnung

$$F \approx H\left(\left.\frac{\partial u}{\partial x}\right|_B - \left.\frac{\partial u}{\partial x}\right|_A\right) = H \left.\frac{\partial^2 u}{\partial x^2}\right|_C \Delta x\,.$$

Bezeichnen wir mit ρ die Dichte der Saite, so muss $F \approx \rho\,\Delta x \left.\dfrac{\partial^2 u}{\partial t^2}\right|_C$ gelten. Für $\Delta x \longrightarrow 0$ erhalten wir die (partielle) Differentialgleichung

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \tag{5.15}$$

mit einer positiven Konstanten $a^2 = \dfrac{H}{\rho}$ und den Randbedingungen

$$u(0,t) = u(\ell,t) = 0\,, \quad t \ge 0\,, \tag{5.16}$$

sowie den Anfangsbedingungen

$$u(x,0) = f(x)\,, \quad \frac{\partial u(x,0)}{\partial t} = g(x)\,, \quad 0 \le x \le \ell\,. \tag{5.17}$$

Wir normieren ℓ auf π und machen den Ansatz (**Fouriersche Methode**)

$$u(x,t) = v(x)w(t)\,.$$

Aus (5.15) folgt

$$v(x)w''(t) = a^2 v''(x)w(t)\,,$$

so dass

$$\frac{w''(t)}{a^2 w(t)} = \frac{v''(x)}{v(x)} = \text{const} =: -\lambda$$

gelten muss. Wir erhalten die gewöhnlichen Differentialgleichungen

$$w''(t) + a^2\lambda w(t) = 0\,, \quad v''(x) + \lambda v(x) = 0\,. \tag{5.18}$$

Die Einbeziehung der Randbedingungen (5.16) liefert für die zweite dieser Differentialgleichungen die Randwertaufgabe

$$v'' + \lambda v = 0\,, \quad v(0) = v(\pi) = 0\,, \tag{5.19}$$

welches eigentlich ein **Eigenwertproblem** ist, da nur solche $\lambda \in \mathbb{R}$ interessieren, für die nichttriviale Lösungen von (5.19) existieren. Die allgemeine Lösung der Differentialgleichung in (5.19) lautet

(a) $c_1 e^{\sqrt{-\lambda}x} + c_2 e^{-\sqrt{-\lambda}x}$ für $\lambda < 0$,

(b) $c_1 x + c_2$ für $\lambda = 0$,

(c) $c_1 \cos\sqrt{\lambda}x + c_2 \sin\sqrt{\lambda}x$ für $\lambda > 0$.

Schreiben wir diese allgemeinen Lösungen in der Form $c_1 f_1(x) + c_2 f_2(x)$, so sind die Randbedingungen in (5.19) gleichbedeutend mit

$$c_1 f_1(0) + c_2 f_2(0) = 0\,, \quad c_1 f_1(\pi) + c_2 f_2(\pi) = 0\,.$$

Es existieren also genau dann nichttriviale Lösungen von (5.19), wenn

$$D = \det \begin{bmatrix} f_1(0) & f_2(0) \\ f_1(\pi) & f_2(\pi) \end{bmatrix} = 0$$

gilt. Für die obigen drei Fälle erhalten wir

(a) $D = e^{-\sqrt{-\lambda}\pi} - e^{\sqrt{-\lambda}\pi} \neq 0\,,$

(b) $D = -\pi \neq 0\,,$

(c) $D = \sin\sqrt{\lambda}\pi = 0 \quad\Longleftrightarrow\quad \lambda = \lambda_n = n^2\,,\ n = 1, 2, \ldots$

Die Zahlen λ_n nennt man **Eigenwerte** von (5.19). Die Funktionen $v_n(x) = \sin nx$ sind die entsprechenden **Eigenfunktionen**. Die dazugehörigen Funktionen $w_n(t)$ (vgl. (5.18)) finden wir als allgemeine Lösungen der Differentialgleichung $w'' + (an)^2 w = 0$, d.h.

$$w_n(t) = A_n \cos(ant) + B_n \sin(ant)$$

mit beliebigen Konstanten $A_n, B_n \in \mathbb{R}$. Somit ist jede Funktion

$$u_n(x,t) = \sin nx\,[A_n \cos(ant) + B_n \sin(ant)]$$

eine Lösung von (5.15) und (5.16). Es bleiben noch die Anfangsbedingungen (5.17) zu erfüllen. Da das Problem (5.15), (5.16) linear und homogen ist, versuchen wir dies durch Superposition der $u_n(x,t)$ zu erreichen:

$$u(x,t) = \sum_{n=1}^{\infty} \sin nx\,[A_n \cos(ant) + B_n \sin(ant)]\ . \tag{5.20}$$

Wir nehmen an, dass die Reihe (5.20) konvergiert und gliedweise differenzierbar ist. Die Randbedingungen (5.16) sind offenbar erfüllt. Wir erhalten

$$\frac{\partial u(x,t)}{\partial t} = \sum_{n=1}^{\infty} \sin nx\,[-A_n an \sin(ant) + B_n an \cos(ant)]\ . \tag{5.21}$$

Aus den Bedingungen (5.17) folgt nun

$$\sum_{n=1}^{\infty} A_n \sin nx = f(x)\,, \quad \sum_{n=1}^{\infty} B_n n \sin nx = \frac{1}{a}\,g(x)\,. \tag{5.22}$$

Bemerkung 5.28 *Sind f'' und g' Funktionen endlicher Variation mit $f(0) = f(\pi) = 0$ sowie $g(0) = g(\pi) = 0$, so gilt*

$$A_n = \frac{2}{\pi}\int_0^{\pi} f(x) \sin nx\,dx = O(n^{-3})\,, \quad B_n = \frac{2}{an\pi}\int_0^{\pi} g(x) \sin nx\,dx = O(n^{-3})\,.$$

Unter den Voraussetzungen der Bemerkung 5.28 konvergieren die Reihen (5.20), (5.21) und (5.22) gleichmäßig sogar für $x, t \in \mathbb{R}$. Die Summen der Reihen (5.22) für $x \notin [0, \pi]$ bezeichnen wir ebenfalls mit $f(x)$ bzw. $\frac{1}{a}\,g(x)$. Damit genügt die Summe der Reihe (5.20) den Rand- und Anfangsbedingungen. Es ist noch zu überprüfen, ob die Differentialgleichung (5.15) erfüllt ist. Aufgrund der gleichmäßigen Konvergenz der entsprechenden Reihe in (5.22) lässt sich diese gliedweise integrieren, so dass

$$g_1(x) = -a \sum_{n=1}^{\infty} B_n \cos nx$$

eine Stammfunktion für $g(x)$ ist. Aus (5.20) folgt nun

$$\begin{aligned} u(x,t) &= \frac{1}{2}\sum_{n=1}^{\infty}\{A_n\,[\sin n(x+at)+\sin n(x-at)] - B_n\,[\cos n(x+at)-\cos n(x-at)]\} \\ &= \frac{1}{2}\left[f(x+at)+f(x-at)+\frac{g_1(x+at)-g_1(x-at)}{a}\right]. \end{aligned}$$

Hieraus folgt

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{1}{2}\left\{a^2\left[f''(x+at)+f''(x-at)\right]+a\left[g'(x+at)-g'(x-at)\right]\right\} = a^2\frac{\partial^2 u(x,t)}{\partial x^2}.$$

5.11.2 Randwertaufgaben

Für stetig differenzierbare Funktionen $a_j : [a,b] \longrightarrow \mathbb{R}$, $j = 0,1$, definieren wir den Operator $L : \mathbf{C}^{(2)}[a,b] \longrightarrow \mathbf{C}[a,b]$ durch

$$(Ly)(x) = y''(x) + a_1(x)y'(x) + a_0(x)y(x).$$

Mit $\mathcal{L} \subset \mathbf{C}^{(2)}[a,b]$ bezeichnen wir den (linearen Teil-) Raum aller Lösungen der homogenen Gleichung

$$Ly = 0, \tag{5.23}$$

d.h.

$$\mathcal{L} = \left\{y \in \mathbf{C}^{(2)}[a,b] : (Ly)(x) = 0 \;\forall\, x \in [a,b]\right\}.$$

Da man die Funktionen $a_j(x)$ über das Intervall $[a,b]$ hinaus stetig differenzierbar fortsetzen kann, ist die Existenz- und Eindeutigkeitsaussage für lineare Differentialgleichungen mit variablen Koeffizienten auf ein Intervall $(a-\varepsilon, b+\varepsilon)$ anwendbar und somit die Existenz und Eindeutigkeit der Lösung des Anfangswertproblems

$$Ly = 0, \quad y(x_0) = y_0,\; y'(x_0) = y_1,$$

für $x_0 \in [a,b]$ gesichert. Damit können wir für jedes $x \in [a,b]$ die **Anfangswertabbildung**

$$A_x : \mathcal{L} \longrightarrow \mathbb{R}^2, \quad y \mapsto \begin{bmatrix} y(x) \\ y'(x) \end{bmatrix}$$

betrachten. $A_x : \mathcal{L} \longrightarrow \mathbb{R}^2$ ist ein linearer Isomorphismus. Außerdem definieren wir die **Transportabbildung**

$$\tau : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \mapsto \begin{bmatrix} y(b) \\ y'(b) \end{bmatrix},$$

wobei $y \in \mathcal{L}$ diejenige Lösung von (5.23) ist, die die Anfangswerte $y(a) = y_0$, $y'(a) = y_1$ hat. Aus dem Diagramm

$$
\begin{tikzcd}
& \mathcal{L} \arrow[dl, "A_a"'] \arrow[dr, "A_b"] & \\
\mathbb{R}^2 \arrow[rr, "\tau"'] & & \mathbb{R}^2
\end{tikzcd}
$$

folgt $\tau = A_b \circ A_a^{-1}$, so dass auch $\tau : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ein linearer Isomorphismus ist.
Wir betrachten nun Randbedingungen der Gestalt

$$R_a y := \alpha_0 y(a) + \alpha_1 y'(a) = 0\,, \quad R_b y := \beta_0 y(b) + \beta_1 y'(b) = 0 \tag{5.24}$$

mit $\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ und $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Für $\gamma = \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} \in \mathbb{R}^2$ bezeichnen wir mit V_γ die Gerade

$$V_\gamma = \left\{ \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} \in \mathbb{R}^2 : \gamma_0 y_0 + \gamma_1 y_1 = 0 \right\}$$

(linearer Teilraum von $\mathbb{R}^2$). Eine Funktion $y \in \mathcal{L}$ genügt genau dann den Randbedingungen (5.24), wenn $\begin{bmatrix} y(a) \\ y'(a) \end{bmatrix} \in V_\alpha$ und $\begin{bmatrix} y(b) \\ y'(b) \end{bmatrix} \in V_\beta$ gilt. Das ist gleichbedeutend damit, dass die Kurve

$$\left\{ \begin{bmatrix} x \\ y(x) \\ y'(x) \end{bmatrix} : x \in [a,b] \right\}$$

(die im erweiterten Phasenraum liegt) in $\{a\} \times V_\alpha$ "startet" und in $\{b\} \times V_\beta$ "endet". Da $\tau(V_\alpha)$ ein linearer Unterraum von $\mathbb{R}^2$ ist, kann es nur dann nichttriviale Lösungen von (5.23), (5.24) geben, wenn $\tau(V_\alpha) = V_\beta$ gilt. Damit kann der Lösungsraum von (5.23), (5.24) nur 0- oder 1-dimensional sein, nämlich gleich

$$A_b^{-1}\left(\tau(V_\alpha) \cap V_\beta\right) = A_a^{-1}(V_\alpha) \cap A_b^{-1}(V_\beta)\,. \tag{5.25}$$

Satz 5.29 *Ist $\{\varphi_1, \varphi_2\}$ ein Fundamentalsystem von Lösungen von* (5.23) *(d.h. eine Basis in $\mathcal{L}$), so besitzt das homogene Randwertproblem* (5.23), (5.24) *genau dann nur die triviale Lösung, wenn*

$$\det \begin{bmatrix} R_a\varphi_1 & R_a\varphi_2 \\ R_b\varphi_1 & R_b\varphi_2 \end{bmatrix} \neq 0 \tag{5.26}$$

gilt.

Bei der Betrachtung inhomogener Randbedingungen

$$R_a y = \alpha_2\,, \quad R_b y = \beta_2 \tag{5.27}$$

gehen die Unterräume V_α und V_β in zu diesen parallele Geraden über.

Folgerung 5.30 *Das inhomogene Randwertproblem* (5.23), (5.27) *ist genau dann eindeutig lösbar, wenn das zugehörige homogene Randwertproblem* (5.23), (5.24) *nur die triviale Lösung besitzt.*

Sturmsche Randwertaufgaben

Wir betrachten für $f \in \mathbf{C}[a,b]$ das Randwertproblem

$$Ly = f\,, \tag{5.28}$$

$$R_a y = \alpha_2\,, \quad R_b y = \beta_2\,. \tag{5.29}$$

Eine Lösung dieses Problems ist offenbar genau dann eindeutig, wenn das entsprechende homogene Problem (5.23), (5.24) nur die triviale Lösung besitzt. Das Problem (5.28), (5.29) lässt sich leicht halbhomogenisieren: Es sei $y^* \in \mathbf{C}^{(2)}[a,b]$ eine Funktion mit $R_a y^* = \alpha_2$ und $R_b y^* = \beta_2$. Wir setzen $z = y - y^*$. Dann ist (5.28), (5.29) äquivalent zu

$$Lz = f - Ly^*\,, \quad R_a z = R_b z = 0\,.$$

Wir multiplizieren die Gleichung (5.28) mit

$$p(t) = \exp\left(\int_a^t a_1(s)\,ds\right)$$

und erhalten

$$py'' + pa_1 y' + pa_0 y = pf$$

bzw. wegen $(py')' = py'' + p'y' = py'' + pa_1 y'$

$$(py')' + pa_0 y = pf\,.$$

Wir können uns somit auf die Betrachtung der sog. **Sturmschen Randwertaufgaben**

$$Ly = f\,, \quad R_a y = R_b y = 0 \tag{5.30}$$

mit

$$Ly = (py')' + qy \quad \text{und} \quad p \in \mathbf{C}^{(2)}[a,b]\,,\ p(x) > 0\,,\ a \le x \le b\,,\ q \in \mathbf{C}^{(1)}[a,b]$$

einschränken. Es seien $\{\varphi_1, \varphi_2\} \subset \mathcal{L}$ ein Fundamentalsystem von Lösungen von $Ly = 0$ und

$$W(x) = \varphi_1(x)\varphi_2'(x) - \varphi_1'(x)\varphi_2(x)$$

die **Wronski-Determinante** dieses Systems. Es gilt dann

$$p(x)W'(x) = -p'(x)W(x)\,,$$

d.h. $[p(x)W(x)]' \equiv 0$, so dass

$$p(x)W(x) = p(a)W(a) \neq 0 \quad \forall\, x \in [a,b]\,. \tag{5.31}$$

Wir schreiben das Quadrat $[a,b]^2$ als Vereinigung der Dreiecke

$$D_1 = \{(x,t) : a \le x \le t \le b\} \quad \text{und} \quad D_2 = \{(x,t) : a \le t \le x \le b\}$$

und definieren die **Greensche Funktion** $G(x,t)$ der Sturmschen Randwertaufgabe (5.30) als

$$G(x,t) = \begin{cases} \dfrac{\varphi_1(x)\varphi_2(t)}{p(a)W(a)} & : \quad (x,t) \in D_1\,, \\ \dfrac{\varphi_1(t)\varphi_2(x)}{p(a)W(a)} & : \quad (x,t) \in D_2\,. \end{cases}$$

Außerhalb der Diagonalen $\{(x,x) : a \le x \le b\}$ existieren offenbar die stetigen partiellen Ableitungen $\partial_1 G := G_x$ und $\partial_{11} G := G_{xx}$, wobei für $a < x < b$ die Beziehungen

$$\partial_1 G(x+0,x) = \frac{\varphi_1(x)\varphi_2'(x)}{p(a)W(a)}\,, \quad \partial_1 G(x-0,x) = \frac{\varphi_1'(x)\varphi_2(x)}{p(a)W(a)}$$

erfüllt sind. Daraus ergibt sich unter Verwendung von (5.31) die "Sprungrelation"

$$\partial_1 G(x+0,x) - \partial_1 G(x-0,x) = \frac{1}{p(x)}\,, \quad a < x < b\,. \tag{5.32}$$

Das Fundamentalsystem $\{\varphi_1, \varphi_2\}$ genüge der Bedingung (5.26). Die Funktionen

$$\psi_1(x) = c_{11}\varphi_1(x) + c_{12}\varphi_2(x)\,, \quad \psi_2(x) = c_{21}\varphi_1(x) + c_{22}\varphi_2(x)$$

mit $c_{11} = R_a\varphi_2$, $c_{12} = -R_a\varphi_1$, $c_{21} = R_b\varphi_2$ und $c_{22} = -R_b\varphi_1$ bilden dann auch ein Fundamentalsystem für $Ly = 0$, da $\det[c_{jk}] \ne 0$ gilt. Außerdem ist $R_a\psi_1 = R_b\psi_2 = 0$. Wir können also im weiteren o.E.d.A. ein Fundamentalsystem $\{\varphi_1, \varphi_2\}$ mit

$$R_a\varphi_1 = R_b\varphi_2 = 0 \tag{5.33}$$

zugrundelegen. Wir setzen

$$\varphi(x) = \int_a^b G(x,t)f(t)\,dt\,, \quad a \le x \le b\,.$$

Dann gilt

$$\varphi(x) = \int_a^x G(x,t)f(t)\,dt + \int_x^b G(x,t)f(t)\,dt$$

und somit

$$\begin{aligned} \varphi'(x) &= G(x,x)f(x) + \int_a^x \partial_1 G(x,t)f(t)\,dt - G(x,x)f(x) + \int_x^b \partial_1 G(x,t)f(t)\,dt \\ &= \int_a^b \partial_1 G(x,t)f(t)\,dt\,. \end{aligned}$$

Für die zweite Ableitung von $\varphi(x)$ erhalten wir auf analoge Weise unter Verwendung der Beziehung (5.32)

$$\begin{aligned} \varphi''(x) &= \partial_1 G(x+0,x)f(x) - \partial_1 G(x-0,x)f(x) + \int_a^b \partial_{11} G(x,t)f(t)\,dt \\ &= \frac{f(x)}{p(x)} + \int_a^b \partial_{11} G(x,t)f(t)\,dt\,. \end{aligned}$$

Aus der Definition von $G(x,t)$ folgt für $x \neq t$

$$p(x)\partial_{11}G(x,t) + p'(x)\partial_1 G(x,t) + q(x)G(x,t) = 0$$

und somit

$$(L\varphi)(x) = f(x)\,.$$

Außerdem gilt

$$\begin{aligned} R_a\varphi &= \int_a^b \left[\alpha_0 G(a,t) + \alpha_1 \partial_1 G(a,t)\right] f(t)\,dt \\ &= \frac{1}{p(a)W(a)} \int_a^b \left[\alpha_0\varphi_1(a) + \alpha_2\varphi_1'(a)\right] \varphi_2(t) f(t)\,dt = 0 \end{aligned}$$

und analog $R_b\varphi = 0$. Damit haben wir folgenden Satz bewiesen.

Satz 5.31 *Es seien $\{\varphi_1, \varphi_2\}$ ein Fundamentalsystem von Lösungen der Gleichung $Ly = 0$, die die Bedingungen $R_a\varphi_1 = R_b\varphi_2 = 0$ und $(R_a\varphi_2)(R_b\varphi_1) \neq 0$ erfüllen. Ferner sei $G(x,t)$ die zugehörige Greensche Funktion. Dann ist das Randwertproblem* (5.30) *eindeutig lösbar, und die Lösung ist gegeben durch*

$$\varphi(x) = \int_a^b G(x,t) f(t)\,dt\,.$$

Beispiel 5.32 *Für die Randwertaufgabe*

$$y'' = f\,, \quad y(0) = y(1) = 0\,,$$

kann man das Fundamentalsystem $\{\varphi_1, \varphi_2\}$ mit $\varphi_1(x) = x$ und $\varphi_2(x) = x - 1$ wählen, so dass $W(x) \equiv 1$ und

$$G(x,t) = \begin{cases} x(t-1) & : \quad 0 \le x \le t \le 1\,, \\ t(x-1) & : \quad 0 \le t \le x \le 1\,, \end{cases}$$

ist. Hieraus erhält man die Lösungsdarstellung

$$\begin{aligned} y(x) = \int_0^1 G(x,t) f(t)\,dt &= (x-1)\int_0^x t f(t)\,dt + x\int_x^1 (t-1) f(t)\,dt \\ &= \int_0^x F(t)\,dt - x\int_0^1 F(t)\,dt\,, \end{aligned}$$

wobei $F(x)$ eine Stammfunktion von $f(x)$ bezeichnet. Betrachten wir die Randwertaufgabe

$$y'' = f\,, \quad y(0) = y'(1) = 0\,,$$

so erhalten wir mit $\varphi_1(x) = x$ und $\varphi_2(x) \equiv 1$ die Lösungsdarstellung

$$y(x) = -\int_0^x t f(t)\,dt - x\int_x^1 f(t)\,dt = \int_0^x F(t)\,dt - xF(1)\,.$$

Beispiel 5.33 *Wir suchen die Lösung des Randwertproblems*

$$y'' - y = 1 =: f(x)\,, \quad y(0) = y'(a) = 0 \quad (a > 0)\,.$$

Mit $\varphi_1(x) = \dfrac{e^x - e^{-x}}{2} = \sinh x$ *und* $\varphi_2(x) = \dfrac{e^{x-a} + e^{a-x}}{2} = \cosh(x-a)$ *erhalten wir*

$$G(x,t) = \begin{cases} -\dfrac{\sinh x\,\cosh(t-a)}{\cosh a} & : \ 0 \le x \le t \le a\,, \\ -\dfrac{\sinh t\,\cosh(x-a)}{\cosh a} & : \ 0 \le t \le x \le a\,, \end{cases}$$

und

$$y(x) = -\frac{\cosh(x-a)}{\cosh a}\int_0^x \sinh t\, f(t)\,dt - \frac{\sinh x}{\cosh a}\int_x^a \cosh(t-a)\, f(t)\,dt = \frac{\cosh(x-a)}{\cosh a} - 1\,.$$

Beispiel 5.34 (Wärmeausbreitung in einem Stab der Länge ℓ)

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}\,, \quad 0 < x < \ell\,,\ t > 0\,,$$

$u(x,t)$ *- Temperatur an der Stelle* x *zum Zeitpunkt* t *,*
$a^2 = \dfrac{k}{c\,\rho}$ *,* k *- Wärmeleitvermögen,* ρ *- Dichte,* c *- Wärmekapazität*

Anfangsbedingung: $u(x,0) = f(x)\,,\ 0 < x < \ell$

Randbedingungen: $u(0,t) = 0\,,\ h\,u(\ell,t) + k\,\dfrac{\partial u(\ell,t)}{\partial x} = 0\,,\ t > 0$

(freie Wärmeausstrahlung am rechten Rand, h *- Wärmeübergangszahl)*

Separationsansatz: $u(x,t) = v(x)w(t) \Longrightarrow$

$$u_n(x,t) = e^{-a^2\lambda_n^2 t}\sin\lambda_n t\,, \quad \tan\lambda_n\ell + \frac{k\lambda_n}{h}\,,\ \lambda_n > 0\,,\ n = 1,2,\ldots$$

5.12 Übungsaufgaben

1. Bestimmen Sie in Abhängigkeit von $\lambda \in \mathbb{R}$ alle Lösungen der Differentialgleichung

$$y'' + \lambda^2 y = 0\,,$$

die den Randbedingungen

(a) $y(0) = 0\,,\ y(\pi) = 1\,,$ (b) $y(0) = y(\pi) = 1$

genügen.

