



TU Clausthal
Institut für Mathematik

Thomas Hanschke

Stochastik II

Prof. Dr. Thomas Hanschke
Institut für Mathematik
Technische Universität Clausthal
Erzstraße 1
38678 Clausthal-Zellerfeld

1. Auflage 2007 (30. April 2007)

Autor: Prof. Dr. Thomas Hanschke
(E-Mail: Hanschke@math.tu-clausthal.de)
Assistenz: Dipl.-Math. Alexander Herzog
(E-Mail: Herzog@math.tu-clausthal.de)
Dipl.-Math. Sylvia Arns
(E-Mail: Sylvia.Arn@math.tu-clausthal.de)
Hendrik Baumann (Cand.-Math.)
(E-Mail: Hendrik.Baumann@tu-clausthal.de)

© Arbeitsgruppe „Stochastische Modelle in den Ingenieurwissenschaften“ am Institut für Mathematik der TU Clausthal, Clausthal-Zellerfeld, 2007

Obwohl bei der Erstellung von Texten und Abbildungen mit großer Sorgfalt vorgegangen wurde, lassen sich Fehler nicht vollständig ausschließen. Es wird für fehlerhafte Angaben und deren Folgen weder eine juristische Verantwortung noch irgendeine Haftung übernommen.

Als Manuskript vervielfältigt.

Inhaltsverzeichnis

Vorwort	5
0.1 Mit der Stochastik II zusammenhängende Vorlesungen	6
13 Charakteristische Funktionen	7
13.1 Komplexwertige Zufallsvariablen	8
13.2 Eigenschaften charakteristischer Funktionen	10
Literaturverzeichnis	23
14 Schwache Konvergenz und zentraler Grenzwertsatz	25
14.1 Vage und schwache Konvergenz	26
14.2 Der Stetigkeitssatz von Lévy	34
14.3 Der zentrale Grenzwertsatz	39
Literaturverzeichnis	49
15 Statistische Auswertung von Zufallsexperimenten	51
15.1 Einleitung	52
15.2 Punktschätzer	52
15.3 Maximum-Likelihood-Schätzer	60
15.4 Konfidenzintervalle	63
15.5 Hypothesentests	69
Literaturverzeichnis	73
16 Bedingte Erwartung und bedingte Verteilung	75
16.1 Bedingte Erwartung	76
16.2 (Regulär) bedingte Verteilung	84
Literaturverzeichnis	97
17 Erneuerungstheorie	99
17.1 Erneuerungsfunktion und Erneuerungsgleichung	100
17.2 Vorwärts- und Rückwärtsrekurrenzzeit	110
17.3 Anwendungsbeispiel 1	114
17.4 Stationärer Zustand	118
17.5 Anwendungsbeispiel 2	135
17.6 Modifizierte Erneuerungsprozesse	137
17.7 Die Überlagerung von Erneuerungsprozessen	139
17.8 Abbrechende Erneuerungsprozesse	141
Literaturverzeichnis	143

18 Regenerative stochastische Prozesse	145
18.1 Einleitung	146
18.2 Zeitstetige regenerative stochastische Prozesse	147
18.3 Zeitdiskrete regenerative stochastische Prozesse	151
Literaturverzeichnis	153
19 Instandhaltungsstrategien	155
19.1 Einleitung	156
19.2 Wiederholung der Zuverlässigkeit	156
19.3 Einführung	157
19.4 Altersunabhängige Erneuerungen	160
19.5 Altersabhängige Erneuerung	161
19.6 Blockerneuerungen	166
Literaturverzeichnis	170
20 Monte–Carlo–Simulation	171
20.1 Einleitung	172
20.2 Erzeugung gleichverteilter Zufallszahlen	174
20.3 Erzeugung nichtgleichverteilter Zufallszahlen	184
20.4 Monte–Carlo–Integration	189
20.5 Stochastische Simulation von Warteschlangen	193
20.6 Statistik regenerativer Prozesse	200
Literaturverzeichnis	203
A Alternative Definition der Laplace-Transformierten	207
B Tabelle der χ^2–Verteilung	211
C Tabelle der Standardnormalverteilung	213
D Quantile der Standardnormalverteilung	215
E Zeichenerklärungen	217
F Literatur	219
G Historie	231

Vorwort

Die Stochastik II ist Teil eines viersemestrigen Zyklus, den ich seit 1993 regelmäßig für Studierende der Mathematik und Informatik an der Technischen Universität Clausthal abhalte. Das Skriptum ist bereits an den speziellen Anforderungen des neuen Bachelor-Studiengangs „Angewandte Mathematik“ und den darauf aufbauenden Master-Studiengängen „Angewandte Mathematik“ und „Operations Research“ ausgerichtet und beinhaltet deshalb neben den weiterführenden Kapiteln der Wahrscheinlichkeitstheorie (charakteristische Funktion, schwache Konvergenz, zentraler Grenzwertsatz, bedingte Erwartung, bedingte Verteilung) und der Einführung in die Theorie der stochastischen Prozesse (Erneuerungstheorie und regenerative Prozesse) auch Anwendungen aus den Bereichen Statistik, Instandhaltung, Monte Carlo-Integration und stochastische Simulation.

Dank der finanziellen Unterstützung durch die ELAN-Initiative des Landes Niedersachsen (elearning academic network Niedersachsen) kann auch dieses Skriptum wieder als Online-Version zur Verfügung gestellt werden:

<http://www.stochastik.tu-clausthal.de/Stochastik2Skript/>

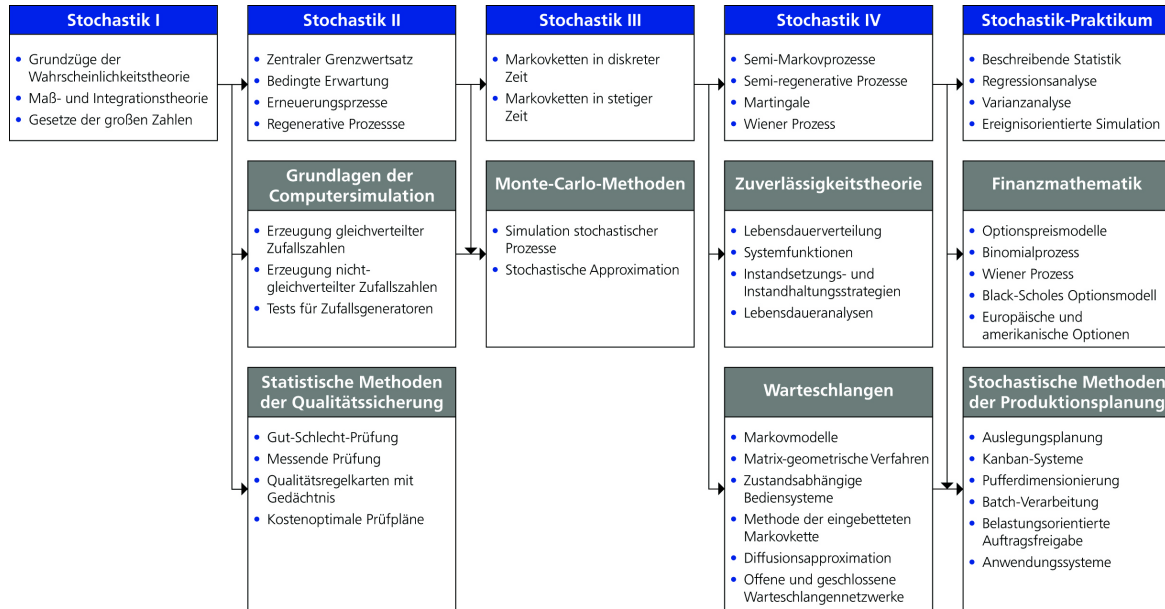
Für die tatkräftige Zuarbeit danke ich Frau Dipl.-Math. Sylvia Arns, Herrn Dipl.-Math. Alexander Herzog und unserem Hilfsassistenten Herrn Hendrik Baumann. Meinem ehemaligen Kollegen Prof. Dr. Joachim Hilgert, der nach einer Vorversion dieses Skriptums im SS 2002 die Stochastik II gelesen hatte, verdanke ich einige wichtige Korrekturhinweise.

Thomas Hanschke

Clausthal, April 2007

0.1 Mit der Stochastik II zusammenhängende Vorlesungen

Die Stochastik II Vorlesung ist Teil einer Reihe weiterführender Veranstaltungen, die man der nachstehenden Grafik entnehmen kann.



(Siehe auch PowerPoint-Präsentation zum Stochastik-Vorlesungsplan.)

Kapitel 13

Charakteristische Funktionen

Der Begriff der charakteristischen Funktion ist grundlegend für die Wahrscheinlichkeitstheorie. Es wird sich zeigen, dass viele mit dem asymptotischen Verhalten von Folgen von Zufallsvariablen zusammenhängende Fragestellungen sehr einfach und elegant mit Hilfe charakteristischer Funktionen beantwortet werden können.

Schlüsselwörter: Komplexwertige Zufallsvariable, Fourier-Transformierte oder charakteristische Funktion, Satz von Bochner, Umkehrformel, Summen stochastisch unabhängiger Zufallsvariablen, Momenten-Berechnung aus charakteristischen Funktionen.

13.1 Komplexwertige Zufallsvariablen

Um den Begriff der charakteristischen Funktion einführen zu können, werden Kenntnisse über komplexwertige Zufallsvariablen $X: \Omega \rightarrow \mathbb{C}$ benötigt. Indem man

$$X = \operatorname{Re} X + i \operatorname{Im} X$$

setzt, wobei $\operatorname{Re} X$ der Realteil von X , $\operatorname{Im} X$ der Imaginärteil von X und i die imaginäre Einheit bedeuten, kann die Behandlung komplexwertiger messbarer Abbildungen dem bereits behandelten Kalkül untergeordnet werden.

13.1 Definition (komplexwertige Zufallsvariable):

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $X: \Omega \rightarrow \mathbb{C}$ eine Abbildung. X heißt komplexwertige Zufallsvariable genau dann, wenn $\operatorname{Re} X$ und $\operatorname{Im} X$ \mathfrak{F} - \mathfrak{B} -messbare Abbildungen von Ω in \mathbb{R} sind.

13.2 Definition (P -integrierbar über Ω , Erwartungswert von X über Ω):

Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und X eine komplexwertige Zufallsvariable über $(\Omega, \mathfrak{F}, P)$. X heißt genau dann P -integrierbar über Ω , wenn $\operatorname{Re} X$ und $\operatorname{Im} X$ P -integrierbar sind. Dann heißt

$$\int_{\Omega} X dP := \int_{\Omega} \operatorname{Re} X dP + i \int_{\Omega} \operatorname{Im} X dP$$

der Erwartungswert von X über Ω . Er wird wie im reellen Fall mit $\mathbf{E}_P[X]$ (oder einfach mit $\mathbf{E}[X]$, wenn das Wahrscheinlichkeitsmaß aus dem Kontext heraus klar ist) bezeichnet.

13.3 Satz:

$(\Omega, \mathfrak{F}, P)$ sei ein Wahrscheinlichkeitsraum und $X, Y: \Omega \rightarrow \mathbb{C}$ seien komplexwertige Zufallsvariablen.

a) Ist X P -integrierbar, dann ist auch \overline{X} P -integrierbar und es gilt:

$$\int_{\Omega} \overline{X} dP = \overline{\int_{\Omega} X dP}.$$

b) Ist X P -integrierbar, so ist auch cX P -integrierbar für jedes $c \in \mathbb{C}$ und es gilt:

$$\int_{\Omega} cX dP = c \int_{\Omega} X dP.$$

c) Sind X und Y P -integrierbar, dann ist auch $X + Y$ P -integrierbar und es gilt:

$$\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y].$$

d) X ist genau dann P -integrierbar, wenn $|X|$ P -integrierbar ist.

e) Ist X P -integrierbar, so gilt

$$\left| \int_{\Omega} X dP \right| \leq \int_{\Omega} |X| dP.$$

f) Sind X und Y stochastisch unabhängig und P -integrierbar, so gilt

$$\mathbf{E}[X \cdot Y] = \mathbf{E}[X] \cdot \mathbf{E}[Y].$$

Beweis:

Die Eigenschaften a) bis c) folgen unmittelbar aus dem Ansatz $X = \operatorname{Re} X + i \operatorname{Im} X$ bzw. für b) außerdem aus dem Ansatz $c = a + ib$.

d) Aufgrund der Definition des P -Integrals von X folgt aus der P -Integrierbarkeit von X die P -Integrierbarkeit von $\operatorname{Re} X$ und $\operatorname{Im} X$. Mit Hilfe der Äquivalenz

$$f \text{ ist } \mu\text{-integrierbar} \iff |f| \text{ ist } \mu\text{-integrierbar}$$

kann aus der P -Integrierbarkeit von $\operatorname{Re} X$ und $\operatorname{Im} X$ auf die P -Integrierbarkeit von $|\operatorname{Re} X|$ und $|\operatorname{Im} X|$ geschlossen werden. Wegen $|X| \leq |\operatorname{Re} X| + |\operatorname{Im} X|$ und der Monotonie des Integrals folgt somit die P -Integrierbarkeit von $|X|$. Die Umkehrung ergibt sich aus den Abschätzungen $|\operatorname{Re} X| \leq |X|$ und $|\operatorname{Im} X| \leq |X|$.

e) Zuerst werden $X(\omega)$ und $\int_{\Omega} X(\omega) dP(\omega)$ in Polarkoordinaten dargestellt:

$$X(\omega) = r(\omega) \cdot e^{i\gamma(\omega)}, \omega \in \Omega, \text{ mit } r(\omega) \in \mathbb{R}^+ \text{ und } 0 \leq \gamma(\omega) < 2\pi,$$

$$\int_{\Omega} X(\omega) dP(\omega) = r' \cdot e^{i\gamma'} \text{ mit } r' \in \mathbb{R}^+ \text{ und } 0 \leq \gamma' < 2\pi$$

(für $X(\omega) \neq 0$, sonst nicht definiert). Unter Einbezug der Eulerschen Formel

$$e^{iz} = \cos z + i \sin z$$

und der Identität

$$|e^{iz}| = |\cos z + i \sin z| = \sqrt{\cos^2 z + \sin^2 z} = 1$$

erhält man

$$\begin{aligned} \left| \int_{\Omega} X(\omega) dP(\omega) \right| &= \left| r' \cdot e^{i\gamma'} \right| = |r'| \cdot |e^{i\gamma'}| = |r'| = r' \\ &= r' \cdot e^{i\gamma'} \cdot e^{-i\gamma'} \\ &= \int_{\Omega} X(\omega) dP(\omega) \cdot e^{-i\gamma'} = \int_{\Omega} X(\omega) \cdot e^{-i\gamma'} dP(\omega) \\ &= \int_{\Omega} r(\omega) \cdot e^{i\gamma(\omega)} \cdot e^{-i\gamma'} dP(\omega) = \int_{\Omega} r(\omega) e^{i(\gamma(\omega) - \gamma')} dP(\omega) \\ &= \int_{\Omega} r(\omega) \cdot \cos(\gamma(\omega) - \gamma') dP(\omega) + \underbrace{\int_{\Omega} r(\omega) \cdot i \cdot \sin(\gamma(\omega) - \gamma') dP(\omega)}_{=0, \text{ da } r' \in \mathbb{R}^+} \\ &\leq \int_{\Omega} r(\omega) dP(\omega) = \int_{\Omega} |X(\omega)| dP(\omega). \end{aligned}$$

f) Es gilt:

$$\begin{aligned}
 \mathbf{E}[X \cdot Y] &= \mathbf{E}[(\operatorname{Re} X + i \operatorname{Im} X) \cdot (\operatorname{Re} Y + i \operatorname{Im} Y)] \\
 &= \mathbf{E}[\operatorname{Re} X \cdot \operatorname{Re} Y - \operatorname{Im} X \cdot \operatorname{Im} Y + i \operatorname{Re} X \operatorname{Im} Y + i \operatorname{Im} X \operatorname{Re} Y] \\
 &= \mathbf{E}[\operatorname{Re} X \cdot \operatorname{Re} Y] - \mathbf{E}[\operatorname{Im} X \cdot \operatorname{Im} Y] + i \mathbf{E}[\operatorname{Re} X \operatorname{Im} Y] + i \mathbf{E}[\operatorname{Im} X \operatorname{Re} Y] \\
 &\stackrel{(*)}{=} \mathbf{E}[\operatorname{Re} X] \cdot \mathbf{E}[\operatorname{Re} Y] - \mathbf{E}[\operatorname{Im} X] \cdot \mathbf{E}[\operatorname{Im} Y] \\
 &\quad + i \mathbf{E}[\operatorname{Re} X] \cdot \mathbf{E}[\operatorname{Im} Y] + i \mathbf{E}[\operatorname{Im} X] \cdot \mathbf{E}[\operatorname{Re} Y] \\
 &= (\mathbf{E}[\operatorname{Re} X] + i \mathbf{E}[\operatorname{Im} X]) \cdot (\mathbf{E}[\operatorname{Re} Y] + i \mathbf{E}[\operatorname{Im} Y]) = \mathbf{E}[X] \cdot \mathbf{E}[Y].
 \end{aligned}$$

((*): Die stochastische Unabhängigkeit der betreffenden Real- und Imaginärteile folgt aus Satz 11.26, indem man $X_1 = (\operatorname{Re} X, \operatorname{Im} X)$ und $X_2 = (\operatorname{Re} Y, \operatorname{Im} Y)$ wählt und für g_1 und g_2 die entsprechenden Projektionsabbildungen einsetzt.) ■

Der Satz von der majorisierten Konvergenz gilt auch im Komplexen.

13.4 Satz (Satz von der majorisierten Konvergenz im Komplexen):

Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X_n)_{n \in \mathbb{N}}$ eine Folge komplexwertiger P -integrierbarer Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. Für alle $n \in \mathbb{N}$ gelte $|X_n| \leq Y$ mit einer P -integrierbaren, \mathfrak{F} - \mathfrak{B} -messbaren Abbildung $Y: \Omega \rightarrow \mathbb{R}^+$ und es existiere $\lim_{n \rightarrow \infty} X_n$ fast überall. Dann ist $\lim_{n \rightarrow \infty} X_n$ P -integrierbar und es gilt:

$$\lim_{n \rightarrow \infty} \int_{\Omega} X_n dP = \int_{\Omega} \lim_{n \rightarrow \infty} X_n dP.$$

Beweis:

Der Beweis ergibt sich aus dem Satz von der majorisierten Konvergenz unter Berücksichtigung der Abschätzungen $|\operatorname{Re} X_n| \leq |X_n|$ und $|\operatorname{Im} X_n| \leq |X_n|$, $n \in \mathbb{N}$. ■

13.2 Eigenschaften charakteristischer Funktionen

13.5 Definition (Fourier–Transformierte, charakteristische Funktion):

Es sei μ ein endliches Maß auf $(\mathbb{R}, \mathfrak{B})$. Dann heißt die durch

$$\varphi_{\mu}(t) := \int_{\mathbb{R}} e^{itx} d\mu, \quad t \in \mathbb{R},$$

definierte Funktion die charakteristische Funktion von μ . Sie ist für alle $t \in \mathbb{R}$ definiert, da aufgrund von $|e^{itx}| = 1$ das Integral stets konvergiert. Wenn X eine reelle Zufallsvariable über $(\Omega, \mathfrak{F}, P)$ ist, dann heißt die durch

$$\varphi_X(t) := \mathbf{E}[e^{itX}] := \int_{\Omega} e^{itX} dP = \int_{\mathbb{R}} e^{itx} dP_X, \quad t \in \mathbb{R},$$

definierte Funktion die Fourier–Transformierte oder charakteristische Funktion von X bzw. P_X .

Es ist stets

$$|\varphi_X(t)| = \left| \int_{\mathbb{R}} e^{itx} dP_X \right| \leq \int_{\mathbb{R}} |e^{itx}| dP_X = \int_{\mathbb{R}} dP_X = 1 \quad \forall t \in \mathbb{R}$$

und nach Satz 13.3 a) gilt außerdem:

$$\varphi_X(-t) = \int_{\mathbb{R}} e^{-itx} dP_X = \int_{\mathbb{R}} \overline{e^{itx}} dP_X = \overline{\int_{\mathbb{R}} e^{itx} dP_X} = \overline{\varphi_X(t)}, \quad \forall t \in \mathbb{R}.$$

13.6 Beispiel:

Einige Beispiele für charakteristische Funktionen

(siehe auch Mathematica-Notebook zu diesem Beispiel)

1. Es sei P_X eine Einpunktverteilung, d.h. es existiert ein $c \in \mathbb{R}$ mit $P(X = c) = 1$. Dann lautet die zugehörige charakteristische Funktion

$$\varphi_X(t) = e^{itc}, \quad t \in \mathbb{R}.$$

2. Es sei X eine Bernoulli-verteilte Zufallsgröße mit $P(X = 1) = p$ und $P(X = 0) = 1 - p$. Die zu X gehörende charakteristische Funktion berechnet sich wie folgt:

$$\varphi_X(t) = \mathbf{E}[e^{itx}] = e^{it \cdot 1} p + e^{it \cdot 0} (1 - p) = 1 + p(e^{it} - 1), \quad t \in \mathbb{R}.$$

3. Eine mit dem Parameter $\lambda > 0$ Poisson-verteilte Zufallsgröße X hat die charakteristische Funktion

$$\varphi_X(t) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{itk} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}, \quad t \in \mathbb{R}.$$

4. Eine auf dem Intervall $[0, 1]$ gleichverteilte Zufallsgröße X habe die Dichte $f(x) = I_{[0,1]}(x)$, $x \in \mathbb{R}$. Dann gilt:

$$\begin{aligned} \varphi_X(t) &= \int_{\mathbb{R}} e^{itx} dP_X = \int_{\mathbb{R}} e^{itx} f(x) dx = \int_{\mathbb{R}} e^{itx} I_{[0,1]}(x) dx = \int_0^1 e^{itx} \cdot 1 dx \\ &= \frac{1}{it} e^{itx} \Big|_0^1 = \frac{1}{it} (e^{it} - 1), \quad t \in \mathbb{R}. \end{aligned}$$

5. Es sei $X \stackrel{d}{=} \mathcal{N}(0, 1)$. Dann gilt:

$$\begin{aligned} \varphi_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-it)^2/2} e^{-t^2/2} dx \\ &= e^{-t^2/2} \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-it)^2/2} dx}_{=1, \text{ Integral von } \mathcal{N}(it, 1)} = e^{-t^2/2}, \quad t \in \mathbb{R}. \end{aligned}$$

6. Für eine mit dem Parameter $\theta > 0$ exponentiell verteilte Zufallsgröße ergibt sich

$$\begin{aligned}\varphi_X(t) &= \int_{\mathbb{R}} e^{itx} dP_X = \int_0^{\infty} e^{itx} \cdot \theta \cdot e^{-\theta x} dx = \theta \int_0^{\infty} e^{x(it-\theta)} dx \\ &= \frac{\theta}{it - \theta} e^{x(it-\theta)} \Big|_0^{\infty} = -\frac{\theta}{it - \theta} = \frac{\theta}{\theta - it}, \quad t \in \mathbb{R}.\end{aligned}$$

7. Für eine mit den Parametern n und θ Erlang-verteilte Zufallsgröße X lässt sich mittels partieller Integration (Übung!) zeigen:

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} dP_X = \int_0^{\infty} e^{itx} \frac{\theta^n x^{n-1} e^{-\theta x}}{(n-1)!} dx = \left[\frac{\theta}{\theta - it} \right]^n, \quad t \in \mathbb{R}.$$

8. Es sei X eine Zufallsvariable mit der Dichte

$$f_X(x) = \begin{cases} \frac{1}{T} \left(1 - \frac{|x|}{T}\right), & \text{für } |x| \leq T \\ 0, & \text{für } |x| > T, \end{cases}$$

die sogenannte Dreiecksverteilung auf $[-T, T]$.

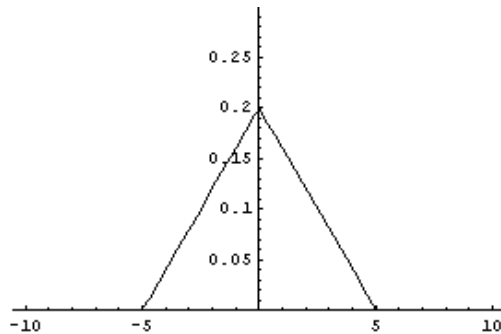


Abbildung 13.1: Dreiecksverteilung mit $T := 5$

Die zugehörige charakteristische Funktion φ_X berechnet sich wie folgt:

$$\begin{aligned}\varphi_X(t) &= \int_{\mathbb{R}} e^{itx} f_X(x) dx \\ &= \frac{1}{T} \left[\int_{-T}^0 \left(1 + \frac{x}{T}\right) \cdot e^{itx} dx + \int_0^T \left(1 - \frac{x}{T}\right) \cdot e^{itx} dx \right], \quad t \in \mathbb{R}.\end{aligned}$$

Es gilt:

$$\begin{aligned}
 \int_{-T}^0 \left(1 + \frac{x}{T}\right) \cdot e^{itx} dx &= \left[\frac{e^{itx}}{it} \left(1 + \frac{x}{T}\right) \right]_{x=-T}^{x=0} - \frac{1}{it} \cdot \frac{1}{T} \int_{-T}^0 e^{itx} dx \\
 &= \frac{1}{it} - 0 - \frac{1}{it} \cdot \frac{1}{T} \left[\frac{e^{itx}}{it} \right]_{x=-T}^{x=0} = \frac{1}{it} + \frac{1}{t^2 T} (1 - e^{-itT}) \quad \text{und} \\
 \int_0^T \left(1 - \frac{x}{T}\right) \cdot e^{itx} dx &= \left[\frac{e^{itx}}{it} \left(1 - \frac{x}{T}\right) \right]_{x=0}^{x=T} + \frac{1}{it} \cdot \frac{1}{T} \int_0^T e^{itx} dx \\
 &= 0 - \frac{1}{it} + \frac{1}{it} \cdot \frac{1}{T} \left[\frac{e^{itx}}{it} \right]_{x=0}^{x=T} = -\frac{1}{it} - \frac{1}{t^2 T} (e^{itT} - 1).
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow \varphi_X(t) &= \frac{1}{T} \cdot \frac{1}{t^2 T} \cdot (2 - e^{itT} - e^{-itT}) = \frac{2}{T^2 t^2} \left(1 - \frac{e^{itT} + e^{-itT}}{2} \right) \\
 &= \frac{2}{T^2 t^2} \left(1 - \frac{1}{2} \left(1 + 1 + \frac{itT}{1} + \frac{-itT}{1} + \frac{(-itT)^2}{2} + \frac{(-itT)^2}{2} + \dots \right) \right) \\
 &= \frac{2}{T^2 t^2} \left(1 - \left(1 + \frac{i^2 t^2 T^2}{2} + \frac{i^4 t^4 T^4}{4!} + \frac{i^6 t^6 T^6}{6!} + \dots \right) \right) \\
 &= \frac{2(1 - \cos(Tt))}{T^2 t^2}, \quad t \in \mathbb{R},
 \end{aligned}$$

wobei wir im letzten Schritt

$$\begin{aligned}
 e^z &= 1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots \\
 \cos z &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots
 \end{aligned}$$

benutzt haben.

13.7 Satz (Analytische Eigenschaften der charakteristischen Funktion):

Es sei μ ein endliches Maß auf $(\mathbb{R}, \mathfrak{B})$ mit der charakteristischen Funktion φ_μ . Dann gilt:

(i) $\varphi_\mu(t)$ ist stetig für alle $t \in \mathbb{R}$.

(ii) φ_μ ist positiv-semidefinit, d.h.

$$\sum_{j=1}^n \sum_{k=1}^n \gamma_j \bar{\gamma}_k \varphi_\mu(t_j - t_k) \geq 0$$

für alle $t_j \in \mathbb{R}$, alle $\gamma_j \in \mathbb{C}$, $j = 1, \dots, n$ und alle $n \in \mathbb{N}$.

(iii) Es gilt $\varphi_\mu(0) = \mu(\mathbb{R})$. Ist μ ein Wahrscheinlichkeitsmaß, so gilt also $\varphi_\mu(0) = 1$.

Beweis:

- (i) Es gilt $e^{i(t+h)x} \rightarrow e^{itx}$ für $h \rightarrow 0$ und es ist $|e^{itx}| = 1$ für alle $t \in \mathbb{R}$. Folglich lässt sich mit Hilfe des Satzes von der majorisierten Konvergenz schließen:

$$\begin{aligned} \lim_{h \rightarrow 0} \varphi_\mu(t+h) &= \lim_{h \rightarrow 0} \int_{\mathbb{R}} e^{i(t+h)x} d\mu(x) = \int_{\mathbb{R}} \lim_{h \rightarrow 0} e^{i(t+h)x} d\mu(x) \\ &= \int_{\mathbb{R}} e^{itx} d\mu(x) = \varphi_\mu(t). \end{aligned}$$

(ii)

$$\begin{aligned} \sum_{j=1}^n \sum_{k=1}^n \gamma_j \bar{\gamma}_k \varphi_\mu(t_j - t_k) &= \sum_{j=1}^n \sum_{k=1}^n \gamma_j \bar{\gamma}_k \int_{\mathbb{R}} e^{i(t_j - t_k)x} d\mu(x) \\ &= \int_{\mathbb{R}} \sum_{j=1}^n \gamma_j e^{it_j x} \cdot \sum_{k=1}^n \bar{\gamma}_k e^{-it_k x} d\mu(x) \quad \left(\text{da } \varphi_\mu(-s) = \overline{\varphi_\mu(s)} \right) \\ &= \int_{\mathbb{R}} \sum_{j=1}^n \gamma_j e^{it_j x} \cdot \overline{\sum_{k=1}^n \gamma_k e^{it_k x}} d\mu(x) \\ &= \int_{\mathbb{R}} \left| \sum_{j=1}^n \gamma_j e^{it_j x} \right|^2 d\mu(x) \geq 0. \end{aligned}$$

$$(iii) \quad \varphi_\mu(0) = \int_{\mathbb{R}} e^{i \cdot 0 \cdot x} d\mu = \int_{\mathbb{R}} 1 d\mu = \mu(\mathbb{R}). \quad \blacksquare$$

In Satz 13.7 wurde gezeigt, welche Eigenschaften charakteristische Funktionen aufweisen. Im Folgenden soll nun umgekehrt beschrieben werden, durch welche Eigenschaften charakteristische Funktionen von Wahrscheinlichkeitsverteilungen gekennzeichnet sind. Es wird sich zeigen, dass dies genau die Eigenschaften aus Satz 13.7 sind, d.h. Satz 13.7 stellt eine „genau dann, wenn“-Beziehung dar.

13.8 Satz (Satz von Bochner):

Eine Funktion $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ ist genau dann charakteristische Funktion eines endlichen Maßes μ auf $(\mathbb{R}, \mathfrak{B})$, wenn sie die Bedingungen (i), (ii) und (iii) aus Satz 13.7 erfüllt.

Beweis:

Siehe M. Loève: „Probability theory“, Van Nostrand, New York, 1963; Kapitel 4, §14: „Distribution functions und characteristic functions“.

Die folgenden Sätze geben Auskunft über Existenz und Eindeutigkeit der Umkehrung der Fourier-Transformation.

13.9 Satz (Umkehrformel von Lévy):

Es sei φ die charakteristische Funktion der Zufallsgröße X . Dann gilt für $-\infty < a < b < \infty$:

$$\lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

$$= P(a < X < b) + \frac{P(X = a) + P(X = b)}{2}.$$

Beweis:

Es wird

$$I(c) := \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} E[e^{itX}] dt$$

gesetzt. Da $[(e^{-ita} - e^{-itb})/it]e^{itX}$ für t auf dem kompakten Intervall $[-c, c]$ beschränkt ist, folgt mit dem Satz von Fubini:

$$\begin{aligned} I(c) &= \frac{1}{2\pi} E \left[\int_{-c}^c \frac{e^{it(X-a)} - e^{it(X-b)}}{it} dt \right] \\ &= \frac{1}{2\pi} E \left[\underbrace{\int_{-c}^c \frac{\cos(t(X-a)) - \cos(t(X-b))}{it} dt}_{=0, \text{ da hier } \int_{-c}^0 \dots dt = \int_0^c \dots dt} \right] \\ &\quad + \frac{1}{2\pi} E \left[\int_{-c}^c \frac{\sin(t(X-a)) - \sin(t(X-b))}{t} dt \right] \\ &= \frac{1}{\pi} E \left[\int_0^c \frac{\sin(t(X-a)) - \sin(t(X-b))}{t} dt \right] =: E[J_c(X)]. \end{aligned}$$

Für den Integralsinus gilt (z.B. Fichtenholz, S.654 oder [Si\(z\) auf functions.wolfram.com](https://functions.wolfram.com)):

$$\int_0^\infty \frac{\sin \alpha t}{t} dt = \int_0^{\operatorname{sgn}(\alpha) \cdot \infty} \frac{\sin x}{x} dx = \begin{cases} \frac{\pi}{2}, & \text{für } \alpha > 0, \\ 0, & \text{für } \alpha = 0, \\ -\frac{\pi}{2}, & \text{für } \alpha < 0. \end{cases}$$

Deswegen folgt

$$\lim_{c \rightarrow \infty} J_c(u) = \begin{cases} 1, & \text{für } a < u < b, \\ \frac{1}{2}, & \text{für } u = a \text{ oder } u = b, \\ 0, & \text{für } u < a \text{ oder } u > b. \end{cases}$$

Da $|J_c(u)| \leq 2$ für alle $-\infty < u$ und $c < \infty$ lässt sich mit Hilfe des Satzes von der majorisierten Konvergenz schließen, dass

$$\begin{aligned} \lim_{c \rightarrow \infty} I(c) &= \lim_{c \rightarrow \infty} E[J_c(X)] = E[\lim_{c \rightarrow \infty} J_c(X)] \\ &= E \left[\frac{1}{2} I_{[X=a \text{ oder } X=b]} + I_{[a < X < b]} \right] \\ &= \frac{P(X = a) + P(X = b)}{2} + P(a < X < b). \end{aligned} \quad \blacksquare$$

13.10 Bemerkung:

Sind a und b Stetigkeitspunkte der Verteilungsfunktion F von X , so reduziert sich die Formel aus Satz 13.9 auf

$$\lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt = F_X(b) - F_X(a).$$

13.11 Satz (Eindeutigkeit):

Es seien μ und ν zwei endliche Maße auf $(\mathbb{R}, \mathfrak{B})$ mit den charakteristischen Funktionen φ_μ und φ_ν . Dann gilt $\varphi_\mu(t) = \varphi_\nu(t)$ für alle $t \in \mathbb{R}$ genau dann, wenn $\mu = \nu$ gilt.

Beweis:

Da μ ein endliches Maß ist, kann es höchstens abzählbar viele $x \in \mathbb{R}$ mit $\mu(\{x\}) > 0$ geben. Damit lassen sich monoton fallende Folgen $(a_j)_{j \in \mathbb{N}}$ und $(b_k)_{k \in \mathbb{N}}$ mit $\lim_{j \rightarrow \infty} a_j = -\infty$ und $\lim_{k \rightarrow \infty} b_k = b \in \mathbb{R}$ finden mit $\mu(\{a_j\}) = 0$ und $\mu(\{b_k\}) = 0$.

Wegen der Stetigkeit von unten ergibt sich weiter

$$\mu((-\infty, b)) = \lim_{j \rightarrow \infty} \mu((a_j, b)) \quad \text{und} \quad \mu((-\infty, b]) = \lim_{k \rightarrow \infty} \mu((-\infty, b_k)).$$

Damit folgt mit Bemerkung 13.10 angewandt auf das Wahrscheinlichkeitsmaß $\frac{1}{\mu(\mathbb{R})}\mu$:

$$\mu((-\infty, b]) = \lim_{k \rightarrow \infty} \lim_{j \rightarrow \infty} \lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita_j} - e^{-itb_k}}{it} \varphi_\mu(t) dt.$$

Analog lässt sich eine entsprechende Formel für $\nu((-\infty, b])$ aufstellen. Da die Borelsche σ -Algebra durch Intervalle der Form $(-\infty, b]$ erzeugt wird, folgt $\mu = \nu$. Die Umkehrung erhält man aus der Eindeutigkeit der charakteristischen Funktion. ■

Da Verteilungsfunktionen und Bildmaße sich wechselseitig eindeutig bestimmen, ergibt sich sofort als Spezialfall des obigen Satzes:

13.12 Satz (Eindeutigkeit):

Es seien X und Y reellwertige Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$ mit den Verteilungsfunktionen F_X und F_Y sowie den charakteristischen Funktionen φ_X und φ_Y . Dann gilt $\varphi_X(t) = \varphi_Y(t)$ für alle $t \in \mathbb{R}$ genau dann, wenn $F_X(t) = F_Y(t)$ für alle $t \in \mathbb{R}$ gilt.

13.13 Satz:

Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und es bezeichne φ die charakteristische Funktion der Zufallsgröße $X: \Omega \rightarrow \mathbb{R}$ mit der Verteilungsfunktion F . Gilt

$$\int_{\mathbb{R}} |\varphi(t)| dt < \infty,$$

dann ist F stetig differenzierbar und es gilt

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi(t) dt \quad \forall x \in \mathbb{R}.$$

Beweis:

Es wird zunächst die Stetigkeit von F gezeigt. Da $\int_{\mathbb{R}} |\varphi(t)| dt < \infty$ vorausgesetzt ist, kann der Satz von der majorisierten Konvergenz angewendet werden und es ergibt sich

$$\lim_{b \rightarrow a+0} \frac{1}{2\pi} \int_{\mathbb{R}} \underbrace{\frac{e^{-ita} - e^{-itb}}{it}}_{\text{beschränkt}} \varphi(t) dt = \frac{1}{2\pi} \int_{\mathbb{R}} \underbrace{\lim_{b \rightarrow a+0} \frac{e^{-ita} - e^{-itb}}{it}}_{=0} \varphi(t) dt = 0,$$

wobei die Beschränktheit von $\frac{e^{-ita} - e^{-itb}}{it}$ wie im Beweis der Umkehrformel 13.9 folgt, wenn man berücksichtigt, dass diese Funktion im Unendlichen verschwindet. Andererseits gilt nach der Umkehrformel:

$$\begin{aligned} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt &= \lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt \\ &= P(a < X < b) + \frac{P(X = a) + P(X = b)}{2}. \end{aligned}$$

Zusammen mit

$$\lim_{b \rightarrow a+} \left(P(a < X < b) + \frac{P(X = a) + P(X = b)}{2} \right) \geq \frac{1}{2} P(X = a) \geq 0$$

ergibt sich $P(X = a) = P(X = a + 0) - P(X = a - 0) = 0$, was beweist, dass F stetig ist.

Aus

$$F(b) - F(a) = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) \frac{e^{-ita} - e^{-itb}}{it} dt \quad \text{für } b > 0$$

folgt

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) \frac{e^{-itx} - e^{-it(x+h)}}{ith} dt \quad \text{für } h > 0 \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-itx} \frac{1 - e^{-ith}}{ith} dt. \end{aligned}$$

Damit wird

$$\begin{aligned} f(x) = F'(x) &= \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-itx} \frac{1 - e^{-ith}}{ith} dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-itx} \cdot \lim_{h \rightarrow 0} \frac{1 - e^{-ith}}{ith} dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-itx} \cdot 1 dt. \end{aligned}$$

Damit ist die Existenz der rechtsseitigen Ableitung von F nachgewiesen. Für $h < 0$ ergibt sich auf dieselbe Weise die linksseitige Ableitung.

Es sind sowohl $\varphi(t)$ (vgl. Satz 13.7) als auch e^{itx} stetig in t . Außerdem gilt

$$\int_{\mathbb{R}} |\varphi(t) \cdot e^{-itx}| dt = \int_{\mathbb{R}} |\varphi(t)| \cdot \underbrace{|e^{-itx}|}_{=1} dt < \infty,$$

da nach Voraussetzung $\int_{\mathbb{R}} |\varphi(t)| dt < \infty$ ist. Damit lässt sich mit Hilfe des Satzes von der majorisierten Konvergenz schließen:

$$\begin{aligned} \lim_{h \rightarrow 0} f(x+h) &= \lim_{h \rightarrow 0} \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-it(x+h)} dt = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) \lim_{h \rightarrow 0} e^{-it(x+h)} dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \varphi(t) e^{-itx} dt = f(x). \end{aligned}$$

■

13.14 Beispiel:

Es sei $\varphi(t) := e^{-\frac{t^2}{2}}$, $t \in \mathbb{R}$. Wegen

$$\int_{\mathbb{R}} \left| e^{-\frac{t^2}{2}} \right| dt = \int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt = 2\pi$$

kann mit Hilfe von Satz 13.13 geschlossen werden:

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{t^2}{2}} \cdot e^{-itx} dt = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{(t+ix)^2}{2}} \cdot e^{\frac{(ix)^2}{2}} dt \\ &= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \cdot \int_{\mathbb{R}} e^{-\frac{(t+ix)^2}{2}} dt}_{=\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}. \end{aligned}$$

Offensichtlich ist $f(x)$ die Dichte der Standard-Normalverteilung (vgl. Beispiel 13.6 (5)).

13.15 Beispiel:

Es sei X eine Zufallsvariable mit der charakteristischen Funktion

$$\varphi_X(t) := \begin{cases} 1 - \frac{|t|}{T}, & \text{für } |t| \leq T \\ 0, & \text{für } |t| > T. \end{cases}$$

Es soll die Dichte $f_X(x)$ von X berechnet werden. Dazu wird Satz 13.13 herangezogen:

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi_X(t) dt \\ &= \frac{1}{2\pi} \int_{-T}^0 \left(1 + \frac{t}{T}\right) e^{-itx} dt + \frac{1}{2\pi} \int_0^T \left(1 - \frac{t}{T}\right) e^{-itx} dt. \end{aligned}$$

Es gilt

$$\begin{aligned} \int_{-T}^0 \left(1 + \frac{t}{T}\right) e^{-itx} dt &= \left[\frac{e^{-itx}}{-ix} \left(1 + \frac{t}{T}\right) \right]_{t=-T}^{t=0} - \frac{1}{-ixT} \int_{-T}^0 e^{-itx} dt \\ &= -\frac{1}{ix} - 0 + \frac{1}{ixT} \left[\frac{e^{-itx}}{-ix} \right]_{t=-T}^{t=0} = -\frac{1}{ix} - \frac{1}{(ix)^2 T} (1 - e^{iTx}) \end{aligned}$$

und

$$\begin{aligned} \int_0^T \left(1 - \frac{t}{T}\right) e^{-itx} dt &= \left[\frac{e^{-itx}}{-ix} \left(1 - \frac{t}{T}\right) \right]_{t=0}^{t=T} - \frac{1}{ixT} \int_0^T e^{-itx} dt \\ &= 0 + \frac{1}{ix} - \frac{1}{ixT} \left[\frac{e^{-itx}}{-ix} \right]_{t=0}^{t=T} = \frac{1}{ix} + \frac{1}{(ix)^2 T} (e^{-iTx} - 1). \end{aligned}$$

Damit wird

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi x^2 T} (2 - e^{iTx} - e^{-iTx}) = \frac{1}{\pi x^2 T} \left(1 - \frac{e^{iTx} + e^{-iTx}}{2}\right) \\ &= \frac{1}{\pi x^2 T} \left(1 - \frac{1}{2} \left(1 + 1 + \frac{iTx}{1} + \frac{-iTx}{1} + \frac{(-iTx)^2}{2} + \frac{(-iTx)^2}{2} + \dots\right)\right) \\ &= \frac{1}{\pi x^2 T} \left(1 - \left(1 + \frac{i^2 T^2 x^2}{2} + \frac{i^4 T^4 x^4}{4!} + \frac{i^6 T^6 x^6}{6!} + \dots\right)\right) \\ &= \frac{1 - \cos(Tx)}{\pi x^2 T}, \quad x \in \mathbb{R}. \end{aligned}$$

13.16 Satz:

Es seien X und Y zwei reelle Zufallsvariablen mit den charakteristischen Funktionen $\varphi_X(t)$ und $\varphi_Y(t)$. Dann gilt:

a) $\varphi_{aX+b}(t) = e^{itb} \varphi_X(at)$ für alle $a, b \in \mathbb{R}$.

b) Sind X und Y stochastisch unabhängig, so gilt:

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

c) Aus $\mathbf{E}[|X^k|] < \infty$ folgt, dass $\varphi_X(t)$ k -mal stetig differenzierbar ist. In diesem Fall gilt:

$$\varphi_X^{(k)}(0) = i^k \mathbf{E}[X^k].$$

d) Aus $\mathbf{E}[|X^k|] < \infty$ folgt

$$\varphi_X(t) = \sum_{j=0}^k \frac{(it)^j \mathbf{E}[X^j]}{j!} + o(|t|^k) \text{ für } t \rightarrow 0.$$

e) Falls $\mathbf{E}[|X|^{k+\delta}] < \infty$ für ein $\delta \in [0, 1]$ gilt

$$\varphi_X(t) = \sum_{j=0}^k \frac{(it)^j \mathbf{E}[X^j]}{j!} + R(t) \text{ mit } |R(t)| \leq \frac{2^{1-\delta} |t|^{k+\delta} \mathbf{E}[|X|^{k+\delta}]}{(1+\delta) \cdot \dots \cdot (k+\delta)}.$$

Beweis:

a) Es gilt:

$$\varphi_{aX+b}(t) = \mathbf{E}[e^{it(aX+b)}] = \mathbf{E}[e^{itaX} e^{itb}] = e^{itb} \mathbf{E}[e^{itaX}] = e^{itb} \varphi_X(at).$$

b) Da mit X und Y auch e^{itX} und e^{itY} stochastisch unabhängig sind, kann mit Hilfe von Satz 13.3 f) geschlossen werden:

$$\varphi_{X+Y}(t) = \mathbf{E}[e^{it(X+Y)}] = \mathbf{E}[e^{itX} e^{itY}] = \mathbf{E}[e^{itX}] \cdot \mathbf{E}[e^{itY}] = \varphi_X(t) \cdot \varphi_Y(t).$$

c) Wir setzen zunächst $k := 1$. Für alle $t, h \in \mathbb{R}$ mit $h \neq 0$ gilt dann

$$\frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \int_{\mathbb{R}} e^{itx} \frac{e^{ihx} - 1}{h} dP_X.$$

Weiter gilt

$$\begin{aligned} |e^{ihx} - 1| &= \sqrt{(\cos hx - 1)^2 + (\sin hx)^2} = \sqrt{2(1 - \cos hx)} \\ &= 2\sqrt{\frac{1 - \cos hx}{2}} = 2 \left| \sin\left(\frac{hx}{2}\right) \right| \leq 2 \left| \frac{hx}{2} \right| = |hx|. \end{aligned}$$

(Im vorletzten Schritt wird die Tatsache $|\sin x| \leq |x|$ verwendet. Für $x = 0$ gilt $\sin x = 0 = x$ und für $x \in [0, \frac{\pi}{2}]$ ist $(\sin x)' = \cos x \leq 1 = (x)'$. Für $x > 1$ ist die Behauptung evident und für $x < 0$ ist sie aus Symmetriegründen dann ebenfalls erfüllt.)