2. Für welche reellen Zahlen λ hat das Randwertproblem

$$y''(x) - 2y'(x) + \lambda y(x) = 0\,, \quad y(0) = y(1) = 0$$

nichttriviale Lösungen?

3. Geben Sie die Lösung des Anfangsrandwertproblems

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}\,, \quad 0 < x < \ell\,,\ t > 0\,,$$

(a) $u(x,0) = f(x)\,,\ 0 < x < \ell\,,\ u(0,t) = u(l,t) = 0\,,\ t > 0\,,$

(b) **(HA)** $u(x,0) = f(x)\,,\ \dfrac{\partial u(0,t)}{\partial x} = 0\,,\ u(l,t) = u_0 = \text{const}\,,\ t > 0\,,\ 0 < x < \ell\,,$

in Form einer Reihendarstellung an (vgl. Beispiel 5.34).

4. Finden Sie eine Reihendarstellung für die Lösung des Anfangsrandwertproblems

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)\,, \quad 0 < x < x_0\,,\ 0 < y < y_0\,,\ t > 0\,,$$

$$u(x,y,0) = f(x,y)\,, \quad t > 0\,,$$

$$u(0,y,t) = u(x_0,y,t) = u(x,0,t) = u(x,y_0,t) = 0\,, \quad 0 < x < x_0\,,\ 0 < y < y_0\,.$$

5. Für welche $\alpha \in \mathbb{R}$ ist das Randwertproblem

$$y'' = x^2\,, \quad \alpha y(0) + y'(0) = 1\,,\ y(1) + \alpha y'(1) = \alpha$$

eindeutig lösbar? Man berechne für diese α die Lösung. Kann man die Lösung mittels einer Greenschen Funktion darstellen?

6. **(HA)** Man löse das Randeigenwertproblem

$$y'' + x^{-1}y' + \lambda x^{-2} y = 0\,, \quad y'(1) = y'\left(e^{2\pi}\right) = 0\,,$$

d.h., man bestimme alle Eigenwerte und Eigenfunktionen dieses Problems.

5.13 Lokale Stabilitätstheorie

Wir hatten gesehen, dass sich jede Lösung $\varphi(t) = [\varphi_j(t)]_{j=1}^n$ des Differentialgleichungssystems

$$\dot{x} = A\,x$$

für eine gegebene Matrix $A \in \mathbb{R}^{n\times n}$ so schreiben lässt, dass jede Komponente $\varphi_j(t)$ eine Linearkombination von Funktionen der Gestalt

$$t^k e^{\alpha t} \cos \beta t\,,\ t^k e^{\alpha t} \sin \beta t$$

ist, wobei $\alpha + \mathbf{i}\beta$ alle Eigenwerte von A mit $\beta \geq 0$ durchläuft.

Im weiteren wollen wir uns einige Vorstellungen über die Lösungskurven **nichtlinearer** Differentialgleichungssysteme erarbeiten. Dazu untersuchen wir vorerst zwei Beispiele.

Wir betrachten das (entkoppelte) lineare System

$$\begin{aligned} \dot{x}_1 &= -x_1 \\ \dot{x}_2 &= -x_2 \\ \dot{x}_3 &= x_3 \end{aligned} \tag{5.34}$$

bzw.

$$\dot{x} = A\,x \quad \text{mit} \quad A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Als Lösung des AWPs $x(0) = x^0$ erhalten wir

$$\varphi(t) = \begin{bmatrix} x_1^0 e^{-t} \\ x_2^0 e^{-t} \\ x_3^0 e^{t} \end{bmatrix} = \begin{bmatrix} e^{-t} & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{t} \end{bmatrix} x^0 .$$

Wir sehen, dass für

$$x^0 \in \operatorname{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} =: E^s$$

gilt

$$\lim_{t \to \infty} \varphi(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Für

$$x^0 \in \operatorname{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} =: E^u$$

gilt dagegen

$$\lim_{t \to -\infty} \varphi(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Man nennt deshalb E^s den **stabilen** und E^u den **instabilen Unterraum** für das System (5.34).

Wir betrachten nun das (nichtlineare) System

$$\begin{aligned} \dot{x}_1 &= -x_1 \\ \dot{x}_2 &= -x_2 + x_1^2 \\ \dot{x}_3 &= x_3 + x_1^2 \end{aligned} \tag{5.35}$$

bzw.

$$\dot{x} = v(x)$$

mit

$$v(x) = \begin{bmatrix} -x_1 \\ -x_2 + x_1^2 \\ x_3 + x_1^2 \end{bmatrix}.$$

Der einzige Gleichgewichtspunkt dieses Systems, d.h. ein Punkt x^* mit $v(x^*) = \Theta$, ist der Punkt $x^* = \Theta$. Nun gilt

$$v'(x) = \begin{bmatrix} -1 & 0 & 0 \\ 2x_1 & -1 & 0 \\ 2x_1 & 0 & 1 \end{bmatrix}$$

und somit $v'(x^*) = v'(\Theta) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A$. Wir bestimmen die Lösung $\psi(t)$ des AWPs $x(0) = x^0$ für das System (5.35). Aus der ersten Gleichung in (5.35) folgt $\psi_1(t) = x_1^0 e^{-t}$, und es bleiben die beiden Gleichungen

$$\dot{x}_2 + x_2 = (x_1^0)^2 e^{-2t}, \quad \dot{x}_3 - x_3 = (x_1^0)^2 e^{-2t}$$

zu lösen. Wir erhalten

$$\psi_2(t) = x_2^0 e^{-t} + (x_1^0)^2 \left(e^{-t} - e^{-2t}\right), \quad \psi_3(t) = x_3^0 e^t + \frac{(x_1^0)^2}{3} \left(e^t - e^{-2t}\right).$$

Offenbar gilt

$$\lim_{t\to\infty} \psi(t) = \Theta \quad \Longleftrightarrow \quad x_3^0 + \frac{(x_1^0)^2}{3} = 0$$

und

$$\lim_{t\to-\infty} \psi(t) = \Theta \quad \Longleftrightarrow \quad x_1^0 = x_2^0 = 0.$$

Man nennt deshalb

$$S := \left\{ c \in \mathbb{R}^3 : c_3 = -\frac{c_1^2}{3} \right\}$$

die **stabile** und

$$U := \left\{ c \in \mathbb{R}^3 : c_1 = c_2 = 0 \right\}$$

die **instabile Mannigfaltigkeit** des Systems (5.35) im Gleichgewichtspunkt Θ. Wir sehen, dass E^s Tangentenebene an S ist und dass $E^u = U$ gilt.

Ferner gilt für $x_3^0 + \dfrac{(x_1^0)^2}{3} = 0$ auch

$$\begin{aligned}\psi_1(t) &= x_1^0 e^{-t}\\ \psi_2(t) &= [x_2^0 + (x_1^0)^2]e^{-t} - (x_1^0)^2 e^{-2t}\\ \psi_3(t) &= -\frac{(x_1^0)^2}{3}\, e^{-2t}\end{aligned}$$

und somit

$$\psi_3(t) + \frac{[\psi_1(t)]^2}{3} = 0\,.$$

Für $x_1^0 = x_2^0 = 0$ ist

$$\begin{aligned}\psi_1(t) &= 0\\ \psi_2(t) &= 0\\ \psi_3(t) &= x_3^0 e^t\,.\end{aligned}$$

Aus $x^0 \in S$ (bzw. $x^0 \in U$) folgt also $\psi(t) \in S$ (bzw. $\psi(t) \in U$) für alle $t \in \mathbb{R}$. (S und U sind invariant bzgl. des Systems (5.35).)

Nun zur allgemeinen Situation: Wir betrachten im $\mathbb{R}^n$ das lineare Differentialgleichungssystem

$$\dot{x} = A\,x \tag{5.36}$$

(d.h. $A \in \mathbb{R}^{n\times n}$). Es bezeichne $w^j = u^j + \mathbf{i}v^j$ ($u^j, v^j \in \mathbb{R}^n$) einen verallgemeinerten Eigenvektor von A zum Eigenwert $\lambda_j = \alpha_j + \mathbf{i}\beta_j$ ($\alpha_j \in \mathbb{R}$, $\beta_j \geq 0$), d.h.

$$\exists\,\ell \in \mathbb{N} : (A - \lambda_j I)^\ell w^j = \Theta\,, \quad w^j \neq \Theta\,.$$

Ferner sei

$$B = \left\{u^1, \ldots, u^k, v^{k+1}, u^{k+1}, \ldots, v^m, u^m\right\}$$

eine Basis in $\mathbb{R}^n$ (d.h. $\beta_1 = \cdots = \beta_k = 0$ und $2(m-k) + k = 2m - k = n$).

Definition 5.35 *Man nennt*

$$E^s = \operatorname{span}\left\{u^j, v^j : \alpha_j < 0\right\}$$

den **stabilen**,

$$E^u = \operatorname{span}\left\{u^j, v^j : \alpha_j > 0\right\}$$

den **instabilen** *und*

$$E^c = \operatorname{span}\left\{u^j, v^j : \alpha_j = 0\right\}$$

den **zentralen Unterraum** *von $\mathbb{R}^n$ für das System (5.36).*

Beispiel 5.36 $A = \begin{bmatrix} -2 & -1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

Beispiel 5.37 $A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Beispiel 5.38 $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Anhand der Darstellung der Lösung eines linearen Systems gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten lassen sich folgende Theoreme ableiten.

Theorem 5.39 *Folgende Aussagen sind äquivalent:*

(a) $\lim_{t\to\infty} e^{tA}x^0 = \Theta \quad \forall\, x^0 \in \mathbb{R}^n$ *und* $\lim_{t\to-\infty} \left|e^{tA}x^0\right| = \infty \quad \forall\, x^0 \in \mathbb{R}^n \setminus \{\Theta\}$.

(b) Alle Eigenwerte der Matrix A haben negativen Realteil.

Theorem 5.40 *Folgende Aussagen sind äquivalent:*

(a) $\lim_{t\to-\infty} e^{tA}x^0 = \Theta \quad \forall\, x^0 \in \mathbb{R}^n$ *und* $\lim_{t\to\infty} \left|e^{tA}x^0\right| = \infty \quad \forall\, x^0 \in \mathbb{R}^n \setminus \{\Theta\}$.

(b) Alle Eigenwerte der Matrix A haben positiven Realteil.

Da die verallgemeinerten Eigenunterräume einer Matrix A bezüglich A invariant sind, ergibt sich

Folgerung 5.41

(a) Ist $x^0 \in E^s$, *so* $e^{tA}x^0 \in E^s \;\forall\, t \in \mathbb{R}$ *und* $\lim_{t\to\infty} e^{tA}x^0 = \Theta$.

(b) Ist $x^0 \in E^u$, *so* $e^{tA}x^0 \in E^u \;\forall\, t \in \mathbb{R}$ *und* $\lim_{t\to-\infty} e^{tA}x^0 = \Theta$.

Theorem 5.42 (Stabile Mannigfaltigkeit) *Das Vektorfeld* $v : M \longrightarrow \mathbb{R}^n$ *(*$M \subset \mathbb{R}^n$ *- Gebiet) sei stetig differenzierbar und* $\Phi : \mathcal{A} \subset \mathbb{R} \times M \longrightarrow M$ *sei der lokale Fluß zu dem autonomen System*

$$\dot{x} = v(x)\,. \tag{5.37}$$

Ferner sei $v(\Theta) = \Theta$, *d.h.* Θ *ist stationärer Punkt von (5.37). Die Matrix* $A := v'(\Theta)$ *habe* k *Eigenwerte mit negativem Realteil und* $n-k$ *Eigenwerte mit positivem Realteil (entspr. ihrer algebraischen Vielfachheit gezählt). Mit* E^s *und* E^u *bezeichnen wir den stabilen und instabilen Unterraum des Systems (5.36). Dann existiert eine* k*-dimensionale differenzierbare Mannigfaltigkeit* S , *zu der* E^s *im Punkt* Θ *tangential ist, so dass*

$$\Phi(t, x^0) \in S \quad \forall\, x^0 \in S, \forall\, t \geq 0 \quad \textit{und} \quad \lim_{t\to\infty} \Phi(t, x^0) = \Theta \quad \forall\, x^0 \in S\,.$$

Ferner existiert eine $n-k$-dimensionale differenzierbare Mannigfaltigkeit U, zu der E^u im Punkt Θ tangential ist, so dass

$$\Phi(t,x^0) \in U \quad \forall\, x^0 \in U, \forall\, t \le 0 \quad \textit{und} \quad \lim_{t\to-\infty} \Phi(t,x^0) = \Theta \quad \forall\, x^0 \in U\,.$$

Genauer: Es existieren differenzierbare Funktionen $h_s : E^s \longrightarrow (E^s)^\perp$, $h_u : E^u \longrightarrow (E^u)^\perp$ und ein $\varepsilon > 0$, so dass

$$S = \{(z,w) \in \mathbb{R}^n : w = h_s(z),\, |z| < \varepsilon\}\,, \quad h_s'(\Theta) = \Theta\,,$$
$$U = \{(z,w) \in \mathbb{R}^n : w = h_u(z),\, |z| < \varepsilon\}\,, \quad h_u'(\Theta) = \Theta\,.$$

Beispiel 5.43 *Wir betrachten das autonome System*

$$\dot{x}_1 = -x_1 - x_2^2\,, \quad \dot{x}_2 = x_2 + x_1^2\,. \tag{5.38}$$

Bemerkung 5.44

(a) Ist x^* Gleichgewichtspunkt von $\dot{x} = v(x)$, so ist nach der Koordinatentransformation $y = x - x^*$ der Punkt $y^* = \Theta$ Gleichgewichtspunkt von $\dot{y} = v(y + x^*)$.

(b) Nach den Überlegungen in Bsp. 5.43 kann man S und U eigentlich nur als lokale stabile bzw. instabile Mannigfaltigkeit des Systems $\dot{x} = v(x)$ bezeichnen. Deshalb die folgende Definition.

Definition 5.45 *Es seien $\Phi(t,x)$ der (lokale) Fluß des Systems $\dot{x} = v(x)$ mit $v(\Theta) = \Theta$ und S bzw. U wie in Theorem 5.42. Unter der* **globalen stabilen** *bzw.* **instabilen Mannigfaltigkeit** *dieses Systems im Punkt Θ versteht man*

$$W^s(\Theta) = \bigcup_{t\le 0} \Phi(t,S) \quad \textit{bzw.} \quad W^u(\Theta) = \bigcup_{t\ge 0} \Phi(t,U)\,.$$

$W^s(\Theta)$ und $W^u(\Theta)$ sind eindeutig bestimmt und invariant bzgl. Φ. Ist $\Phi(0,x) \in W^s$ (bzw. W^u), so existieren $T \in \mathbb{R}$, $\varepsilon > 0$ und $\alpha > 0$ mit

$$|\Phi(t,x)| \le \varepsilon\, e^{-\alpha t} \quad (\text{bzw. } \varepsilon\, e^{\alpha t}) \quad \forall\, t \ge T \quad (\text{bzw. } \le T)\,.$$

Theorem 5.46 *Hat $A = v'(\Theta)$ k Eigenwerte mit negativem, j Eigenwerte mit positivem und $m = n - k - j > 0$ Eigenwerte mit verschwindendem Realteil, so existiert eine m-dimensionale differenzierbare (zentrale) Mannigfaltigkeit $W^c(0)$, die zum zentralen Unterraum E^c von $\dot{x} = A\,x$ im Punkt Θ tangential und invariant bzgl. des Flusses $\Phi(t,x)$ ist.*

Beispiel 5.47 $\dot{x}_1 = x_1^2\,, \quad \dot{x}_2 = -x_2\,.$

Dieses Beispiel zeigt, dass die zentrale Mannigfaltigkeit i.a. nicht eindeutig bestimmt ist.

Wir studieren im weiteren ein (i.a. nichtlineares) autonomes Differentialgleichungssystem

$$\dot{x} = v(x) \tag{5.39}$$

mit $v(\Theta) = \Theta$ und ein lineares System (mit konstanten Koeffizienten)

$$\dot{y} = A\,y\,. \tag{5.40}$$

Beispiel 5.48 *Wir betrachten*

$$\begin{aligned}\dot{x}_1 &= -x_1\,,\\ \dot{x}_2 &= -x_2 + x_3^2\,,\\ \dot{x}_3 &= x_3\,,\end{aligned}$$

und

$$\begin{aligned}\dot{y}_1 &= -y_1\,,\\ \dot{y}_2 &= -y_2\,,\\ \dot{y}_3 &= y_3\,.\end{aligned}$$

Definition 5.49 *Man nennt (5.39) und (5.40) in einer Umgebung von Θ zueinander* **topologisch äquivalent***, wenn es offene Umgebungen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^n$ des Punktes Θ und einen einen Homöomorphismus $H : U \longrightarrow V$ gibt, der Orbits von (5.39) in U in Orbits von (5.40) in V überführt. Man nennt (5.39) und (5.40) zueinander* **topologisch konjugiert***, wenn H zusätzlich die Parametrisierung durch die Zeit t beibehält.*

Ist z.B. (5.39) auch ein lineares System $\dot{x} = B\,x$ und gilt $B = T^{-1}A\,T$, so erhalten wir für eine Lösung von (5.39)

$$\varphi(t) = e^{tB}x^0 = T^{-1}e^{tA}T\,x^0\,,$$

also für $\psi(t) = T\,\varphi(t)$, $y^0 = T\,x^0$,

$$\psi(t) = e^{tA}y^0\,,$$

d.h. eine Lösung von (5.39). Anders geschrieben sieht das so aus:

$$T\,e^{tB}x^0 = e^{tA}T\,x^0\,.$$

Also sind $\dot{x} = B\,x$ und $\dot{y} = A\,y$ zueinander topologisch konjugiert.

Theorem 5.50 (Hartman–Grobman) *Es sei $v : M \longrightarrow \mathbb{R}^n$ stetig differenzierbar mit $v(\Theta) = \Theta$, und $A = v'(\Theta)$ habe <u>keinen</u> Eigenwert mit verschwindendem Realteil. Mit $\Phi(t, x)$ bezeichnen wir den entsprechenden (lokalen) Fluß für das System (5.39). Dann existiert ein Homöomorphismus $H : U \longrightarrow V$ (vgl. Def. 5.49), so dass für alle $x^0 \in U$ ein offenes Intervall $(-\delta, \delta) \subset \mathbb{R}$ existiert mit*

$$H(\Phi(t, x^0)) = e^{tA}H(x^0) \quad \forall\, t \in (-\delta, \delta)\,,$$

d.h. (5.39) und (5.40) (mit $A = v'(\Theta)$) sind zueinander topologisch konjugiert.

Beispiel 5.51 *Im Beispiel 5.48 ist ein solcher Homöomorphismus gegeben durch*

$$H(x) = \begin{bmatrix} x_1 \\ x_2 - \dfrac{x_3^2}{3} \\ x_3 \end{bmatrix}.$$

Definition 5.52 *Es seien* $v : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *und* $\mathcal{L} : M \longrightarrow \mathbb{R}$ *stetig differenzierbar. Unter* $\dot{\mathcal{L}}(x)$ *verstehen wir die Ableitung von* $\mathcal{L}$ *entlang der Lösungskurven von (5.39), d.h.*

$$\dot{\mathcal{L}}(x) = \frac{d}{dt}\mathcal{L}\left(\Phi(t,x)\right)\Big|_{t=0} = \mathcal{L}'\left(\Phi(0,x)\right)\dot{\Phi}(0,x) = \mathcal{L}'(x)v(x)\,.$$

Definition 5.53 *Ein Punkt* $x^* \in M$ *mit* $v(x^*) = \Theta$ *heißt* **stabiler** *Gleichgewichtspunkt, wenn für jedes* $\varepsilon > 0$ *ein* $\delta > 0$ *existiert, so dass*

$$|\Phi(t,x) - x^*| < \varepsilon \quad \forall\, x \in U_\delta(x^*)\,,\ \forall\, t \geq 0\,.$$

Er heißt **asymptotisch stabil**, *wenn er stabil ist und wenn ein* $\eta > 0$ *existiert, so dass*

$$\lim_{t\to\infty}\Phi(t,x) = x^* \quad \forall\, x \in U_\eta(x^*)\,.$$

Der Punkt x^* *heißt* **instabil**, *wenn er nicht stabil ist.*

Theorem 5.54 *Es seien* $v : M \longrightarrow \mathbb{R}^n$ *und* $\mathcal{L} : M \longrightarrow \mathbb{R}$ *stetig differenzierbar sowie* $v(x^*) = \Theta$ *für ein* $x^* \in M$. *Ferner seien* $\mathcal{L}(x^*) = 0$ *und* $\mathcal{L}(x) > 0$ *für alle* $x \in M \setminus \{x^*\}$.

(a) Gilt $\dot{\mathcal{L}}(x) \leq 0 \quad \forall\, x \in M$, *so ist* x^* *stabil.*

(b) Gilt $\dot{\mathcal{L}}(x) < 0 \quad \forall\, x \in M \setminus \{x^*\}$, *so ist* x^* *asymptotisch stabil.*

(c) Gilt $\dot{\mathcal{L}}(x) > 0 \quad \forall\, x \in M \setminus \{x^*\}$, *so ist* x^* *instabil.*

($\mathcal{L}$ *heißt* **Ljapunov-Funktion**.*)*

Beispiel 5.55 $\dot{x}_1 = -x_2^3\,, \quad \dot{x}_2 = x_1^3$

Beispiel 5.56

$$\begin{aligned}
\dot{x}_1 &= -2x_2 + x_2x_3 - x_1^3 \\
\dot{x}_2 &= x_1 - x_1x_3 - x_2^3 \\
\dot{x}_3 &= x_1x_2 - x_3^3
\end{aligned}$$

Beispiel 5.57 *Für eine stetige Funktion* $q : \mathbb{R} \longrightarrow \mathbb{R}$ *mit* $y\,q(y) > 0$ *(*$y \neq 0$*) betrachten wir die Differentialgleichung*

$$\ddot{y} + q(y) = 0\,.$$

Definition 5.58 *Seien jetzt* $M \subset \mathbb{R}^{2n}$ *und* $H : M \longrightarrow \mathbb{R}$ *zweimal stetig differenzierbar. Wir schreiben* $H(x,y)$ *mit* $x, y \in \mathbb{R}^n$. *Ein System*

$$\dot{x} = \frac{\partial H}{\partial y}\,, \quad \dot{y} = -\frac{\partial H}{\partial x} \tag{5.41}$$

heißt **Hamiltonsches System** *mit* n *Freiheitsgraden.* $H(x,y)$ *heißt* **Hamilton-Funktion** *oder* **totale Energie** *des Systems (5.41).*

Satz 5.59 *Die totale Energie des Systems (5.41) bleibt konstant entlang der Lösungskurven von (5.41).*

Das folgt aus der Tatsache, dass entlang einer Lösungskurve

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\,\dot{x} + \frac{\partial H}{\partial y}\,\dot{y} = \frac{\partial H}{\partial x}\frac{\partial H}{\partial y} - \frac{\partial H}{\partial y}\frac{\partial H}{\partial x} = 0$$

gilt.

Beispiel 5.60 *Die Gleichung des ungedämpften nichtlinearen Pendels*

$$\ddot{x} + \sin x = 0$$

ist äquivalent zu dem Hamilton–System

$$\dot{x} = y\,, \quad \dot{y} = -\sin x$$

mit der Hamilton–Funktion

$$H(x,y) = \frac{y^2}{2} + 1 - \cos x\,,$$

die wir bereits im Beispiel des Abschnitts 5.6 betrachtet haben.

Beispiel 5.61 *Jedes* **Newtonsche System** $\ddot{x} = f(x)$ *ist ein Hamilton-System mit der totalen Energie*

$$H(x,y) = \frac{y^2}{2} - \int_0^x f(s)\,ds\,.$$

Definition 5.62 *Sei* $M \subset \mathbb{R}^n$ *und sei* $\Phi : M \longrightarrow \mathbb{R}$ *zweimal stetig differenzierbar. Das System*

$$\dot{x} = -\mathrm{grad}\,\Phi(x) \tag{5.42}$$

heißt **Gradientensystem** *auf* M .

Ein Punkt $x^0 \in M$ ist genau dann Gleichgewichtspunkt von (5.42), wenn $\mathrm{grad}\,\Phi(x^0) = \Theta$ gilt, d.h. wenn x^0 ein sogenannter kritischer oder singulärer Punkt von Φ ist. In regulären Punkten $x^0 \in M$ von Φ (d.h. $\mathrm{grad}\,\Phi(x^0) \neq \Theta$) steht $\mathrm{grad}\,\Phi(x^0)$ senkrecht auf der Niveaulinie $\big\{x \in M : \Phi(x) = \Phi(x^0)\big\}$.

Sei x^0 kritischer (singulärer) Punkt von Φ . Wir setzen $\mathcal{L}(x) = \Phi(x) - \Phi(x^0)$. Dann gilt für das System (5.42)

$$\dot{\mathcal{L}}(x) = [\mathrm{grad}\,\Phi(x)]^T[-\mathrm{grad}\,\Phi(x)] = -\left|\mathrm{grad}\,\Phi(x)\right|^2 < 0$$

für alle $x \neq x^0$ aus einer Umgebung von x^0 , falls x^0 ein strenges lokales Extremum ist. Wir erhalten

Satz 5.63 *Ein strenges lokales Minimum von* $\Phi(x)$ *ist ein asymptotisch stabiler Gleichgewichtspunkt des Gradientensystems (5.42).*

5.13.1 Die van der Pol'sche Gleichung

Wir betrachten eine Gleichung der Gestalt

$$\ddot{x} + f(x)\dot{x} + g(x) = 0 \tag{5.43}$$

mit gegebenen stetig differenzierbaren Funktionen $f, g : \mathbb{R} \longrightarrow \mathbb{R}$, wobei $g(0) = 0$ sei. Die Gleichung (5.43) ist äquivalent zu dem System

$$\dot{x}_1 = x_2 - F(x_1)\,, \quad \dot{x}_2 = -g(x_1)\,, \tag{5.44}$$

wobei $F(x_1) = \int_0^{x_1} f(s)\,ds$. Ist nämlich $[x_1(t), x_2(t)]^T$ eine Lösung von (5.44), so gilt

$$\ddot{x}_1 = \dot{x}_2 - f(x_1)\dot{x}_1 = -g(x_1) - f(x_1)\dot{x}_1\,,$$

so dass $x(t) = x_1(t)$ Lösung von (5.43) ist. Ist umgekehrt $x(t) =: x_1(t)$ Lösung von (5.43), so folgt

$$\dot{x}_1(t) = -F(x_1(t)) - \int_0^t g(x_1(s))\,ds + \mathrm{const} =: -F(x_1(t)) + x_2(t)$$

mit

$$\dot{x}_2(t) = -g(x_1(t))\,.$$

Wir nehmen nun an, dass für ein $\delta > 0$ die Bedingungen

$$G(x) := \int_0^x g(s)\,ds > 0\,,\ g(x)F(x) > 0\,, \quad 0 < |x| < \delta\,, \tag{5.45}$$

erfüllt sind, und setzen $\mathcal{L}(x_1, x_2) = \dfrac{x_2^2}{2} + G(x_1)$. Dann ist

$$\dot{\mathcal{L}}(x_1, x_2) = g(x_1)\dot{x}_1 + x_2\dot{x}_2 = g(x_1)[x_2 - F(x_1)] + x_2[-g(x_1)] = -g(x_1)F(x_1) < 0$$

für $0 < |x_1| < \delta$. D.h. der Nullpunkt ist (asymptotisch) stabiler Gleichgewichtspunkt. Gilt in der zweiten Ungleichung von (5.45) das umgekehrte Ungleichheitszeichen, so ist er instabiler Gleichgewichtspunkt. Wenden wir dieses Resultat auf die van der Pol'sche Gleichung

$$\ddot{x} + \gamma(x^2 - 1)\dot{x} + x = 0$$

an, so erhalten wir

$$g(x) = x\,, \quad G(x) = \frac{x^2}{2}\,, \quad f(x) = \gamma(x^2 - 1)\,, \quad F(x) = \gamma\left(\frac{x^3}{3} - x\right)$$

und somit

$$g(x)F(x) = \gamma\, x^2 \left(\frac{x^2}{3} - 1\right) > 0 \ \ (< 0) \quad \text{für} \quad 0 < |x| < \sqrt{3}\,,$$

wenn $\gamma < 0$ (> 0) ist.

5.13.2 Das System Dampfmaschine – Fliehkraftregler

Auf jeden Massepunkt wirken die Zentrifugalkraft $m\theta^2 \sin\varphi$ und die Schwerkraft $m\,g$ sowie die Reibung $b\,\dot\varphi$. Die Differentialgleichung der Dampfmaschine lautet

$$\mathcal{J}\,\dot\omega = \mathcal{M}_1 - \mathcal{M}\,,$$

wobei $\mathcal{J}$ das Trägheitsmoment und ω die Winkelgeschwindigkeit des Schwungrades, $\mathcal{M}_1$ das Drehmoment der Dampfkraft und $\mathcal{M}$ das Drehmoment, das die Last (z.B. ein Förderkorb) auf das Schwungrad ausübt, bezeichnen. Die Dampfmaschine und der Regler sind über das Übersetzungsverhältnis n ($\theta = n\omega$) und über die Gleichung

$$\mathcal{M}_1 = \mathcal{F}_1 + k(\cos\varphi - \cos\varphi^*)$$

miteinander gekoppelt. Es folgt

$$m\,\ddot\varphi = m\,n^2\omega^2 \sin\varphi\,\cos\varphi - m\,g\,\sin\varphi - b\,\dot\varphi\,, \quad \mathcal{J}\,\dot\omega = k\,\cos\varphi - \mathcal{F}$$

mit $\mathcal{F} = \mathcal{M} - \mathcal{F}_1 + k\,\cos\varphi^*$. Dieses Differentialgleichungssystem ist äquivalent zu dem System

$$\begin{aligned} \dot x_1 &= x_2 \\ \dot x_2 &= n^2 x_3^2 \sin x_1 \cos x_1 - g\,\sin x_1 - \frac{b}{m}\,x_2 \\ \dot x_3 &= \frac{k}{\mathcal{J}}\,\cos x_1 - \frac{\mathcal{F}}{\mathcal{J}}\,, \end{aligned}$$

welches wir in der Form

$$\begin{aligned} \dot x_1 &= x_2 \\ \dot x_2 &= (\alpha\,x_3^2 \cos x_1 - \beta)\sin x_1 - \gamma\,x_2 \\ \dot x_3 &= \delta\,\cos x_1 - \eta \end{aligned}$$

schreiben. Wir verwenden dabei die Bezeichnungen

$$x_1 = \varphi\,,\ x_2 = \dot\varphi\,,\ x_3 = \omega\,,\ \alpha = n^2\,,\ \beta = g\,,\ \gamma = \frac{b}{m}\,,\ \delta = \frac{k}{\mathcal{J}}\,,\ \eta = \frac{\mathcal{F}}{\mathcal{J}}\,.$$

Der Phasenraum dieses Systems ist somit gleich $M = \left(0, \frac{\pi}{2}\right) \times \mathbb{R} \times (0,\infty) \subset \mathbb{R}^3$. Der stationäre Punkt $x^* \in M$ genügt den Gleichungen

$$\cos x_1^* = \frac{\eta}{\delta} = \frac{\mathcal{F}}{k}\,, \quad x_2^* = 0\,, \quad x_3^* = \sqrt{\frac{\beta}{\alpha\,\cos x_1^*}} = \sqrt{\frac{\beta\,\delta}{\alpha\,\eta}} = \frac{1}{n}\sqrt{\frac{g\,k}{\mathcal{F}}}\,. \tag{5.46}$$

Die Ableitung des Geschwindigkeitsfeldes in diesem stationären Punkt ist gleich

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -\alpha(x_3^*)^2 \sin^2 x_1^* & -\gamma & 2\alpha\,x_3^* \cos x_1^* \sin x_1^* \\ -\delta\,\sin x_1^* & 0 & 0 \end{bmatrix}.$$

Für das charakteristische Polynom dieser Matrix erhalten wir

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I - A) \\ &= \lambda^2(\lambda+\gamma) + 2\alpha\,\delta\,x_3^*\cos x_1^*\sin^2 x_1^* + \alpha(x_3^*)^2\lambda\sin^2 x_1^* \\ &= \lambda^3 + \gamma\,\lambda^2 + \alpha(x_3^*)^2\,(\sin x_1^*)^2\,\lambda + 2\alpha\,\delta\,x_3^*\cos x_1^*\sin^2 x_1^*\,. \end{aligned}$$

Man nennt ein Polynom stabil, wenn alle seine Nullstellen negativen Realteil besitzen. Wir betrachten ein Polynom

$$p(\lambda) = \lambda^3 + a\,\lambda^2 + b\,\lambda + c \quad \text{mit} \quad a,b,c > 0 \tag{5.47}$$

und stellen die Frage, wann dieses Polynom stabil ist:

1. Wir schreiben $p(\lambda) = (\lambda + a)(\lambda^2 + b) - a\,b + c$ und nehmen an, dass $\lambda = \mathbf{i}\rho\,,\ \rho \in \mathbb{R}\setminus\{0\}$ Nullstelle ist. Dann gilt
$$0 = (\mathbf{i}\rho + a)(-\rho^2 + b) - ab + c\,, \quad \text{d.h.} \quad b - \rho^2 = 0 \quad \text{und} \quad a\,b = c\,.$$
Ist umgekehrt $c - a\,b = 0$, so ist $\mathbf{i}\sqrt{b}$ rein imaginäre Nullstelle.

2. Sei $a\,b < c$. Wir lassen a und b gegen Null streben, und zwar so, dass diese Ungleichung erhalten bleibt. Das Polynom $p(\lambda) = \lambda^3 + c$ hat neben $-c^{\frac{1}{3}}$ die Nullstellen
$$c^{\frac{1}{3}}\left(\cos\frac{\pi}{3} \pm \mathbf{i}\sin\frac{\pi}{3}\right)$$
rechts von der imaginären Achse.

3. Sei nun $a\,b > c$. Wir betrachten $c \longrightarrow 0$. Das Polynom $\lambda(\lambda^2 + a\,\lambda + b)$ hat die Nullstellen $\lambda_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}$ und $\lambda_3 = 0$. Die Nullstelle $\lambda_3 = 0$ geht für kleines $c > 0$ in eine negative Nullstelle über, da $\lambda_1\lambda_2\lambda_3 = -c < 0$.

Also: Das Polynom (5.47) ist genau dann stabil, wenn $a\,b > c$ gilt. Für den Fliehkraftregler bedeutet das

$$\gamma\,\alpha(x_3^*)^2\sin^2 x_1^* > 2\alpha\,\delta\,x_3^*\sin^2 x_1^*\cos x_1^*\,,$$

d.h.

$$x_3^*\gamma > 2\delta\,\cos x_1^* = 2\eta\,.$$

Diese Bedingung ist äquivalent zu

$$\frac{b\,\mathcal{J}}{m} > \frac{2\mathcal{F}}{\omega^*}\,. \tag{5.48}$$

Die Größe

$$\nu = \left|\frac{d\omega^*}{d\mathcal{F}}\right|$$

heißt Ungleichförmigkeit des Laufes der Dampfmaschine (ν – Geschwindigkeit der Änderung von ω^* bei Änderung der Last $\mathcal{M}$). Aus der Gleichung $(\omega^*)^2\mathcal{F} = \text{const}$ (siehe (5.46)) folgt

$$2\omega^*\frac{d\omega^*}{d\mathcal{F}}\,\mathcal{F} + (\omega^*)^2 = 0\,,$$

d.h.

$$\frac{d\omega^*}{d\mathcal{F}} = -\frac{\omega^*}{2\mathcal{F}}\,.$$

Somit ist die Stabilitätsbedingung (5.48) äquivalent zu

$$\frac{b\,\mathcal{J}}{m}\,\nu > 1\,. \tag{5.49}$$

Technische Entwicklungen im 19. Jahrhundert, die eigentlich die Maschinen verbessern sollten, wirkten der Bedingung (5.49) entgegen:

- Im Zusammenhang mit der Steigerung der Kapazität der Maschinen wurde das Gewicht der Schieber vergrößert und deshalb die Masse der Kugeln am Regler erhöht.
- Verbesserte Oberflächen führten zu einer Verringerung der Reibung.
- Das Trägheitsmoment des Schwungrades wurde zur Erhöhung der Arbeitsgeschwindigkeit der Maschinen verkleinert.
- Man strebte danach, die Abhängigkeit der Geschwindigkeit von der Last zu verringern, was zu einer Verkleinerung der Ungleichförmigkeit des Laufes führte.

5.14 Globale Stabilitätstheorie

Definition 5.64 *Es seien $M_1, M_2 \subset \mathbb{R}^n$ zwei zusammenhängende offene Mengen und $v^j : M_j \longrightarrow \mathbb{R}^n$, $j = 1, 2$, zwei stetig differenzierbare Geschwindigkeitsfelder. Die Systeme*

$$\dot{x} = v^1(x) \tag{5.50}$$

und

$$\dot{y} = v^2(y) \tag{5.51}$$

heißen zueinander **topologisch äquivalent***, wenn ein Homöomorphismus $H : M_1 \longrightarrow M_2$ existiert, der Integralkurven von (5.50) auf Integralkurven von (5.51) abbildet. D.h., sind $\Phi(t,x)$ und $\Psi(\tau,y)$ die (lokalen) Flüsse von (5.50) bzw. (5.51), so existiert für jedes $x \in M_1$ eine bezüglich τ streng monoton wachsende stetige Funktion $t(x,\tau)$, so dass*

$$H(\Phi(t(x,\tau),x)) = \Psi(\tau, H(x)) \tag{5.52}$$

für alle $x \in M_1$ und alle $\tau \in \big(a_{H(x)}, b_{H(x)}\big)$ gilt.

Beispiel 5.65 *$M = (0,\infty)$, $\dot{x} = x^2$, $x(0) = x^0 > 0$. Der Fluß dieses Systems ist gegeben durch*

$$\Phi(t,x) = \frac{x}{1-xt}\,, \quad -\infty < t < \frac{1}{x}\,.$$

Die Funktion $\tau_x(t) = \tau(x,t) = t + \dfrac{x^2 t}{1-x\,t}$ bildet $-\infty < t < \dfrac{1}{x}$ eineindeutig auf $-\infty < \tau < \infty$ ab, wobei $\tau(x,0) = 0$ gilt. Mit $t_x(\tau) = t(x,\tau)$ bezeichnen wir die entsprechende Umkehrabbildung. Die Gleichung

$$\dot{y} = \frac{y^2}{1+y^2}$$

ist global integrierbar, und es gilt für $\psi(\tau) = \Phi(t(x,\tau),x)$ *die Beziehung*

$$\psi'(\tau) = \frac{\partial\Phi(t,x)}{\partial t}\frac{1}{\partial\tau/\partial t} = \frac{[\Phi(t,x)]^2}{1+\dfrac{(1-xt)x^2+x^3t}{(1-xt)^2}} = \frac{[\psi(\tau)]^2}{1+[\psi(\tau)]^2}\,.$$

Es gilt also (5.52) mit $H : (0,\infty) \longrightarrow (0,\infty)\,,\ x \mapsto x\,,$ *als Identität.*

Theorem 5.66 *Zu jedem System (5.50) existiert ein System (5.51), welches global integrierbar und topologisch äquivalent zu (5.50) ist.*

Wir betrachten deshalb im weiteren nur global integrierbare Systeme

$$\dot{x} = v(x) \tag{5.53}$$

und das dazugehörige dynamische System $\Phi : \mathbb{R} \times M \longrightarrow M$, wobei $v : M \longrightarrow \mathbb{R}^n$ als stetig differenzierbar vorausgesetzt wird.

Definition 5.67 *Ein Punkt* $p \in M$ *heißt* ω-**Grenzpunkt** (α-**Grenzpunkt**) *des Orbits* $\Gamma = \Gamma_x = \{\Phi(t,x) : t \in \mathbb{R}\}$, *wenn eine Folge* $\{t_j\}_{j=1}^{\infty} \subset \mathbb{R}$ *mit* $\lim_{j\to\infty} t_j = +\infty$ ($\lim_{j\to\infty} t_j = -\infty$) *und*

$$\lim_{j\to\infty} \Phi(t_j,x) = p$$

existiert. Die Menge aller ω- *bzw.* α-*Grenzpunkte heißt* ω- *bzw.* α-*Grenzmenge von* Γ *und wird mit* $\omega(\Gamma)$ *bzw.* $\alpha(\Gamma)$ *bezeichnet. Die Menge* $\omega(\Gamma) \cup \alpha(\Gamma)$ *heißt Grenzmenge von* Γ .

Theorem 5.68 *Die* α- *und* ω-*Grenzmengen sind abgeschlossene Teilmengen von* M . *Ist* $K \subset M$ *kompakt und* $\Gamma \subset K$, *so sind* $\alpha(\Gamma)$ *und* $\omega(\Gamma)$ *nicht leer, zusammenhängend und kompakt.*

Theorem 5.69 *Aus* $p \in \alpha(\Gamma)$ *folgt* $\Gamma_p \subset \alpha(\Gamma)$, *und aus* $p \in \omega(\Gamma)$ *folgt* $\Gamma_p \subset \omega(\Gamma)$.

Insbesondere sind also $\alpha(\Gamma)$ und $\omega(\Gamma)$ bezüglich des Flusses Φ invariante Teilmengen des Phasenraumes, d.h. z.B.

$$\Phi(t,y) \in \alpha(\Gamma) \quad \forall y \in \alpha(\Gamma)\,,\ \forall t \in \mathbb{R}\,.$$

Hat also z.B. ein Orbit nur einen ω-Grenzpunkt, so ist dieser ein Gleichgewichtspunkt.

Definition 5.70 *Unter einer Umgebung einer Menge* $A \subset M$ *verstehen wir eine offene Menge* $U \subset M$ *mit* $A \subset U$. *Wir schreiben*

$$\Phi(t,x) \longrightarrow A \quad \text{für} \quad t \longrightarrow \infty\,,$$

wenn

$$\lim_{t\to\infty} \operatorname{dist}(\Phi(t,x),A) = \lim_{t\to\infty} \inf\{|\Phi(t,x)-y| : y \in A\} = 0\,.$$

Eine abgeschlossene, bezüglich Φ *invariante Menge* $A \subset M$ *heißt* **attraktive Menge**, *wenn eine Umgebung* $U \subset M$ *von* A *existiert, so dass* $\Phi(t,x) \longrightarrow A$ *für* $t \longrightarrow \infty$ *und* $\forall x \in U$. *Eine attraktive Menge heißt* **Attraktor**, *wenn sie einen dichten Orbit enthält, d.h. einen Orbit, dessen Abschließung* A *umfaßt.*

Beispiel 5.71 *Wir betrachten das System*

$$\begin{aligned} \dot{x} &= -y + x(1 - x^2 - y^2)\,, \\ \dot{y} &= \phantom{-y+{}} x + y(1 - x^2 - y^2) \end{aligned}$$

und verwenden zu dessen Untersuchung die Polarkoordinaten

$$x = r\cos\varphi\,,\ \ y = r\sin\varphi\,.$$

Es folgt

$$\begin{aligned} \dot{r} &= r(1 - r^2)\,, \\ \dot{\varphi} &= 1\,. \end{aligned}$$

Der Koordinatenursprung ist Gleichgewichtspunkt. Der Einheitskreis $\mathbb{T} = \{r = 1\}$ ist ein Orbit mit $\alpha(\mathbb{T}) = \omega(\mathbb{T}) = \mathbb{T}$ und ω-Grenzmenge für alle anderen Orbits außer dem Koordinatenursprung.

Beispiel 5.72

$$\begin{aligned} \dot{x} &= -y + x(1 - x^2 - y^2 - z^2)\,, \\ \dot{y} &= \phantom{-y+{}} x + y(1 - x^2 - y^2 - z^2)\,, \\ \dot{z} &= 0\,. \end{aligned}$$

Beispiel 5.73

$$\begin{aligned} \dot{x} &= -y + x(1 - x^2 - y^2)\,, \\ \dot{y} &= \phantom{-y+{}} x + y(1 - x^2 - y^2)\,, \\ \dot{z} &= \alpha > 0\,. \end{aligned}$$

Definition 5.74 *Ein periodischer Orbit (auch Zyklus genannt) Γ heißt* **stabil***, wenn für jedes $\varepsilon > 0$ eine Umgebung U von Γ existiert, so dass*

$$\mathrm{dist}(\Phi(t,x),\Gamma) < \varepsilon \quad \forall t \geq 0\,,\ \forall x \in U\,.$$

Sonst heißt Γ **instabil***. Der periodische Orbit Γ heißt* **asymptotisch stabil***, wenn er stabil ist und wenn eine Umgebung U von Γ existiert, so dass*

$$\lim_{t\to\infty} \mathrm{dist}(\Phi(t,x),\Gamma) = 0 \quad \forall x \in U$$

gilt.

Für eine gewisse Umgebung V von Γ definieren wir die lokale stabile bzw. instabile Mannigfaltigkeit von Γ durch

$$S(\Gamma) = \left\{x \in V : \lim_{t\to\infty} \mathrm{dist}(\Phi(t,x),\Gamma) = 0\right\}$$

bzw.

$$U(\Gamma) = \left\{x \in V : \lim_{t\to-\infty} \mathrm{dist}(\Phi(t,x),\Gamma) = 0\right\} .$$

Die entsprechenden globalen Mannigfaltigkeiten sind dann gegeben durch

$$W^s(\Gamma) = \bigcup_{t\le 0} \{\Phi(t,x) : x \in S(\Gamma)\} \quad \text{bzw.} \quad W^u(\Gamma) = \bigcup_{t\ge 0} \{\Phi(t,x) : x \in U(\Gamma)\} .$$

Beispiel 5.75

$$\begin{aligned} \dot{x} &= -y + x(1-x^2-y^2) , \\ \dot{y} &= \phantom{-y+{}} x + y(1-x^2-y^2) , \\ \dot{z} &= z . \end{aligned}$$

Es existiert ein (instabiler) periodischer Orbit

$$\Gamma = \{(\cos t, \sin t, 0) : t \in \mathbb{R}\} .$$

Der Koordinatenurpsrung ist Gleichgewichtspunkt. Weitere invariante Mengen sind die z-Achse, die x-y–Ebene und der Zylinder

$$\left\{(x,y,z) : x^2+y^2 = 1\right\} = \{(r,\varphi,z) : r = 1\} .$$

Es gilt

$$W^s(\Gamma) = \left\{(x,y,0) : x^2+y^2 > 0\right\} \quad \textit{und} \quad W^u(\Gamma) = \left\{(x,y,z) : x^2+y^2 = 1\right\} .$$

Definition 5.76 *Einen periodischen Orbit Γ nennt man* **Grenzzyklus**, *wenn er α- oder ω-Grenzmenge eines anderen Orbits ist. Gilt $\Gamma = \omega(\Gamma_x)$ für alle x aus einer Umgebung von Γ, so heißt Γ* **ω-Grenzzyklus** *oder* **stabiler Grenzzyklus**. *Ist $\Gamma = \alpha(\Gamma_x)$ für alle x aus einer Umgebung von Γ, so heißt Γ* **α-Grenzyklus** *bzw.* **instabiler Grenzzyklus**.

Ist also Γ ein stabiler Grenzzyklus, so ist er ein asymptotisch stabiler Zyklus und zugleich ein Attraktor.