Aus $|e^{ihx} - 1| \leq |hx|$ wiederum folgt sofort $\left| e^{itx} \frac{e^{ihx} - 1}{h} \right| \leq |x|$.

Nach Voraussetzung ist $|X|$ P_X -integrierbar. Mit Hilfe des Satzes von der majorisierten Konvergenz, der auch für komplexwertige Zufallsvariable gültig ist (siehe 13.4), kann deshalb geschlossen werden:

$$\begin{aligned} \varphi_X^{(1)}(t) &= \lim_{h \rightarrow 0} \frac{\varphi_X(t+h) - \varphi_X(t)}{h} = \lim_{h \rightarrow 0} \int_{\mathbb{R}} e^{itx} \frac{e^{ihx} - 1}{h} dP_X \\ &= \int_{\mathbb{R}} \left[\lim_{h \rightarrow 0} e^{itx} \frac{e^{ihx} - 1}{h} \right] dP_X = i \int_{\mathbb{R}} x e^{itx} dP_X. \end{aligned}$$

Damit wird

$$\varphi_X^{(1)}(0) = \int_{\mathbb{R}} ix \cdot e^0 dP_X = i\mathbf{E}[X].$$

Weiter lässt sich induktiv von k auf $k + 1$ schließen:

$$\begin{aligned}
 \varphi_X^{(k+1)}(t) &= \lim_{h \rightarrow 0} \frac{\varphi_X^{(k)}(t+h) - \varphi_X^{(k)}(t)}{h} \\
 &= \lim_{h \rightarrow 0} i^k \cdot \int_{\mathbb{R}} \frac{x^k \cdot (e^{i(t+h)x} - e^{itx})}{h} dP_X \\
 &= i^k \cdot \int_{\mathbb{R}} x^k e^{itx} \lim_{h \rightarrow 0} \frac{e^{ihx} - 1}{h} dP_X \\
 &= i^{k+1} \cdot \int_{\mathbb{R}} x^{k+1} e^{itx} dP_X.
 \end{aligned}$$

Und damit:

$$\varphi_X^{(k+1)}(0) = i^{k+1} \cdot \int_{\mathbb{R}} x^{k+1} e^0 dP_X = i^{k+1} \mathbf{E}[X^{k+1}].$$

d) Die Behauptung folgt aus dem Taylor'schen Satz.

e) • Folgende Gleichung lässt sich induktiv nachweisen:

$$e^{it} - \sum_{j=0}^k \frac{(it)^j}{j!} = i^{k+1} \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_2} e^{it_1} dt_1 \dots dt_k dt_{k+1}.$$

(Für $k = 1$ lässt sich die Gleichung leicht nachrechnen und für $k \rightarrow k + 1$ schreibt man die rechte Seite als $i \int_0^t i^{k+1} \int_0^{t_{k+2}} \dots \int_0^{t_2} e^{it_1} dt_1 \dots dt_{k+1} dt_{k+2}$ und setzt nach Induktionsvoraussetzung $e^{it_{k+1}} - \sum_{j=0}^k \frac{(it_{k+2})^j}{j!}$ für $i^{k+1} \int_0^{t_{k+2}} \dots \int_0^{t_2} e^{it_1} dt_1 \dots dt_{k+1} dt_{k+2}$ ein. Die verbleibende Gleichung $e^{it} - \sum_{j=0}^{k+1} \frac{(it)^j}{j!} = i \int_0^t e^{it_{k+1}} - \sum_{j=0}^k \frac{(it_{k+2})^j}{j!} dt_{k+1}$ lässt sich ebenfalls leicht nachrechnen.)

- Es ist $|\sin t| \leq |t|^\delta$ für $\delta \in [0, 1]$, denn für $|t| \geq 1$ ist $|t|^\delta \geq 1$ und somit ist die Ungleichung offensichtlich erfüllt und für $|t| < 1$ gilt $|t| \leq |t|^\delta$ und außerdem natürlich $|\sin t| \leq |t|$. Damit lässt sich nun folgende Abschätzung aufstellen:

$$\begin{aligned}
 |e^{it} - 1| &= \sqrt{(\cos t - 1)^2 + (\sin t)^2} = \sqrt{2(1 - \cos t)} \\
 &= 2\sqrt{\frac{1 - \cos t}{2}} = 2 \left| \sin \left(\frac{t}{2} \right) \right| \leq 2^{1-\delta} |t|^\delta \quad \text{für } \delta \in [0, 1].
 \end{aligned}$$

Hieraus ergibt sich:

$$\begin{aligned}
 \left| e^{it} - \sum_{j=0}^k \frac{(it)^j}{j!} \right| &= \left| \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_3} \int_0^{t_2} e^{it_1} dt_1 \dots dt_{k+1} \right| \\
 &= \left| \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_3} \frac{1}{i} e^{it_2} - \frac{1}{i} dt_2 \dots dt_{k+1} \right| \\
 &= \left| \frac{1}{i} \right| \cdot \left| \int_0^t \int_0^{t_{k+1}} \dots \int_0^{t_3} e^{it_2} - 1 dt_2 \dots dt_{k+1} \right| \\
 &\leq \int_0^{|t|} \int_0^{t_{k+1}} \dots \int_0^{t_3} |e^{it_2} - 1| dt_2 \dots dt_{k+1} \\
 &\leq 2^{1-\delta} \int_0^{|t|} \int_0^{t_{k+1}} \dots \int_0^{t_3} (t_2)^\delta dt_2 \dots dt_{k+1} \\
 &= \frac{2^{1-\delta} |t|^{k+\delta}}{(1+\delta) \dots (k+\delta)}.
 \end{aligned}$$

Durch Anwendung dieser Abschätzung auf $\mathbf{E} \left[e^{itX} - \sum_{j=0}^k \frac{(itX)^j}{j!} \right]$ folgt dann die Behauptung:

$$\begin{aligned}
 \varphi_X(t) &= \mathbf{E}[e^{itX}] = \mathbf{E} \left[\sum_{j=0}^k \frac{(itX)^j}{j!} + e^{itX} - \sum_{j=0}^k \frac{(itX)^j}{j!} \right] \\
 &= \sum_{j=0}^k \frac{(it)^j \mathbf{E}[X^j]}{j!} + \underbrace{\mathbf{E} \left[e^{itX} - \sum_{j=0}^k \frac{(itX)^j}{j!} \right]}_{=: R(t)}
 \end{aligned}$$

mit

$$|R(t)| \leq \mathbf{E} \left[\frac{2^{1-\delta} |tX|^{k+\delta}}{(1+\delta) \dots (k+\delta)} \right] = \frac{2^{1-\delta} |t|^{k+\delta} \mathbf{E}[|X|^{k+\delta}]}{(1+\delta) \dots (k+\delta)}.$$

■

Satz 13.16 hat zahlreiche inner-mathematische Anwendungen, einige davon sollen hier notiert werden:

- a) Im Beispiel 13.14 5. wurde gezeigt, dass eine standardnormalverteilte Zufallsgröße X die charakteristische Funktion

$$\varphi_X(t) = e^{-t^2/2}, \quad t \in \mathbb{R},$$

besitzt. Wie man sich leicht klar machen kann, ist $Y := \sigma X + \mu$ normalverteilt mit den Parametern μ und σ , wobei $\mu \in \mathbb{R}$ und $\sigma > 0$ seien. Für die zugehörige charakteristische Funktion gilt dann aufgrund von Satz 13.16 a)

$$\varphi_Y(t) = \varphi_{\sigma X + \mu}(t) = e^{it\mu} \varphi_X(\sigma t) = e^{it\mu} \cdot e^{-(\sigma t)^2/2}, \quad t \in \mathbb{R}.$$

b) Summe von Poisson-verteilten Zufallsgrößen:

Es seien X und Y reellwertige und stochastisch unabhängige Zufallsvariablen mit $X \stackrel{d}{=} P(\lambda)$ und $Y \stackrel{d}{=} P(\mu)$. Dann gilt nach Satz 13.12 $X + Y \stackrel{d}{=} P(\lambda + \mu)$, denn es ist

$$\varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t) = e^{\lambda(e^{it}-1)} e^{\mu(e^{it}-1)} = e^{(\lambda+\mu)(e^{it}-1)}, \quad t \in \mathbb{R}.$$

c) Summe von normalverteilten Zufallsgrößen:

Diesmal seien X und Y reellwertige und stochastisch unabhängige Zufallsvariablen mit $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma)$ und $Y \stackrel{d}{=} \mathcal{N}(\mu', \sigma')$. Dann gilt $X + Y \stackrel{d}{=} \mathcal{N}(\mu + \mu', \sqrt{\sigma^2 + \sigma'^2})$ und

$$\begin{aligned} \varphi_{X+Y}(t) &= \varphi_X(t) \cdot \varphi_Y(t) = e^{it\mu} \cdot e^{-(\sigma t)^2/2} \cdot e^{it\mu'} \cdot e^{-(\sigma' t)^2/2} \\ &= e^{it(\mu+\mu')} \cdot e^{-(\sigma^2 + \sigma'^2)t^2/2}, \quad t \in \mathbb{R}. \end{aligned}$$

d) Summe von exponentiell verteilten Zufallsgrößen:

Es seien X_1, \dots, X_n unabhängige und jeweils mit demselben Parameter $\theta > 0$ exponentiell verteilten Zufallsgrößen. Dann gilt $\sum_{i=1}^n X_i \stackrel{d}{=} \text{Erlang}(n, \theta)$, denn es ist

$$\varphi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n \varphi_{X_i}(t) = [\varphi_{X_1}(t)]^n = \left[\frac{\theta}{\theta - it} \right]^n, \quad t \in \mathbb{R}.$$

Literatur zu Kapitel 13

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- Y.S. CHOW/H. TEICHER:
Probability Theory - Independence, Interchangeability, Martingales,
Springer-Verlag, New York, 1988.
ISBN: 0387406077
- G.H. FICHTENHOLZ:
Differential- und Integralrechnung II,
VEB Deutscher Verlag der Wissenschaften, Berlin, 1974.
ISBN: 3817112793
- M. FISZ:
Wahrscheinlichkeitsrechnung und mathematische Statistik,
VEB, Deutscher Verlag der Wissenschaften 1989.
ISBN: 3326000790
- P. P. SPIES:
Grundlagen stochastischer Modelle,
Hanser, München, 1982.
ISBN: 3446137114

Kapitel 14

Schwache Konvergenz und zentraler Grenzwertsatz

Der in diesem Kapitel behandelte zentrale Grenzwertsatz besagt, dass die geeignet normierte Summe einer großen Zahl stochastisch unabhängiger und identisch verteilter Zufallsvariablen annähernd standardnormal verteilt ist, was die Sonderstellung der Normalverteilung in der Stochastik erklärt (sogenannter zentraler Grenzwertsatz). Um den zentralen Grenzwertsatz zu beweisen, werden zunächst die Begriffe der vagen und schwachen Konvergenz eingeführt und einige Zusammenhänge mit charakteristischen Funktionen behandelt.

Schlüsselwörter: Vage und schwache Konvergenz, Konvergenz in Verteilung, Auswahlssatz von Helly, Stetigkeitssatz von Lévy, zentraler Grenzwertsatz, Satz von de Moivre-Laplace, Satz von Berry-Esseen.

14.1 Vage und schwache Konvergenz

In diesem Abschnitt wird eine weitere Konvergenzart für Zufallsvariablen vorgestellt, die sogenannte schwache Konvergenz bzw. Konvergenz in Verteilung. Dafür wird zunächst das folgende einführende Beispiel betrachtet:

Es mögen X_1, \dots, X_n reellwertige, identisch verteilte und stochastisch unabhängige Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$ bezeichnen. Es wird $M_n := \max \{X_1, \dots, X_n\} \forall n \in \mathbb{N}$ und $F(x) := P(X_1 \leq x)$ gesetzt. Eine einfache Überlegung zeigt

$$\{M_n \leq x\} \iff \{X_1 \leq x, \dots, X_n \leq x\},$$

so dass gilt

$$\begin{aligned} P(M_n \leq x) &= P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) \\ &= (F(x))^n, \quad x \in \mathbb{R}, \quad n \in \mathbb{N}. \end{aligned}$$

Es wird nun speziell $X_1 \stackrel{d}{=} \text{Exp}(\lambda)$, mit $\lambda > 0$ gewählt. Gefragt wird nach der Verteilungsfunktion von $Y_n := M_n + \lambda^{-1} \ln(n)$ für $n \rightarrow \infty$:

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{Y_n}(x) &= \lim_{n \rightarrow \infty} (F(x + \lambda^{-1} \ln(n)))^n = \lim_{n \rightarrow \infty} \left(1 - e^{-(\lambda x + \ln(n))}\right)^n \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{-e^{-\lambda x}}{n}\right)^n = e^{-e^{-\lambda x}} \quad \forall x \in \mathbb{R}. \end{aligned}$$

Da die Grenzfunktion wieder eine stetige Verteilungsfunktion ist, legt dieses Beispiel nahe, die Konvergenz einer Folge $(P_n)_{n \in \mathbb{N}}$ von Wahrscheinlichkeitsverteilungen auf $(\mathbb{R}, \mathfrak{B})$ gegen eine Limesverteilung P vermöge

$$\lim_{n \rightarrow \infty} P_n(B) = P(B) \quad \forall B \in \mathfrak{B}$$

festzulegen, was jedoch unzweckmäßig ist, wie ein anderes Beispiel zeigt. Dazu werden die Einpunktverteilungen

$$P_n(B) := I_B\left(\frac{1}{n}\right) \quad \forall B \in \mathfrak{B}, \quad n \in \mathbb{N}$$

betrachtet. Diese Verteilungen approximieren offensichtlich diejenige Verteilung, die der Menge $\{0\}$ die Wahrscheinlichkeit 1 zuordnet. Folglich möchte man als Limesverteilung $P(B) := I_B(0) \forall B \in \mathfrak{B}$ haben. Doch $P_n(B) \rightarrow P(B)$ ist bereits für $B := (-\infty, 0]$ nicht erfüllt, denn es gilt

$$P_n((-\infty, 0]) = 0 \quad \forall n \in \mathbb{N} \quad \text{und} \quad P((-\infty, 0]) = 1.$$

Betrachtet man anstelle der Verteilungen P_n die zugehörigen Verteilungsfunktionen $F_n = I_{[\frac{1}{n}, \infty)}$, so ergibt sich hier

$$I_{[\frac{1}{n}, \infty)} \rightarrow I_{(0, \infty)} \quad \text{für } n \rightarrow \infty.$$

Jedoch ist $I_{(0, \infty)}$ keine Verteilungsfunktion. Die Schwierigkeit resultiert hier offenbar — im Gegensatz zum vorangegangenen Beispiel — aus der Tatsache, dass ein Randpunkt von $B = (-\infty, 0]$ bzgl. P eine positive Wahrscheinlichkeit besitzt. Deswegen muss der Begriff der schwachen Konvergenz bzw. Verteilungskonvergenz etwas allgemeiner gefasst werden.

Im Folgenden werden Maße μ auf $(\mathbb{R}, \mathfrak{B})$ betrachtet. Im Hinblick auf zukünftige Anwendungen wird vorausgesetzt, dass μ lokal endlich ist, d.h. dass $\mu(K) < \infty$ für alle kompakten Mengen $K \subset \mathbb{R}$ ist. Mit C_b werde die Menge aller stetigen Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)| < \infty$ bezeichnet und mit C_0 werde die Menge aller stetigen Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ mit kompaktem Träger $T := \overline{\{x \mid f(x) \neq 0\}}$ bezeichnet. Für eine beliebige Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ bezeichne $C(f)$ die Menge ihrer Stetigkeitspunkte.

14.1 Definition (vage Konvergenz, schwache Konvergenz):

Es seien μ sowie die Elemente der Folge $(\mu_n)_{n \in \mathbb{N}}$ lokal endliche Maße auf $(\mathbb{R}, \mathfrak{B})$. Man sagt $(\mu_n)_{n \in \mathbb{N}}$ konvergiert vage gegen μ , kurz $\mu_n \xrightarrow{v} \mu$, falls

$$\int_{\mathbb{R}} f \, d\mu_n \longrightarrow \int_{\mathbb{R}} f \, d\mu \quad \forall f \in C_0.$$

Gilt diese Beziehung sogar für alle $f \in C_b$ und ist $\mu(\mathbb{R}) < \infty$, so konvergiert $(\mu_n)_{n \in \mathbb{N}}$ schwach gegen μ , kurz $\mu_n \xrightarrow{w} \mu$ (vom englischen „weak convergence“).

14.2 Definition (straff):

Es seien die Elemente der Folge $(\mu_n)_{n \in \mathbb{N}}$ lokal endliche Maße auf $(\mathbb{R}, \mathfrak{B})$. $(\mu_n)_{n \in \mathbb{N}}$ heißt straff, falls

$$\lim_{N \rightarrow \infty} \sup_{n \in \mathbb{N}} \mu_n(\mathbb{R} \setminus [-N, N]) = 0$$

gilt.

Den Zusammenhang zwischen vager und schwacher Konvergenz verdeutlicht der folgende Satz.

14.3 Satz:

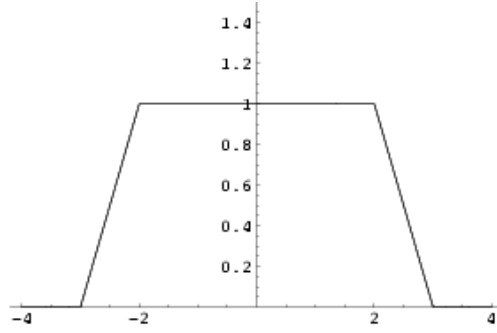
Es seien μ sowie die Elemente der Folge $(\mu_n)_{n \in \mathbb{N}}$ lokal endliche Maße auf $(\mathbb{R}, \mathfrak{B})$ mit $\mu_n \xrightarrow{v} \mu$ für $n \rightarrow \infty$. Dann sind folgende Aussagen äquivalent:

- a) $\mu_n \xrightarrow{w} \mu$ für $n \rightarrow \infty$.
- b) $\lim_{n \rightarrow \infty} \mu_n(\mathbb{R}) = \mu(\mathbb{R}) < \infty$.
- c) $(\mu_n)_{n \in \mathbb{N}}$ ist straff.

Beweis:

„a) \Rightarrow b)“: Da die konstante Funktion $I_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 1 \, \forall x \in \mathbb{R}$ zu C_b gehört, impliziert $\mu_n \xrightarrow{w} \mu$ für $n \rightarrow \infty$:

$$\mu_n(\mathbb{R}) = \int_{\mathbb{R}} d\mu_n = \int_{\mathbb{R}} I_{\mathbb{R}} \, d\mu_n \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} I_{\mathbb{R}} \, d\mu = \int_{\mathbb{R}} d\mu = \mu(\mathbb{R}).$$


 Abbildung 14.1: $g_N(x)$ für $N := 2$

„b) \Rightarrow c)“: Es sei $g_N \in C_0$ gegeben durch

$$g_N(x) := \begin{cases} 1 & , \quad |x| \leq N \\ N+1-|x| & , \quad |x| \in (N, N+1) \\ 0 & , \quad |x| \geq N+1. \end{cases}$$

Unter Benutzung der Abschätzungen

$$I_{[-N,N]} \leq g_N \leq I_{[-N-1,N+1]} \quad \text{bzw.} \quad I_{\mathbb{R} \setminus [-N-1,N+1]} \leq 1 - g_N \leq I_{\mathbb{R} \setminus [-N,N]}$$

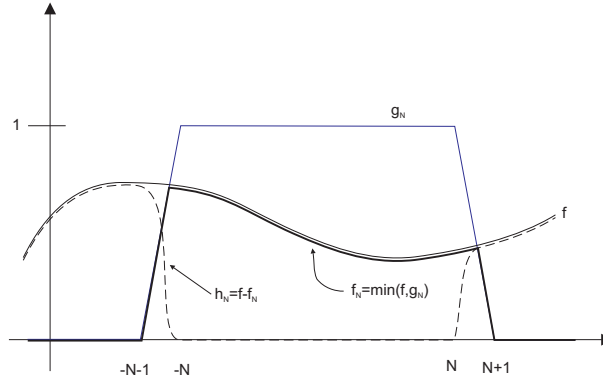
und den gegebenen Voraussetzungen folgt dann für alle hinreichend großen N :

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mu_n(\mathbb{R} \setminus [-N-1, N+1]) &= \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} I_{\mathbb{R} \setminus [-N-1, N+1]} d\mu_n \\ &\leq \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} (1 - g_N) d\mu_n \\ &= \lim_{n \rightarrow \infty} \mu_n(\mathbb{R}) - \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_N d\mu_n \\ &= \mu(\mathbb{R}) - \int_{\mathbb{R}} g_N d\mu \\ &= \int_{\mathbb{R}} (1 - g_N) d\mu \\ &\leq \int_{\mathbb{R}} I_{\mathbb{R} \setminus [-N, N]} d\mu \\ &= \mu(\mathbb{R} \setminus [-N, N]). \end{aligned}$$

Da $\mu(\mathbb{R} \setminus [-N, N]) \rightarrow 0$ für $N \rightarrow \infty$, folgt die Behauptung.

„c) \Rightarrow a)“: Für beliebiges $f \in C_b$, wobei o.B.d.A. $0 \leq f \leq 1$ angenommen werden darf, sei nun $f_N := \min(f, g_N) \in C_0$ und $h_N := f - f_N$. Wegen $f_N \leq f = f_N + h_N$ und $0 \leq h_N \leq I_{\mathbb{R} \setminus [-N, N]}$ für alle N folgt zunächst

$$0 \leq \int_{\mathbb{R}} h_N d\mu_n \leq \int_{\mathbb{R}} I_{\mathbb{R} \setminus [-N, N]} d\mu_n = \mu_n(\mathbb{R} \setminus [-N, N]) \rightarrow 0 \quad \text{für } N \rightarrow \infty.$$


 Abbildung 14.2: Graphische Darstellung der Funktionen f_N und h_N

Wir rechnen weiter:

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n &= \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n - \lim_{n \rightarrow \infty} \mu_n(\mathbb{R} \setminus [-N, N]) \\
 &\leq \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} (f - h_N) \, d\mu_n \quad (h_N \leq I_{\mathbb{R} \setminus [-N, N]}) \\
 &= \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} f_N \, d\mu_n \quad (f_N = f - h_N) \\
 &= \int_{\mathbb{R}} f_N \, d\mu \quad (f_N \in C_0 \text{ und } (\mu_n)_{n \in \mathbb{N}} \text{ konv. vage}) \\
 &\leq \int_{\mathbb{R}} f \, d\mu \quad (f = f_N + h_N \text{ und } 0 \leq h_N \leq 1) \\
 &= \int_{\mathbb{R}} f_N \, d\mu + \int_{\mathbb{R}} h_N \, d\mu \\
 &\leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} f_N \, d\mu_n + \int_{\mathbb{R}} h_N \, d\mu \\
 &\leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n + \mu(\mathbb{R} \setminus [-N, N]).
 \end{aligned}$$

Wegen $\mu(\mathbb{R} \setminus [-N, N]) \rightarrow 0$ für $N \rightarrow \infty$ folgt hieraus

$$\int_{\mathbb{R}} f \, d\mu = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n,$$

was zu zeigen war. ■

Es ist wichtig zu wissen, dass man sich beim Nachweis von schwacher Konvergenz im Fall endlicher Maße auf gleichmäßig stetige und beschränkte Funktionen f zurückziehen kann, wie der folgende Satz zeigt.