Beispiel 5.77

$$\begin{aligned} \dot{x} &= -y + x(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} , \\ \dot{y} &= \phantom{-y+{}} x + y(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}} . \end{aligned}$$

Dieses System ist äquivalent zu

$$\dot{r} = r^3 \sin \frac{1}{r}, \quad \dot{\varphi} = 1 .$$

Periodische Orbits sind

$$\Gamma_{(n)} = \left\{ \sqrt{x^2 + y^2} = \frac{1}{n\pi} \right\}, \quad n = 1, 2, 3, \ldots$$

Dabei sind die $\Gamma_{(2n)}$, $n = 1, 2, \ldots$, *stabil und die* $\Gamma_{(2n-1)}$, $n = 1, 2, \ldots$, *instabil.*

Bemerkung 5.78 *Für ebene Systeme gilt folgendes: Ist ein periodischer Orbit* Γ ω*-Grenzmenge eines Orbits außerhalb von* Γ *, so ist* Γ ω*-Grenzmenge für alle* Γ_x *mit* x *in einer "äußeren" Umgebung von* Γ *. Außerdem windet sich jeder dieser Orbits* Γ_x *für* $t \longrightarrow \infty$ *um* Γ *herum, und zwar in dem Sinne, dass jede Gerade senkrecht zu* Γ *zu unendlich vielen Zeitpunkten* t_n *von* Γ_x *geschnitten wird, wobei* $t_n \longrightarrow \infty$ *für* $n \longrightarrow \infty$ *. (Gilt auch für "innen", und gilt auch für* α*-Grenzmenge.)*

Beispiel 5.79 *Das System*

$$\dot{x} = y, \quad \dot{y} = x + x^2$$

ist ein Hamiltonsches System mit $H(x, y) = \frac{y^2}{2} - \frac{x^2}{2} - \frac{x^3}{3}$ *. Die Orbits liegen also auf den Kurven* $y^2 - x^2 - \frac{2}{3} x^3 = \text{const}$ *. Stationäre Punkte sind der Koordinatenursprung und der Punkt* $(-1, 0)$ *. Die Kurve*

$$y^2 = x^2 + \frac{2}{3} x^3$$

geht durch die Punkte $(-\frac{3}{2}, 0)$ *und* $(0, 0)$ *und enthält vier Orbits, nämlich*

$$\Gamma = \left\{ (x, y) : y^2 = x^2 + \frac{2}{3} x^3 , \ -\frac{3}{2} \le x < 0 \right\},$$

den Koordinatenursprung und zwei weitere. Dabei gilt $\omega(\Gamma) = \alpha(\Gamma) = (0, 0)$ *, d.h.* Γ *gehört sowohl zur stabilen als auch zur instabilen Mannigfaltigkeit des Koordinatenursprungs.*

Definition 5.80 *Es seien* Γ_{x^0} *ein periodischer Orbit der (primen) Periode* $T > 0$ *und* Σ *die Hyperebene, die durch* x^0 *geht und senkrecht auf* Γ_{x^0} *steht. Dann existieren (wegen der Stetigkeit des Flusses) eine Umgebung* U *von* x^0 *und eine (eindeutig bestimmte) stetig differenzierbare Funktion* $\tau : U \longrightarrow \mathbb{R}$ *, so dass* $\tau(x^0) = T$ *(Periode von* Γ*) und* $\Phi(\tau(x), x) \in \Sigma$ *für alle* $x \in U$ *. Die Abbildung* $P : U \cap \Sigma \longrightarrow \Sigma$ *mit* $P(x) = \Phi(\tau(x), x)$ *wird* **Poincaré-Abbildung** *genannt.*

Wir betrachten nochmals

$$\dot{r} = r(1 - r^2), \ \dot{\varphi} = 1, \quad r(0) = r_0 > 0, \ \varphi(0) = \varphi_0 .$$

Es folgt

$$r(t) = \frac{1}{\sqrt{1 + \left(\frac{1}{r_0^2} - 1 \right) e^{-2t}}}, \quad \varphi(t) = t + \varphi_0 .$$

Für die Poincarè-Abbildung im Punkt $(r,\varphi) = (1,\varphi_0)$ des Einheitskreises gilt:

$$P(r_0) = \frac{1}{\sqrt{1+\left(\frac{1}{r_0^2}-1\right)e^{-4\pi}}}\,, \quad P(1) = 1\,,$$

$$P'(r_0) = e^{-4\pi} r_0^{-3}\left[1+\left(\frac{1}{r_0^2}-1\right)e^{-4\pi}\right]^{-3/2}\,,$$

d.h.

$$P'(1) = e^{-4\pi} < 1\,.$$

Wir betrachten jetzt ebene Systeme. Es seien $x^0 \in \Gamma$, Γ ein periodischer Orbit und Σ die Gerade senkrecht zu Γ im Punkt x^0. Der Punkt x^0 zerlegt Σ in Σ^+ und Σ^-, wobei Σ^+ ins Äußere von Γ zeige. Mit s bezeichnen wir den vorzeichenbehafteten Abstand von x^0 für die Punkte auf Σ, wobei $s > 0$ für Punkte aus Σ^+ und $s < 0$ für Punkte aus Σ^- gelte. Wir können jetzt die Poincarè-Abbildung mit $P(s)$ bezeichnen. Sie ist für $|s| < \delta$ ($\delta > 0$ hinreichend klein) erklärt, wobei $P(0) = 0$ gilt. Mit $d(s)$ bezeichnen wir die Verschiebungsfunktion $d(s) := P(s) - s$. Dann ist $d(0) = 0$ und $d'(s) = P'(s) - 1$. Aus dem Mittelwertsatz der Differentialrechnung folgt $d(s) = d(s) - d(0) = d'(\xi)s$. Da $d'(s)$ stetig ist, hat für hinreichend kleine $|s|$ die Größe $d'(\xi)$ dasselbe Vorzeichen wie $d'(0)$, d.h. für hinreichend kleine $|s|$ gilt:

- Ist $d'(0) < 0$ (also $P'(0) < 1$), so ist $d(s) < 0$ für $s > 0$ und $d(s) > 0$ für $s < 0$, und somit ist Γ ein stabiler Grenzzyklus.

- Ist $d'(0) > 0$ (also $P'(0) > 1$), so ist $d(s) > 0$ für $s > 0$ und $d(s) < 0$ für $s < 0$, und somit ist Γ ein instabiler Grenzzyklus.

Bemerkung 5.81 *Ist Γ_{x^0} ein periodischer Orbit in $\mathbb{R}^2$ der Periode T und $P(s)$ die wie oben definierte Poincaré-Abbildung bzgl. der Normalen zu Γ_{x^0} im Punkt x^0, so gilt*

$$P'(0) = \exp\left\{\int_0^T \nabla\cdot v\left(\Phi(t,x^0)\right)\,dt\right\}.$$

Ist also

$$\int_0^T \nabla\cdot v\left(\Phi(t,x^0)\right)\,dt < 0 \quad (>0)\,,$$

so ist Γ_{x^0} ein stabiler (instabiler) Grenzzyklus.

5.15 Übungsaufgaben

1. Ermitteln Sie für das System

$$\begin{aligned}\dot x_1 &= -x_1\,,\\ \dot x_2 &= x_1^2 + x_2\,,\end{aligned}$$

d.h.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ x_1^2 \end{bmatrix},$$

die stabile und instabile Mannigfaltigkeit im Nullpunkt. Hinweis: Man betrachte eine äquivalente Integralgleichung und löse diese mittels sukzessiver Approximation. Mit der Startnäherung $u_1^0 = u_2^0 = 0$ bleibt die entstehende Folge ab dem dritten Iterationsschritt stehen. Vgl. auch Bsp. 5.43.

2. Zeigen Sie die Gültigkeit der in Beispiel 5.51 formulierten Aussage. Ermitteln Sie unter Verwendung dieser Tatsache die stabile und instabile Mannigfaltigkeit des nichtlinearen Systems im Beispiel 5.48.

3. Diskutieren Sie das Stabilitätsverhalten der Gleichgewichtspunkte des Systems $\dot{x} = v(x)$. Hinweis: Benutzen Sie das Hartman-Grobman-Theorem.

   (a) $v(x) = \begin{bmatrix} x_1^2 - x_2^2 - 1 \\ 2\,x_2 \end{bmatrix}$ (b) $v(x) = \begin{bmatrix} x_2 - x_1^2 + 2 \\ 2\,x_2^2 - 2\,x_1 x_2 \end{bmatrix}$

   (c) $v(x) = \begin{bmatrix} -4\,x_1 - 2\,x_2 + 4 \\ x_1 x_2 \end{bmatrix}$

4. Zeigen Sie mit Hilfe einer Ljapunov-Funktion der Gestalt $\mathcal{L}(x_1, x_2) = c_1 x_1^2 + c_2 x_2^2$, dass der Nullpunkt für das System

$$\dot{x} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} x + F(x)$$

   im Falle

   (a) $F(x) = \begin{bmatrix} -x_1^3 - x_1 x_2^2 \\ -x_2^3 - x_1^2 x_2 \end{bmatrix}$ asymptotisch stabil, (b) $F(x) = \begin{bmatrix} x_1^3 + x_1 x_2^2 \\ x_2^3 + x_1^2 x_2 \end{bmatrix}$ instabil,

   (c) $F(x) = \begin{bmatrix} -x_1 x_2 \\ x_1^2 \end{bmatrix}$ stabil, aber nicht asymptotisch stabil ist.

Kapitel 6

Einführung in die Wahrscheinlichkeitsrechnung

6.1 Der Begriff des Wahrscheinlichkeitsraumes

- Ausgangspunkt: beliebig oft wiederholbares Experiment (Versuch)

- Die Menge Ω aller, aber sich gegenseitig ausschließender Ausgänge des Experiments nennen wir **Raum der Elementarereignisse**. Die Elemente $\omega \in \Omega$ sind somit die **Elementarereignisse**.

- Beispiele:

  - Werfen einer Münze: $\Omega = \{w, z\}$
  - Würfeln: $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - Werfen einer Münze, bis zum ersten Mal das Wappen fällt: $\Omega = \{w, zw, zzw, \ldots\}$

- Man nennt Ω einen **diskreten Raum von Elementarereignissen**, wenn er aus höchstens abzählbar vielen Elementen besteht.

- Die Elemente der Potenzmenge $\mathcal{P}(\Omega)$ von Ω heißen **zufällige Ereignisse**. Ω ist das **sichere Ereignis**, $\Omega \setminus A$ das zu A **komplementäre Ereignis**. $A \in \mathcal{P}(\Omega)$ und $B \in \mathcal{P}(\Omega)$ sind **unvereinbar**, wenn $A \cap B = \emptyset$ gilt.

- Beispiel: Beim zweimaligen Werfen eines Würfels besteht Ω aus den $6^2 = 36$ Elementen $(i, j) \in \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$. Die zufälligen Ereignisse

$$A = \{(i, j) \in \Omega : i + j \leq 3\} \quad \text{und} \quad B = \{(i, 6) : i = 1, 2, 3, 4, 5, 6\}$$

  sind unvereinbar. Für $C = \{(i, j) \in \Omega : j = 2, 4, 6\}$ gilt $A \cap C = \{(1, 2)\}$.

- Es seien Ω ein diskreter Raum von Elementarereignissen und $P : \Omega \longrightarrow [0,1]$ eine Funktion mit $\sum\limits_{\omega\in\Omega} P(\omega) = 1$. Wir definieren $P : \mathcal{P}(\Omega) \longrightarrow [0,1]$ durch

$$P(\emptyset) = 0 \quad \text{und} \quad P(A) = \sum_{\omega\in A} P(\omega) \quad \text{für} \quad A \in \mathcal{P}(\Omega) \setminus \{\emptyset\} \ .$$

Dann gilt

  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
  - $P(\Omega \setminus A) = 1 - P(A)$,
  - $P\left(\bigcup\limits_{i=1}^{n} A_i\right) = \sum\limits_{i=1}^{n} P(A_i)$, falls $A_i \cap A_j = \emptyset$ für $i \neq j$.

- **Klassische Definition der Wahrscheinlichkeit**
Besteht Ω aus n Elementen und gilt $P(\omega) = \dfrac{1}{n}\ \forall\, \omega \in \Omega$ (diskrete Gleichverteilung), so ist

$$P(A) = \frac{\text{Anzahl der Elemente von } A}{n} \ .$$

Man nennt $\Omega = \{a_1, \ldots, a_n\}$ auch **Grundgesamtheit** und $\{a_{j_1}, \ldots, a_{j_k}\}$ eine **Stichprobe** vom Umfang k aus der Grundgesamtheit Ω.

Wir betrachten eine Stichprobe **ohne** Zurücklegen. Dann muss $k \leq n$ sein, und die Anzahl der möglichen Stichproben vom Umfang k ist gleich

$$n(n-1)(n-2)\cdots(n-k+1) =: (n)_k \ .$$

Damit ist z.B. die Wahrscheinlichkeit dafür, dass die ersten zwei Elemente einer Stichprobe vom Umfang k die Elemente a_1 und a_2 sind, gleich $\dfrac{2(n-2)_{k-2}}{(n)_k} = \dfrac{2}{n(n-1)}$, die Wahrscheinlichkeit dafür, dass an der ersten Stelle a_1 und an der zweiten Stelle a_2 steht, gleich $\dfrac{(n-2)_{k-2}}{(n)_k} = \dfrac{1}{n(n-1)}$.

Die Anzahl der möglichen Stichproben **mit** Zurücklegen vom Umfang k ist gleich n^k. Damit ist die Wahrscheinlichkeit dafür, dass in einer solchen Stichprobe vom Umfang $k \leq n$ alle Elemente verschieden sind, gleich

$$\frac{(n)_k}{n^k} \ .$$

Die Anzahl der Stichproben **ohne** Zurücklegen vom Umfang k, die die gleichen Elemente enthalten, ist gleich $k!$, so dass die Anzahl der Stichproben **ohne** Zurücklegen vom Umfang k mit unterschiedlicher Zusammensetzung gleich

$$\frac{(n)_k}{k!} = \binom{n}{k} = \text{Anzahl der Kombinationen zu } k \text{ aus } n \text{ Elementen}$$

ist.

- **Die hypergeometrische Verteilung**

  In einer Urne mögen n Kugeln liegen, von denen m schwarz und $n-m$ weiß sind. Wie groß ist die Wahrscheinlichkeit $P_{m,n}(m_1,k)$ dafür, dass in einer Stichprobe vom Umfang k ohne Zurücklegen genau m_1 schwarze Kugeln sind? Es gibt $\binom{n}{k}$ Stichproben mit "unterschiedlicher" (die Kugeln seien z.B. durchnumeriert) Zusammensetzung, $\binom{m}{m_1}$ Möglichkeiten m_1 Kugeln aus den m schwarzen Kugeln und $\binom{n-m}{k-m_1}$ Kugeln aus den $n-m$ weißen Kugeln auszuwählen. Somit gilt

$$P_{m,n}(m_1,k) = \frac{\binom{m}{m_1}\binom{n-m}{k-m_1}}{\binom{n}{k}}\,.$$

  $\{P_{m,n}(0,k), P_{m,n}(1,k), \ldots, P_{m,n}(k,k)\}$ nennt man hypergeometrische Verteilung. Es ergibt sich die Identität

$$\sum_{m_1=0}^{k}\binom{m}{m_1}\binom{n-m}{k-m_1} = \binom{n}{k}\,.$$

  Wie groß ist die Wahrscheinlichkeit dafür, dass auf einem Lotto-Schein für das Gewinnspiel "6 aus 49" die 6 Richtigen stehen? Antwort:

$$P_{6,49}(6,6) = \frac{1}{\binom{49}{6}} \approx 7,2\cdot 10^{-8}$$

- **Das Bernoulli-Schema**

  Ein Versuch habe zwei mögliche Ausgänge (z.B. 0 - Misserfolg, 1 - Erfolg). Es sei $p \in (0,1)$ die Wahrscheinlichkeit für einen Erfolg in einem Versuch. Dann ist die Wahrscheinlichkeit $P(k,r)$ dafür, dass bei r unabhängigen Versuchen genau k erfolgreich sind, gleich

$$P(k,r) = \binom{r}{k} p^k (1-p)^{r-k}\,.$$

  Insbesondere gilt

$$P(1,1) = p \quad \text{und} \quad \sum_{k=0}^{r} P(k,r) = 1\,.$$

  $\{P(0,r), P(1,r), \ldots, P(r,r)\}$ heißt **Binominalverteilung**.

Im weiteren sei Ω ein beliebiger Raum von Elementarereignissen.

Definition 6.1 *Ein System $\mathcal{F}$ von Teilmengen von Ω, d.h. $\mathcal{F} \subset \mathcal{P}(\Omega)$, nennt man eine* **$\sigma$-Algebra** *(oder Borelsches Feld), wenn folgende Axiome erfüllt sind:*

(Σ1) $\Omega \in \mathcal{F}$,

(Σ2) $A_n \in \mathcal{F}$, $n = 1,2,3,\ldots \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$,

(Σ3) $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$.

Das Paar $\langle\Omega,\mathcal{F}\rangle$ nennt man dann einen **messbaren Raum** *und die Elemente aus $\mathcal{F}$ (messbare)* **Ereignisse**. *Wird* $(\Sigma 2)$ *ersetzt durch*

$(A2)$ $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$,

so spricht man lediglich von einer **Algebra**.

Beispiel 6.2 *Es seien $\Omega = [0,1] \subset \mathbb{R}$ und $\mathcal{F}$ die Menge aller Teilmengen von $[0,1]$, die sich als Vereinigung endlich vieler Intervalle schreiben lassen. $\mathcal{F}$ ist eine Algebra, aber keine σ-Algebra. Der Durchschnitt aller σ-Algebren, die $\mathcal{F}$ enthalten, ist die kleinste σ-Algebra, die $\mathcal{F}$ enthält, und wird* **Borelsche σ-Algebra** *genannt. Die Elemente dieser σ-Algebra heißen* **Borelmengen** *oder* **Borel-messbare Mengen** *des Intervalls $[0,1]$.*

Definition 6.3 *Ist $\mathcal{A} \subset \mathcal{P}(\Omega)$, so nennt man die kleinste σ-Algebra, die $\mathcal{A}$ enthält, die von $\mathcal{A}$* **erzeugte σ-Algebra** *und bezeichnet sie mit $\sigma(\mathcal{A})$.*

Für eine σ-Algebra $\mathcal{F}$ gilt wegen $(\Sigma 1)$ und $(\Sigma 3)$ stets $\emptyset \in \mathcal{F}$, und wegen $(\Sigma 2)$ und $(\Sigma 3)$ folgt aus $B_n \in \mathcal{F}$, $n = 1, 2, 3, \ldots$,

$$\bigcap_{n=1}^{\infty} B_n = \Omega \setminus \bigcup_{n=1}^{\infty} (\Omega \setminus B_n) \in \mathcal{F}\,.$$

$\emptyset$ heißt **unmögliches** Ereignis, Ω **sicheres** Ereignis und $\overline{A} := \Omega \setminus A$ das zu $A \in \mathcal{F}$ **komplementäre** Ereignis. Gilt $A \cap B = \emptyset$, so nennt man die Ereignisse A und B **unvereinbar**.

Definition 6.4 *Eine reellwertige Funktion $P : \mathcal{F} \longrightarrow \mathbb{R}$ auf einer σ-Algebra $\mathcal{F}$ eines messbaren Raumes $\langle\Omega,\mathcal{F}\rangle$ wird* **Wahrscheinlichkeitsverteilung** *oder* **Wahrscheinlichkeitsmaß** *auf $\mathcal{F}$ genannt, wenn folgende Axiome erfüllt sind:*

$(W1)$ $P(A) \geq 0 \ \forall\, A \in \mathcal{F}$,

$(W2)$ $P(\Omega) = 1$,

$(W3)$ $A_n \in \mathcal{F}$, $n = 1, 2, 3, \ldots$, $A_i \cap A_j = \emptyset$ $(i \neq j)$ $\implies P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$.

Das Tripel $\langle\Omega,\mathcal{F},P\rangle$ heißt dann **Wahrscheinlichkeitsraum**. *P nennt man auch eine* **Mengenfunktion**. *Das Axiom (W3) kennzeichnet die sog.* **σ-Additivität** *der Mengenfunktion P.*

Bemerkung 6.5 $(W3)$ *ist äquivalent zur* **Additivität** *und dem* **Stetigkeitsaxiom**

$(W3')$ $B_n \in \mathcal{F}$, $B_{n+1} \subset B_n$, $n = 1, 2, \ldots$, $B = \bigcap_{n=1}^{\infty} B_n \implies \lim_{n\to\infty} P(B_n) = P(B)$.

Bemerkung 6.6 *Man spricht von einer Wahrscheinlichkeitsverteilung P auf einer Algebra $\mathcal{F}$, wenn man statt $(W3)$ verlangt, dass*

$(AW3)$ $A_n \in \mathcal{F}$, $n = 1, 2, \ldots$, $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, $A_i \cap A_j = \emptyset$ $(i \neq j)$ $\implies P(A) = \sum_{n=1}^{\infty} P(A_n)$.

6.2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Beispiel 6.7 *Beim dreimaligen Werfen einer Münze betrachten wir die Ereignisse*

$$A = \{\textit{Das Wappen fällt genau 1 Mal.}\} = \{(wzz), (zwz), (zzw)\}$$

und

$$B = \{\textit{Die Anzahl der gefallenen Wappen ist ungerade.}\} = A \cup \{(www)\}$$

mit den Wahrscheinlichkeiten $P(A) = 3/8$ *bzw.* $P(B) = 1/2$. *Wie groß ist die Wahrscheinlichkeit für das Eintreten des Ereignisses* A *unter der Bedingung, dass* B *eingetreten ist?*

Antwort: 3/4

Definition 6.8 *Es seien* $\langle \Omega, \mathcal{F}, P \rangle$ *ein Wahrscheinlichkeitsraum und* $A, B \in \mathcal{F}$. *Ist* $P(B) > 0$, *so ist die* **bedingte Wahrscheinlichkeit** *von* A *unter der Bedingung, dass* B *eintrat, gleich*

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

Die Ereignisse A *und* B *heißen* **unabhängig**, *wenn*

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

d.h. $P(A|B) = P(A)$ *(falls* $P(B) > 0$*), gilt. Die Ereignisse* $A_1, \ldots, A_n \in \mathcal{F}$ *nennt man* **vollständig unabhängig**, *wenn für beliebige* $1 \le i_1 < i_2 < \cdots < i_r \le n$ *die Beziehung*

$$P\left(\bigcap_{k=1}^{r} A_{i_k}\right) = \prod_{k=1}^{r} P(A_{i_k})$$

gilt.

Beispiel 6.9 *Zweimaliges Werfen einer Münze,* $A = \{(wz), (ww)\}$, $B = \{(wz), (zz)\}$. *Dann gilt* $P(A) = P(B) = 1/2$, $A \cap B = \{(wz)\}$ *und* $P(A \cap B) = 1/4 = P(A)P(B)$, *d.h.,* A *und* B *sind unabhängig.*

Beispiel 6.10 *Werfen eines Tetraeders mit einer roten, einer blauen, einer grünen und einer rot/blau/grünen Seite auf eine Ebene. Mit* R, B *und* G *seien die Ereignisse bezeichnet, dass eine Seite mit roter, blauer bzw. grüner Farbe auf der Ebene liegt. Es folgt* $P(R) = P(B) = P(G) = 1/2$, $P(R \cap B) = P(R \cap G) = P(B \cap G) = 1/4$, *aber* $P(R \cap B \cap G) = 1/4 \neq P(R)P(B)P(G)$. *Also ist paarweise Unabhängigkeit nicht hinreichend für vollständige Unabhängigkeit.*

Satz 6.11 *Es seien* $\langle \Omega, \mathcal{F}, P \rangle$ *ein Wahrscheinlichkeitsraum und* $A, B_1, \ldots, B_n \in \mathcal{F}$, $B_i \cap B_j = \emptyset$ $(i \neq j)$, $P(B_j) > 0 \ \forall j = 1, \ldots, n$ *und* $A \subset \bigcup_{j=1}^{n} B_j$. *Dann gilt die Formel von der* **totalen Wahrscheinlichkeit**

$$P(A) = \sum_{j=1}^{n} P(B_j) P(A|B_j). \tag{6.1}$$

Ist also $P(A) > 0$, so folgt die **Bayessche Formel**

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(B_j)P(A|B_j)}{\sum_{k=1}^{n} P(B_k)P(A|B_k)}. \tag{6.2}$$

Dieser Satz bleibt auch für $n = \infty$ gültig!

Beispiel 6.12 *Zwei Betriebe produzieren das gleiche Erzeugnis mit den Ausschussraten P_1 bzw. P_2, wobei der zweite Betrieb die k-fache Produktionsmenge im Vergleich zum ersten habe. (Die Produkte beider Betriebe werden gemischt.) Wie groß ist die Wahrscheinlichkeit dafür, dass ein Produkt aus der zweiten Fabrik ist, wenn es Ausschuss ist.*

Beispiel 6.13 *Radioaktiver Zerfall. Für $t \in \mathbb{R}$, $u > 0$ und $j \in \{0, 1, 2, 3, \ldots\}$ sei $B_j(t, u)$ das Ereignis, dass im Zeitintervall $[t, t + u)$ genau j Zusammenstöße von Teilchen stattfinden. Wir machen folgende Annahmen:*

(a) $B_j(v, t)$ *und* $B_k(v + t, u)$ *sind unabhängig,*

(b) $P(B_1(t, \delta)) = \lambda\delta + o(\delta)$,

(c) $\sum_{k=2}^{\infty} P(B_k(t, \delta)) = o(\delta)$.

Wir setzen $p_k(t) = P(B_k(v, t))$ und erhalten für $k \geq 1$ aus (6.1)

$$\begin{aligned}
p_k(t + \Delta t) &= \sum_{j=0}^{k} p_j(t) P(B_k(v, t + \Delta t)|B_j(v, t)) \\
&= \sum_{j=0}^{k} p_j(t) P(B_{k-j}(v + t, \Delta t)) \\
&= o(\Delta t) + p_{k-1}(t)P(B_1(v + t, \Delta t)) + p_k(t)P(B_0(v + t, \Delta t)) \\
&= o(\Delta t) + p_{k-1}(t)[\lambda\Delta t + o(\Delta t)] + p_k(t)[1 - \lambda\Delta t - o(\Delta t)],
\end{aligned}$$

so dass

$$\frac{p_k(t + \Delta t) - p_k(t)}{\Delta t} = \lambda p_{k-1}(t) - \lambda p_k(t) + \frac{o(\Delta t)}{\Delta t}.$$

Es folgt

$$p_k'(t) = \lambda[p_{k-1}(t) - p_k(t)], \quad k = 1, 2, 3 \ldots \tag{6.3}$$

Weiterhin gilt wegen $B_0(v, t + \Delta t) = B_0(v, t) \cap B_0(v + t, \Delta t)$ die Formel

$$p_0(t + \Delta t) = p_0(t)[1 - \lambda\Delta t - o(\Delta t)],$$

so dass

$$p_0'(t) = -\lambda p_0(t)\,. \tag{6.4}$$

Offenbar ist $p_0(0) = P(B_0(v,0)) = 1$ *und* $p_k(0) = 0$, $k = 1, 2, \ldots$ *Somit folgt aus* (6.4)

$$p_0(t) = p_0(0)e^{-\lambda t} = e^{-\lambda t}$$

und aus (6.3)

$$p_1'(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}\,,$$

so dass $p_1(t) = \lambda t e^{-\lambda t}$. *Induktiv kann man nun zeigen, dass*

$$p_k(t) = \frac{(\lambda t)^k}{k!}\, e^{-\lambda t}\,, \quad k = 0, 1, 2, \ldots\,, \tag{6.5}$$

gilt. Man sagt, dass $B_k(v,t)$ **Poisson-verteilt** *ist mit dem* **Parameter** λt. *Bei jedem Zusammenstoß zweier Teilchen entstehe mit der Wahrscheinlichkeit* $p \in (0,1)$ *ein Paar* μ*-Mesonen. Wie groß ist die Wahrscheinlichkeit dafür, dass pro Zeiteinheit k Paare von* μ*-Mesonen entstehen?*

6.3 Zufallsgrößen und Verteilungsfunktionen

Im weiteren seien $\langle \Omega, \mathcal{F}, P \rangle$ ein Wahrscheinlichkeitsraum und $\mathcal{B} \subset \mathcal{P}(\mathbb{R})$ die σ-Algebra der Borelmengen in $\mathbb{R}$.

Definition 6.14 *Eine Funktion* $f : \Omega \longrightarrow \mathbb{R}$ *heißt* **messbar**, *wenn* $f^{-1}(B) \in \mathcal{F}$ *für jede Borelmenge* $B \in \mathcal{B}$ *gilt. Da durch* $P_f(A) := P(f^{-1}(A))$ *eine Wahrscheinlichkeitsverteilung auf* $\mathcal{B}$ *definiert wird, nennt man* f *auch* **Zufallsgröße**. *Mit*

$$F(x) = F_f(x) = P_f((-\infty, x)) = P(\{\omega \in \Omega : f(\omega) < x\}$$

bezeichnen wir die **Verteilungsfunktion** *der Zufallsgröße* f. *Häufig verwendet man die Schreibweise* $F(x) = P(f < x)$.

Beispiel 6.15 *In einem Bernoulli-Schema mit der Erfolgswahrscheinlichkeit* p *betrachten wir als Elementarereignisse die Stichproben vom Umfang* n, *d.h.* $\Omega = \{(i_1, \ldots, i_n) : i_k \in \{0,1\}\}$. *Die Zufallsgröße*

$$f(i_1, \ldots, i_n) = \sum_{k=1}^{n} i_k = \textit{Anzahl der Erfolge}$$

nennt man **binominal verteilt** *mit den Parametern* p *und* n. *Es gilt*

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

und somit

$$F(x) = \sum_{k<x} \binom{n}{k} p^k (1-p)^{n-k}\,.$$

Dabei gilt $F(x) = 0$ *für* $x \leq 0$ *und* $F(x) = 1$ *für* $x > n$.

Beispiel 6.16 *Zufälliges Werfen eines Punktes in das Intervall* $[a,b]$, $\Omega = [a,b]$, $\mathcal{F}$ *- Borelmengen von* $[a,b]$. *Wir haben*

$$P([a,x)) = \frac{x-a}{b-a},$$

und für $f(x) = x$

$$F(x) = \begin{cases} 0 & : \quad x \le a, \\ \dfrac{x-a}{b-a} & : \quad a < x \le b, \\ 1 & : \quad b < x. \end{cases}$$

(**Gleichverteilung** *auf* $[a,b]$)

Eigenschaften einer Verteilungsfunktion:

(V1) $x_1 \le x_2 \implies F(x_1) \le F(x_2)$,

(V2) $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$,

(V3) $\lim\limits_{x\to x_0-0} = F(x_0)$ (linksseitige Stetigkeit) $\forall\, x_0 \in \mathbb{R}$.

Satz 6.17 *Zu jeder Funktion* $F : \mathbb{R} \longrightarrow \mathbb{R}$ *mit den Eigenschaften* $(V1)$–$(V3)$ *existieren ein Wahrscheinlichkeitsraum* $\langle \Omega, \mathcal{F}, P\rangle$ *und eine auf ihm definierte Zufallsgröße* f, *so dass* $F(x) = P(f < x)$ *für alle* $x \in \mathbb{R}$ *gilt.*

Normalverteilung oder Gauß-Verteilung:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(u-a)^2}{2\sigma^2}}\, du, \quad \sigma > 0,\ a \in \mathbb{R}$$

Cauchy-Verteilung:

$$F(x) = \frac{1}{\pi} \int_{-\infty}^{x} \frac{dt}{1+t^2}$$

Poisson-Verteilung:

$$P(f = m) = \frac{\mu^m}{m!}\, e^{-\mu}, \quad F(x) = \sum_{m<x} \frac{\mu^m}{m!}\, e^{-\mu}$$

Es gibt drei Typen von Verteilungsfunktionen:

- **Diskrete** Verteilungsfunktionen entsprechen zufälligen Größen, die höchsten abzählbar unendlich viele Werte annehmen. Dabei gilt

  $$P(f = x) = F(x+0) - F(x).$$

- **Absolut stetige** Verteilungsfunktionen lassen sich in der Form
$$F(x) = \int_{-\infty}^{x} g(t)\,dt$$
mit einer Borel-messbaren Funktion $g : \mathbb{R} \longrightarrow \mathbb{R}$ schreiben. Dabei gilt für $B \in \mathcal{B}$
$$P(f \in B) = \int_B g(t)\,dt$$
(vgl. Absch. 6.4). Die Funktion $g(t)$ nennt man **Wahrscheinlichkeitsdichte** .

- Die **singulären** Verteilungsfunktionen sind zwar stetig, haben aber die Eigenschaft, dass die Menge ihrer Wachstumspunkte das Lebesgue-Maß Null hat, d.h.
$$m\,\{x \in \mathbb{R} : F(x+\varepsilon) - F(x-\varepsilon) > 0 \;\forall\, \varepsilon > 0\} = 0\,.$$

6.4 Integration messbarer Funktionen

Es sei $\langle \Omega, \mathcal{F}, P \rangle$ ein Wahrscheinlichkeitsraum.

Definition 6.18 *Mit* $\chi_A : \Omega \longrightarrow \mathbb{R}$ *bezeichnen wir die* **charakteristische Funktion** *des Ereignisses* $A \in \mathcal{F}$, *d.h.*
$$\chi_A(\omega) = \begin{cases} 1 & : \quad \omega \in A\,, \\ 0 & : \quad \omega \notin A\,. \end{cases}$$
Diese ist offenbar eine Zufallsgröße. Unter einer **messbaren Treppenfunktion** *verstehen wir eine Zufallsgröße der Gestalt*
$$f(\omega) = \sum_{k=1}^{n} x_k \chi_{A_k}(\omega) \tag{6.6}$$
mit $x_k \in \mathbb{R}$, $A_k \in \mathcal{F}$, $A_k \cap A_j = \emptyset$ $(k \neq j)$, $\bigcup_{k=1}^{n} A_k = \Omega$. *Integral der Funktion* (6.6) *nennt man die Zahl*
$$\int_\Omega f\,dP = \int_\Omega f(\omega)\,P(d\omega) = \sum_{k=1}^{n} x_k P(A_k)\,.$$
(*Diese Definition ist korrekt, d.h. unabhängig von der Darstellung* (6.6) *der Zufallsgröße* f .) *Analog ist für* $A \in \mathcal{F}$
$$\int_A f\,dP = \int_\Omega f(\omega)\chi_A(\omega)\,P(d\omega) = \sum_{k=1}^{n} x_k P(A \cap A_k)\,.$$

Lemma 6.19 *Zu jeder messbaren, nichtnegativen und beschränkten Funktion* $f : \Omega \longrightarrow \mathbb{R}$ *existiert eine Folge* $\{f_n(\omega)\}_{n=1}^{\infty}$ *messbarer Treppenfunktionen, so dass* $f_n(\omega) \uparrow f(\omega)$ *für* $n \longrightarrow \infty$ *und für alle* $\omega \in \Omega$ *gilt. Ist* $\{g_n(\omega)\}_{n=1}^{\infty}$ *eine weitere solche Folge, so gilt*
$$\lim_{n\to\infty} \int_\Omega f_n\,dP = \lim_{n\longrightarrow\infty} \int_\Omega g_n\,dP\,. \tag{6.7}$$

Definition 6.20 *Den Grenzwert in* (6.7) *bezeichnet man mit*

$$\int_\Omega f\,dP\,.$$

Lemma 6.21 *Ist* $f : \Omega \longrightarrow \mathbb{R}$ *eine beschränkte messbare Funktion, so sind auch* $f^+(\omega) := \max\{0, f(\omega)\}$ *und* $f^-(\omega) := \max\{0, -f(\omega)\}$ *messbar. Dabei gilt* $f = f^+ - f^-$.

Definition 6.22 *Unter dem* **Integral** *einer beliebigen beschränkten messbaren Funktion* $f : \Omega \longrightarrow \mathbb{R}$ *versteht man die Zahl*

$$\int_\Omega f\,dP = \int_\Omega f^+\,dP - \int_\Omega f^-\,dP\,.$$

Eigenschaften des Integrals:

1. $A_n \in \mathcal{F}\,,\ A_i \cap A_j = \emptyset\ (i \neq j),\ \bigcup_{n=1}^{\infty} A_n = A \implies \int_A f\,dP = \sum_{n=1}^{\infty} \int_{A_n} f\,dP\,,$
2. $\int_A f\,dP + \int_A g\,dP = \int_A (f+g)\,dP\,,$
3. $a \int_A f\,dP = \int_A (af)\,dP \quad \forall\, a \in \mathbb{R}\,,$
4. $f(\omega) \le g(\omega)\ \forall\, \omega \in A \implies \int_A f\,dP \le \int_A g\,dP\,.$

6.5 Numerische Charakteristika von Zufallsgrößen

Sind F die Verteilungsfunktion der Zufallsgröße f und $g : \mathbb{R} \longrightarrow \mathbb{R}$ Borel-messbar, so erhält man

$$\int_\Omega g(f(\omega))P(d\omega) = \int_{\mathbb{R}} g(x)P_f(dx) =: \int_{\mathbb{R}} g(x)\,dF(x)\,. \tag{6.8}$$

Letzteres Integral nennt man ein Lebesgue-Stieltjes-Integral. Dabei gilt: Ist g eine stetige Funktion, so ist

$$\int_{\mathbb{R}} g(x)\,dF(x) = \lim_{a\to-\infty, b\to\infty} \lim_{\delta\to 0} \sum_{k=0}^{N-1} g(\xi_k)\left[F(x_{k+1}) - F(x_k)\right]\,,$$

wobei der innere Grenzwert über alle Zerlegungen $a = x_0 < x_1 < \cdots < x_N = b$ mit $\delta = \max\{|x_k - x_{k-1}| : k = 1, \ldots, N\} \longrightarrow 0$ des Intervalls $[a, b]$ und beliebige $\xi_k \in [x_k, x_{k+1}]$ zu bestimmen ist.

Definition 6.23 *Unter dem* **Erwartungswert** *einer Zufallsgröße* $f : \Omega \longrightarrow \mathbb{R}$ *(oder dem ersten Moment) versteht man die Zahl*

$$Mf = \int_\Omega f(\omega)P(d\omega) = \int_{\mathbb{R}} x\,dF(x)\,.$$

Im Fall einer diskreten Verteilung ist dieses Integral gleich

$$Mf = \sum_{k=1}^{\infty} x_k P(f = x_k) = \sum_{k=1}^{\infty} x_k \left[F(x_k + 0) - F(x_k)\right] .$$

Beispiel 6.24 *Sei f Bernoulli-verteilt, d.h. $P(f = 0) = q$, $P(f = 1) = p = 1 - q$. Es folgt $Mf = \sum_{k=0}^{1} kP(f = k) = 0 \cdot q + 1 \cdot p = p$. Ist g die Anzahl der Erfolge im Bernoulli-Schema (vgl. Bsp. 6.15) mit n Versuchen, so ist $g = \sum_{k=1}^{n} f_k$, wobei die f_k wie f Bernoulli-verteilt sind. Es folgt (Eigenschaften der Integrale!) $Mg = \sum_{k=1}^{n} Mf_k = n \cdot p$.*
Probe:

$$\begin{aligned} Mg &= \sum_{k=0}^{n} kP(g = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-(k-1)} = np \end{aligned}$$

Beispiel 6.25 *Ist $F(x) = \int_{-\infty}^{x} h(t)\, dt$, so folgt aus (6.8) (g(x)=x)*

$$Mf = \int_{-\infty}^{\infty} x\, dF(x) = \int_{-\infty}^{\infty} xh(x)\, dx .$$

Sei f normal verteilt, d.h.

$$h(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)^2}{2\sigma^2}} .$$

Dann ist

$$\begin{aligned} Mf &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a) e^{-\frac{(x-a)^2}{2\sigma^2}} dx + \frac{a}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2\sigma^2}} dt + a = a . \end{aligned}$$

Beispiel 6.26 *Sei f Poisson-verteilt mit dem Parameter μ. Dann ist*

$$Mf = \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \mu \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} = \mu .$$

Definition 6.27 *Sei* $B \in \mathcal{F}$ *mit* $P(B) > 0$. *Wir betrachten den Wahrscheinlichkeitsraum* $\langle \Omega, \mathcal{F}, P_B \rangle$ *mit* $P_B(A) = P(A|B)\ \forall A \in \mathcal{F}$.

$$M_B f = M(f|B) = \int_\Omega f(\omega) P_B(d\omega)$$

heißt dann **bedingter Erwartungswert** *der Zufallsgröße* $f : \Omega \longrightarrow \mathbb{R}$ *bzgl. des Ereignisses* B.

Folgerung 6.28 *Ist* $f = \sum_{k=1}^{n} x_k \chi_{A_k}$, *so folgt*

$$M_B f = \sum_{k=1}^{n} x_k P(A_k|B) = \sum_{k=1}^{n} x_k \frac{P(A_k \cap B)}{P(B)} = \frac{1}{P(B)} \int_B f(\omega) P(d\omega).$$

Hieraus ergibt sich die allgemeine Formel

$$M_B f = \frac{1}{P(B)} \int_B f(\omega) P(d\omega)$$

für jede Zufallsgröße $f : \Omega \longrightarrow \mathbb{R}$.

Folgerung 6.29 *Aus* $B_n \in \mathcal{F}$, $B_i \cap B_j = \emptyset\ (i \neq j)$, $P(B_n) > 0$ *und* $\bigcup_{n=1}^{\infty} B_n = \Omega$ *folgt*

$$Mf = \int_\Omega f(\omega) P(d\omega) = \sum_{n=1}^{\infty} \int_{B_n} f(\omega) P(d\omega) = \sum_{n=1}^{\infty} P(B_n) M(f|B_n).$$

(Analogon zur Formel über die totale Wahrscheinlichkeit)

Folgerung 6.30 $F(x|B) = P_B(f < x) = P(f < x|B)$ *ist die Verteilungsfunktion der Zufallsgröße* f *auf* $\langle \Omega, \mathcal{F}, P_B \rangle$ *und heißt* **bedingte Verteilungsfunktion** *von* f *unter der Bedingung* B *im Wahrscheinlichkeitsraum* $\langle \Omega, \mathcal{F}, P \rangle$.

Beispiel 6.31 *Die Zeit der ausfallfreien Arbeit eines technischen Gerätes sei eine Zufallsgröße* f *mit der Verteilungsfunktion* $F(x)$, $F(0) = P(f < 0) = 0$. *Es sei bekannt, dass das Gerät bereits die Zeit* a *arbeitet. Welche Verteilung hat die restliche Lebensdauer* f_a *und wie groß ist der Erwartungswert von* f_a? *Für* $x \geq 0$ *ist*

$$\begin{aligned} F_a(x) &= P(f_a < x) = P(f < x + a | f \geq a) = 1 - P(f \geq x + a | f \geq a) \\ &= 1 - \frac{P(f \geq x + a)}{P(f \geq a)} = 1 - \frac{1 - F(x + a)}{1 - F(a)} = \frac{F(x + a) - F(a)}{1 - F(a)}. \end{aligned}$$

Ferner ist nach Folgerung 6.28

$$\begin{aligned} Mf_a &= M_{\{f \geq a\}}(f - a) = \frac{1}{P(f \geq a)} \int_{\{f \geq a\}} (f - a)\, dP \\ &= \frac{1}{1 - F(a)} \int_a^\infty (x - a)\, dF(x) = \frac{1}{1 - F(a)} \int_0^\infty t\, dF(t + a). \end{aligned}$$

In vielen Situationen gilt

$$P(f \geq x) = e^{-\mu x}\,, \quad \mu > 0\,, \quad x \geq 0\,.$$

Dann folgt

$$F(x) = 1 - e^{-\mu x} = \mu \int_0^x e^{-\mu t} dt\,, \quad x \geq 0\,,$$

und

$$P(f_a \geq x) = 1 - F_a(x) = \frac{P(f \geq x + a)}{P(f \geq a)} = \frac{e^{-\mu(x+a)}}{e^{-\mu a}} = e^{-\mu x}\,.$$

Definition 6.32 *Unter der* **Varianz** *einer Zufallsgröße* $f : \Omega \longrightarrow \mathbb{R}$ *versteht man die Zahl*

$$Df = M(f - Mf)^2\,.$$

Sie ist ein Maß für die Streuung der Wahrscheinlichkeitsverteilung. Die Größe $\sqrt{Df}$ *heißt* **Standardabweichung**.

Folgerung 6.33

(a) $Df = M(f^2 - 2f(Mf) + (Mf)^2) = Mf^2 - 2(Mf)(Mf) + (Mf)^2 = Mf^2 - (Mf)^2\,,$

(b) $F(x) = \int_{-\infty}^x g(t)\,dt \implies Df = \int_{-\infty}^\infty (x - Mf)^2 g(x)\,dx\,,$

(c) $Df = \min\{M(f - b)^2 : b \in \mathbb{R}\}\,,$ *denn*

$$M(f - b)^2 = Mf^2 - 2b(Mf) + b^2 = Mf^2 - (Mf)^2 + (Mf - b)^2\,.$$

Beispiel 6.34 *Sei* f *mit den Parametern* a *und* σ *normal verteilt. Aus Beispiel 6.25 ergibt sich*

$$\begin{aligned} Df &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^\infty t^2 e^{-\frac{t^2}{2}} dt \\ &= -\frac{\sigma^2}{\sqrt{2\pi}} \left[t e^{-\frac{t^2}{2}} \right]_{-\infty}^\infty + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{t^2}{2}} dt = \sigma^2\,. \end{aligned}$$

Beispiel 6.35 *Sei* f *Poisson-verteilt mit dem Parameter* μ. *Aus Beispiel 6.26 folgt*

$$\begin{aligned} Df &= Mf^2 - (Mf)^2 = \sum_{k=0}^\infty k^2 \frac{\mu^k e^{-\mu}}{k!} - \mu^2 \\ &= \sum_{k=1}^\infty \frac{k(k-1) + k}{k!} \mu^k e^{-\mu} - \mu^2 \\ &= \sum_{k=2}^\infty \frac{\mu^k}{(k-2)!} e^{-\mu} + \sum_{k=1}^\infty \frac{\mu^k}{(k-1)!} e^{-\mu} - \mu^2 \\ &= \mu^2 + \mu - \mu^2 = \mu\,. \end{aligned}$$

Beispiel 6.36 *Sei f Bernoulli-verteilt. Wegen $f^2 = f$ folgt nach Beispiel 6.24*

$$Df = Mf - (Mf)^2 = p - p^2 = p(1-p) = pq\,.$$

Für identisch Bernoulli-verteilte Zufallsgrößen $f_1, \ldots, f_n$ ergibt sich somit für die binomial verteilte Zufallsgröße $f = \sum_{k=1}^{n} f_k$

$$Df = \sum_{k=1}^{n} Df_k = npq\,.$$

Dabei hat man zu berücksichtigen, dass

$$\begin{aligned} D(f_1+f_2) &= M(f_1+f_2)^2 - (M(f_1+f_2))^2 \\ &= Mf_1^2 + 2M(f_1f_2) + Mf_2^2 - (Mf_1)^2 - 2(Mf_1)(Mf_2) - (Mf_2)^2\,, \end{aligned}$$

$$P(f_1f_2 = 1) = P(f_1 = 1)P(f_2 = 1) = p^2\,, \quad \textit{also} \quad M(f_1f_2) = p^2 = (Mf_1)(Mf_2)\,,$$

und somit

$$D(f_1+f_2) = Df_1 + Df_2 = 2pq\,.$$

Definition 6.37 *Es seien $f_1, \ldots, f_n$ Zufallsgrößen auf $\langle \Omega, \mathcal{F}, P \rangle$. Man nennt dann die Abbildung $f : \Omega \longrightarrow \mathbb{R}^n$ mit*

$$f(\omega) = [f_1(\omega), \ldots, f_n(\omega)]^T$$

einen zufälligen Vektor. Bezeichnen wir mit $\mathcal{B}^n$ die σ-Algebra der Borelmengen im $\mathbb{R}^n$, so sind $P_f(B) = P(f \in B)$, $B \in \mathcal{B}^n$, das Verteilungsgesetz und

$$F(x_1, \ldots, x_n) = F_f(x_1, \ldots, x_n) = P(f_1 < x_1, \ldots, f_n < x_n)$$

die Verteilungsfunktion des zufälligen Vektors f . Die Zufallsgrößen $f_1, \ldots, f_n$ heißen (stochastisch) **unabhängig***, wenn*

$$P(f_1 \in B_1, \ldots, f_n \in B_n) = P(f_1 \in B_1) \cdots P(f_n \in B_n) \quad \forall B_j \in \mathcal{B}$$

gilt.