14.4 Satz:

Es seien μ und die Elemente der Folge $(\mu_n)_{n \in \mathbb{N}}$ endliche Maße auf $(\mathbb{R}, \mathfrak{B})$. Wenn für jede gleichmäßig stetige und beschränkte Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$

$$\int_{\mathbb{R}} f \, d\mu_n \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu$$

gilt, dann gilt auch schon $\mu_n \xrightarrow{w} \mu$ für $n \rightarrow \infty$.

Beweis:

Es sei $f \in C_b$ mit $\sup_{x \in \mathbb{R}} |f(x)| = M < \infty$ und $\varepsilon > 0$. Es werden eine Zahl $a \in \mathbb{R}$ mit

$\int_{|x| \geq a} d\mu(x) \leq \frac{\varepsilon}{2M}$ sowie stetige Funktionen f_* , f^* mit folgenden Eigenschaften gewählt:

$$\begin{aligned} f_*(x) &:= f^*(x) := f(x) && \text{für } |x| \leq a \\ -M &\leq f_*(x) \leq f(x) && \text{für } a \leq |x| \leq a+1 \\ f(x) &\leq f^*(x) \leq M && \text{für } a \leq |x| \leq a+1 \\ f_*(x) &= -M && \text{für } |x| \geq a+1 \\ f^*(x) &= M && \text{für } |x| \geq a+1. \end{aligned}$$

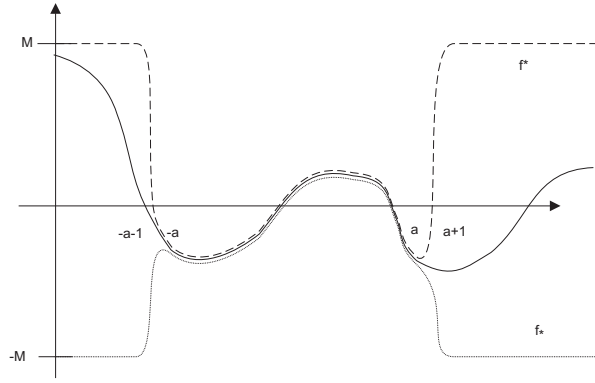


Abbildung 14.3: Graphische Darstellung der Funktionen f_* und f^*

Da jede auf der kompakten Menge $K := [-a-1, a+1]$ stetige Funktion auf K auch gleichmäßig stetig ist und f^* und f_* außerhalb von K konstant sind, sind f^* und f_* gleichmäßig stetig (vgl. Aumann/Haupt: „Einführung in die reelle Analysis“, Band I, S. 139). Da $f^*(x) = f_*(x) = f(x)$ für $|x| \leq a$ ist, gilt:

$$\int_{\mathbb{R}} (f^* - f_*) d\mu \leq 2M \int_{|x| \geq a} d\mu(x) \leq \varepsilon.$$

Weiter folgt

$$\begin{aligned} \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} f d\mu_n &\geq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} f_* d\mu_n = \int_{\mathbb{R}} f_* d\mu \\ &= \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} (f - f_*) d\mu \\ &\geq \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} (f^* - f_*) d\mu \geq \int_{\mathbb{R}} f d\mu - \varepsilon. \end{aligned}$$

Und entsprechend

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n &\leq \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} f^* \, d\mu_n = \int_{\mathbb{R}} f^* \, d\mu \\
 &= \int_{\mathbb{R}} f \, d\mu + \int_{\mathbb{R}} (f^* - f) \, d\mu \\
 &\leq \int_{\mathbb{R}} f \, d\mu + \int_{\mathbb{R}} (f^* - f_*) \, d\mu \leq \int_{\mathbb{R}} f \, d\mu + \varepsilon.
 \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f \, d\mu_n = \int_{\mathbb{R}} f \, d\mu.$$

■

14.5 Satz:

Es seien P und die Elemente der Folge $(P_n)_{n \in \mathbb{N}}$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathfrak{B})$, F und die Elemente der Folge $(F_n)_{n \in \mathbb{N}}$ bezeichnen die zugehörigen Verteilungsfunktionen. Dann gilt:

$$P_n \xrightarrow[n \rightarrow \infty]{w} P \iff F_n(x) \xrightarrow[n \rightarrow \infty]{} F(x) \quad \forall x \in C(F).$$

Beweis:

Es wird zunächst $P_n \xrightarrow[n \rightarrow \infty]{w} P$ und $x \in C(F)$ angenommen. Zu $\varepsilon > 0$ sei $\delta > 0$ so gewählt, dass $|F(x) - F(y)| \leq \varepsilon$ für alle x, y mit $|x - y| \leq \delta$ gilt. Ferner werden $\underline{f}, \bar{f} \in C_b$ definiert durch $0 \leq \underline{f}, \bar{f} \leq 1$ und

$$\underline{f}(y) := \begin{cases} 1, & \text{für } y \leq x - \delta \\ 0, & \text{für } y \geq x, \end{cases} \quad \bar{f}(y) := \begin{cases} 1, & \text{für } y \leq x \\ 0, & \text{für } y \geq x + \delta. \end{cases}$$

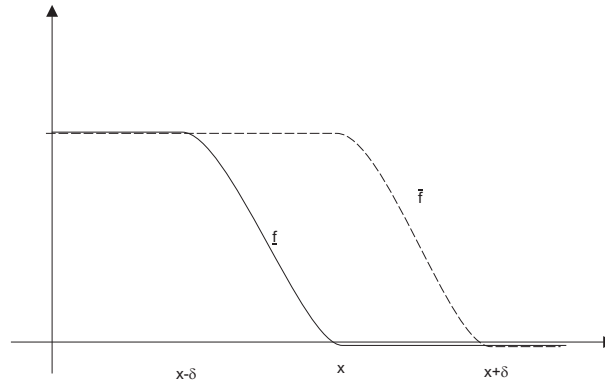


Abbildung 14.4: Graphische Darstellung der Funktionen \underline{f} und \bar{f}

Dann folgt

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} F_n(x) &= \limsup_{n \rightarrow \infty} P_n((-\infty, x]) = \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} I_{(-\infty, x]}(y) dP_n(y) \\
 &\leq \limsup_{n \rightarrow \infty} \int_{\mathbb{R}} \bar{f}(y) dP_n(y) = \int_{\mathbb{R}} \bar{f}(y) dP(y) \\
 &\leq \int_{\mathbb{R}} I_{(-\infty, x+\delta]}(y) dP(y) = P((-\infty, x+\delta]) = F(x+\delta) \leq F(x) + \varepsilon, \\
 \liminf_{n \rightarrow \infty} F_n(x) &= \liminf_{n \rightarrow \infty} P_n((-\infty, x]) = \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} I_{(-\infty, x]}(y) dP_n(y) \\
 &\geq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}} \underline{f}(y) dP_n(y) = \int_{\mathbb{R}} \underline{f}(y) dP(y) \\
 &\geq \int_{\mathbb{R}} I_{(-\infty, x-\delta]}(y) dP(y) = P((-\infty, x-\delta]) = F(x-\delta) \geq F(x) - \varepsilon.
 \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt $\lim_{n \rightarrow \infty} F_n(x) = F(x)$.

Sei nun umgekehrt $F_n(x) \xrightarrow[n \rightarrow \infty]{} F(x) \forall x \in C(F)$ vorausgesetzt. Es wird zunächst gezeigt, dass $P_n \xrightarrow[n \rightarrow \infty]{v} P$ gilt. Dazu wird eine beliebige Funktion $f \in C_0$ mit kompaktem Träger $K := [a, b]$, $a < b$, betrachtet. Da f gleichmäßig stetig ist, lässt sich zu jedem $\varepsilon > 0$ ein $\delta > 0$ angeben, so dass $|f(x) - f(y)| \leq \varepsilon$ ist für alle x, y mit $|x - y| \leq \delta$. Es werden $x_i \in C(F)$, $1 \leq i \leq k$ so gewählt, dass $a = x_1 < x_2 < \dots < x_{k+1} = b$ und $\max_{1 \leq i \leq k} (x_{i+1} - x_i) < \delta$ ist. Dann gilt:

$$\begin{aligned}
 H_n &:= \int_{[a,b]} f dP_n - \int_{[a,b]} f dP \\
 &= \sum_{i=1}^k \left\{ \left[\int_{[x_i, x_{i+1}]} f(x) dP_n(x) - \int_{(x_i, x_{i+1}]} f(x_i) dP_n(x) \right] \right. \\
 &\quad + \left[\int_{[x_i, x_{i+1}]} f(x_i) dP_n(x) - \int_{(x_i, x_{i+1}]} f(x_i) dP(x) \right] \\
 &\quad \left. + \left[\int_{[x_i, x_{i+1}]} f(x_i) dP(x) - \int_{(x_i, x_{i+1}]} f(x) dP(x) \right] \right\} \\
 &= \sum_{i=1}^k \left\{ \int_{(x_i, x_{i+1}]} [f(x) - f(x_i)] dP_n(x) + \int_{(x_i, x_{i+1}]} [f(x_i) - f(x)] dP(x) \right. \\
 &\quad \left. + f(x_i)[F_n(x_{i+1}) - F_n(x_i) - F(x_{i+1}) + F(x_i)] \right\},
 \end{aligned}$$

woraus

$$|H_n| \leq \varepsilon + \varepsilon + \sum_{i=1}^k |f(x_i)| \cdot |F_n(x_{i+1}) - F_n(x_i) - F(x_{i+1}) + F(x_i)|$$

bzw.

$$|H_n| \longrightarrow 2\varepsilon \quad \text{für } n \rightarrow \infty$$

gefolgert werden kann. Da $\varepsilon > 0$ beliebig gewählt war, folgt $P_n \xrightarrow[n \rightarrow \infty]{v} P$. Um $P_n \xrightarrow[n \rightarrow \infty]{w} P$ nachzuweisen, wird Satz 14.3 b) angewandt. Aus $\lim_{x \rightarrow +\infty} F_n(x) = \lim_{x \rightarrow +\infty} F(x) = 1 \quad \forall n \in \mathbb{N}$ folgt $P_n(\mathbb{R}) = 1 \quad \forall n \in \mathbb{N}$ sowie $P(\mathbb{R}) = 1$ und damit $\lim_{n \rightarrow \infty} P_n(\mathbb{R}) = P(\mathbb{R}) < \infty$. ■

Ansichts von Satz 14.5 definiert man:

14.6 Definition (schwache Konvergenz, Konvergenz in Verteilung):

Es seien X und die Elemente der Folge $(X_n)_{n \in \mathbb{N}}$ reellwertige Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. Die zugehörigen Verteilungsfunktionen seien F und $(F_n)_{n \in \mathbb{N}}$. Man sagt, die Folge $(X_n)_{n \in \mathbb{N}}$ konvergiert schwach oder in Verteilung gegen X , in Zeichen $X_n \xrightarrow{i.V.} X$, falls

$$F_n(x) \xrightarrow[n \rightarrow \infty]{} F(x) \quad \text{für alle } x \in C(F)$$

gilt.

Es stellt sich die Frage, wie sich der Begriff der Konvergenz in Verteilung den bereits behandelten Begriffen der stochastischen und der fast sicheren Konvergenz unterordnet. Dazu wird der folgende Satz betrachtet.

14.7 Satz:

Es seien X und die Elemente der Folge $(X_n)_{n \in \mathbb{N}}$ reellwertige Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. Dann gilt

$$X_n \xrightarrow[n \rightarrow \infty]{p} X \quad \Rightarrow \quad X_n \xrightarrow[n \rightarrow \infty]{i.V.} X.$$

Beweis:

Aufgrund von Satz 14.4 kann man sich beim Nachweis der schwachen Konvergenz auf gleichmäßig stetige beschränkte Funktionen f zurückziehen. Es gibt zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit der Eigenschaft

$$|f(x) - f(y)| \leq \varepsilon \quad \text{für alle } x, y \in \mathbb{R} \text{ mit } |x - y| \leq \delta.$$

Für $n \in \mathbb{N}$ bezeichne $A_n := \{\omega \mid |X_n(\omega) - X(\omega)| \leq \delta\}$, und wir setzen wieder $M := \sup_{x \in \mathbb{R}} |f(x)|$. Dann gilt:

$$\begin{aligned} |\mathbf{E}[f(X_n)] - \mathbf{E}[f(X)]| &= \left| \int_{\Omega} f \circ X_n \, dP - \int_{\Omega} f \circ X \, dP \right| \\ &\leq \int_{A_n} |f \circ X_n - f \circ X| \, dP + \int_{\bar{A}_n} |f \circ X_n - f \circ X| \, dP \\ &\leq \varepsilon \int_{A_n} dP + 2M \int_{\bar{A}_n} dP \\ &\leq \varepsilon + 2M P(\bar{A}_n) = \varepsilon + 2M \cdot P(|X_n - X| > \delta). \end{aligned}$$

Aufgrund der Voraussetzung $X_n \xrightarrow[n \rightarrow \infty]{p} X$ gilt $P(|X_n - X| > \delta) \rightarrow 0$ für $n \rightarrow \infty$ und damit

$$\lim_{n \rightarrow \infty} |\mathbf{E}[f(X_n)] - \mathbf{E}[f(X)]| \leq \varepsilon$$

für jedes $\varepsilon > 0$. Aus $\lim_{n \rightarrow \infty} |\mathbf{E}[f(X_n)] - \mathbf{E}[f(X)]| = 0$ aber folgt $P_{X_n} \xrightarrow[n \rightarrow \infty]{w} P_X$ und aufgrund von Satz 14.5 auch $F_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} F_X(x)$ für alle $x \in C(F_X)$. ■

14.8 Satz:

Es seien X und die Elemente der Folge $(X_n)_{n \in \mathbb{N}}$ reellwertige Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. Außerdem sei X P -f.s. konstant. Dann gilt:

$$X_n \xrightarrow[n \rightarrow \infty]{i.V.} X \Rightarrow X_n \xrightarrow[n \rightarrow \infty]{p} X.$$

Beweis:

Es wird $P(X = a) = 1$ mit $a \in \mathbb{R}$ angenommen. Für ein $\varepsilon > 0$ wird $f \in C_b$ mit $f \geq 0$, $f(a) = 0$ sowie $f(x) = 1 \ \forall x \in \mathbb{R}$ mit $|x - a| \geq \varepsilon$ gewählt. Dann gilt für alle $n \in \mathbb{N}$:

$$\begin{aligned} P(|X_n - X| \geq \varepsilon) &= P(|X_n - a| \geq \varepsilon) = \int_{\mathbb{R}} I_{\mathbb{R} \setminus (a-\varepsilon, a+\varepsilon)} dP_{X_n} \\ &\leq \int_{\mathbb{R}} f dP_{X_n}. \end{aligned}$$

Nach Voraussetzung gilt:

$$\int_{\mathbb{R}} f dP_{X_n} \xrightarrow[n \rightarrow \infty]{} \int_{\mathbb{R}} f dP_X = f(a) = 0.$$

Damit folgt

$$P(|X_n - X| \geq \varepsilon) \rightarrow 0 \text{ für } n \rightarrow \infty. \quad \blacksquare$$

14.2 Der Stetigkeitssatz von Lévy

Dieser Abschnitt beschäftigt sich mit dem Stetigkeitssatz von P. Lévy. Der Satz besagt, dass man Verteilungskonvergenz mit Hilfe charakteristischer Funktionen überprüfen kann. Zur Vorbereitung benötigt man einen Auswahlssatz von Helly. Der Auswahlssatz von Helly bezieht sich auf sogenannte beschränkte Folgen $(\mu_n)_{n \in \mathbb{N}}$ lokal endlicher Maße auf $(\mathbb{R}, \mathfrak{B})$. Darunter versteht man Folgen $(\mu_n)_{n \in \mathbb{N}}$ lokal endlicher Maße auf $(\mathbb{R}, \mathfrak{B})$ mit der Eigenschaft

$$\sup_{n \in \mathbb{N}} \mu_n(K) < \infty$$

für jede kompakte Menge $K \subset \mathbb{R}$. Zum Beweis des Auswahlssatzes von Helly werden zunächst weitere Kriterien für schwache bzw. vage Konvergenz zur Verfügung gestellt. Das Kriterium für schwache Konvergenz stellt eine Teilaussage des Portmanteau-Theorems dar.

14.9 Satz:

Für endliche Maße μ, μ_1, μ_2, \dots auf \mathbb{R} gilt

$$\mu_n \xrightarrow{w} \mu \quad \Leftrightarrow \quad \lim_{n \rightarrow \infty} \mu_n(C) = \mu(C) \text{ für alle } C \in \mathfrak{B} \text{ mit } \mu(\partial C) = 0.$$

Beweis:

Der Beweis ist technisch aufwendig. Daher soll hier nur auf den Beweis in „Wahrscheinlichkeitstheorie“ von P. Gänssler und W. Stute (Sätze 1.12.5 und 8.4.9) verwiesen werden.

14.10 Definition:

Ist μ ein Maß auf \mathbb{R} und $I \subseteq \mathbb{R}$ ein Intervall, so wird die Einschränkung von μ auf I definiert durch

$$\mu^I(A) = \mu(A \cap I) \quad (A \in \mathfrak{B}).$$

14.11 Satz:

Es seien μ, μ_1, μ_2, \dots lokal endliche Maße auf \mathbb{R} . Dann sind paarweise äquivalent:

- (a): $\mu_n \xrightarrow{v} \mu$.
- (b): $\mu_n^I \xrightarrow{w} \mu^I$ für alle beschränkten Intervalle I mit $\mu(\partial I) = 0$.
- (c): $\lim_{n \rightarrow \infty} \mu_n(C) = \mu(C)$ für alle beschränkten $C \in \mathfrak{B}$ mit $\mu(\partial C) = 0$.

Beweis:

Folgt aus Satz 14.9 und der Definition der vagen Konvergenz. ■

14.12 Satz (Auswahlsatz von Helly):

Jede beschränkte (und straffe) Folge $(\mu_n)_{n \in \mathbb{N}}$ lokaler endlicher Maße auf $(\mathbb{R}, \mathfrak{B})$ enthält eine vage (schwach) konvergente Teilfolge.

Beweis:

Es sei F_n maßdefinierende Funktion von μ_n mit $F_n(0) = 0$, d.h.

$$F_n(x) = \mu_n((0, x]) \text{ für } x > 0 \quad \text{und} \quad F_n(x) = -\mu_n((x, 0]) \text{ für } x < 0.$$

Da $(0, x]$ bzw. $(-x, 0]$ in kompakten Mengen enthalten sind, gilt wegen der Beschränktheit von $(\mu_n)_{n \in \mathbb{N}}$

$$M(x) := \sup_{n \in \mathbb{N}} F_n(x) < \infty \quad (\text{für alle } x \in \mathbb{R}).$$

Der Beweis gliedert sich in vier Schritte

- (i) Konstruktion einer Folge $(n_k)_{k \in \mathbb{N}}$ natürlicher Zahlen, so dass F_{n_k} auf einer dichten Teilmenge von \mathbb{R} für $k \rightarrow \infty$ gegen eine Funktion F konvergiert.
- (ii) Rechtsseitig stetige Fortsetzung von F auf \mathbb{R} zu einer maßdefinierenden Funktion des Grenzmaßes μ .
- (iii) Nachweis von $\mu_n \xrightarrow{v} \mu$.

(iv) Beweis der eingeklammerten Aussage.

Zu (i): Es wird eine abzählbare, aber dichte Teilmenge $T = \{x_1, x_2, x_3, \dots\} \subseteq \mathbb{R}$ (also etwa $T = \mathbb{Q}$) betrachtet. Die Menge $\{F_n(x_1) : n \in \mathbb{N}\}$ ist durch $M(x_1)$ beschränkt, es existiert also eine unendliche (geordnete) Teilmenge $I_1 \subseteq \mathbb{N}$ mit

$$F_n(x_1) \xrightarrow{n \rightarrow \infty, n \in I_1} z_1.$$

Genauso ist die Menge $\{F_n(x_2) : n \in I_1\}$ durch $M(x_2)$ beschränkt und daher existiert eine unendliche Teilmenge $I_2 \subseteq I_1 \setminus \{n_1\}$ für $n_1 := \min I_1$, so dass

$$F_n(x_2) \xrightarrow{n \rightarrow \infty, n \in I_2} z_2.$$

Dieses Verfahren kann fortgesetzt werden und es ergibt sich eine Folge von Mengen $(I_n)_{n \in \mathbb{N}}$ mit

$$\mathbb{N} \supset I_1 \supset I_2 \supset I_3 \supset \dots \quad \text{und} \quad n_1 = \min I_1 < n_2 = \min I_2 < n_3 = \min I_3 < \dots$$

und eine Folge $(z_j)_{j \in \mathbb{N}}$, so dass für alle $j \in \mathbb{N}$

$$F_n(x_j) \xrightarrow{n \rightarrow \infty, n \in I_j} z_j =: F(x_j)$$

gilt. Speziell ist dann auch für alle $j \in \mathbb{N}$

$$\lim_{k \rightarrow \infty} F_{n_k}(x_j) = F(x_j).$$

Zu (ii): Die $F_n(x)$ sind in x monoton nicht fallend, somit ist auch $F(x)$ auf T monoton wachsend. Daher kann $F(x)$ rechtsseitig stetig auf \mathbb{R} durch die Definition

$$F(x) := \inf\{\mu(\{t\}) : t \in T \cap [x, \infty)\}, \quad x \in \mathbb{R} \setminus T$$

zu einer maßdefinierenden Funktion fortgesetzt werden. Für alle Stetigkeitspunkte $x \in C(F)$ gilt dabei $F_{n_k}(x) \xrightarrow{k \rightarrow \infty} F(x)$, da die Differenz

$$|F_{n_k}(x) - F(x)| \leq |F_{n_k}(x) - F_{n_k}(x_j)| + |F_{n_k}(x_j) - F(x_j)| + |F(x_j) - F(x)|$$

für genügend große k und genügend nahe bei x gelegene $x_j \in T$, $x_j > x$ beliebig klein wird. Es sei nun $\mu := \mu_F$ das von F definierte Maß.

Zu (iii): Da $C(F)$ dicht in \mathbb{R} ist, gilt für $b \in C(F)$

$$\limsup_{k \rightarrow \infty} \mu_{n_k}(\{b\}) \leq \inf_{a \in C(F), a < b} \lim_{k \rightarrow \infty} (F_{n_k}(b) - F_{n_k}(a)) = \inf_{a \in C(F), a < b} (F(b) - F(a)) = 0,$$

also $\mu_{n_k}(\{b\}) \xrightarrow{k \rightarrow \infty} 0$.