Die Zufallsgrößen $f_1, \ldots, f_n$ sind genau dann unabhängig, wenn die Beziehung

$$F_f(x_1, \ldots, x_n) = F_{f_1}(x_1) \cdots F_{f_n}(x_n) \quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n$$

erfüllt ist.

Für das Produkt zweier unabhängiger Zufallsgrößen gilt

$$M(f_1f_2) = (Mf_1)(Mf_2)\,.$$

Dies ergibt sich aus folgenden Überlegungen: Aus der Unabhängigkeit folgt $P_f(dx_1, dx_2) = P_{f_1}(dx_1)P_{f_2}(dx_2)$ und somit

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1,x_2)P_f(dx_1,dx_2) &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} g(x_1,x_2)P_{f_1}(dx_1)\right]P_{f_2}(dx_2)\\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} g(x_1,x_2)\,dF(x_1)\right]\,dF(x_2)\,.\end{aligned}$$

Nun braucht man nur noch $g(x_1, x_2) = x_1x_2$ zu wählen.

Für die Verteilungsfunktion $F_{f_1+f_2}(x)$ der Summe zweier unabhängiger Zufallsgrößen folgt

$$\begin{aligned}F_{f_1+f_2}(x) &= P(f_1+f_2<x) = \int\int_{x_1+x_2<x} P_f(dx_1,dx_2)\\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{x-x_1} P_{f_2}(dx_2)\right]P_{f_1}(dx_1)\\ &= \int_{-\infty}^{\infty} F_{f_2}(x-x_1)P_{f_1}(dx_1) = \int_{-\infty}^{\infty} F_{f_2}(x-t)\,dF_{f_1}(t)\,.\end{aligned}$$

Die Verteilungsfunktion der Summe zweier unabhängiger Zufallsgrößen ist also gleich der Faltung der Verteilungsfunktionen F_{f_1} und F_{f_2}. Ist $F_{f_2}(t) = \int_{-\infty}^{t} \varphi_2(\tau)\,d\tau$, so folgt

$$\begin{aligned}F_{f_1+f_2}(x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{x-t} \varphi_2(\tau)\,d\tau\,dF_{f_1}(t)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x} \varphi_2(s-t)\,ds\,dF_{f_1}(t)\\ &= \int_{-\infty}^{x}\int_{-\infty}^{\infty} \varphi_2(s-t)\,dF_{f_1}(t)\,ds =: \int_{-\infty}^{x} \varphi(s)\,ds\,.\end{aligned}$$

Ist außerdem $F_{f_1}(t) = \int_{-\infty}^{t} \varphi_1(\tau)\,d\tau$, so ist

$$\varphi(s) = \int_{-\infty}^{\infty} \varphi_2(s-t)\varphi_1(t)\,dt\,.$$

Definition 6.38 *Eine Zufallsgröße f heißt* **normiert** *(bzw. standardisiert), wenn $Mf = 0$ und $Df = 1$ gilt. Ist f eine Zufallsgröße mit $0 < Df < \infty$ und $|Mf| < \infty$, so ist*

$$f_1 = \frac{f - Mf}{\sqrt{Df}}$$

die entsprechende normierte Zufallsgröße. Unter dem **Korrelationskoeffizienten** *zweier Zufallsgrößen f und g versteht man die Zahl*

$$\rho(f,g) = M(f_1g_1)\,,$$

wobei f_1 und g_1 die entsprechenden normierten Zufallsgrößen sind.

Folgerung 6.39 *Es gilt stets* $|\rho(f,g)| \le 1$, *denn es ist*

$$0 \le D(f_1 \pm g_1) = M(f_1 \pm g_1)^2 = 2 \pm 2\rho(f,g),$$

woraus $\pm\rho(f,g) \le 1$ *folgt.*

Folgerung 6.40 *Sind* f *und* g *unabhängige Zufallsgrößen, so ist* $\rho(f,g) = 0$, *denn es ist* $M(f_1 g_1) = (Mf_1)(Mg_1) = 0$.

Folgerung 6.41 *Es ist* $|\rho(f,g)| = 1$ *genau dann, wenn reelle Zahlen* $a \ne 0$ *und* b *existieren, so dass* $P(g = af + b) = 1$ *gilt.*

Beweis. Es seien $P(g = af + b) = 1$ und $Mf = \alpha$, $\sqrt{Df} = \beta$. Dann gilt, da $D(af + b) = a^2 Df = a^2\beta^2$,

$$\rho(f,g) = M\left(\frac{f-\alpha}{\beta} \cdot \frac{af + b - a\alpha - b}{|a|\beta}\right) = M\left(\frac{a}{|a|}\frac{(f-\alpha)^2}{\beta^2}\right) = \operatorname{sgn} a\,.$$

Es sei jetzt $\rho(f,g) = 1$. Es folgt $D(f_1 - g_1) = 2 - 2\rho(f,g) = 0$ und somit $P((f_1 - g_1)^2 > 0) = 0$, also $P(f_1 = g_1) = 1$. Im Fall $\rho(f,g) = -1$ verwende man $D(f_1 + g_1) = 0$. □

Die Zufallsgrößen f und g heißen positiv (negativ) korreliert, wenn $\rho(f,g) > 0$ (< 0).

Beispiel 6.42 *Das gesendete Signal eines Senders sei eine Zufallsgröße* f. *Infolge von Störungen empfängt der Teilnehmer das Signal* $g = \alpha f + \Delta$. *Dabei sind* α *der Verstärkungsfaktor und* Δ *die Wirkung der Störungen. Es seien* f *und* Δ *unabhängige Zufallsgrößen mit* $Mf = a$, $Df = 1$, $M\Delta = 0$, $D\Delta = \sigma^2$. *Verwendet man*

$$\begin{aligned} D(\alpha f + \Delta) &= M(\alpha f + \Delta)^2 - \alpha^2 a^2 = \alpha^2 M f^2 + \sigma^2 - \alpha^2 a^2 \\ &= \alpha^2(1 + a^2) + \sigma^2 - \alpha^2 a^2 = \alpha^2 + \sigma^2, \end{aligned}$$

so folgt

$$\rho(f,g) = M\left((f-a)\frac{\alpha f + \Delta - \alpha a}{\sqrt{\alpha^2+\sigma^2}}\right) = \frac{\alpha}{\sqrt{\alpha^2+\sigma^2}} M(f-a)^2 = \frac{\alpha}{\sqrt{\alpha^2+\sigma^2}}\,.$$

Ist nun $\sigma >> \alpha$, *so ist* $\rho(f,g) \approx 0$, *d.h.* g *ist praktisch unabhängig von* f. *Ist* $\sigma << \alpha$, *so ist* $\rho(f,g) \approx 1$, *und* f *ist aus* g *leicht wieder reproduzierbar.*

6.6 Das Gesetz der großen Zahlen

Lemma 6.43 (Tschebyscheffsche Ungleichung) *Ist* $P(f \ge 0) = 1$, *so gilt für jedes* $\varepsilon > 0$ *die Ungleichung*

$$P(f \ge \varepsilon) \le \frac{Mf}{\varepsilon}\,.$$

Beweis. Wir definieren $\Omega_\varepsilon = \{\omega \in \Omega : f(\omega) \geq \varepsilon\}$ und erhalten

$$\begin{aligned} Mf &= \int_\Omega f(\omega)P(d\omega) = \int_{\Omega_0} f(\omega)P(d\omega) \\ &\geq \int_{\Omega_\varepsilon} f(\omega)P(d\omega) \geq \varepsilon \int_{\Omega_\varepsilon} P(d\omega) = \varepsilon P(f \geq \varepsilon)\,. \end{aligned}$$

□

Satz 6.44 (Schwaches Gesetz der großen Zahlen) *Es seien f_j, $j = 1, 2, \dots$, unabhängig identisch Bernoulli-verteilte Zufallsgrößen mit $P(f_j = 1) = p$ und $P(f_j = 0) = q = 1 - p$, und es sei $S_n = f_1 + \cdots + f_n$ (die Anzahl der Erfolge in den ersten n Versuchen). Dann gilt für jedes $\varepsilon > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) = 0\,.$$

Beweis. Unter Verwendung von Lemma 6.43 und Beispiel 6.36 erhält man

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) &= P\left(\left[\frac{S_n}{n} - p\right]^2 \geq \varepsilon^2\right) \\ &\leq \frac{M\left(\frac{S_n}{n} - p\right)^2}{\varepsilon^2} = \frac{M(S_n - np)^2}{\varepsilon^2 n^2} = \frac{npq}{\varepsilon^2 n^2} \\ &= \frac{pq}{\varepsilon^2 n} \longrightarrow 0 \quad (n \longrightarrow \infty)\,. \end{aligned}$$

□

Satz 6.45 (Starkes Gesetz der großen Zahlen) *Mit den Voraussetzungen des Satzes 6.44 gilt für jedes $\varepsilon > 0$*

$$\lim_{n\to\infty} P\left(\sup_{k\geq n}\left|\frac{S_k}{k} - p\right| \geq \varepsilon\right) = 0\,.$$

Beweis. Aus Lemma 6.43 folgt

$$\begin{aligned} P\left(\sup_{k\geq n}\left|\frac{S_k}{k} - p\right| \geq \varepsilon\right) &= P\left(\bigcup_{k=n}^{\infty}\left\{\left|\frac{S_k}{k} - p\right| \geq \varepsilon\right\}\right) \\ &\leq \sum_{k=n}^{\infty} P\left(\left|\frac{S_k}{k} - p\right| \geq \varepsilon\right) = \sum_{k=n}^{\infty} P\left(\left[\frac{S_k}{k} - p\right]^4 \geq \varepsilon^4\right) \\ &\leq \sum_{k=n}^{\infty} \frac{M(S_k - kp)^4}{k^4\varepsilon^4}\,. \end{aligned}$$

Aus

$$\begin{aligned}
M(S_k - kp)^4 &= M\left[\sum_{j=1}^{k}(f_j - p)\right]^4 \\
&= \sum_{j=1}^{k} M(f_j - p)^4 + 6 \sum_{i<j\le k} M\left[(f_i - p)^2(f_j - p)^2\right] \\
&= k(pq^4 + p^4q) + 3(k-1)k(pq)^2 \le k + (k-1)k = k^2
\end{aligned}$$

folgt

$$P\left(\sup_{k\ge n}\left|\frac{S_k}{k} - p\right| \ge \varepsilon\right) \le \frac{1}{\varepsilon^4}\sum_{k=n}^{\infty}\frac{1}{k^2} \longrightarrow 0 \quad (n \longrightarrow \infty)\,.$$

□

Folgerung 6.46 *Es seien $f : [0,1] \longrightarrow \mathbb{R}$ stetig und $\varepsilon > 0$ beliebig. Unter Verwendung von Folgerung 6.29 ergibt sich*

$$\begin{aligned}
M\left|f\left(\frac{S_n}{n}\right) - f(p)\right| &= M\left(\left|f\left(\frac{S_n}{n}\right) - f(p)\right|; \left|\frac{S_n}{n} - p\right| < \varepsilon\right)\cdot P\left(\left|\frac{S_n}{n} - p\right| < \varepsilon\right) \\
&\quad + M\left(\left|f\left(\frac{S_n}{n}\right) - f(p)\right|; \left|\frac{S_n}{n} - p\right| \ge \varepsilon\right)\cdot P\left(\left|\frac{S_n}{n} - p\right| \ge \varepsilon\right) \\
&\le \sup_{|x|<\varepsilon}|f(p+x) - f(p)| + c\,P\left(\left|\frac{S_n}{n} - p\right| \ge \varepsilon\right)\,.
\end{aligned}$$

Es folgt

$$\limsup_{n\to\infty} M\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \le \sup_{|x|<\varepsilon}|f(p+x) - f(p)|$$

und somit

$$\lim_{n\to\infty} Mf\left(\frac{S_n}{n}\right) = f(p) \quad \textit{glm. bzgl.} \quad p \in [0,1]\,.$$

Folgerung 6.47 *Wegen $P(S_n = k) = \binom{n}{k}p^k(1-p)^{n-k}$ ist*

$$Mf\left(\frac{S_n}{n}\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}\,.$$

Aus Folgerung 6.46 ergibt sich somit

$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} \longrightarrow f(x)$$

glm. bzgl. $x \in [0,1]$. (Weierstraßscher Approximationssatz, Bernsteinsche Polynome)

Approximation der Binomialverteilung: Wir setzen

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad \text{und} \quad \Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt$$

(Dichte und Verteilungsfunktion der Standardnormalverteilung) und

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

Satz 6.48 (de Moivre-Laplace) *Es seien f_j , $j = 1, 2, \ldots$, unabhängig identisch Bernoulli-verteilte Zufallsgrößen mit $P(f_i = 1) = p$, $P(f_i = 0) = q = 1-p$, $S_n = \sum_{j=1}^{n} f_j$ und $S_n^* = \dfrac{S_n - np}{\sqrt{npq}}$ (standardisierte Binomialverteilung) sowie $a < b$. Dann gilt*

$$\lim_{n \longrightarrow \infty} P(a \le S_n^* \le b) = \Phi(b) - \Phi(a) .$$

Man kann die Binomialverteilung auch durch die Poissonverteilung approximieren, und zwar bei kleinen Erfolgswahrscheinlichkeiten. Es gilt sogar etwas allgemeiner

Satz 6.49 *Sind $f_1, \ldots, f_n$ unabhängige Bernoulli-verteilte Zufallsgrößen mit $P(f_i = 1) = p_i$ und $\mu = p_1 + \cdots + p_n$, $f = f_1 + \cdots + f_n$, so ist*

$$\sum_{k=0}^{n} \left| P(f = k) - \frac{\mu^k}{k!}\, e^{-\mu} \right| \le 2 \sum_{i=1}^{n} p_i^2 .$$

Folgerung 6.50 *Sind $0 \le \widetilde{p}_n \le 1$ und $\lim_{n\to\infty} n\widetilde{p}_n = \mu$, so gilt*

$$\lim_{n\to\infty} b_{n,\widetilde{p}_n}(k) = \frac{\mu^k}{k!}\, e^{-\mu} .$$

Beispiel 6.51 *Es soll der Prozentsatz der Wähler einer Partei geschätzt werden. Werden n Wähler befragt und sind darunter S_n Wähler dieser Partei, so sei S_n/n die Schätzung für die Wahrscheinlichkeit p , dass ein zufällig befragter Wähler ein Wähler dieser Partei ist. Wie groß muss n sein, damit die Wahrscheinlichkeit eines Irrtums von mehr als 1% nicht größer als 0.05 ist?*

Es ist also die Bedingung $P(-0.01 \le S_n/n - p \le 0.01) \approx 0.95$ zu erfüllen. Nach Satz 6.48 bedeutet dies

$$\begin{aligned} P\left(\frac{-0.01n}{\sqrt{npq}} \le S_n^* \le \frac{0.01n}{\sqrt{npq}}\right) &\approx \Phi\left(\frac{0,01n}{\sqrt{npq}}\right) - \Phi\left(\frac{-0.01n}{\sqrt{npq}}\right) \\ &= 2\Phi\left(\frac{0.01n}{\sqrt{npq}}\right) - 1 \approx 0.95 , \end{aligned}$$

also $\Phi(0.01n/\sqrt{npq}) \approx 0.975$. Aus $\Phi(1.96) \approx 0.975$ folgt $0.01\sqrt{n}/\sqrt{pq} \approx 1,96$, d.h. $n \approx pq \cdot 10^4 \cdot (1.96)^2$. Wegen $pq = p(1-p) \le 0.25$ sind $n = 2500 \cdot (1.96)^2 \approx 9600$ befragte Personen ausreichend. Ist bekannt, dass $p \le 0.1$ gilt, so genügen $n = 0.1 \cdot 0.9 \cdot 10^4 \cdot (1.96)^2 \approx 3450$ Befragungen.

Beispiel 6.52 *Die Wahrscheinlichkeit, heute Geburtstag zu haben ist gleich* $p = 1/365$. *Die Zahl der Geburtstagskinder unter* 20 *Personen, die sich in einem Raum befinden, ist dann praktisch Poisson-verteilt mit* $\mu = 20/365$.

Beispiel 6.53 *Von einem Produkt sei ein kleiner Anteil* $p = 0.015$ *bereits unmittelbar nach der Herstellung defekt. Wie groß muss man* n *wählen, damit ein Karton mit* n *(unabhängigen) Stück dieses Produktes mit Wahrscheinlichkeit* ≥ 0.8 *mindestens* 100 *intakte Exemplare enthält.*

Wir suchen also das kleinstmögliche n *mit*

$$0.8 \leq \sum_{k=0}^{n-100} b_{n,p}(k) =: r_n \,.$$

Mit $\mu_n = np$ *ist nach Folgerung 6.50*

$$r_n \approx \sum_{k=0}^{n-100} \frac{\mu_n^k}{k!}\, e^{-\mu_n} \,.$$

Rechnerisch findet man, dass $n = 102$ *die kleinste natürliche Zahl mit* $r_n \geq 0.8$ *ist.*

In der Menge der Verteilungsfunktionen $\{F, G, \ldots\}$ definieren wir die Metrik

$$d(F, G) = \sup\{|F(x) - G(x)| : x \in \mathbb{R}\} \,,$$

wofür wir auch $d(f, g)$ oder $d(f, G)$ schreiben, falls $F(x) = P(f < x)$ und $G(x) = P(g < x)$.

Satz 6.54 (Zentraler Grenzwertsatz) *Sind* $f_1, f_2, \ldots$ *unabhängig identisch verteilte Zufallsgrößen mit* $0 < \sigma^2 := Df_i < \infty$ *und*

$$S_n = f_1 + \cdots + f_n \,, \quad S_n^* = \frac{S_n - nMf_1}{\sigma\sqrt{n}} \,,$$

so gilt $d(S_n^*, \Phi) \longrightarrow 0$ $(n \longrightarrow \infty)$, *wobei*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \,.$$

6.7 Schätzungen

Beispiel 6.55 *Es sei die Zahl der Fische in einem Teich zu schätzen. Dazu werden* W *Fische gefangen, markiert und wieder eingesetzt. Von* n *gefangenen Fischen seien nun* x *markiert. Wir können die hypergeometrische Verteilung*

$$P_{W,N}(x, n) = \frac{\binom{W}{x}\binom{N-W}{n-x}}{\binom{N}{n}}$$

verwenden, bei der N unbekannt ist. Zum Schätzen dieser Zahl verwenden wir einen **Maximum-Likelihood-Ansatz**, *d.h. wir nehmen das $\widehat{N}$ als Schätzwert, für das die Wahrscheinlichkeit $P_{W,\widehat{N}}(x,n)$ maximal wird. Aus*

$$\frac{P_{W,N}(x,n)}{P_{W,N-1}(x,n)} = \frac{(N-W)(N-n)}{N(N-W-n+x)}$$

folgt, dass $P_{W,N}(x,n) > P_{W,N-1}(x,n)$ genau dann gilt, wenn $Wn > Nx$ ist. Wir wählen also für $\widehat{N}$ die größte ganze Zahl $\leq Wn/x$.

Verallgemeinerung:

- X - nichtleere, höchstens abzählbare Menge, der **Stichprobenraum** (im Beispiel 6.55 ist $X = \{0,1,2,\ldots,n\}$),
- $\{P_\theta : \theta \in \Theta\}$ - Familie von Wahrscheinlichkeitsmaßen auf X (im Beispiel 6.55 ist $P_\theta(x) = P_{W,\theta}(x,n)$, $\Theta = \{0,1,2,\ldots\}$),
- $g(\theta)$ - zu schätzende Funktion (im Beispiel 6.55 ist $g(\theta) = \theta$).
- $L_x : \Theta \longrightarrow \mathbb{R}$ mit $L_x(\theta) = P_\theta(x)$ heißt **Likelihood-Funktion**.
- Ein $\widehat{\theta}(x)$ mit $L_x(\widehat{\theta}(x)) = \sup\{L_x(\theta) : \theta \in \Theta\}$ heißt **Maximum-Likelihood-Schätzung** von θ und $g(\widehat{\theta}(x))$ Maximum-Likelihood-Schätzung von $g(\theta)$.

Beispiel 6.56 *Wir schätzen die Erfolgswahrscheinlichkeit im Bernoulli-Schema mit n Versuchen und x Erfolgen: Es ist $\theta = p$ und*

$$L_x(p) = \binom{n}{x} p^x (1-p)^{n-x}\,.$$

Wir verwenden die sog. log-Likelihood-Funktion

$$\mathcal{L}_x(p) = \ln L_x(p) = \ln \binom{n}{x} + x \ln p + (n-x)\ln(1-p)\,.$$

Es folgt

$$\frac{d}{dp}\mathcal{L}_x(p) = \frac{x}{p} - \frac{n-x}{1-p} \quad \text{und} \quad \frac{d^2}{dp^2}\mathcal{L}_x(p) = -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}\,,$$

woraus sich die Maximum-Likelihood-Schätzung $\widehat{p} = \dfrac{x}{n}$ ergibt.

Für $T : X \longrightarrow \mathbb{R}$ und $\theta \in \Theta$ definieren wir den Erwartungswert von T bzgl. P_θ als

$$M_\theta T = \sum_{x \in X} T(x) P_\theta(x)\,.$$

Eine Schätzung $\widehat{g} : X \longrightarrow \mathbb{R}$ heißt **erwartungstreue Schätzung** von $g(\theta)$, wenn $M_\theta \widehat{g} = g(\theta)$ gilt. Speziell heißt $\widehat{\theta}$ erwartungstreue Schätzung von θ, wenn $M_\theta \widehat{\theta} = \theta$. Die Größe

$$b(\theta, \widehat{g}) = M_\theta \widehat{g} - g(\theta)$$

nennt man **Bias** der Schätzung $\widehat{g}$. Im Beispiel 6.56 ist $\widehat{\theta}(x) = \frac{x}{n}$. Da x binomialverteilt mit der Erfolgswahrscheinlichkeit θ ist, gilt $M_\theta \widehat{\theta} = \theta$, d.h. der vorliegende Schätzer ist erwartungstreu.

Beispiel 6.57 *Es seien $x_1, \dots, x_n$ unabhängige Messungen einer Größe μ mit dem Erwartungswert $Mx_i = \mu$ und der Varianz $Dx_i = \sigma^2$. Wegen*

$$M\left(\frac{x_1 + \cdots + x_n}{n}\right) = \frac{Mx_1 + \cdots + Mx_n}{n} = \mu$$

ist

$$\widehat{g}_1(x) = \frac{x_1 + \cdots + x_n}{n} =: \overline{x}\,, \quad x = (x_1, \dots, x_n)\,,$$

eine erwartungstreue Schätzung für μ. Als Schätzer für σ^2 verwenden wir

$$\widehat{g}_2(x) = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \overline{x})^2 =: s^2\,.$$

Es gilt wegen der Unabhängigkeit der x_i

$$M(\overline{x} - \mu)^2 = D\overline{x} = \frac{1}{n^2} \sum_{k=1}^{n} Dx_k = \frac{\sigma^2}{n}$$

und somit

$$\begin{aligned} M(x_i - \overline{x})^2 &= M[x_i - \mu - (\overline{x} - \mu)]^2 \\ &= M(x_i - \mu)^2 - \frac{2}{n} \sum_{j=1}^{n} M[(x_i - \mu)(x_j - \mu)] + M(\overline{x} - \mu)^2 \\ &= \sigma^2 - \frac{2}{n}\,\sigma^2 + \frac{1}{n}\,\sigma^2 = \left(1 - \frac{1}{n}\right)\sigma^2\,. \end{aligned}$$

Es folgt

$$Ms^2 = \frac{1}{n-1}\, n \left(1 - \frac{1}{n}\right)\sigma^2 = \sigma^2\,,$$

so dass die Schätzung s^2 erwartungstreu ist.