Es sei nun I ein beliebiges beschränktes Intervall mit den Randpunkten $a < b$ und $\mu(\{a, b\}) = 0$. Dann ist für alle $t \in I \cap C(F)$

$$\mu_{n_k}((a, t] \cap I) = F_{n_k}(t) - F_{n_k}(a),$$

und mit $k \rightarrow \infty$ folgt

$$\mu^I((a, t]) = F(t) - F(a),$$

was nach Satz 14.5 $\mu_{n_k}^I \xrightarrow{w} \mu^I$ impliziert. Da I ein beliebiges beschränktes Intervall mit $\mu(\partial I) = 0$ war, folgt mit Satz 14.11

$$\mu_{n_k} \xrightarrow{v} \mu.$$

Zu (iv): Die eingeklammerte Aussage folgt aus dem bereits Bewiesenen und Satz 14.3. ■

14.13 Satz (Stetigkeitssatz von Lévy):

Es bezeichne $(X_n)_{n=1}^\infty$ eine Folge reellwertiger Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$ mit den Verteilungsfunktionen $(F_n)_{n \in \mathbb{N}}$ und den charakteristischen Funktionen $(\varphi_n)_{n \in \mathbb{N}}$. Dann gilt:

- a) Gilt $X_n \xrightarrow{i.V.} X$ und ist φ die charakteristische Funktion von X , so ist $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t) \quad \forall t \in \mathbb{R}$.
- b) Gilt $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t) \quad \forall t \in \mathbb{R}$ und ist φ in einer Umgebung von $t = 0$ stetig, dann konvergiert $(X_n)_{n \in \mathbb{N}}$ in Verteilung, d.h. es gibt eine Verteilungsfunktion F , so dass $\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad \forall x \in C(F)$ und φ die charakteristische Funktion bzgl. F ist.

Beweis:

- a) Zunächst wird angenommen, dass $X_n \xrightarrow{i.V.} X$ für $n \rightarrow \infty$ gilt. Da sowohl der Realteil als auch der Imaginärteil der Funktion $x \rightarrow e^{itx}$ für jedes $t \in \mathbb{R}$ beschränkte stetige Funktionen darstellen, impliziert $P_{X_n} \xrightarrow{w} P_X$:

$$\varphi_n(t) = \int_{\mathbb{R}} e^{itx} dP_{X_n}(x) \longrightarrow \int_{\mathbb{R}} e^{itx} dP_X(x) = \varphi(t) \quad (n \rightarrow \infty).$$

- b) Es sei nun $\lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t) \quad \forall t \in \mathbb{R}$ und φ sei in einer Umgebung von $t = 0$ stetig. Es wird zunächst gezeigt, dass dann die Folge der Wahrscheinlichkeitsverteilungen $(P_{X_n})_{n \in \mathbb{N}}$ straff ist (vgl. Definition 14.2). Allgemein gilt für eine reellwertige Zufallsvariable Y über $(\Omega, \mathfrak{F}, P)$ mit der charakteristischen Funktion φ_Y :

$$\begin{aligned} \frac{1}{u} \int_0^u [1 - \operatorname{Re} \varphi_Y(v)] dv &= \frac{1}{u} \int_0^u (1 - \mathbf{E}[\cos(vY)]) dv \\ &= \frac{1}{u} \int_0^u \left(1 - \int_{\mathbb{R}} \cos(vy) dP_Y(y) \right) dv \\ &= \frac{1}{u} \int_{\mathbb{R}} \int_0^u (1 - \cos(vy)) dv dP_Y(y) \quad (\text{Satz von Fubini}) \\ &= \frac{1}{u} \int_{\mathbb{R}} \left[u - \frac{1}{y} \sin(uy) \right] dP_Y(y) \\ &= \mathbf{E} \left[1 - \frac{\sin(uY)}{uY} \right]. \end{aligned}$$

Indem man $Z := uY$ und $\beta := \inf_{|x| \geq 1} (1 - (\sin x)/x)$ setzt, erhält man

$$\begin{aligned} \mathbf{E} \left[1 - \frac{\sin(uY)}{uY} \right] &= \mathbf{E} \left[1 - \frac{\sin Z}{Z} \right] = \int_{\mathbb{R}} \left(1 - \frac{\sin x}{x} \right) dP_Z(x) \\ &\geq \int_{|x| \geq 1} 1 - \frac{\sin x}{x} dP_Z(x) \geq \beta \int_{|x| \geq 1} dP_Z(x) = \beta \cdot P(|Z| \geq 1) \\ &= \beta \cdot P\left(|Y| \geq \frac{1}{u}\right). \end{aligned}$$

Folglich lässt sich mit Hilfe des Satzes von der majorisierten Konvergenz schließen:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \beta \cdot P\left(|X_n| \geq \frac{1}{u}\right) &\leq \frac{1}{u} \limsup_{n \rightarrow \infty} \int_0^u [1 - \operatorname{Re} \varphi_n(v)] dv \\ &= \frac{1}{u} \int_0^u [1 - \operatorname{Re} \varphi(v)] dv. \end{aligned}$$

Da $\lim_{u \rightarrow 0} \varphi(u) = 1$ ist (aufgrund der Stetigkeit von φ in einer Umgebung von $t = 0$), folgt mit Hilfe der Regel von l'Hospital:

$$\lim_{u \rightarrow 0} \limsup_{n \rightarrow \infty} P\left(|X_n| > \frac{1}{u}\right) = \lim_{u \rightarrow 0} \limsup_{n \rightarrow \infty} P_{X_n}\left(\mathbb{R} \setminus \left[-\frac{1}{u}, \frac{1}{u}\right]\right) = 0,$$

weshalb die Folge $(P_{X_n})_{n \in \mathbb{N}}$ straff ist. Nach dem Auswahlssatz von Helly enthält die Folge $(P_{X_n})_{n \in \mathbb{N}}$ eine schwach konvergente Teilfolge $(P_{X_{n_k}})_{k \in \mathbb{N}}$ mit $P_{X_{n_k}} \rightarrow P_X$ für $k \rightarrow \infty$. Deshalb gilt

$$\varphi(u) = \lim_{k \rightarrow \infty} \int_{\mathbb{R}} e^{iux} dP_{X_{n_k}}(x) = \int_{\mathbb{R}} e^{iux} dP_X(x),$$

so dass φ die charakteristische Funktion von P_X bzw. X ist. Da P_X durch φ eindeutig bestimmt ist, besitzt jede schwach konvergente Teilfolge von $(P_{X_n})_{n \in \mathbb{N}}$ denselben Grenzwert P_X . ■

Bemerkung:

Lässt man die Bedingung, dass φ in einer Umgebung von $t = 0$ stetig ist, fallen, so gilt die Aussage des Satzes bereits nicht mehr, wie nachfolgendes Beispiel zeigt.

14.14 Beispiel:

Es seien X_n Zufallsvariablen, die auf $[-n, n]$ gleichverteilt sind. Für $|x| \leq n$ gilt dann

$$F_n(x) = \int_{-n}^x \frac{1}{2n} dy = \frac{1}{2} + \frac{x}{2n}.$$

Die zugehörige charakteristische Funktion lautet:

$$\begin{aligned}\varphi_n(t) &= \int_{-n}^n \frac{1}{2n} e^{itx} dx \\ &= \frac{1}{2itn} (e^{itn} - e^{-itn}) \\ &= \frac{1}{2itn} (\cos(nt) + i \sin(nt) - \cos(nt) + i \sin(nt)) \\ &= \frac{1}{nt} \sin(nt).\end{aligned}$$

Führt man den Grenzübergang $n \rightarrow \infty$ durch, ergibt sich:

$$\varphi(t) = \begin{cases} 1 & , \quad t = 0 \\ 0 & , \quad t \neq 0. \end{cases}$$

Offensichtlich ist $\varphi(t)$ im Punkt $t = 0$ unstetig. Beim Grenzübergang $n \rightarrow \infty$ ergibt sich für F_n aber auch keine Verteilungsfunktion:

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} \left(\frac{1}{2} + \frac{x}{2n} \right) = \frac{1}{2}.$$

14.3 Der zentrale Grenzwertsatz

Der im Folgenden behandelte zentrale Grenzwertsatz trifft Aussagen über die Verteilung der Summe von unabhängigen und identisch verteilten Zufallsvariablen.

14.15 Satz (Zentraler Grenzwertsatz):

Lévy¹/Mises²

Es sei $(X_n)_{n \in \mathbb{N}}$ eine i.i.d. Folge von reellwertigen Zufallsvariablen über demselben Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit $0 < \mathbf{Var}[X_1] < \infty$. Dann gilt

$$P \left(\frac{S_n - n \cdot \mathbf{E}[X_1]}{\sigma[X_1] \cdot \sqrt{n}} \leq x \right) \xrightarrow{n \rightarrow \infty} \Phi(x) \quad \forall x \in \mathbb{R},$$

wobei $S_n := \sum_{i=1}^n X_i$, $\sigma[X_1] := \sqrt{\mathbf{Var}[X_1]}$ und $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung sei.

Beweis:

Es wird

$$Y_k := \frac{X_k - \mathbf{E}[X_k]}{\sigma[X_k]}$$

definiert. Offensichtlich ist $\mathbf{E}[Y_k] = 0$ und $\mathbf{Var}[Y_k] = \mathbf{E}[Y_k^2] = 1$. Mit Hilfe der Taylorentwicklung lässt sich die charakteristische Funktion von Y_k schreiben als (vgl. Satz 13.16 d))

$$\varphi_{Y_k}(t) = 1 + \frac{it\mathbf{E}[Y_k]}{1!} + \frac{(it)^2\mathbf{E}[Y_k^2]}{2!} + o(t^2) = 1 - \frac{t^2}{2} + o(t^2).$$

¹Paul Lévy, 15.09.1886 – 15.12.1971, Professor für Mathematik in Paris

²Richard von Mises, 19.04.1883 – 1953, studierte in Wien Mathematik und war später Professor für Mathematik an Universitäten in Berlin und Straßburg. Vor dem ersten Weltkrieg emigrierte er in die USA.

Es folgt

$$\varphi_n(t) = \varphi_{\sum_{k=1}^n Y_k}(t) = \left(1 - \frac{t^2}{2} + o(t^2)\right)^n$$

und weiterhin

$$\varphi_{\sum_{k=1}^n Y_k/\sqrt{n}}(t) = \varphi_n\left(\frac{t}{\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n.$$

Mit $e^x = \lim_{n \rightarrow \infty} (1 + \frac{x}{n})^n$ folgt

$$\varphi(t) := \lim_{n \rightarrow \infty} \varphi_n\left(\frac{t}{\sqrt{n}}\right) = e^{-\frac{1}{2}t^2},$$

d.h. $\varphi(t)$ ist die charakteristische Funktion einer standardnormalverteilten Zufallsvariable. Nach Satz 14.13 folgt damit die Behauptung. ■

Ein Spezialfall des zentralen Grenzwertsatzes ist der Satz von de Moivre–Laplace:

14.16 Satz (Satz von de Moivre–Laplace):

Es sei $(S_n)_{n \in \mathbb{N}}$ eine Folge von stochastisch unabhängigen Zufallsgrößen mit der Eigenschaft $S_n \stackrel{d}{=} \mathcal{B}(n, p) \quad \forall n \in \mathbb{N}, 0 < p < 1$. Dann gilt:

$$P\left(\frac{S_n - np}{\sqrt{n \cdot p \cdot (1-p)}} \leq x\right) \xrightarrow{n \rightarrow \infty} \Phi(x) \quad \forall x \in \mathbb{R},$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung bedeutet.

Beweis:

In Satz 14.15 wählen wir $(X_n)_{n \in \mathbb{N}}$ mit $P(X_n = 1) = p$ und $P(X_n = 0) = 1 - p \quad \forall n \in \mathbb{N}$. Dann gilt: $\mathbf{E}[X_n] = p$, $\sigma[X_n] = \sqrt{p(1-p)}$ und $S_n \stackrel{d}{=} \mathcal{B}(n, p) \quad \forall n \in \mathbb{N}$. ■

(Siehe auch Mathematica-Notebook zum zentralen Grenzwertsatz.)

Die Frage, wie schnell sich die Verteilung der Zufallsvariable $\frac{S_n - n\mathbf{E}[X_1]}{\sigma[X_1] \cdot \sqrt{n}}$ der Standard-Normalverteilung Φ nähert, beantwortet der Satz von Berry und Esseen.

14.17 Satz (Berry–Esseen):

Es sei $(X_k)_{k \in \mathbb{N}}$ eine i.i.d. Folge von reellwertigen Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. Ferner seien

$$\mu := \mathbf{E}[X_1], \quad \sigma^2 := \mathbf{Var}[X_1], \quad \gamma := \mathbf{E}\left[\left|\frac{X_1 - \mu}{\sigma}\right|^3\right] < \infty \quad \text{und} \quad S_n := \sum_{i=1}^n X_i.$$

Dann gilt

$$\sup_{x \in \mathbb{R}} \left| P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) - \Phi(x) \right| \leq \frac{C\gamma}{\sqrt{n}}$$

mit einer von n unabhängigen Konstanten C ($C \approx 2.8741$).

Für den Beweis des Satzes von Berry–Esseen benötigt man einige Hilfsmittel, die zunächst in Form der Sätze 14.19 bis 14.21 zur Verfügung gestellt werden.

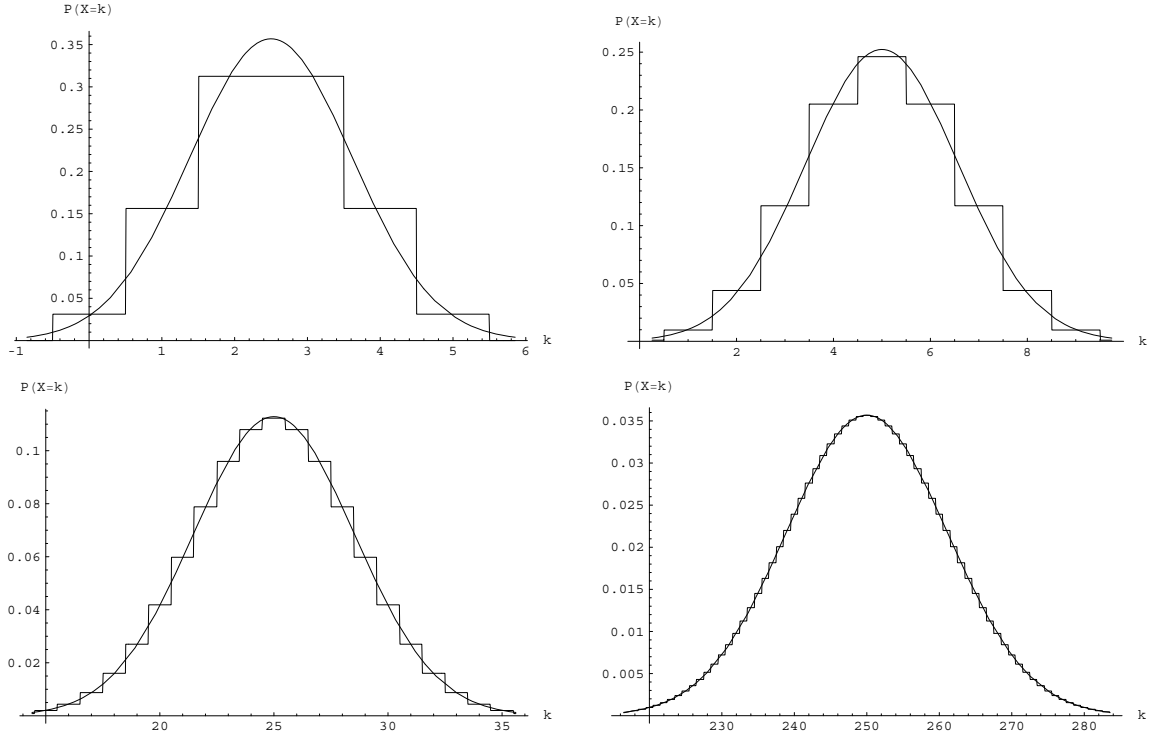


Abbildung 14.5: Binomialverteilung mit den Parametern $p = 0.5$ und $n = 5, 10, 50$ und 500 und jeweils überlagert von der Dichte einer Normalverteilung mit den Parametern $\mu = n \cdot p$ und $\sigma^2 = n \cdot p \cdot (1 - p)$

14.18 Satz:

Es sei F eine Verteilungsfunktion und G eine rellwertige differenzierbare Funktion mit $\lim_{x \rightarrow -\infty} G(x) = 0$, $\lim_{x \rightarrow +\infty} G(x) = 1$ und $\sup_{x \in \mathbb{R}} |G'(x)| \leq M > 0$. Dann existiert eine reelle Zahl a , so dass für alle $T > 0$:

$$\left| \int_{\mathbb{R}} \Delta(x+a) \frac{1 - \cos(Tx)}{Tx^2} dx \right| \geq 2M \cdot \delta \left[\frac{\pi}{2} - \frac{6}{T\delta} \right],$$

wobei $\Delta(x) := F(x) - G(x)$ und $\delta := \left(\frac{1}{2M}\right) \sup_{x \in \mathbb{R}} |F(x) - G(x)|$ sind.

Beweis:

Wir definieren $\alpha := \sup_{x \in \mathbb{R}} |F(x) - G(x)|$. Da $\Delta := F - G$ beschränkt ist und für $x \rightarrow \pm\infty$ verschwindet, existiert ein $b \in \mathbb{R}$ mit entweder $|\Delta(b-0)| = \alpha$ oder $|\Delta(b)| = \alpha$. Da F monoton nicht fallend ist, kommen nur die Fälle $F(b-0) - G(b) = -\alpha$ und $F(b) - G(b) = \alpha$ in Betracht. Es wird $F(b-0) - G(b) = -\alpha$ angenommen und $a := b - \delta$ gesetzt. Der Mittelwertsatz liefert für G und das Intervall $[a+x, b]$:

$$G(b) - G(a+x) = G'(\xi)[b - (a+x)] = G'(\xi)[b - (b - \delta + x)] = G'(\xi)(\delta - x) \leq M(\delta - x)$$

für ein $\xi \in (a+x, b)$. Folglich gilt für alle x mit $|x| < \delta$:

$$\begin{aligned}
 F(x+a) - G(x+a) &= F(x+a) + G(b) - G(a+x) - G(b) \\
 &\leq F(x+a) - G(b) + M(\delta-x) \\
 &\leq F(b-0) - G(b) + M(\delta-x) = -\alpha + M(\delta-x) \\
 &= -2M\delta + M(\delta-x) = -M(x+\delta).
 \end{aligned}$$

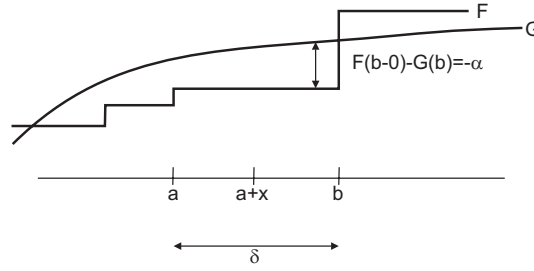


Abbildung 14.6: Darstellung von $\alpha = \sup_{x \in \mathbb{R}} |F(x) - G(x)|$, Fall 1

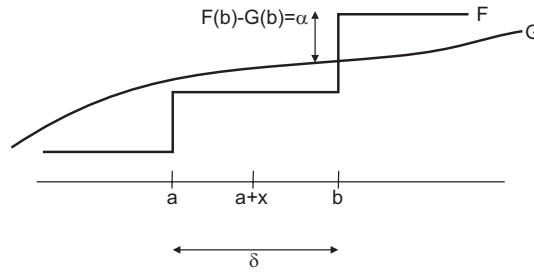


Abbildung 14.7: Darstellung von $\alpha = \sup_{x \in \mathbb{R}} |F(x) - G(x)|$, Fall 2

Damit folgt

$$\begin{aligned}
 \int_{-\delta}^{\delta} \Delta(x+a) \frac{1 - \cos(Tx)}{Tx^2} dx &\leq -M \int_{-\delta}^{\delta} (x+\delta) \frac{1 - \cos(Tx)}{Tx^2} dx \\
 &= \underbrace{-M \int_{-\delta}^{\delta} x \frac{1 - \cos(Tx)}{Tx^2} dx}_{=0 \text{ aus Symmetriegründen}} - M\delta \int_{-\delta}^{\delta} \frac{1 - \cos(Tx)}{Tx^2} dx \\
 &= -2M\delta \int_0^{\delta} \frac{1 - \cos(Tx)}{Tx^2} dx = -\alpha \int_0^{\delta} \frac{1 - \cos(Tx)}{Tx^2} dx. \quad (*)
 \end{aligned}$$

Aufgrund der Voraussetzungen gilt außerdem

$$\int_{|x|>\delta} \Delta(x+a) \frac{1 - \cos(Tx)}{Tx^2} dx \leq \alpha \int_{|x|>\delta} \frac{1 - \cos(Tx)}{Tx^2} dx. \quad (**)$$

Weiter ist

$$\begin{aligned}
 \int_{-\infty}^{\infty} \Delta(x+a) \cdot \frac{1 - \cos(Tx)}{Tx^2} dx &= \int_{-\delta}^{\delta} \Delta(x+a) \cdot \frac{1 - \cos(Tx)}{Tx^2} dx \\
 &\quad + \int_{|x|>\delta} \Delta(x+a) \cdot \frac{1 - \cos(Tx)}{Tx^2} dx \\
 &\leq \underbrace{\alpha \int_{|x|>\delta} \frac{1 - \cos(Tx)}{Tx^2} dx}_{\text{wegen (**)}} - \underbrace{\alpha \int_0^{\delta} \frac{1 - \cos(Tx)}{Tx^2} dx}_{\text{wegen (*)}} \\
 &= \alpha \left(2 \int_{\delta}^{\infty} \frac{1 - \cos(Tx)}{Tx^2} dx - \int_0^{\delta} \frac{1 - \cos(Tx)}{Tx^2} dx \right) \\
 &= \alpha \left(3 \int_{\delta}^{\infty} \frac{1 - \cos(Tx)}{Tx^2} dx - \int_0^{\infty} \frac{1 - \cos(Tx)}{Tx^2} dx \right) \\
 &= \alpha \frac{1}{T} \left(3 \int_{\delta T}^{\infty} \frac{1 - \cos z}{z^2} dz - \int_0^{\infty} \frac{1 - \cos z}{z^2} dz \right) \\
 &\leq \alpha \left(\frac{6}{\delta T} - \frac{\pi}{2} \right),
 \end{aligned}$$

wobei im letzten Schritt die Beziehungen

$$\int_0^{\infty} \frac{1 - \cos(z)}{z^2} dz = \frac{\pi}{2} \quad (\text{partielle Integration und Integral-Sinus})$$

und

$$\int_{\delta T}^{\infty} \frac{1 - \cos(z)}{z^2} dz \leq 2 \cdot \int_{\delta T}^{\infty} \frac{1}{z^2} dz = \frac{2}{\delta T}$$

benutzt wurden. Da der Ausdruck $\alpha \left(\frac{6}{\delta T} - \frac{\pi}{2} \right)$ für hinreichend großes T negativ ist, folgt

$$\left| \int_{\mathbb{R}} \Delta(x+a) \frac{1 - \cos(Tx)}{Tx^2} dx \right| \geq -\alpha \left(\frac{6}{\delta T} - \frac{\pi}{2} \right) = 2M \cdot \delta \left[\frac{6}{\delta T} - \frac{\pi}{2} \right].$$

Der Fall $F(b) - G(b) = \alpha$ wird analog behandelt. ■

14.19 Satz (Esseen):

Es seien X und Y reellwertige Zufallsvariablen über einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit den Verteilungsfunktionen F_X und F_Y , den charakteristischen Funktionen φ_X und φ_Y und den Erwartungswerten $\mathbf{E}[X] < \infty$ und $\mathbf{E}[Y] < \infty$. Ferner sei F_Y stetig

differenzierbar auf \mathbb{R} mit $\sup_{x \in \mathbb{R}} |F'_Y(x)| \leq M > 0$. Dann gilt für alle $T > 0$:

$$\sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)| \leq \frac{2}{\pi} \int_0^T \frac{|\varphi_X(u) - \varphi_Y(u)|}{u} du + \frac{24M}{\pi T}.$$

Beweis:

Es gilt für eine reellwertige Zufallsvariable Z mit Verteilungsfunktion F_Z und Erwartungswert $\mathbf{E}[Z] < \infty$

$$\mathbf{E}[Z] = \int_{\mathbb{R}^+} (1 - F_Z(x)) dx - \int_{\mathbb{R}^-} F_Z(x) dx.$$

Wegen der Abschätzungen

$$|F_X(x) - F_Y(x)| \leq |F_X(x)| + |F_Y(x)| \quad \text{für } x < 0$$

und

$$|F_X(x) - F_Y(x)| = |1 - F_Y(x) - (1 - F_X(x))| \leq |1 - F_Y(x)| + |1 - F_X(x)| \quad \text{für } x \geq 0$$

ist deshalb

$$\int_{\mathbb{R}} |F_X(x) - F_Y(x)| dx < \infty.$$

Deswegen kann der Ausdruck $\varphi_X(t) - \varphi_Y(t)$ partiell integriert werden:

$$\begin{aligned} \varphi_X(t) - \varphi_Y(t) &= \mathbf{E}[e^{itX}] - \mathbf{E}[e^{itY}] = \int_{\Omega} e^{itX} dP - \int_{\Omega} e^{itY} dP \\ &= \int_{\mathbb{R}} e^{itx} dP_X(x) - \int_{\mathbb{R}} e^{itx} dP_Y(x) = \int_{\mathbb{R}} e^{itx} d(P_X(x) - P_Y(x)) \\ &= -it \int_{\mathbb{R}} [F_X(x) - F_Y(x)] e^{itx} dx. \quad (\text{Übung}) \end{aligned}$$

Hieraus folgt mit $\Delta(x) := F_X(x) - F_Y(x)$ und $a \in \mathbb{R}$:

$$\begin{aligned} \frac{\varphi_X(t) - \varphi_Y(t)}{-it} e^{-ita} &= \int_{\mathbb{R}} \Delta(x) e^{it(x-a)} dx \\ &= \int_{\mathbb{R}} \Delta(y+a) e^{ity} dy. \end{aligned}$$

Unter Zuhilfenahme des Satzes von Fubini lässt sich weiter schließen:

$$\begin{aligned}
 \frac{1}{2\pi} \int_{-T}^T \frac{\varphi_X(t) - \varphi_Y(t)}{-it} e^{-ita} \left(1 - \frac{|t|}{T}\right) dt &= \frac{1}{2\pi} \int_{-T}^T \int_{\mathbb{R}} \Delta(y+a) e^{ity} \left(1 - \frac{|t|}{T}\right) dy dt \\
 &= \int_{\mathbb{R}} \int_{-T}^T \frac{1}{2\pi} e^{ity} \left(1 - \frac{|t|}{T}\right) \Delta(y+a) dt dy \\
 &= \int_{\mathbb{R}} \frac{1 - \cos(Ty)}{\pi T y^2} \Delta(y+a) dy,
 \end{aligned}$$

wobei im letzten Schritt das Ergebnis aus Beispiel 13.6 Nr.8 benutzt wurde. Die linke Seite dieser Formel stellt offensichtlich die Ableitung von $(F_X * G)(a) - (F_Y * G)(a)$ dar, wobei G die Verteilung mit der charakteristischen Funktion

$$\varphi(t) = \left(1 - \frac{|t|}{T}\right) \cdot I_{[-T, T]}(t), \quad t \in \mathbb{R},$$

und der Dichte

$$f(x) = \frac{1 - \cos(Tx)}{\pi x^2 T}, \quad x \in \mathbb{R},$$

ist (siehe Beispiel 13.15). Durch die Faltung von $F_X - F_Y$ mit der auf $[-T, T]$ konzentrierten Verteilung G wird gewissermaßen eine Glättung von F_X und F_Y herbeigeführt, weshalb Satz 14.19 oft auch als „Glättungslemma“ bezeichnet wird. Die spezielle Form von G ist dabei unerheblich, wesentlich für das Weitere ist, dass φ für $|t| > T$ verschwindet.

Aufgrund von Satz 14.18 gilt nun:

$$\begin{aligned}
 \alpha \left(\frac{\pi}{2} - \frac{6}{\delta T} \right) &\leq \left| \int_{\mathbb{R}} \Delta(y+a) \cdot \frac{1 - \cos(Ty)}{Ty^2} dy \right| \\
 &= \frac{1}{2} \left| \int_{-T}^T \frac{\varphi_X(u) - \varphi_Y(u)}{-iu} e^{-iua} \left(1 - \frac{|u|}{T}\right) du \right| \\
 &\leq \frac{1}{2} \int_{-T}^T \left| \frac{\varphi_X(u) - \varphi_Y(u)}{u} \right| du = \int_0^T \frac{|\varphi_X(u) - \varphi_Y(u)|}{u} du,
 \end{aligned}$$

woraus mit $\delta := \frac{\alpha}{2M}$ die Ungleichung

$$\frac{2}{\pi} \int_0^T \frac{|\varphi_X(u) - \varphi_Y(u)|}{u} du + \frac{24M}{\pi T} \geq \alpha$$

folgt. ■

14.20 Satz:

Für komplexe Zahlen x, y gilt:

$$|x^n - y^n| \leq n |x - y| a^{n-1},$$

mit $a := \max(|x|, |y|)$.

Beweis:

Aus der Identität

$$x^n - y^n = (x - y)x^{n-1} + y(x - y)x^{n-2} + \dots + y^{n-1}(x - y)$$

folgt sofort

$$\begin{aligned} |x^n - y^n| &\leq |(x - y)x^{n-1}| + |y(x - y)x^{n-2}| + \dots + |y^{n-1}(x - y)| \\ &\leq n \cdot |x - y| \cdot a^{n-1}. \end{aligned}$$

■

14.21 Satz:

Sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $X: \Omega \rightarrow \mathbb{R}$ eine reellwertige Zufallsvariable mit $\mathbf{E}[X] = 0$, $\mathbf{E}[X^2] = 1$ und $\mathbf{E}[|X|^3] < \infty$. Dann gilt für die charakteristische Funktion φ_X von X :

$$\left| \varphi_X^n \left(\frac{t}{\sqrt{n}} \right) - e^{-\frac{1}{2}t^2} \right| \leq \frac{\mathbf{E}[|X|^3] \cdot |t|^3}{24\sqrt{n}} \left(\frac{3}{\sqrt{n}} |t| + 4 \right) \exp \left(-\frac{5}{18} \frac{n-1}{n} t^2 \right)$$

für alle $|t| \leq \frac{4}{3\mathbf{E}[|X|^3]} \sqrt{n}$. Ist $n \geq 10$ (und $|t| \leq \frac{4}{3\mathbf{E}[|X|^3]} \sqrt{n}$) so gilt

$$\left| \varphi_X^n \left(\frac{t}{\sqrt{n}} \right) - e^{-\frac{1}{2}t^2} \right| \leq \frac{\mathbf{E}[|X|^3] \cdot |t|^3}{24\sqrt{n}} \cdot (|t| + 4) \cdot \exp \left(-\frac{1}{4}t^2 \right).$$

Beweis:

Satz 13.16 impliziert mit $\mathbf{E}[X] = 0$, $\mathbf{E}[X^2] = 1$ und $k := 2$, $\delta := 1$:

$$\left| \varphi_X(t) - 1 + \frac{1}{2}t^2 \right| \leq \frac{1}{6} \cdot \mathbf{E}[|X|^3] \cdot |t|^3.$$

Wegen $\mathbf{E}[X^2] = 1$ ist $\mathbf{E}[|X|^3] \geq 1$. Folglich ist $\frac{4}{3\mathbf{E}[|X|^3]} < \sqrt{2}$. Mit Hilfe der Ungleichungen

$$1 - x \leq e^{-x} \leq 1 - x + \frac{1}{2}x^2 \quad \forall x > 0$$

schließt man, dass für $|t| \leq \frac{4}{3\mathbf{E}[|X|^3]}$ gilt

$$|\varphi_X(t)| \leq 1 - \frac{1}{2}t^2 + \frac{1}{6} \cdot \mathbf{E}[|X|^3] |t|^3 = 1 - \frac{1}{2}t^2 + \frac{1}{6} \cdot \mathbf{E}[|X|^3] \cdot |t| \cdot |t|^2 \leq 1 - \frac{5}{18}t^2 \leq e^{-\frac{5}{18}t^2} \quad (14.1)$$

und

$$\left| \varphi_X(t) - e^{-\frac{1}{2}t^2} \right| \leq \left| \varphi_X(t) - 1 + \frac{1}{2}t^2 \right| + \left| e^{-\frac{1}{2}t^2} - 1 + \frac{1}{2}t^2 \right| \quad (14.2)$$

$$\leq \frac{1}{6} \cdot \mathbf{E}[|X|^3] \cdot |t|^3 + \frac{1}{8}t^4. \quad (14.3)$$

Wenn man Satz 14.20 auf $x := \varphi_X\left(\frac{t}{\sqrt{n}}\right)$ und $y = \exp\left(-\frac{1}{2}\frac{t^2}{n}\right)$ anwendet, erhält man

$$\left| \varphi_X^n\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{1}{2}t^2} \right| \leq n \left| \varphi_X\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{1}{2}\frac{t^2}{n}} \right| \cdot \left(\max\left(\left| \varphi_X\left(\frac{t}{\sqrt{n}}\right) \right|, \left| e^{-\frac{1}{2}\frac{t^2}{n}} \right| \right) \right)^{n-1}$$

Aufgrund von (14.1) gilt:

$$\max\left(\left| \varphi_X\left(\frac{t}{\sqrt{n}}\right) \right|, \left| e^{-\frac{1}{2}\frac{t^2}{n}} \right| \right) \leq e^{-\frac{5}{18}\frac{t^2}{n}}$$

Wir fassen zusammen und erhalten aufgrund von (14.1) und (14.2):

$$\left| \varphi_X^n\left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{1}{2}t^2} \right| \leq n \cdot \left[\frac{\mathbf{E}[|X|^3]}{6} \left| \frac{t}{\sqrt{n}} \right|^3 + \frac{1}{8} \left(\frac{t}{\sqrt{n}} \right)^4 \right] \cdot e^{-\frac{5}{18}\frac{n-1}{n}t^2} \quad (14.4)$$

$$= |t|^3 \cdot \frac{\mathbf{E}[|X|^3]}{24\sqrt{n}} \left(4 + \frac{3}{\sqrt{n}} |t| \right) \cdot e^{-\frac{5}{18}\frac{n-1}{n}t^2}. \quad (14.5)$$

Wegen $\frac{3}{\sqrt{10}} < 1$ ist für $n \geq 10$ (und $|t| \leq \frac{4\sqrt{n}}{3\mathbf{E}[|X|^3]}$) offensichtlich

$$|t|^3 \cdot \frac{\mathbf{E}[|X|^3]}{24\sqrt{n}} (|t| + 4) \cdot e^{-\frac{1}{4}t^2}$$

eine obere Schranke für den Ausdruck (14.4), womit alles gezeigt ist. ■

Mit Hilfe dieser Sätze lässt sich nun der Satz 14.17 von Berry und Esseen beweisen.

Beweis (des Satzes von Berry-Esseen):

Ohne Beschränkung der Allgemeinheit kann für die Zufallsvariablen in Satz 14.17 $\mathbf{E}[X_1] = 0$ und $\mathbf{Var}[X_1] = \mathbf{E}[X_1^2] = 1$ angenommen werden. Andernfalls ersetzt man die Zufallsvariablen X_k durch die Variablen $X'_k := \frac{X_k - \mathbf{E}[X_k]}{\sqrt{\mathbf{Var}[X_k]}}$, $k = 1, 2, \dots$. Um zu der Abschätzung aus Satz 14.17

zu gelangen, wird Satz 14.19 mit $X := S_n := \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$ und $Y \stackrel{d}{=} \mathcal{N}(0, 1)$ angewandt. Die dazugehörigen charakteristischen Funktionen lauten

$$\varphi_X(t) = \varphi_{\frac{\sum_{i=1}^n X_i}{\sqrt{n}}}(t) = \varphi_{X_1}^n\left(\frac{t}{\sqrt{n}}\right), \quad t \in \mathbb{R}$$

und

$$\varphi_Y(t) = e^{-\frac{1}{2}t^2}, \quad t \in \mathbb{R}.$$

Da

$$F_Y'(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}, \quad t \in \mathbb{R},$$

ist, kann $M := \frac{1}{\sqrt{2\pi}}$ gewählt werden. Da $\mathbf{E}[|X|^3] \geq 1$ gilt und die Abschätzung von Berry-Esseen für $\sqrt{n} \leq 3$ trivialerweise richtig ist, kann $n \geq 10$ angenommen werden und die zweite Abschätzung aus Satz 14.21 verwendet werden.

$$\begin{aligned} \sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)| &\leq \frac{2}{\pi} \int_0^T \frac{|\varphi_X(u) - \varphi_Y(u)|}{u} du + \frac{24M}{\pi T} \\ &= \frac{2}{\pi} \int_0^T \frac{\left| \varphi_{X_1}\left(\frac{u}{\sqrt{n}}\right) - e^{-\frac{1}{2}u^2} \right|}{u} du + \frac{24M}{\pi T} \\ &\leq \frac{2}{\pi} \int_0^T \frac{\mathbf{E}[|X|^3]}{24\sqrt{n}} \cdot \frac{(|u| + 4)e^{-\frac{1}{4}u^2}}{u} du + \frac{24M}{\pi T}. \end{aligned}$$

Mit $T := \frac{4}{3\mathbf{E}[|X|^3]} \sqrt{n}$ und $M := \frac{1}{\sqrt{2\pi}}$ erhält man schließlich

$$\begin{aligned} \sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)| &\leq \frac{2}{\pi} \int_0^T \frac{\mathbf{E}[|X|^3]}{24\sqrt{n}} \cdot \frac{(|u| + 4)e^{-\frac{1}{4}u^2}}{u} du + \frac{24 \cdot 3 \cdot \mathbf{E}[|X|^3]}{\pi \sqrt{2\pi} \cdot 4 \cdot \sqrt{n}} \\ &= \frac{\mathbf{E}[|X|^3]}{\pi \sqrt{n}} \left[\frac{1}{12} \int_0^T (u^3 + 4u^2) e^{-\frac{1}{4}u^2} du + 9\sqrt{\frac{2}{\pi}} \right] \\ &\leq \frac{\mathbf{E}[|X|^3]}{\pi \sqrt{n}} \left[\frac{1}{12} \int_0^\infty (u^3 + 4u^2) e^{-\frac{1}{4}u^2} du + 9\sqrt{\frac{2}{\pi}} \right] \\ &= \frac{\mathbf{E}[|X|^3]}{\pi \sqrt{n}} \left[\frac{1}{12} \cdot 8 \cdot (1 + \sqrt{\pi}) + 9\sqrt{\frac{2}{\pi}} \right] \\ &= \frac{\mathbf{E}[|X|^3]}{\sqrt{n}} \cdot C, \end{aligned}$$

wobei $C := \frac{2}{3} \cdot \frac{1+\sqrt{\pi}}{\pi} + \frac{9}{\pi} \sqrt{\frac{2}{\pi}} \approx 2.874103874$ ist. ■

14.22 Bemerkung:

Auch wenn die Berry-Esseen-Konstante C durch entsprechende Rechnungen verbessert werden kann (vgl. z.B. Gänsler/Stute, „Wahrscheinlichkeitstheorie“), ist eine Verbesserung der Konvergenzordnung $\frac{1}{\sqrt{n}}$ ohne zusätzliche Voraussetzungen nicht möglich. Man betrachte dazu eine i.i.d. Folge $(X_n)_{n \in \mathbb{N}}$ reellwertiger Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$ mit $P(X_n = 1) = P(X_n = -1) = \frac{1}{2}$. Für diesen Fall gilt:

$$P(S_{2n} = 0) = P\left(\sum_{i=1}^{2n} X_i = 0\right) = \binom{2n}{n} \left(\frac{1}{2}\right)^{2n}, \quad n \in \mathbb{N}.$$

Die Stirling'sche Formel besagt

$$\binom{2n}{n} \left(\frac{1}{2}\right)^{2n} \sim \frac{1}{\sqrt{n\pi}} \quad \text{für } n \rightarrow \infty,$$

so dass aufgrund der Symmetrie von S_{2n}

$$P(S_{2n} \leq 0) = \frac{1}{2}(1 + P(S_{2n} = 0)) = \frac{1}{2} + \frac{1}{2\sqrt{n\pi}} + o\left(n^{-\frac{1}{2}}\right) \quad \text{für } n \rightarrow \infty$$

folgt. Damit wird

$$\sup_{x \in \mathbb{R}} \left| P(S_{2n} \leq (2n)^{\frac{1}{2}}x) - \Phi(x) \right| \geq \left| P(S_{2n} \leq 0) - \frac{1}{2} \right| \sim \frac{1}{2\sqrt{n\pi}} \quad \text{für } n \rightarrow \infty.$$

Literatur zu Kapitel 14

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- G. ALSMEYER:
Erneuerungstheorie,
Teubner-Verlag, Stuttgart, 1991.
ISBN: 3519027305
 - G. ALSMEYER:
Wahrscheinlichkeitstheorie,
Skripten zur Mathematischen Statistik, Nr. 30, Institut für Mathematische Statistik der
Westfälischen Wilhelms-Universität Münster, 3. Auflage 2003.
 - G. AUMANN/O. HAUPT:
Einführung in die reelle Analysis, Band I,
de Gruyter, Bln., 1981.
ISBN: 3110019701
 - M. A. BERGER:
An Introduction to Probability and Stochastic Processes,
Springer-Verlag, New York, 1992.
ISBN: 3540977848
 - P. BILLINGSLEY:
Probability and Measure,
2nd edition, John Wiley and Sons, New York, 1986.
ISBN: 0471007102
 - Y.S. CHOW/H. TEICHER:
Probability Theory - Independence, Interchangeability, Martingales,
Springer-Verlag, New York, 1988.
ISBN: 0387406077
-

- P. GÄNSSLER UND W. STUTE:
Wahrscheinlichkeitstheorie,
Springer–Verlag, Berlin, 1977.
ISBN: 3540084185
- E. HENZE:
Einführung in die Maßtheorie,
Bibl. Institut, Mannheim, 1971.
ISBN: 341100505X
- R. MATHAR/D. PFEIFFER:
Stochastik für Informatiker,
Teubner, 1990.
ISBN: 3519022400
- P. P. SPIES:
Grundlagen stochastischer Modelle,
Hanser, München, 1982.
ISBN: 3446137114

Kapitel 15

Statistische Auswertung von Zufallsexperimenten

Die Statistik beschäftigt sich mit dem Schätzen und Testen von Parametern stochastischer Modelle. Oft existiert eine Messreihe und man interessiert sich z.B. für den Erwartungswert der Grundgesamtheit. Da sich aus der Messreihe zunächst nur der empirische Mittelwert berechnen lässt, stellt sich z.B. die Frage, wie groß die Diskrepanz zwischen dem theoretischen Wert und dem Schätzwert ist. Diese und ähnliche Fragestellungen sind Gegenstand der mathematischen Statistik.

Schlüsselwörter: Statistisches Modell, Stichprobe, (Punkt-)Schätzer, erwartungstreu, Konsistenz, starke Konsistenz, verzerrt, unverzerrt, Bias, gleichmäßig bester Schätzer, Cramér-Rao-Schranke, α -Quantil, Konfidenzintervall, Hypothesentest.

15.1 Einleitung

Allgemein hat man es bei der statistischen Auswertung eines Zufallsexperiments mit folgender Situation zu tun: In einem System, das durch einen Maßraum $(\Omega, \mathfrak{F}, P)$ beschrieben wird, wird ein Experiment durchgeführt. Die erhobenen Daten betrachtet man als eine Realisierung einer zum Experiment gehörenden Zufallsvariable $X: \Omega \rightarrow \mathfrak{X}$, wobei $(\mathfrak{X}, \mathfrak{A})$ ein geeigneter Messraum ist. Es wird angenommen, dass die unbekannte Verteilung P_X von X einer gewissen Familie $\mathfrak{P} := (P_\theta)_{\theta \in \Theta}$ von Verteilungen auf $(\mathfrak{X}, \mathfrak{A})$ angehört, d.h. es gibt ein $\theta_0 \in \Theta$, für das P_{θ_0} das Bildmaß P_X von P unter X ist. Die explizite Gestalt von $(\Omega, \mathfrak{F}, P)$ tritt dabei gänzlich in den Hintergrund und man bezeichnet $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta \in \Theta})$ als ein statistisches Modell für die Zufallsvariable X .

15.1 Beispiel:

In Kapitel 4 hatten wir uns im Rahmen der statistischen Qualitätssicherung mit der sogenannten Gut–Schlecht–Prüfung beschäftigt. Es wird ein Los vom Umfang N angeliefert und eine Stichprobe vom Umfang n entnommen. Wir vereinbaren

$$X_i := \begin{cases} 1 & \text{falls } i\text{-tes Stück fehlerhaft} \\ 0 & \text{falls } i\text{-tes Stück fehlerfrei} \end{cases} \quad (i = 1, \dots, n)$$

und setzen $X := (X_1, \dots, X_n)$. Unter der Annahme, dass die X_i stochastisch unabhängig und identisch $B(1, \theta)$ -verteilt sind mit unbekanntem Parameter $\theta \in [0, 1]$, genügt X einer $\otimes_{i=1}^{100} B(1, \theta)$ -Verteilung. Damit erhalten wir als statistisches Modell für X :

$$(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta \in \Theta}) = \left(\{0, 1\}^n, \mathfrak{B}(\{0, 1\}^n), (\otimes_{i=1}^{100} B(1, \theta))_{\theta \in [0, 1]} \right).$$

Man könnte aber auch ebensogut die Fehleranzahl

$$Y := \sum_{i=1}^n X_i$$

beobachten. Da $Y \stackrel{d}{=} B(n, \theta)$, lautet das statistische Modell für Y :

$$(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta \in \Theta}) = (\{0, 1, \dots, n\}, \mathfrak{B}(\{0, 1, \dots, n\}), (B(n, \theta))_{\theta \in [0, 1]}).$$

In der Regel hat man es mit einem Einzelexperiment zu tun, das n -mal unter gleichwertigen Bedingungen durchgeführt wird, wobei die Daten x_1, \dots, x_n erhoben werden. Diese Daten werden als Realisierung stochastisch unabhängiger und identisch verteilter Zufallsvariablen $X_1, \dots, X_n: \Omega \rightarrow \mathfrak{X}$ aufgefasst. In diesem Fall gilt $P_{X_1} = \dots = P_{X_n}$ und das statistische Modell für $X := (X_1, \dots, X_n)$ hat die Form $(\mathfrak{X}^n, \mathfrak{A}^{\otimes n}, (P_\theta^{\otimes n})_{\theta \in \Theta})$, wobei $\mathfrak{A}^{\otimes n}$ und $P_\theta^{\otimes n}$ die n -fache Produkt- σ -Algebra bzw. das n -fache Produktmaß von $P_\theta = P_{X_1}$ bezeichnen. Die Zufallsvariable $X := (X_1, \dots, X_n)$ heißt dann Stichprobe vom Umfang n .