Beispiel 6.58 *Die Standdardabweichung $\sqrt{np(1-p)}$ einer binomialverteilten Zufallsgröße ist* **nicht** *erwartungstreu schätzbar, da für jeden Schätzer $\widehat{g}$*

$$M_p\widehat{g} = \sum_{x=0}^{n} \widehat{g}(x) \binom{n}{x} p^x (1-p)^{n-x} \neq \sqrt{np(1-p)}$$

gilt, da

$$\frac{d}{dp}\, M_p\widehat{g}\big|_{p=0} < \infty\,, \quad \textit{aber} \quad \frac{d}{dp}\sqrt{np(1-p)}\big|_{p=0} = \infty\,.$$

Aus der Tschebyscheffschen Ungleichung (Lemma 6.43) folgt

$$\frac{M(f-Mf)^2}{\varepsilon^2} \geq P(|f-Mf|^2 \geq \varepsilon^2) = P(|f-Mf| \geq \varepsilon)\,,$$

d.h.

$$P(|f-Mf| \geq \varepsilon) \leq \frac{Df}{\varepsilon^2}\,.$$

Wir suchen für das Beispiel 6.56 ein $d > 0$, so dass

$$P_p(\widehat{p}-d \leq p \leq \widehat{p}+d) \geq 0.95 =: 1-\varepsilon\,.$$

Wegen $M_p\widehat{p} = p$ und $D\widehat{p} = \dfrac{p(1-p)}{n} \leq \dfrac{1}{4n}$ gilt

$$P_p(|\widehat{p}-p| \geq d) \leq \frac{p(1-p)}{nd^2} \leq \frac{1}{4nd^2} \leq \varepsilon$$

genau dann, wenn

$$d \geq \frac{1}{\sqrt{4n\varepsilon}} = \sqrt{\frac{5}{n}}\,.$$

Sind z.B. $n = 50$ und $k = 20$, so sind $\widehat{p} = 0.4$ und $d = 1/\sqrt{10} \approx 0.32$. Wir würden also als ein sog. **Konfidenzintervall** das Intervall $[0.08, 0.72]$ erhalten, für welches

$$P_p(p \in [0.08, 0.72]) \geq 0.95$$

gilt. Dieses erscheint uns als viel zu groß.

Definition 6.59 *Ist* $\{K(x) : x \in X\}$ *eine Familie von Teilmengen von* $g(\Theta)$ *und gilt für ein* $\alpha > 0$

$$P_\theta(\{x \in X : g(\theta) \in K(x)\}) \geq 1-\alpha \quad \forall \theta \in \Theta\,, \tag{6.9}$$

so heißt $K(x)$ **Konfidenzbereich** *für* $g(\theta)$ *zum* **Konfidenzniveau** $1-\alpha$ *zur Beobachtung* x.

Wir setzen

$$\widetilde{K} := \{(x,\theta) \in X \times \Theta : \theta \in K(x)\}$$

und

$$A(\theta) = \{x \in X : (x,\theta) \in \widetilde{K}\} = \{x \in X : \theta \in K(x)\}\,.$$

Dann gilt also $\theta \in K(x)$ genau dann, wenn $(x,\theta) \in \widetilde{K}$, und dies ist genau dann der Fall, wenn $x \in A(\theta)$. Somit ist die Ungleichung (6.9) äquivalent zu

$$P_\theta(A(\theta)) \geq 1-\alpha \quad \forall \theta \in \Theta\,. \tag{6.10}$$

Man verfolgt nun das Ziel, $K(x)$ für jedes feste x möglichst klein zu wählen. $\widetilde{K}$ ist klein, wenn die Mengen $A(\theta)$ möglichst wenige x enthalten. Um aber (6.10) erfüllen zu können, wählt man x mit möglichst großem $P_\theta(x)$:

- Anordnen von $X = \{x_1, x_2, \ldots\}$ so, dass

$$P_\theta(x_1) \geq P_\theta(x_2) \geq \ldots ,$$

- $A(\theta) = \{x_1, \ldots, x_k\}$ so, dass k minimal und

$$P_\theta(A(\theta)) = \sum_{j=1}^{k} P_\theta(x_j) \geq 1 - \alpha .$$

Konfidenzintervalle für die Erfolgswahrscheinlichkeit im Bernoulli-Schema:
Wir haben $b_{n,p}(x) = \binom{n}{x} p^x (1-p)^{n-x}$ und somit

$$\frac{b_{n,p}(x)}{b_{n,p}(x-1)} = \frac{n-x+1}{x} \frac{p}{1-p} > 1$$

genau dann, wenn $x < (n+1)p$. Wir definieren $A(p) = \{a, a+1, \ldots, b\}$ so, dass $a = a(p)$ maximal mit $P(f < a) < \alpha/2$ und $b = b(p)$ minimal mit $P(f > b) < \alpha/2$ (f binomialverteilt).

Lemma 6.60 *Es seien f $b_{n,p}$-verteilt und $x < n$ sowie $\mathcal{F}(p) = P_p(f \leq x)$. Dann ist $\mathcal{F}$ stetig und streng monoton fallend, und es gilt $\mathcal{F}(0) = 1$, $\mathcal{F}(1) = 0$.*

Beweis. Es seien $p_1 < p_2$, $p_2 p_3 = p_1$ und $f_{i1}, \ldots, f_{in}$ $(i = 2, 3)$ zwei unabhängige Bernoulli-Folgen mit $P(f_{ij} = 1) = p_i$ $(i = 2, 3; j = 1, \ldots, n)$ sowie $f_{1j} = f_{2j} f_{3j}$. Dann ist f_{1j} eine Bernoulli-Folge mit $P(f_{1j} = 1) = p_1$. Es folgt

$$\{f_{21} + \cdots + f_{2n} \leq x\} \subset \{f_{11} + \cdots + f_{1n} \leq x\}$$

und somit

$$\mathcal{F}(p_1) = P(f_{11} + \cdots + f_{1n} \leq x) \geq P(f_{21} + \cdots + f_{2n} \leq x) = \mathcal{F}(p_2) .$$

Das Ereignis $\{f_{21} + \cdots + f_{2n} = n, f_{31} + \cdots + f_{3n} \leq x\}$ hat positive Wahrscheinlichkeit und ist in $\{f_{11} + \cdots + f_{1n} \leq x\}$ enthalten, aber nicht in $\{f_{21} + \cdots + f_{2n} \leq x\}$. Also gilt $\mathcal{F}(p_1) > \mathcal{F}(p_2)$. □

Wir definieren $p_o(n) = 1$ und $p_o(x)$ für $x < n$ so, dass $P_{p_o(x)}(f \leq x) = \alpha/2$ (f binomialverteilt), $p_u(0) = 0$ und $P_{p_u(x)}(f \geq x) = \alpha/2$ für $x > 0$. Dann ist $x \geq a(p)$ äquivalent zu $P_p(f \leq x) = P_p(f < x+1) \geq \alpha/2$ und somit zu $p \leq p_o(x)$. Analog ist $x \leq b(p)$ äquivalent zu $p_u(x) \leq p$. Es folgt aus der Definition der Mengen $A(p)$

$$K(x) = [p_u(x), p_o(x)] \quad \text{und} \quad P_p(A(p)) > 1 - \alpha .$$

Also: $\{K(x)\}$ ist ein System von Konfidenzintervallen zum Niveau $1 - \alpha$.

Wir betrachten jetzt **absolut stetige Verteilungsfunktionen**

$$F_\theta(x) = \int_{-\infty}^{x} \varphi_\theta(t)\, dt .$$

Da die Differenz $F_\theta(x + \varepsilon) - F_\theta(x) \approx \varphi_\theta(x)\varepsilon$ maximal wird, wenn $\varphi_\theta(x)$ maximal wird, können wir als Likelihood-Funktion $L_x(\theta) = \varphi_\theta(x)$ nehmen.

1. Wir betrachten Maximum-Likelihood-Schätzer für die Normalverteilung. Die Zufallsgrößen $f_1, \ldots, f_n$ seien unabhängig $N(\mu, , \sigma^2)$-verteilt, $f = (f_1, \ldots, f_n)$. Wir haben $\theta = (\mu, \sigma^2)$,

$$\varphi_\theta(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^n (x_j - \mu)^2\right)$$

und

$$\mathcal{L}_x(\theta) = \ln L_x(\theta) = -n\ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}\sum_{j=1}^n (x_j - \mu)^2 .$$

(a) μ unbekannt, $\sigma^2 = \sigma_0^2$, $\theta = \mu$:

$$\mathcal{L}_x'(\widehat{\mu}) = \frac{1}{\sigma_0^2}\left(\sum_{j=1}^n x_j - n\widehat{\mu}\right) = 0 \;\Rightarrow\; \widehat{\mu} = \frac{1}{n}\sum_{j=1}^n x_j ,$$

$$\mathcal{L}_x''(\widehat{\mu}) = -\frac{n}{\sigma_0^2} < 0 , \quad M_\theta\widehat{\mu} = \mu ,$$

d.h. erwartungstreu.

(b) σ^2 unbekannt, $\mu = \mu_0$, $\theta = \sigma^2$:

$$\mathcal{L}_x(\theta) = -\frac{n}{2}\ln(2\pi\theta) - \frac{1}{2\theta}\sum_{j=1}^n (x_j - \mu_0)^2 ,$$

$$\mathcal{L}_x'(\widehat{\theta}) = -\frac{n}{2\widehat{\theta}} + \frac{1}{2\widehat{\theta}^2}\sum_{j=1}^n (x_j - \mu_0)^2 = 0 \;\Rightarrow\; \widehat{\theta} = \frac{1}{n}\sum_{j=1}^n (x_j - \mu_0)^2 ,$$

$$\mathcal{L}_x''(\widehat{\theta}) = \frac{n}{2\widehat{\theta}^2} - \frac{1}{\widehat{\theta}^3}\sum_{j=1}^n (x_j - \mu_0)^2 = \frac{n}{2\widehat{\theta}^2} - \frac{n}{\widehat{\theta}^2} = -\frac{n}{2\widehat{\theta}^2} < 0 ,$$

$$M_\theta\widehat{\theta} = \frac{1}{n}\sum_{j=1}^n \sigma^2 = \sigma^2 = \theta ,$$

d.h. erwartungstreu.

(c) μ und $\sigma^2 =: \eta$ unbekannt, $\theta = (\mu, \eta)$:

$$\mathcal{L}_x(\mu, \eta) = -\frac{n}{2}\ln(2\pi\eta) - \frac{1}{2\eta}\sum_{j=1}^n (x_j - \mu)^2 ,$$

$$\frac{\partial}{\partial\mu}\mathcal{L}_x(\widehat{\mu}, \widehat{\eta}) = \frac{1}{\widehat{\eta}}\left(\sum_{j=1}^n x_j - n\widehat{\mu}\right) = 0 ,$$

$$\frac{\partial}{\partial\eta}\mathcal{L}_x(\widehat{\mu},\widehat{\eta}) = -\frac{n}{2\widehat{\eta}} + \frac{1}{2\widehat{\eta}^2}\sum_{j=1}^n (x_j-\widehat{\mu})^2 = 0$$

$$\Rightarrow\ \widehat{\mu} = \frac{1}{n}\sum_{j=1}^n x_j\,,\quad \widehat{\eta} = \frac{1}{n}\sum_{j=1}^n (x_j-\widehat{\mu})^2\,,$$

$$A = \frac{\partial^2}{\partial\mu^2}\mathcal{L}_x(\widehat{\mu},\widehat{\eta}) = -\frac{n}{\widehat{\eta}}\,,$$

$$C = \frac{\partial^2}{\partial\eta^2}\mathcal{L}_x(\widehat{\mu},\widehat{\eta}) = \frac{n}{2\widehat{\eta}^2} - \frac{1}{\widehat{\eta}^2}\sum_{j=1}^n (x_j-\widehat{\mu})^2 = -\frac{n}{2\widehat{\eta}^2}\,,$$

$$B = \frac{\partial^2}{\partial\eta\partial\mu}\mathcal{L}_x(\widehat{\mu},\widehat{\eta}) = -\frac{1}{\widehat{\eta}^2}\left(\sum_{j=1}^n x_j - n\widehat{\mu}\right) = 0\,,$$

$$\Rightarrow\ AC - B^2 = \frac{n^2}{2\widehat{\eta}^3} > 0\,,$$

$M_\theta\widehat{\mu} = \mu$, d.h. erwartungstreu, aber $M_\theta\widehat{\eta} = \dfrac{n-1}{n}\eta$ (vgl. Beispiel 6.57), d.h. nicht erwartungstreu.

2. Wir suchen Konfidenzintervalle für die Normalverteilung. Es seien f_1 und f_2 unabhängig $N(0,\sigma_j^2)$-verteilt sowie $f = f_1 + f_2$. Es ist

$$\begin{aligned}
\varphi(x) &= \int_{-\infty}^{\infty} \varphi_1(x-t)\varphi_2(t)\,dt \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2\sigma_1^2}}\, e^{-\frac{t^2}{2\sigma_2^2}}\,dt \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[\frac{(x-t)^2}{\sigma_1^2}+\frac{t^2}{\sigma_2^2}\right]}\,dt \\
&= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, e^{-\frac{x^2}{2\sigma^2}}\,\frac{\sigma_1\sigma_2}{\sigma}\,dz \\
&= \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{x^2}{2\sigma^2}}
\end{aligned}$$

mit $\sigma = \sqrt{\sigma_1^2+\sigma_2^2}$ und der Substitution $z = t\,\dfrac{\sigma}{\sigma_1\sigma_2} - x\,\dfrac{\sigma_2}{\sigma\sigma_1}$. Also ist f $N(0,\sigma^2)$-verteilt. Ist f_j $N(\mu_j,\sigma_j^2)$-verteilt, so ist $f_j-\mu_j$ $N(0,\sigma_j^2)$-verteilt, also f $N(\mu,\sigma^2)$-verteilt mit $\mu = \mu_1+\mu_2$. Induktiv folgt, dass $f_1+\cdots+f_n$ $N(\mu,\sigma^2)$-verteilt ist mit $\mu = \mu_1+\cdots+\mu_n$ und $\sigma^2 = \sigma_1^2+\cdots+\sigma_n^2$. Es seien nun μ unbekannt, σ^2 bekannt. Aus obigen Überlegungen folgt, dass mit $\widehat{\mu} = \dfrac{x_1+\cdots+x_n}{n}$ die Zufallsgröße

$$f = \sqrt{n}\,\frac{\widehat{\mu}-\mu}{\sigma}$$

$N(0,1)$-verteilt ist. Aus $\Phi(1.96) = 0.975$ folgt

$$P_\mu(|f| \le 1.96) = 2\Phi(1.96) - 1 = 0.95\,.$$

Wir setzen

$$K(x) = \left\{\mu \in \mathbb{R} : |\widehat{\mu} - \mu| \le 1.96 \cdot \frac{\sigma}{\sqrt{n}}\right\}\,.$$

Es folgt, dass $\mu \in K(x)$ äquivalent ist zu $|f| \le 1.96$, d.h.

$$P_\mu(\mu \in K(x)) = 0.95\,.$$

Also: $K(x)$ ist Konfidenzintervall zum Niveau 0.95 zur Beobachtung $x = (x_1, \ldots, x_n)$.

3. Die **Methode der kleinsten Quadrate**:

   - n Messpunkte (t_i, x_i), $i = 1, \ldots, n$.
   - Man schätze $x(t) = \alpha + \beta t$ oder $x(t) = \alpha + \beta t + \gamma t^2$ oder allgemein

     $$x(t) = \varphi(\theta_1, \ldots, \theta_p, t)$$

     (θ_j - zu schätzende Parameter).
   - $Q(\theta_1, \ldots, \theta_p) = \sum_{j=1}^{n} [x_j - \varphi(\theta_1, \ldots, \theta_p, t_j)]^2 \longrightarrow \min$.
   - Wir betrachten den Fall der Geraden:

     $$Q(\alpha, \beta) = \sum_{j=1}^{n} [x_j - \alpha - \beta t_j]^2\,,$$

     $$\frac{\partial Q}{\partial \alpha} = 2\sum_{j=1}^{n} (\beta t_j + \alpha - x_j)\,,$$

     $$\frac{\partial Q}{\partial \beta} = 2\sum_{j=1}^{n} t_j(\beta t_j + \alpha - x_j)\,,$$

     $$\Rightarrow \; \widehat{\beta}\overline{t} + \widehat{\alpha} - \overline{x} = 0\,, \quad \widehat{\beta}\,\langle t, t\rangle + \widehat{\alpha}\overline{t} - \langle t, x\rangle = 0$$

     mit

     $$\overline{t} = \frac{1}{n}\sum_{j=1}^{n} t_j\,,\; \overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j\,,\; \langle t, x\rangle = \frac{1}{n}\sum_{j=1}^{n} t_j x_j\,.$$

     Es folgt

     $$\widehat{\beta} = \frac{\overline{t}\,\overline{x} - \langle t, x\rangle}{\overline{t}^2 - \langle t, t\rangle}\,, \quad \widehat{\alpha} = \overline{x} - \widehat{\beta}\overline{t}\,.$$

Die Gerade $x = \widehat{\alpha} + \widehat{\beta}t$ nennt man **Regressionsgerade**. $f = [f_1, \ldots, f_n]^T$ sei der Vektor aus den Messfehlern, die unabhängig $N(0, \sigma^2)$-verteilt seien, d.h. $x_j = \alpha + \beta t_j + f_j$. Die x_j sind dann $N(\alpha + \beta t_j, \sigma^2)$-verteilt, so dass die Verteilungsdichte von $x = [x_1, \ldots, x_n]^T$ die Gestalt

$$\varphi_{\alpha,\beta}(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j - \alpha - \beta t_j)^2\right)$$

hat. $\varphi_{\alpha,\beta}(x)$ wird maximal, wenn $Q(\alpha, \beta)$ minimal wird, d.h. $\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta})$ ist Maximum-Likilihood-Schätzer.

6.8 Tests

6.8.1 Die Problemstellung

Es wird eine Zufallsgröße f mit einer unbekannten Verteilungsfunktion $F_\theta(x) = P_\theta(f < x)$ beobachtet. Anhand des beobachteten Wertes x von f soll entschieden werden, ob P_θ zu einer bestimmten Menge von Verteilungen gehört oder nicht.

Es seien X die Menge der möglichen Werte von f (**Stichprobenraum**) und $\{P_\theta : \theta \in \Theta\}$ die Menge der in Frage kommenden Verteilungen. Eine nichtleere Teilmenge $H \subset \Theta$ sei die **Hypothese** und $K = \Theta \backslash H$ die entsprechende **Alternative**. Gesucht ist eine Entscheidungsregel dafür, wie man aus der Beobachtung x schließt, dass $\theta \in H$ (d_H) oder $\theta \in K$ (d_K). Mit

$$R = \{x \in X : x \longrightarrow d_K\}$$

bezeichnen wir den Verwerfungsbereich des Tests (kritischer Bereich). Wir unterscheiden zwischen einem **Fehler erster Art**: $\theta \in H$, aber d_K , und einem **Fehler zweiter Art**: $\theta \in K$, aber d_H . Im allgemeinen werden eine Funktion $T(x)$ und ein kritischer Wert t angegeben, so dass $R = \{x \in X : T(x) \geq t\}$. Die Funktion $\beta : \Theta \longrightarrow [0, \infty)$ mit $\beta(\theta) = P_\theta(f \in R)$ (die **Verwerfungswahrscheinlichkeit** unter P_θ) heißt **Gütefunktion** des Tests. Man sagt, der Test hat das **Niveau** α , wenn $\beta(\theta) \leq \alpha$ für alle $\theta \in H$ gilt, d.h. wenn die Wahrscheinlichkeit des Fehlers erster Art maximal α ist. Ist $\theta \in K$, so heißt $\beta(\theta)$ die **Macht** des Tests in θ . Ist $\beta(\theta)$ nahe bei 1 , so ist die Wahrscheinlichkeit $1 - \beta(\theta)$ eines Fehlers zweiter Art klein.

Beispiel 6.61 *(„tea tasting lady“, Versuchsanordnung nach Neyman)* $X = \{0, 1, \ldots, n\}$, $\theta = p$, $\Theta = [\frac{1}{2}, 1]$, $H = \{\frac{1}{2}\}$, $K = (\frac{1}{2}, 1]$, $R = \{x \in X : x \geq t\}$. *Als Gütefunktion erhalten wir*

$$\beta(p) = \beta_{t,n}(p) = P_p(f \geq t)\,,$$

wobei $P_p(f = x) = b_{n,p}(x) = \binom{n}{x}p^x(1-p)^{n-x}$. *Das Niveau sei* $\alpha = 0.05$, *und es sei* $n = 5$. *Da* $b_{5,\frac{1}{2}}(5) = (\frac{1}{2})^5 = \frac{1}{32}$ *und* $b_{5,\frac{1}{2}}(4) = 5(\frac{1}{2})^5 = \frac{5}{32}$, *ist* $\beta_{4,5}(\frac{1}{2}) = \frac{6}{32} > \frac{1}{20}$, *also muss* $t = 5$ *gewählt werden. Für* $p > \frac{1}{2}$ *gibt* $\beta_{t,n}(p)$ *die Wahrscheinlichkeit an, mit der die Lady ihre Fähigkeit nachweisen kann. Offenbar gilt*

$$\lim_{p \to \frac{1}{2}+0} (1 - \beta_{t,n}(p)) = 1 - \beta_{t,n}\left(\frac{1}{2}\right) \geq 1 - \alpha\,.$$

Somit ist es nicht möglich, eine vorgegebene Schranke $\alpha' < 1-\alpha$ für die Wahrscheinlichkeit eines Fehlers zweiter Art für alle $p \in K$ zu unterschreiten. Es ergibt sich damit die Notwendigkeit, sich zu entscheiden, welche Abweichungen von der Hypothese man erkennen will. Soll die Wahrscheinlichkeit eines Fehlers zweiter Art für $p \geq p_0 (> \frac{1}{2})$ kleiner als α' sein, so ergibt sich ein minimales notwendiges n. ($\beta_{t,n}(p_0) \geq 1-\alpha'$, wobei t bei gegebenem n durch α festgelegt ist.)

Eine Möglichkeit, die Wahrscheinlichkeit eines Fehlers zweiter Art zu verkleinern: Die zu beobachtende Größe sei binomialverteilt, $n = 5$, $H = \{p = \frac{1}{2}\}$, $K = \{p = \frac{3}{4}\}$, Niveau $\alpha = 0.05$, $R = \{5\}$. Es folgt für die Wahrscheinlichkeit eines Fehlers erster Art

$$P_{\frac{1}{2}}(5) = 2^{-5} = \frac{1}{32} < \frac{1}{20} = \alpha\,,$$

aber

$$P_{\frac{1}{2}}(5) + P_{\frac{1}{2}}(4) = \frac{6}{32} > \frac{1}{20}\,,$$

so dass 4 nicht zu R hinzugenommen werden kann.

Trick: $x = 5 \longrightarrow$ Hypothese wird verworfen, $x \in \{0,1,2,3\} \longrightarrow$ Hypothese wird akzeptiert, $x = 4 \longrightarrow$ Man macht ein zusätzliches Zufallsexperiment, völlig unabhängig von f, mit den möglichen Ausgängen 0 und 1. Ist das Ergebnis 1, so wird die Hypothese verworfen, sonst akzeptiert. Wir bezeichnen die Wahrscheinlichkeit einer 1 in diesem Experiment mit $\varphi(4)$. Die Wahrscheinlichkeit eines Fehlers erster Art ist dann

$$P_{\frac{1}{2}}(5) + P_{\frac{1}{2}}(4)\varphi(4) = [1 + 5\varphi(4)] \cdot \frac{1}{32}\,.$$

Für $\varphi(4) = \frac{3}{25}$ ergibt sich $[1+5\varphi(4)] \cdot \frac{1}{32} = \frac{1}{20}$. Die Größe $\beta(\frac{3}{4}) = b_{5,\frac{3}{4}}(5) + \varphi(4) b_{5,\frac{3}{4}}(4)$ hat sich auch vergrößert und somit $1 - \beta(\frac{3}{4})$ verkleinert. $\varphi(x)$ ist die Wahrscheinlichkeit, mit der die Hypothese bei der Beobachtung x verworfen wird. Man nennt einen solchen Test einen **randomisierten Test**: $\varphi(5) = 1$, $\varphi(4) = \frac{3}{25}$, und $\varphi(x) = 0$ für $x \in \{0,1,2,3\}$. Somit ist der Test durch φ beschrieben.

Man nennt eine Hypothese bzw. Alternative einfach, wenn sie nur aus einer Verteilung P_H bzw. P_K besteht. Wir betrachten einen Test einer einfachen Hypothese P_H gegen eine einfache Alternative P_K und stellen die Frage, ob es unter allen Tests φ mit dem Niveau α einen gibt, der die Macht $\beta_\varphi(K)$ maximiert. Ist $\varphi(x) = 1$, so leistet dieses x einen Beitrag $P_H(x)$ zum Niveau und einen Beitrag $P_K(x)$ zur Macht. Da wir für das Niveau durch α beschränkt sind, wählen wir $\varphi(x) = 1$ für solche x, für die $q(x) = P_K(x)/P_H(x)$ groß ist.

Definition 6.62 *Ein Test φ^* heißt* **Neyman-Pearson-Test**, *wenn eine Zahl c^*, $0 \leq c^* \leq \infty$, existiert, so dass*

$$\varphi^*(x) = \begin{cases} 1 \quad , & P_K(x) > c^* P_H(x)\,, \\ \\ 0 \quad , & P_K(x) < c^* P_H(x)\,. \end{cases}$$

Auf $\{x \in X : P_K(x) = c^* P_H(x)\}$ kann φ^* beliebige Werte $\gamma(x)$ mit $0 \leq \gamma(x) \leq 1$ annehmen. Wir nennen einen Test φ_1 schärfer als φ_2, wenn $\beta_{\varphi_1}(K) > \beta_{\varphi_2}(K)$.

Im weiteren sei $P_H(x) + P_K(x) > 0$ für alle $x \in X$. (Andere x hätten keinen Einfluss auf die Irrtumswahrscheinlichkeiten.)

Satz 6.63 (Neyman-Pearson) *Für das Testen einer einfachen Hypothese gegen eine einfache Alternative gilt:*

(a) Ist φ^* ein Neyman-Pearson-Test, so ist φ^* mindestens so scharf wie alle anderen Tests φ mit $\beta_\varphi(H) \le \beta_{\varphi^*}(H)$.

(b) Zu gegebenem $\alpha \in [0,1]$ existiert ein Neyman-Pearson-Test φ^* mit $\beta_{\varphi^*}(H) = \alpha$. Dabei kann man φ^* auf $\{x \in X : P_K(x) = c^* P_H(x)\}$ konstant wählen.

Beweis. (a) Wir setzen $A := \{x \in X : \varphi^*(x) > \varphi(x)\}$ und $B := \{x \in X : \varphi^*(x) < \varphi(x)\}$. Es folgt $\varphi^*(x) > 0$ und somit $P_K(x) \ge c^* P_H(x)$ für alle $x \in A$, sowie $\varphi^*(x) < 1$, also

$$P_K(x) \le c^* P_H(x)$$

für alle $x \in B$. Wir erhalten

$$\begin{aligned}
\beta_{\varphi^*}(K) - \beta_\varphi(K) &= \sum_{x \in X} [\varphi^*(x) - \varphi(x)] P_K(x) = \sum_{x \in A \cup B} [\varphi^*(x) - \varphi(x)] P_K(x) \\
&\ge \sum_{x \in A \cup B} [\varphi^*(x) - \varphi(x)] c^* P_H(x) = c^* \sum_{x \in X} [\varphi^*(x) - \varphi(x)] P_H(x) \\
&= c^* [\beta_{\varphi^*}(H) - \beta_\varphi(H)] \ge 0 .
\end{aligned}$$

(b) Für $\alpha = 0$ setzen wir $c^* = \infty$. Für $P_H(x) > 0$ folgt dann $P_K(x) < c^* P_H(x)$, also $\varphi^*(x) = 0$ und somit $\beta_{\varphi^*}(H) = 0$. Sei jetzt $\alpha > 0$. Wir definieren $\alpha(c) = P_H(q(x) > c)$ für $c \ge 0$. Dann ist $\alpha(c)$ rechtsseitig stetig, und es gilt $\alpha(c - 0) = P_H(q(x) \ge c)$ sowie $\alpha(c) \longrightarrow 0$ monoton fallend für $c \longrightarrow \infty$. Es sei $c^* = \inf\{c \ge 0 : \alpha(c) \le \alpha\}$. Dann ist $\alpha(c^* - 0) \ge \alpha \ge \alpha(c^*)$. Wir setzen

$$\gamma^* = \begin{cases} 0 & , \quad \alpha(c^* - 0) = \alpha(c^*) , \\ \frac{\alpha - \alpha(c^*)}{\alpha(c^* - 0) - \alpha(c^*)} & , \quad \text{sonst} , \end{cases}$$

und definieren $\varphi^*(x) = \gamma^*$ für $x \in X$ mit $P_K(x) = c^* P_H(x)$. Es folgt

$$\beta_{\varphi^*}(H) = P_H(q(x) > c^*) + \gamma^* P_H(q(x) = c^*) = \alpha(c^*) + \gamma^* [\alpha(c^* - 0) - \alpha(c^*)] = \alpha .$$

□

Wir betrachten nochmals das Beispiel 6.61 der „tea tasting lady“. Es sei vorerst $p_H = \frac{1}{2}$ und $p_K > \frac{1}{2}$. Dann ist

$$q(x) = \frac{P_K(x)}{P_H(x)} = \frac{\binom{n}{x} p_K^x (1 - p_K)^{n-x}}{\binom{n}{x} (\frac{1}{2})^n} = 2^n p_K^x (1 - p_K)^{n-x}$$

streng monoton wachsend. Es folgt, dass $\{x : q(x) > c\}$ von der Gestalt $\{t, t+1, \ldots, n\}$ ist (falls $= \emptyset$, so $t := n+1$). Wir setzen also $\varphi(x) = 1$ für $x \geq t$ und $\varphi(x) = 0$ für $x < t-1$. Ist $q(t-1) < c$, so $\varphi(t-1) = 0$. Ist $q(t-1) = c$, so $\varphi(t-1) = \gamma \in [0,1]$. Wir folgern

$$\beta_\varphi(H) = P_H(x \geq t) + \gamma P_H(x = t-1)\,.$$

Es sei nun $\alpha > 0$ vorgegeben. Es soll ein schärfster Test zum Niveau α gefunden werden. Wir bestimmen t und γ aus $\beta_\varphi(H) = \alpha$. Nach dem Satz von Neyman-Pearson ist φ schärfster Test zum Niveau α. Da p_K in dieser Gleichung nicht vorkommt, ist φ auch schärfster Test gegen die zusammengesetzte Alternative $K = \{p : p \in (\frac{1}{2}, 1]\}$.

6.8.2 Tests für absolut stetige Verteilungen

Die Zufallsgrößen $f_1, \ldots, f_n$ seien unabhängig $N(\mu, \sigma^2)$-verteilt mit unbekanntem μ und σ^2. Für festes μ_0 sei zu testen, ob $\mu = \mu_0$ oder $\mu \neq \mu_0$ gilt, d.h. wir haben $\vartheta = (\mu, \sigma^2)$,

$$H = \{(\mu_0, \sigma^2) : \ \sigma^2 > 0\} \quad \text{und} \quad K = \{(\mu, \sigma^2) \in \mathbb{R}^2 : \ \mu \neq \mu_0\,, \ \sigma^2 > 0\}\,.$$

Die Dichtefunktion des Zufallsvektors $f = [f_1, \ldots, f_n]^T$ ist gegeben durch

$$\phi_\vartheta(x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \mu)^2\right) = L_x(\vartheta)\,.$$

Aus Stetigkeitsgründen und wegen der Dichtheit von K in $\Theta = H \cup K$ gilt offenbar

$$\sup\{L_x(\vartheta) : \ \vartheta \in K\} = \sup\{L_x(\vartheta) : \ \vartheta \in \Theta\}\,.$$

Aus 1.(b),(c), Abschnitt 6.7 folgt

$$\sup\{L_x(\vartheta) : \ \vartheta \in K\} = \varphi_{(\widehat{\mu}, \widehat{\sigma}^2)}(x) \quad \text{und} \quad \sup\{L_x(\vartheta) : \ \vartheta \in H\} = \varphi_{(\mu_0, \widetilde{\sigma}^2)}(x)$$

mit

$$\widehat{\mu} = \frac{1}{n}\sum_{k=1}^n x_k\,, \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^n (x_k - \widehat{\mu})^2\,, \quad \widetilde{\sigma}^2 = \frac{1}{n}\sum_{k=1}^n (x_k - \mu_0)^2\,.$$

Somit ist

$$\varphi_{(\widehat{\mu}, \widehat{\sigma}^2)}(x) = \left(\frac{1}{\sqrt{2\pi}\widehat{\sigma}}\right)^n e^{-\frac{n}{2}} \quad \text{und} \quad \varphi_{(\mu_0, \widetilde{\sigma}^2)}(x) = \left(\frac{1}{\sqrt{2\pi}\widetilde{\sigma}}\right)^n e^{-\frac{n}{2}}\,.$$

Also gilt

$$q(x) := \frac{\sup\{L_x(\vartheta) : \ \vartheta \in K\}}{\sup\{L_x(\vartheta) : \ \vartheta \in H\}} = \left(\frac{\widetilde{\sigma}}{\widehat{\sigma}}\right)^n\,.$$

Wir nennen $\varphi : X \to [0,1]$ einen **Likelihood-Quotienten-Test**, wenn ein $c \in \mathbb{R}$ existiert, so dass

$$\{x \in X : \ q(x) > c\} \subset \{x \in X : \ \varphi(x) = 1\}$$

und

$$\{x \in X : \ q(x) < c\} \subset \{x \in X : \ \varphi(x) = 0\}$$

gilt. Wir setzen

$$T(x) := \frac{\sqrt{n}(\widehat{\mu} - \mu_0)}{s(x)} \tag{6.11}$$

mit $s(x) = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_k - \widehat{\mu})^2}$. Dann ist

$$\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2} = 1 + \frac{(\widehat{\mu} - \mu_0)^2}{\widehat{\sigma}^2} = 1 + \frac{1}{n-1}\,|T(x)|^2 .$$

Damit ist für ein geeignetes $t \in \mathbb{R}$ die Bedingung $q(x) > c$ äquivalent zu $|T(x)| > t$. Für einen Likelihood-Quotienten-Test φ existiert also ein t mit

$$\varphi(x) = \begin{cases} 1 & , \quad |T(x)| > t\,, \\ 0 & , \quad |T(x)| < t\,. \end{cases}$$

Um zu einem gegebenen Niveau $\alpha > 0$ die Zahl t zu finden, muss man die Verteilung von $T(f)$ unter der Hypothese kennen.

Definition 6.64 *Die Verteilung der Summe der Quadrate von k unabhängigen $N(0,1)$-verteilten Zufallsgrößen nennt man* χ^2_k**-Verteilung** *bzw. eine* χ^2**-Verteilung** *mit k Freiheitsgraden. Sind f und g unabhängige Zufallsgrößen, wobei f $N(0,1)$-verteilt und g χ^2_k-verteilt sind, so sagt man, dass die Zufallsgröße*

$$h = \frac{f}{\sqrt{g/k}}$$

eine t_k**-Verteilung** *bzw. eine* t**-Verteilung** *mit k Freiheitsgraden besitzt.*

Die Dichte $\psi_k(x)$ der χ^2_k-Verteilung ist gleich

$$\psi_k(x) = \frac{1}{2^{k/2}\Gamma(k/2)}\, x^{\frac{k}{2}-1} e^{-\frac{x}{2}}\,, \quad x > 0\,.$$

Satz 6.65 *Sind $f_1, \ldots, f_n$ unabhängige $N(\mu_0, \sigma^2)$-verteilte Zufallsgrößen, so ist $T(f)$ mit $T(x)$ aus (6.11) t_{n-1}-verteilt. Die Dichte $\tau_{n-1}(x)$ von $T(f)$ ist gegeben durch*

$$\tau_{n-1}(x) = \frac{\Gamma(n/2)}{\Gamma((n-1)/2)\Gamma(1/2)\sqrt{n-1}} \left(1 + \frac{x^2}{n-1}\right)^{-n/2} .$$

Die t_1-Verteilung ist die **Cauchy-Verteilung**. Ihre Dichte ist gleich (siehe Abschnitt 6.3)

$$\tau_1(x) = \frac{1}{\pi(1+x^2)}\,.$$

Die allgemeine t_k-Verteilung stammt von dem britischen Statistiker W. S. Gosset, der unter dem Pseudonym „Student“ publizierte, weshalb die t-Verteilung auch **Studentsche Verteilung** heißt.

Die Verwerfungswahrscheinlichkeit unter der Hypothese ist für unseren Test nun gleich $\beta(\varphi) = P(|T(x)| > t)$. Man nennt die Zahl $t_{n-1,\beta}$ mit $P(T \leq t_{n-1,\beta}) = \beta$ das (untere) β-**Quantil** der t_{n-1}-Verteilung. Bestimmt man nun aus einer Tabelle der t_{n-1}-Verteilung die Zahl $t = t_{n-1,1-\alpha/2}$, so ist wegen der Symmetrie der t_{n-1}-Verteilung bzgl. des Nullpunktes

$$P(|T(x)| > t) = \alpha\,.$$

Unser Test hat dann das Niveau α. Für $k > 25$ ist die t_k-Verteilung fast identisch mit der $N(0,1)$-Verteilung, so dass man in diesen Fällen auch Tabellen für die Standard-Normalverteilung verwendet.

Den hier betrachteten Test nennt man einen zweiseitigen t-**Test**, da sowohl für kleine als auch für große Werte $T(x)$ die Hypothese verworfen wird.

6.9 Übungsaufgaben

1. Man zeige, dass der Durchschnitt beliebig vieler σ-Algebren wieder eine σ-Algebra ist.

2. Beweisen Sie unter Verwendung der drei Axiome (W1)–(W3) einer Wahrscheinlichkeitsverteilung auf einem messbaren Raum $\langle\Omega, \mathcal{F}\rangle$ folgende weitere Eigenschaften ($A, B, A_n \in \mathcal{F}$, $\overline{A} := \Omega \setminus A$):

   (a) $P(\emptyset) = 0$,

   (b) $P(\overline{A}) = 1 - P(A)$,

   (c) $A \subset B \implies P(A) \leq P(B)$,

   (d) $P(A) \leq 1$,

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

   (f) $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$.

3. Man zeige:

   (a) Sind $A, B \in \mathcal{F}$ unabhängig, so sind es auch A und $\overline{B}$.

   (b) Sind A und B_1 sowie A und B_2 jeweils unabhängig und gilt $B_1 \cap B_2 = \emptyset$, dann sind auch A und $B_1 \cup B_2$ unabhängig.

4. Ein (gut aussehender) Mann kommt bei jeder Frau mit der Wahrscheinlichkeit p an. Berechnen Sie die Wahrscheinlichkeit dafür, dass

   (a) die erste Frau ihn abblitzen lässt, die zweite aber anspringt,

   (b) die zweite Frau anspringt, nachdem ihn die erste hatte abblitzen lassen,

   (c) er nicht mehr als zwei Frau anmachen muss, damit mindestens eine anspringt.

5. Eine Tüte mit 40 Kirschen enthalte 10 madige. Wie groß ist die Wahrscheinlichkeit dafür, dass von 10 zufällig ausgewählten Kirschen keine madig ist?

6. Wie groß ist die Wahrscheinlichkeit, bei 3 von 10 Würfen eine 6 zu würfeln?

7. Wieviele Rosinen muss man in den Teig für 10 Brötchen geben, damit mit 99% Wahrscheinlichkeit jedes Brötchen mindestens eine Rosine enthält?

8. Ein Prüfer hat 18 Standardfragen, von denen er 6 zufällig gewählte stellt. Ein Student kennt die Antworten zu 10 Fragen. Wie groß ist die Wahrscheinlichkeit, die Prüfung zu bestehen, wenn er dazu mindestens 3 Fragen beantworten muss?

9. Die Ausschussrate der Produkte eines Betriebes sei 2%. Ein defektes Teil werde mit 95% Wahrscheinlichkeit aussortiert. Mit 1% Wahrscheinlichkeit sei ein aussortiertes Teil nicht defekt. Man bestimme die Wahrscheinlichkeit dafür, dass ein nicht aussortiertes Teil einwandfrei ist.

10. Bei gleichen Symptomen seien 3 Krankheiten K_1, K_2 und K_3 möglich. Im Mittel liege in 5 (bzw. 1 bzw. 2) von 8 Fällen die Krankheit K_1 (bzw. K_2 bzw. K_3) vor. Als Diagnosehilfe wird ein Test durchgeführt, der bei Vorliegen von K_1 (K_2 bzw. K_3) mit der Wahrscheinlichkeit 0.2 (bzw. 0.3 bzw. 0.9) positiv ausfällt.

    (a) Wie groß ist die Wahrscheinlichkeit, dass der Test positiv ausfällt? Hinweis: totale Wahrscheinlichkeit

    (b) Wie groß ist die Wahrscheinlichkeit für das Vorliegen der einzelnen 3 Krankheiten für einen Patienten, bei dem der Test negativ ausfiel? Hinweis: Satz von Bayes

11. Drei Männer lieben die gleiche Frau und wollen das Problem durch ein Duell "lösen". Sie sind aber unterschiedlich gute Schützen. Sie treffen jeweils mit 100%, 80% und 50%. Zum Ausgleich vereinbaren sie, dass sie in der Reihenfolge (50%, 80%, 100%) dran sind und reihum jeder sein Ziel frei wählen kann, bis nur noch einer übrig bleibt. Wer hat bei optimaler Strategie aller Beteiligten die größten Chancen? Dazu nehmen wir an, dass die Männer mit der 80- und 100-prozentigen Treffsicherheit jeweils auf den besten verbliebenen Schützen zielen. Wie groß ist die Wahrscheinlichkeit für das Überleben des 50-prozentigen Schützen, wenn er zuerst

    (a) auf den mit 100-prozentiger Treffsicherheit zielt,

    (b) auf den mit 80-prozentiger Treffsicherheit zielt,

    (c) mit Absicht daneben schießt (wir nehmen an, dass er das mit 100% schafft).

    Berechnen Sie auch die Überlebenswahrscheinlichkeiten der anderen Schützen für die beste Wahl des 50-prozentigen Schützen.

12. Ein Los in einer bestimmten Lotterie kostet 5 €. Von den Gesamteinnahmen werden 40% für Steuern und den Eigenbedarf der Lotteriegesellschaft verwendet und 60% in Form von Gewinnen wieder ausgeschüttet. Wie groß ist der Erwartungswert des Gewinnes nach dem Kauf eines zufällig ausgewählten Loses?

13. Die Zufallsgröße X habe die Verteilungsfunktion F. Geben Sie die Verteilungsfunktionen der Zufallsgrößen $aX + b\,(a > 0, b \in \mathbb{R}$ Konstanten) und X^2 an.

14. Bestimmen Sie Erwartungswert und Varianz einer auf $[0,1]$ gleichverteilten Zufallsgröße.

15. Es seien f_j , $j = 1,2$, zwei unabhängige Zufallsgrößen mit
$$P(f_j = k) = (1-p)^k p\,, \quad k = 0,1,2,\ldots\,,\ j = 1,2\,,$$
wobei $0 < p < 1$. Bestimmen Sie die Verteilung von $\left[\begin{array}{cc} f_1 & f_2 \end{array}\right]^T$ und von $f = \max\{f_1, f_2\}$.

16. Es seien $f_j : \Omega \longrightarrow (0,\infty)$ unabhängige, identisch verteilte Zufallsgrößen. Berechnen Sie den Erwartungswert der Zufallsgröße $\dfrac{f_1}{f_1 + \cdots + f_n}$.

17. Für das wiederholte Würfeln mit zwei Würfeln berechne man die Wahrscheinlichkeit dafür, dass die Summe “11” vor der Summe “8” fällt.

18. Die Kosten für die Herstellung eines Artikels seien gleich $K > 0$, der Erlös beim Verkauf dieses Artikels gleich $E > K$. Die Nachfrage für diesen Artikel sei eine Zufallsgröße f mit $P(f = k) = p_k$, $k = 0,1,2,\ldots$ Mit $G(a)$ bezeichnen wir den Gewinn, wenn a Exemplare des Artikels hergestellt werden. Für welches a ist der Erwartungswert von $G(a)$ maximal?

19. Berechnen Sie näherungsweise die Wahrscheinlichkeit dafür, dass bei 600-maligem Würfeln mindestens 90 und höchstens 100 Sechsen auftreten.

20. Es seien $f_1, \ldots, f_n$ unabhängige und gleichverteilte Zufallsgrößen in $\{1,2,\ldots,b\}$ mit unbekanntem $b \in \{1,2,\ldots\}$. Man gebe ein Konfidenzintervall für b zum Niveau $1-\alpha$ anhand der Beobachtung von $f = \max\{f_1,\ldots,f_n\}$ an.

21. Es sei f die Anzahl der Misserfolge in einem Bernoulli-Experiment vor dem m-ten Erfolg. Man gebe einen Maximum-Likelihood-Schätzer für die Erfolgswahrscheinlichkeit an.

22. Ein Händler testet beim Kauf eines Paketes mit 100 Glühbirnen, ob dieses weniger als 10 defekte Glühbirnen enthält, indem er 10 Glühbirnen prüft und das Paket nur dann annimmt, wenn alle 10 Glühbirnen in Ordnung sind. Berechnen Sie das Niveau dieses Tests.

23. Geben Sie für $n = 6$ einen randomisierten Neyman-Pearson-Test der “tea testing Lady” mit dem Niveau 0,02 an.

24. $f = (f_1,\ldots,f_n)$ sei Bernoulli-verteilt mit der Erfolgswahrscheinlichkeit p . Wie groß muss man n wählen, um einen Test φ der Hypothese $p = p_H = 0,2$ gegen die Alternative $p = p_K = 0,8$ mit $\beta_H(\varphi) \le 0,05$ und $\beta_K(\varphi) \ge 0,95$ zur Verfügung zu haben. Existiert solch ein Test für beliebige $p_H, p_K \in (0,1)$ mit $p_H \ne p_K$?

25. Für eine mit dem Parameter $\lambda > 0$ Poisson-verteilte Zufallsgröße f gebe man den schärfsten nichtrandomisierten Neyman-Pearson-Test der Hypothese $\lambda = 2$ gegen die Alternative $\lambda = 0,5$ zum Niveau $\alpha = 0,2$ an. Gibt es einen nichtrandomisierten schärferen Test zum gleichen Niveau?

26. In einer Kreisscheibe vom Radius $R > 0$ wird ein Punkt zufällig gewählt (Gleichverteilung). Man bestimme die Dichte der Verteilung seines Abstandes vom Mittelpunkt des Kreises.

27. Die Verteilung von $f = (f_1, f_2)$ sei eine Gleichverteilung in $\big\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\big\}$. Man bestimme die Verteilung von $f_1 + f_2$.

Index