15.2 Punktschätzer

Mit Hilfe der beobachteten Daten $X(\omega) = x = (x_1, \dots, x_n)$ sollen jetzt Aussagen über die zugehörige Verteilung, d.h. über den Parameter $\theta_0 \in \Theta$ oder eine von θ_0 abhängige Kenngröße $\varrho(\theta_0)$ getroffen werden.

15.2 Beispiel:

Sei $X \stackrel{d}{=} \text{Exp}(\lambda)$. In diesem Fall ist $\Theta = (0, \infty)$ und $\theta = \lambda$. Als $\varrho(\theta)$ kommen z.B. $\varrho(\theta) := \mathbf{E}[X] = \theta^{-1}$ oder $\varrho(\theta) := \mathbf{Var}[X] = \theta^{-2}$ in Betracht.

Sei $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$. In diesem Fall ist $\Theta = \mathbb{R} \times (0, \infty)$ und $\theta = (\mu, \sigma)$. $\varrho(\theta)$ könnte z.B. der Variationskoeffizient sein, d.h. $\varrho(\theta) := \frac{\sigma}{\mu}$.

In diesem Zusammenhang von Bedeutung ist der Begriff des Schätzers:

15.3 Definition (Schätzer, Schätzwert):

Es sei $(\mathfrak{X}, \mathfrak{A}, (P_\theta)_{\theta \in \Theta})$ ein statistisches Modell und $\varrho: \Theta \rightarrow \mathbb{R}^\ell$, $\ell \in \mathbb{N}$, eine Abbildung. Dann heißt jede $\mathfrak{A} - \mathfrak{B}^\ell$ -messbare Abbildung $T: \mathfrak{X} \rightarrow \mathbb{R}^\ell$ ein (Punkt-)Schätzer für die Kenngröße $\varrho(\theta) \in \mathbb{R}^\ell$. Die Werte $T(x)$ von T heißen Schätzwerte für $\varrho(\theta)$.

Um sinnvolle Schätzer $\varrho(\theta)$ zu vermeiden, werden an T gewisse Anforderungen gestellt.

1. Forderung:

Für einen hinreichend großen Stichprobenumfang n sollte $T_n = T(X_1, \dots, X_n)$ den gesuchten Wert $\varrho(\theta)$ möglichst gut approximieren. Deswegen definiert man:

15.4 Definition (schwach konsistent, stark konsistent):

Eine Folge $(T_n)_{n \in \mathbb{N}}$ von Punktschätzern für $\varrho(\theta)$ heißt schwach konsistent, wenn

$$T_n \xrightarrow[n \rightarrow \infty]{P_\theta} \varrho(\theta) \quad \forall \theta \in \Theta.$$

Eine Folge $(T_n)_{n \in \mathbb{N}}$ von Punktschätzern heißt stark konsistent, wenn

$$T_n \xrightarrow{P_\theta\text{-f.s.}} \varrho(\theta) \quad \forall \theta \in \Theta.$$

15.5 Definition (Stichprobenmittelwert, Stichprobenmomente):

Falls $X := (X_1, \dots, X_n)$ eine Stichprobe vom Umfang n ist, nennt man

$$\bar{X} := \bar{X}_n := \frac{1}{n}(X_1 + \dots + X_n)$$

den Stichprobenmittelwert,

$$\begin{aligned} S^2 := S_n^2 &:= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - 2X_i \bar{X}_n + \bar{X}_n^2) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\bar{X}_n \underbrace{\frac{1}{n} \sum_{i=1}^n X_i}_{=\bar{X}_n} + \frac{1}{n} \cdot n \cdot \bar{X}_n^2 = \frac{1}{n} \left(\sum_{i=1}^n X_i^2 \right) - \bar{X}_n^2 \end{aligned}$$

die Stichprobenvarianz und

$$m_n^{(k)} := \frac{1}{n} \sum_{i=1}^n X_i^k$$

das k -te Stichprobenmoment, $k \in \mathbb{N}$. Offensichtlich ist $m_n^{(1)} = \bar{X}_n$.

15.6 Bemerkung:

Mit Hilfe des starken Gesetzes der großen Zahlen kann man sofort schließen, dass die Folgen $(\bar{X}_n)_{n \in \mathbb{N}}$, $(S_n^2)_{n \in \mathbb{N}}$ und $(m_n^{(k)})_{n \in \mathbb{N}}$ stark konsistente Folgen von Punktschätzern für $\mathbf{E}[X_1]$, $\mathbf{Var}[X_1]$ und $\mathbf{E}[X_1^k]$ darstellen (sofern die entsprechenden Momente existieren, d.h. sofern $\mathbf{E}[|X_1|] < \infty$, $\mathbf{Var}[X_1] < \infty$ bzw. $\mathbf{E}[|X_1^k|] < \infty$ gilt).

2. Forderung:

Bei hinreichend oftmaliger unabhängiger Anwendung des Schätzers T soll der Durchschnitt der Schätzungen den gesamten Wert $\varrho(\theta)$ beliebig gut approximieren.

15.7 Definition (erwartungstreu, unverzerrt, verzerrt, Bias, Verzerrung):

Ein Punktschätzer $T: \mathfrak{X} \rightarrow \mathbb{R}^\ell$ für die Kenngröße $\varrho(\theta) \in \mathbb{R}^\ell$ heißt erwartungstreu oder unverzerrt, wenn gilt

$$\mathbf{E}_\theta [T(X)] = \int_{\mathbb{R}^n} T(x_1, \dots, x_n) dP_\theta(x_1, \dots, x_n) = \varrho(\theta) \quad \forall \theta \in \Theta$$

Andernfalls heißt T verzerrt (englisch: „biased“).

$$b(\theta, T) := \mathbf{E}_\theta [T(X)] - \varrho(\theta)$$

heißt der Bias (oder die Verzerrung) von T .

15.8 Bemerkung:

Ist $b(\theta, T) > 0$, so wird $\varrho(\theta)$ durch T im Mittel überschätzt, andernfalls unterschätzt.

15.9 Beispiel:

1. Es sei $X := (X_1, \dots, X_n)$ eine Stichprobe vom Umfang n und $\mathbf{E}[|X_1|] < \infty$. Dann ist \bar{X}_n ein erwartungstreuer Schätzer für $\mu := \mathbf{E}[X_1]$, denn es gilt:

$$\mathbf{E}[\bar{X}_n] = \mathbf{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \mathbf{E}\left[\sum_{j=1}^n X_j\right] = \frac{1}{n} (n \cdot \mu) = \mu.$$

2. Es sei $X := (X_1, \dots, X_n)$ eine Stichprobe vom Umfang n mit $\mathbf{E}[X_1] = \mu$ und $\mathbf{Var}[X_1] = \sigma^2 < \infty$. Dann ist S_n^2 ein verzerrter Schätzer für σ^2 .

$$\begin{aligned} \mathbf{E}[S_n^2] &= \mathbf{E}\left[\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2\right] = \frac{1}{n} \sum_{j=1}^n \left(\mathbf{E}[X_j^2 - \mu^2] - \mathbf{E}[\bar{X}_n^2 - \mu^2]\right) \\ &= \frac{1}{n} \sum_{j=1}^n \left(\mathbf{E}[X_j^2] \underbrace{- 2\mathbf{E}[X_j]\mu + \mu^2}_{=-\mu^2} - (\mathbf{E}[\bar{X}^2] \underbrace{- 2\mathbf{E}[\bar{X}]\mu + \mu^2}_{=-\mu^2})\right) \\ &= \frac{1}{n} \sum_{j=1}^n (\mathbf{E}[(X_j - \mu)^2] - \mathbf{E}[(\bar{X} - \mu)^2]) = \frac{1}{n} \cdot n (\sigma^2 - \mathbf{Var}[\bar{X}]) \\ &= \sigma^2 - \mathbf{Var}\left[\frac{1}{n} \sum_{j=1}^n X_j\right] = \sigma^2 - \frac{1}{n^2} \cdot n\sigma^2 \\ &= \sigma^2 - \frac{1}{n}\sigma^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2. \end{aligned}$$

Die Verzerrung ist in diesem Fall $b(\theta, T) = \mathbf{E}[S_n^2] - \sigma^2 = -\frac{1}{n}\sigma^2$ und legt folgende Modifikation nahe:

$$\tilde{S}_n^2 := \frac{n}{n-1} S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Man erkennt sofort, dass \tilde{S}_n^2 ein erwartungstreuer Punktschätzer für σ^2 ist.

3. Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum sowie $X := (X_1, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$ und $Y := (Y_1, \dots, Y_n): \Omega \rightarrow \mathbb{R}^n$ bezeichnen zwei Stichproben vom Umfang n (jeweils i.i.d.) mit $\mathbf{Var}[X_1] < \infty$ und $\mathbf{Var}[Y_1] < \infty$. Außerdem nehmen wir an, dass die Zufallsvariablen X_i und Y_j für $i \neq j$ stochastisch unabhängig sind. Dann heißt

$$C_n := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)$$

die Stichprobenkovarianz von X und Y .

15.10 Satz:

Die in Beispiel 15.9 beschriebene Stichprobenkovarianz ist ein erwartungstreuer Schätzer für $\mathbf{COV}_\theta[X_1, Y_1]$.

Beweis:

Als statistisches Modell für (X, Y) verwenden wir $((\mathbb{R}^2)^n, (\mathfrak{B}^2)^n, (P_\theta^{\otimes n})_{\theta \in \Theta})$, d.h. P_θ ist die gemeinsame Verteilung von X_1 und Y_1 . Da $\mathbf{COV}_\theta[X_1, Y_1] = \mathbf{COV}_\theta[X_1 - \mathbf{E}[X_1], Y_1 - \mathbf{E}[Y_1]]$ ist, kann o.B.d.A. angenommen werden, dass $\mathbf{E}_\theta[X_1] = \mathbf{E}_\theta[Y_1] = 0$ sei. Damit ergibt sich $\mathbf{COV}_\theta[X_1, Y_1] = \mathbf{E}[X_1 \cdot Y_1]$. Es gilt

$$\begin{aligned} \mathbf{E}_\theta[(n-1) \cdot C_n] &= \mathbf{E}_\theta \left[\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n) \right] = \sum_{i=1}^n \mathbf{E}_\theta [(X_i - \bar{X}_n)(Y_i - \bar{Y}_n)] \\ &= n \cdot \mathbf{E}_\theta \left[\left(\left(1 - \frac{1}{n}\right)X_1 - \frac{1}{n}X_2 - \dots - \frac{1}{n}X_n \right) \left(\left(1 - \frac{1}{n}\right)Y_1 - \frac{1}{n}Y_2 - \dots - \frac{1}{n}Y_n \right) \right] \\ &= n \left(\left(1 - \frac{1}{n}\right)^2 \mathbf{E}_\theta[X_1 \cdot Y_1] + \frac{1}{n^2} \sum_{i=2}^n \mathbf{E}_\theta[X_i \cdot Y_i] \right) \\ &= n \left(\left(1 - \frac{1}{n}\right)^2 \mathbf{COV}_\theta[X_1, Y_1] + \frac{n-1}{n^2} \mathbf{COV}_\theta[X_1, Y_1] \right) \\ &= (n-1) \mathbf{COV}_\theta[X_1, Y_1]. \end{aligned}$$

■

Um die Güte eines Schätzers $T(X)$ für $\varrho(\theta)$ messen zu können, betrachtet man die mittlere quadratische Abweichung

$$\mathcal{R}(\theta, T) := \mathbf{E}_\theta [(T(X) - \varrho(\theta))^2], \quad \theta \in \Theta.$$

Es gilt

$$\begin{aligned} \mathcal{R}(\theta, T) &= \mathbf{E}_\theta [T(X)^2 - 2\varrho(\theta)T(X) + (\varrho(\theta))^2] \\ &= \mathbf{E}_\theta [T(X)^2] - 2\varrho(\theta)\mathbf{E}_\theta[T(X)] + (\varrho(\theta))^2 \\ &= \mathbf{E}_\theta [T(X)^2] - (\mathbf{E}_\theta[T(X)])^2 + (\mathbf{E}_\theta[T(X)])^2 - 2\varrho(\theta)\mathbf{E}_\theta[T(X)] + (\varrho(\theta))^2 \\ &= \mathbf{Var}_\theta[T(X)] + (\mathbf{E}_\theta[T(X)] - \varrho(\theta))^2 \\ &= \mathbf{Var}_\theta[T(X)] + (b(\theta, T))^2, \end{aligned}$$

d.h. für erwartungstreue Schätzer $T(X)$ ist

$$\mathcal{R}(\theta, T) = \mathbf{Var}_\theta [T(X)] \quad \forall \theta \in \Theta.$$

Man definiert deshalb:

15.11 Definition (erwartungstreuer Schätzer mit minimaler Varianz, Wirkungsgrad):
 $T(X)$ heißt erwartungstreuer Schätzer für $\varrho(\theta)$ mit (gleichmäßig) minimaler Varianz (englisch: „uniformly minimum variance unbiased estimator“; kurz UMVU) oder auch gleichmäßig bester erwartungstreuer Schätzer, falls er erwartungstreu ist und

$$\mathbf{Var}_\theta [T(X)] \leq \mathbf{Var}_\theta [S(X)] \quad \forall \theta \in \Theta$$

für jeden erwartungstreuen Schätzer $S(X)$ für $\varrho(\theta)$ gilt.

Das Verhältnis $\frac{\mathbf{Var}_\theta [T^{(1)}]}{\mathbf{Var}_\theta [T^{(2)}]}$ heißt Wirkungsgrad von $T^{(1)}$ in Bezug auf $T^{(2)}$.

15.12 Beispiel:

Es sei $X := (X_1, \dots, X_n)$, $X_i \stackrel{d}{=} R(0, \theta)$, $i = 1, \dots, n$. Für $\varrho(\theta) = \theta$ seien die beiden folgenden Schätzer gegeben:

$$\begin{aligned} T_n^{(1)}(X) &:= \frac{2}{n} \sum_{i=1}^n X_i \\ T_n^{(2)}(X) &:= \frac{n+1}{n} \cdot \max\{X_1, \dots, X_N\}. \end{aligned}$$

- Beide Schätzer sind erwartungstreu:

$$\mathbf{E}_\theta [T_n^{(1)}] = \mathbf{E}_\theta \left[\frac{2}{n} \sum_{i=1}^n X_i \right] = \frac{2}{n} \cdot n \cdot \mathbf{E}_\theta [X_1] = 2 \cdot \frac{\theta}{2} = \theta.$$

Setzt man $M_n := \max\{X_1, \dots, X_n\}$, so gilt:

$$P_\theta(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \begin{cases} 0 & , x < 0 \\ \left(\frac{x}{\theta}\right)^n & , x \in [0, \theta] \\ 1 & , x > \theta. \end{cases}$$

Dabei besitzt M_n die Dichte

$$f_n(x) := \begin{cases} n\theta^{-n} \cdot x^{n-1} & , x \in [0, \theta] \\ 0 & , \text{sonst.} \end{cases}$$

Damit folgt $\mathbf{E}_\theta [T_n^{(2)}] = \mathbf{E}_\theta \left[\frac{n+1}{n} \cdot M_n \right] = \frac{n+1}{n} \mathbf{E}_\theta [M_n]$ und

$$\begin{aligned} \mathbf{E}_\theta [M_n] &= \int_0^\theta n\theta^{-n} \cdot x^n dx = n\theta^{-n} \cdot \int_0^\theta x^n dx \\ &= n\theta^{-n} \cdot \left[\frac{1}{n+1} x^{n+1} \right]_0^\theta = \frac{n}{n+1} \cdot \theta. \end{aligned}$$

Damit gilt also $\mathbf{E}_\theta [T_n^{(2)}] = \frac{n+1}{n} \mathbf{E}_\theta [M_n] = \theta$.

- Für die Varianz der beiden Schätzer ergibt sich:

$$\begin{aligned}
 \mathbf{Var}_\theta [T_n^{(1)}] &= \frac{4}{n} \cdot \mathbf{Var}[X_1] = \frac{4}{n} \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}. \\
 \mathbf{Var}_\theta [T_n^{(2)}] &= \left(\frac{n+1}{n}\right)^2 \cdot \mathbf{Var}_\theta [M_n] \\
 &= \left(\frac{n+1}{n}\right)^2 \cdot [\mathbf{E}_\theta [(M_n)^2] - (\mathbf{E}_\theta [M_n])^2] \\
 &= \left(\frac{n+1}{n}\right)^2 \cdot \left[\int_0^\theta x^2 \cdot n\theta^{-n} \cdot x^{n-1} dx - \left(\frac{n}{n+1}\right)^2 \cdot \theta^2 \right] \\
 &= \left(\frac{n+1}{n}\right)^2 \cdot \left[n\theta^{-n} \cdot \int_0^\theta x^{n+1} dx - \left(\frac{n}{n+1}\right)^2 \cdot \theta^2 \right] \\
 &= \left(\frac{n+1}{n}\right)^2 \cdot \left[n\theta^{-n} \cdot \left[\frac{1}{n+2} x^{n+2} \right]_0^\theta - \left(\frac{n}{n+1}\right)^2 \cdot \theta^2 \right] \\
 &= \left(\frac{n+1}{n}\right)^2 \cdot \left[\frac{n}{n+2} \cdot \theta^2 - \left(\frac{n}{n+1}\right)^2 \cdot \theta^2 \right] \\
 &= \left[\frac{(n+1)^2 - n \cdot (n+2)}{n \cdot (n+2)} \right] \cdot \theta^2 \\
 &= \frac{\theta^2}{n(n+2)}.
 \end{aligned}$$

Damit ist also $\mathbf{Var}_\theta [T_n^{(2)}] < \mathbf{Var}_\theta [T_n^{(1)}]$ für $n \geq 2$.

Es stellt sich nun die Frage, wie man erkennt, ob ein gegebener erwartungstreuer Schätzer für $\varrho(\theta)$ erwartungstreuer Schätzer mit minimaler Varianz ist. Eine hinreichende Antwort darauf liefert der folgende Satz.

15.13 Satz (Cramér-Rao):

Es seien $X := (X_1, \dots, X_n)$ eine Stichprobe aus reellwertigen Zufallsvariablen und $(\mathbb{R}^n, \mathfrak{B}^n, (P_\theta^{\otimes n})_{\theta \in \Theta})$ das zugehörige statistische Modell. Es sei $\Theta \subseteq \mathbb{R}$ ein offenes Intervall und die Verteilungsfunktion der Zufallsvariablen X_1 habe außerdem eine Dichte $f(x, \theta)$, die die folgenden Eigenschaften besitzt:

1. Die Ableitung $\frac{\partial f(x, \theta)}{\partial \theta}$ existiert für alle $\theta \in \Theta$ und für fast alle $x \in \mathbb{R}$,
2. $\mathbf{E}_\theta \left[\frac{\partial \ln f(X_1, \theta)}{\partial \theta} \right] = 0 \quad \forall \theta \in \Theta$ und
3. $I(\theta) := \mathbf{E}_\theta \left[\left(\frac{\partial \ln f(X_1, \theta)}{\partial \theta} \right)^2 \right] < \infty \quad \forall \theta \in \Theta$.

Ist $T(X_1, \dots, X_n): \mathbb{R}^n \rightarrow \mathbb{R}$ ein erwartungstreuer Schätzer für $\varrho(\theta)$ mit

$$\sum_{i=1}^n \mathbf{E}_\theta \left[T(X_1, \dots, X_n) \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right] = 1 \quad \forall \theta \in \Theta,$$

so gilt

$$\mathbf{Var}_\theta [T(X_1, \dots, X_n)] \geq \frac{1}{n \cdot I(\theta)} \quad \forall \theta \in \Theta.$$

Die Größe $I(\theta)$ wird Fisher-Information von P_θ genannt und der Quotient $\frac{1}{n \cdot I(\theta)}$ heißt die Cramér-Rao-Schranke.

Beweis:

Unter Berücksichtigung der angegebenen Regularitätsbedingungen erhält man

$$\begin{aligned} 0 &\leq \mathbf{E}_\theta \left[\left(T(X_1, \dots, X_n) - \theta - \frac{1}{n \cdot I(\theta)} \sum_{i=1}^n \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right)^2 \right] \\ &= \mathbf{E}_\theta \left[(T(X_1, \dots, X_n) - \theta)^2 - 2 \left((T(X_1, \dots, X_n) - \theta) \cdot \frac{1}{n \cdot I(\theta)} \sum_{i=1}^n \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right) \right. \\ &\quad \left. + \frac{1}{n^2 (I(\theta))^2} \sum_{k,l=1}^n \frac{\partial \ln f(X_k, \theta)}{\partial \theta} \frac{\partial \ln f(X_l, \theta)}{\partial \theta} \right] \\ &= \mathbf{E}_\theta [(T(X_1, \dots, X_n) - \theta)^2] - 2 \mathbf{E}_\theta \left[T(X_1, \dots, X_n) \cdot \frac{1}{n \cdot I(\theta)} \sum_{i=1}^n \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right] \\ &\quad + \frac{2\theta}{n \cdot I(\theta)} \sum_{i=1}^n \mathbf{E}_\theta \left[\frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right] + \frac{1}{n^2 (I(\theta))^2} \sum_{k,l=1}^n \mathbf{E}_\theta \left[\frac{\partial \ln f(X_k, \theta)}{\partial \theta} \frac{\partial \ln f(X_l, \theta)}{\partial \theta} \right] \\ &= \mathbf{Var}_\theta [T(X_1, \dots, X_n)] - 2 \frac{1}{n \cdot I(\theta)} + \frac{1}{n^2 (I(\theta))^2} \cdot n \cdot I(\theta), \end{aligned}$$

woraus

$$\mathbf{Var}_\theta [T(X_1, \dots, X_n)] \geq \frac{1}{n \cdot I(\theta)}$$

folgt. ■

15.14 Bemerkung:

Für erwartungstreue Schätzer $T(X)$ für θ ist Bedingung

$$\sum_{i=1}^n \mathbf{E}_\theta \left[T(X_1, \dots, X_n) \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right] = 1 \quad \forall \theta \in \Theta$$

im Satz von Cramér-Rao stets erfüllt.

Beweis:

Es seien $x := (x_1, \dots, x_n)$ und $f(x, \theta) := \prod_{i=1}^n f(x_i, \theta)$. Nach dem Satz von der majorisierten Konvergenz gilt für erwartungstreue Schätzer mit Dichte $f(x, \theta)$ die folgende Vertauschbarkeit von Integration und Differentiation:

$$\begin{aligned} \frac{\partial}{\partial \theta} \mathbf{E}_\theta [T(X)] &= \frac{\partial}{\partial \theta} \int_{\mathbb{R}^n} T(x_1, \dots, x_n) dP_\theta(x_1, \dots, x_n) \\ &= \frac{\partial}{\partial \theta} \int_{\mathbb{R}^n} T(x_1, \dots, x_n) f(x_1, \dots, x_n, \theta) d(x_1, \dots, x_n) \\ &= \int_{\mathbb{R}^n} T(x_1, \dots, x_n) \frac{\partial}{\partial \theta} f(x_1, \dots, x_n, \theta) d(x_1, \dots, x_n) \end{aligned}$$

Damit lässt sich nun die eigentliche Behauptung beweisen:

$$\begin{aligned}
 & \sum_{i=1}^n \mathbf{E}_\theta \left[T(X) \frac{\partial \ln f(X_i, \theta)}{\partial \theta} \right] \\
 &= \mathbf{E}_\theta \left[T(X) \frac{\partial}{\partial \theta} \underbrace{\sum_{i=1}^n \ln f(X_i, \theta)}_{=\ln \prod_{i=1}^n f(X_i, \theta) = \ln f(X, \theta)} \right] \\
 &= \int_{\mathbb{R}^n} T(x) \underbrace{\frac{\partial}{\partial \theta} \ln f(x, \theta)}_{=\frac{1}{f(x, \theta)} \cdot \frac{\partial}{\partial \theta} f(x, \theta)} dP_\theta(x_1, \dots, x_n) \\
 &= \int_{\mathbb{R}^n} T(x) \left(\frac{\partial}{\partial \theta} f(x, \theta) \right) \underbrace{\frac{1}{f(x, \theta)} f(x_1, \dots, x_n)}_{=1} d(x_1, \dots, x_n) \\
 &\stackrel{(*)}{=} \frac{\partial}{\partial \theta} \underbrace{\int_{\mathbb{R}^n} T(x) f(x, \theta) d(x_1, \dots, x_n)}_{=\mathbf{E}_\theta[T(X)] = \theta, \text{ da } T \text{ erwartungstreu ist}} \\
 &= \frac{\partial}{\partial \theta} \theta = 1.
 \end{aligned}$$

((*): Verwendung oben genannter Vertauschbarkeit von Integration und Differentiation.) ■

15.15 Beispiel (Stichprobenmittelwert bei normalverteilten Zufallsvariablen):

Es seien $X_1 \stackrel{d}{=} \mathcal{N}(\theta, \sigma)$ und P_θ die Normalverteilung mit Erwartungswert θ und Varianz σ^2 . Die zugehörige Verteilungsfunktion hat die Dichte

$$f(x, \theta) := \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\theta)^2}{2\sigma^2}} \quad \forall (x, \theta) \in \mathbb{R} \times \mathbb{R}.$$

Hieraus folgt:

$$\frac{\partial \ln f(x, \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left(-\frac{(x-\theta)^2}{2\sigma^2} + \ln \frac{1}{\sqrt{2\pi} \cdot \sigma} \right) = \frac{x-\theta}{\sigma^2}.$$

Damit ergibt sich:

$$\mathbf{E}_\theta \left[\frac{\partial \ln f(X_1, \theta)}{\partial \theta} \right] = \mathbf{E}_\theta \left[\frac{X_1 - \theta}{\sigma^2} \right] = 0 \quad \forall \theta \in \mathbb{R}.$$

Die Fisher-Information von P_θ ist dann:

$$I(\theta) = \mathbf{E}_\theta \left[\left(\frac{\partial \ln f(X_1, \theta)}{\partial \theta} \right)^2 \right] = \mathbf{E}_\theta \left[\left(\frac{X_1 - \theta}{\sigma^2} \right)^2 \right] = \frac{1}{\sigma^4} \mathbf{Var}[X_1] = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}.$$

Der Stichprobenmittelwert \bar{X}_n liefert nach Beispiel 15.9 einen erwartungstreuen Schätzer für $\mathbf{E}[X_1] = \theta$. Da einerseits $\mathbf{Var}_\theta[\bar{X}_n] = \frac{1}{n} \mathbf{Var}_\theta[X_1] = \frac{\sigma^2}{n}$ ist und andererseits $\frac{1}{n \cdot I(\theta)} = \frac{\sigma^2}{n}$ gilt, ist \bar{X}_n gleichmäßig bester erwartungstreuer Schätzer für $\varrho(\theta) = \theta$.

Das Konzept des erwartungstreuen Punktschätzers lässt sich in folgender Weise abschwächen:

15.16 Definition (asymptotisch erwartungstreu):

Eine Folge $(T_n)_{n \in \mathbb{N}}$ von Punktschätzern für $\varrho(\theta)$ heißt asymptotisch erwartungstreu, wenn

$$\lim_{n \rightarrow \infty} \mathbf{E}_\theta [T_n] = \varrho(\theta) \quad \forall \theta \in \Theta$$

gilt. Entsprechend lassen sich die Begriffe der Wirksamkeit und des Wirkungsgrads ausdehnen.

15.17 Beispiel:

Die Stichprobenvarianz $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ ist ein asymptotisch erwartungstreuer Schätzer für $\varrho(\theta) = \mathbf{Var}_\theta [X_1]$.

Einen Zusammenhang zwischen den Begriffen „asymptotisch erwartungstreu“ und „schwach konsistent“ stellt der folgende Satz her.

15.18 Satz:

Es sei $(T_n)_{n \in \mathbb{N}}$ eine Folge von Punktschätzern mit

$$\lim_{n \rightarrow \infty} \mathbf{E}_\theta [(T_n - \varrho(\theta))^2] = 0 \quad \forall \theta \in \Theta,$$

dann ist $(T_n)_{n \in \mathbb{N}}$ schwach konsistent für $\varrho(\theta)$.

Beweis:

Durch Anwendung der Rechenregeln für den Erwartungswert ergibt sich:

$$\begin{aligned} \mathbf{E}_\theta [(T_n - \varrho(\theta))^2] &= \int_{\mathbb{R}} P_\theta((T_n - \varrho(\theta))^2 > x) dx \\ &\geq \int_0^{\varepsilon^2} P_\theta((T_n - \varrho(\theta))^2 \geq x) dx \\ &\geq \varepsilon^2 P((T_n - \varrho(\theta))^2 \geq \varepsilon^2) \\ &= \varepsilon^2 P(|T_n - \varrho(\theta)| \geq \varepsilon). \end{aligned}$$

Mit der Voraussetzung $\lim_{n \rightarrow \infty} \mathbf{E}_\theta [(T_n - \varrho(\theta))^2] = 0 \quad \forall \theta \in \Theta$ folgt dann die Konsistenz. ■

15.3 Maximum-Likelihood-Schätzer

Das Maximum-Likelihood-Prinzip ist ein allgemeines statistisches Konzept zur Konstruktion von Schätzfunktionen. Wir erläutern es an einem einführenden Beispiel:

15.19 Beispiel:

In einer Urne mögen sich 9 Kugeln (rote und schwarze) befinden. Wieviele rote und schwarze es sind, ist nicht bekannt. Aus dieser Urne werden nun drei Kugeln ohne Zurücklegen gezogen. Damit ist festgelegt, dass die hypergeometrische Verteilung dem Experiment zugrunde liegt.

Es mögen zwei rote und eine schwarze Kugel gezogen worden sein. Damit ergeben sich folgende Parameter:

$$\begin{aligned} N &:= 9 \\ R &:= \text{unbekannte Anzahl der roten Kugeln} \\ n &:= 3 \\ k &:= 2 \end{aligned}$$

Da zwei rote Kugeln gezogen wurden, muss $R \geq 2$ gelten. Da außerdem eine schwarze Kugel gezogen wurde, kann R maximal 8 betragen. Prinzipiell sind also alle Werte zwischen 2 und 8 für R möglich, aber offensichtlich nicht gleichwahrscheinlich.

Um die Anzahl der roten Kugeln zu schätzen, werden die zum Ereignis $\{k = 2\}$ und den Bedingungen $R = 2, \dots, 8$ gehörenden Wahrscheinlichkeiten berechnet und miteinander verglichen:

R	0	1	2	3	4	5	6	7	8	9
$Hg_{9,R,3}(2)$	0	0	0.083	0.214	0.357	0.476	0.536	0.5	0.333	0

Das Ereignis $\{k = 2\}$ tritt also am wahrscheinlichsten auf, wenn $R = 6$ gilt. Deswegen wird 6 als Schätzwert für die Anzahl der roten Kugeln verwendet.

15.20 Bemerkung:

Etwas formaler lässt sich das obige Schätzverfahren wie folgt zusammenfassen:

Es seien $F_\theta(x)$ eine vorgegebene Wahrscheinlichkeitsverteilung, die von dem unbekannten Parameter $\theta \in \Theta$ abhängt. Wenn F_θ stetig verteilt ist, dann sei $f_\theta(x)$ eine zugehörige Dichte, ansonsten sei $f_\theta(x) := F_\theta(x) - F_\theta(x-0)$. Bei der einmaligen Ausführung des Zufallsexperimentes sei das Ereignis x_0 eingetreten. Dann ist

$$T(\theta) := \theta \in \Theta \text{ mit } f_\theta(x_0) \geq f_{\theta'}(x_0) \quad \forall \theta' \in \Theta$$

ein Schätzer für den unbekannten Parameter θ .

Dieses Schätzverfahren hat jedoch den Nachteil, dass es sich nur auf die Auswertung eines einzigen Zufallsexperimentes stützt. Genauere Schätzungen ergeben sich, wenn man das Experiment mehrfach wiederholt:

15.21 Beispiel:

In dem Versuch aus Beispiel 15.19 seien die drei Kugeln wieder zurückgelegt worden und danach seien die Kugeln in der Urne neu durchgemischt worden.

Beim erneuten Ziehen von drei Kugel, wurden nun $k_2 := 0$ rote Kugeln gezogen. Nur auf diesem Ergebnis basierend, würde man $R = 0$ vermuten, da $Hg_{9,0,3}(0) = 1$, und für $R > 0$ gilt $Hg_{9,R,3}(0) < 1$.

Um zu einem besseren Schätzwert für R zu gelangen, multipliziert man die Werte $Hg_{9,R,3}(2)$ und $Hg_{9,R,3}(0)$ punktweise:

R	0	1	2	3	4	5	6	7	8	9
$Hg_{9,R,3}(2) \cdot Hg_{9,R,3}(0)$	0	0	0.035	0.051	0.043	0.023	0.006	0	0	0

Damit ergibt sich $T(R) = 3$ als Schätzer für R .

Die Verallgemeinerung dieses Schätzverfahrens für den Parameter θ auf die n -malige Durchführung des Zufallsexperimentes nennt man die Maximum-Likelihood-Methode.

15.22 Definition (Maximum-Likelihood-Methode):

Es sei $F_\theta(x)$ eine vorgegebene Wahrscheinlichkeitsverteilung, die von dem unbekannten Parameter $\theta \in \Theta$ abhängt. Bei der n -maligen Ausführung des Zufallsexperimentes seien die Ereignisse $x := (x_1, \dots, x_n)$ eingetreten. Wenn F_θ stetig verteilt ist, dann sei $f_\theta(x)$ die zugehörige Dichte, ansonsten sei $f_\theta(x) := F(x) - F(x-0)$.

Dann heißt

$$L(\theta, x) := \prod_{i=1}^n f_\theta(x_i)$$

die Likelihood-Funktion und der Schätzer

$$T_n(\theta) := \theta \in \Theta \text{ mit } L(\theta, x) \geq L(\theta', x) \quad \forall \theta' \in \Theta$$

für den Parameter θ heißt der zugehörige Maximum-Likelihood-Schätzer.

15.23 Beispiel (Maximum-Likelihood-Schätzer für die Normalverteilung):

Die Zufallsvariable X sei normalverteilt mit den unbekannten Parametern μ und σ^2 . Zu schätzen ist $\theta := (\mu, \sigma) \in \Theta := \mathbb{R} \times \mathbb{R}^+$.

Mit der Dichte

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

der Normalverteilung ergibt sich folgende Likelihood-Funktion:

$$L(\theta, x) := \prod_{i=1}^n f_\theta(x_i) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right).$$

Aufgrund der Monotonie des Logarithmus liegt das Maximum von $L(\theta, x)$ an der Stelle, an der auch $\ln L(\theta, x)$ sein Maximum hat. Es ist

$$\ln L(\theta, x) = -n \cdot \ln \sqrt{2\pi} - n \cdot \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

Zur Bestimmung der Extrema werden die partiellen Ableitungen betrachtet:

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu}(\theta, x) &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \stackrel{!}{=} 0 \iff \mu = \frac{1}{n} \sum_{i=1}^n x_i, \\ \frac{\partial \ln L}{\partial \sigma}(\theta, x) &= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 \stackrel{!}{=} 0 \iff \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = S_n^2. \end{aligned}$$

Damit ergibt sich als Maximum-Likelihood-Schätzer: $T_n(\mu, \sigma^2) := (\bar{X}_n, S_n^2)$.

15.24 Beispiel (Maximum-Likelihood-Schätzer für die Poisson-Verteilung):

Die Zufallsvariable X sei Poisson-verteilt mit dem Parameter $\theta := \lambda \in \mathbb{R}^+$.

Mit $f_\theta(x_i) = \frac{\theta^{x_i}}{x_i!} e^{-\theta}$ ergibt sich folgende Likelihood-Funktion:

$$L(\theta, x) = \frac{\theta^{x_1 + \dots + x_n}}{x_1! \cdots x_n!} e^{-n\theta}.$$

Mit derselben Argumentation wie in Beispiel 15.23 ergibt sich zunächst:

$$\ln L(\theta, x) = \left(\sum_{i=1}^n x_i \right) \cdot \ln \theta - \ln(x_1! \cdots x_n!) - n\theta.$$

Wir differenzieren:

$$\frac{\partial \ln L}{\partial \theta}(\theta, x) = \sum_{i=1}^n \frac{x_i}{\theta} - n \stackrel{!}{=} 0 \iff \theta = \frac{1}{n} \sum_{i=1}^n x_i.$$

Damit ergibt sich als Maximum-Likelihood-Schätzer: $T_n(\lambda) := \bar{X}_n$.

15.4 Konfidenzintervalle

Konsistenz und Erwartungstreue sind mathematische Konzepte, die geeignet sind, um Punktschätzer qualitativ miteinander zu vergleichen. Will man die Güte einer Schätzung auch quantitativ erfassen, braucht man weitere Hilfsmittel, die wir zunächst anhand von Beispielen erläutern wollen.

Dazu betrachten wir noch einmal das unter 15.1 genannte Beispiel der statistischen Qualitätssicherung. Es sei $X := (X_1, \dots, X_n)$ das Ergebnis der Gut-Schlechtprüfung, d.h.

$$P(X_i = 1) = \theta \quad \text{und} \quad P(X_i = 0) = 1 - \theta \quad (i = 1, \dots, n).$$

Dann ist die Fehlerzahl

$$Y_n = \sum_{i=1}^n X_i$$

binomialverteilt mit den Parametern n und θ .

Der Satz von de Moivre-Laplace besagt, dass die Zufallsvariable

$$\frac{Y_n - n\theta}{\sqrt{n\theta(1-\theta)}}$$

für $n \rightarrow \infty$ näherungsweise standard-normalverteilt ist, so dass für jedes $\varepsilon > 0$ näherungsweise gilt:

$$P_\theta \left(-\varepsilon \leq \frac{Y_n - n\theta}{\sqrt{n\theta(1-\theta)}} \leq \varepsilon \right) \approx \Phi(\varepsilon) - \Phi(-\varepsilon) = 2\Phi(\varepsilon) - 1.$$

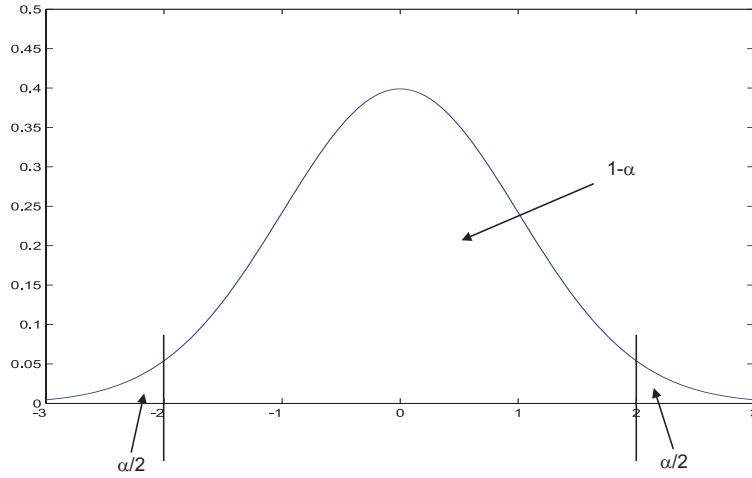


Abbildung 15.1: Darstellung der Dichte der Standardnormalverteilung

Weiter ist

$$\begin{aligned}
 \varepsilon &\leq \frac{Y_n - n\theta}{\sqrt{n\theta(1-\theta)}} \leq \varepsilon \\
 \Leftrightarrow |Y_n - n\theta| &\leq \varepsilon \cdot \sqrt{n\theta(1-\theta)} \\
 \Leftrightarrow (Y_n - n\theta)^2 &\leq \varepsilon^2 \cdot n\theta(1-\theta) \\
 \Leftrightarrow (Y_n)^2 - 2n\theta Y_n + n^2\theta^2 &\leq \varepsilon^2 \cdot n\theta - \varepsilon^2 \cdot n\theta^2 \\
 \Leftrightarrow g(\theta) := (\varepsilon^2 + n) \cdot \theta^2 - (2Y_n + \varepsilon^2)\theta + \frac{1}{n}(Y_n)^2 &\leq 0.
 \end{aligned}$$

Die Nullstellen von $g(\theta)$ lauten

$$\theta_{1,2} = \frac{1}{n + \varepsilon^2} \left[Y_n + \frac{\varepsilon^2}{2} \mp \varepsilon \sqrt{\frac{Y_n(n - Y_n)}{n} + \frac{\varepsilon^2}{4}} \right]$$

und wir erkennen, dass die Bedingung $g(\theta) \leq 0$ für alle $\theta \in [\theta_1, \theta_2]$ erfüllt ist. Indem man nun ε so wählt, dass

$$2\Phi(\varepsilon) - 1 = 1 - \alpha \quad \text{d.h.} \quad \Phi(\varepsilon) = 1 - \frac{\alpha}{2}$$

wird, erhält man

$$P_\theta(\{y \in \mathbb{R}^+ | \gamma_1(y) \leq \theta \leq \gamma_2(y)\}) \geq 1 - \alpha,$$

wobei

$$\gamma_{1,2}(y) := \frac{1}{n + \varepsilon^2} \cdot \left[y + \frac{\varepsilon^2}{2} \mp \varepsilon \sqrt{\frac{y \cdot (n - y)}{n} + \frac{\varepsilon^2}{4}} \right].$$

Die durch die Gleichung $\Phi(\varepsilon) = 1 - \alpha$, $\alpha \in (0, 1)$, eindeutig bestimmte Größe $\varepsilon = z_{1-\frac{\alpha}{2}}$ wird das $(1 - \alpha)$ -Quantil der Standard-Normalverteilung genannt, welches den einschlägigen statistischen Zahlentafeln entnommen werden kann. Gängige Zahlenwerte für α bzw. $z_{1-\frac{\alpha}{2}}$ sind:

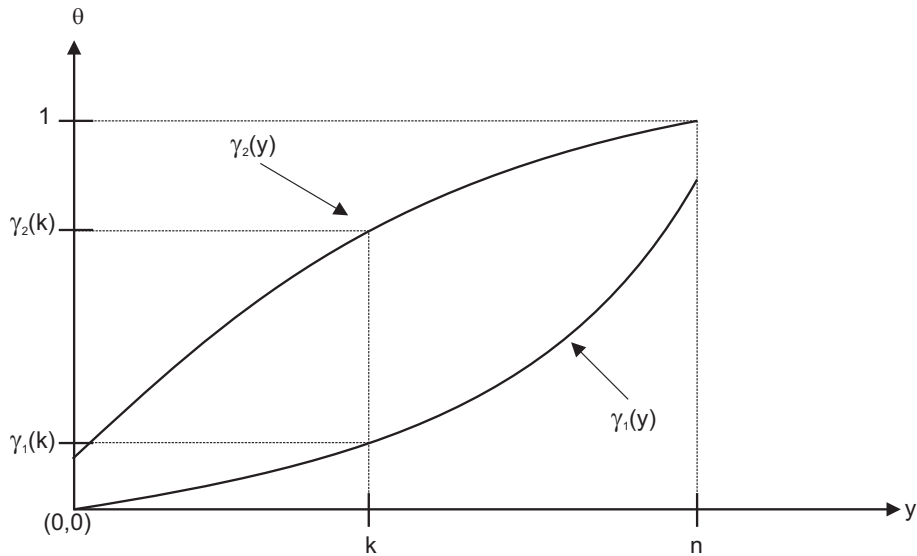


Abbildung 15.2: Konfidenzintervall

α	$z_{1-\frac{\alpha}{2}}$
0.1	1.645
0.05	1.960
0.01	2.576

Wir wollen nun unsere Ergebnisse mathematisch präzisieren:

15.25 Definition (α -Quantil):

Es seien X eine reellwertige Zufallsvariable mit Verteilungsfunktion F und $\alpha \in (0, 1)$. Dann heißt die Zahl

$$z_\alpha = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$$

das α -Quantil von X bzw. von F .

15.26 Definition (Konfidenzintervall, Konfidenzkoeffizient, Konfidenzniveau):

Es sei $X := (X_1, \dots, X_n)$ eine Stichprobe und $(P_\theta)_{\theta \in \Theta}$ das statistische Modell. $T_1(X)$, $T_2(X)$ seien zwei reelle Zufallsvariable mit $T_1(X) \leq T_2(X)$.

$[T_1(X), T_2(X)]$ ist ein Konfidenzintervall für $\varrho(\theta)$ mit Konfidenzkoeffizienten $\gamma \in (0, 1)$ (oder auch zum Konfidenzniveau γ), falls

$$P(T_1(X) \leq \varrho(\theta) \leq T_2(X)) \geq \gamma \quad \forall \theta \in \Theta$$

gilt.

15.27 Bemerkung:

Prinzipiell lässt sich jedes Konfidenzniveau γ erreichen, wenn man $T_1(X) \rightarrow -\infty$ sowie $T_2(X) \rightarrow \infty$ gehen lässt. Interessant ist also die Frage, wie klein das Intervall $[T_1(X), T_2(X)]$ zu einem vorgegebenen Konfidenzniveau gewählt werden kann oder alternativ wie klein $\gamma' \in (0, 1)$ mit

$$P(T_1(X) \leq \varrho(\theta') \leq T_2(X)) \leq \gamma' \quad \forall \theta \in \Theta$$

für $\theta \neq \theta'$ gewählt werden kann.

15.28 Definition (Kennfunktion):

Es sei $[T_1^{(\theta)}, T_2^{(\theta)}]$ ein Konfidenzbereich für $\varrho(\theta)$ zum Konfidenzniveau γ . Dann heißt

$$K: \Theta \times \Theta \longrightarrow [0, 1], \quad K(\theta, \theta') := P\left(T_1^{(\theta)}(X) \leq \varrho(\theta') \leq T_2^{(\theta)}(X)\right)$$

die Kennfunktion von $[T_1^{(\theta)}, T_2^{(\theta)}]$.

15.29 Definition (unverfälscht, gleichmäßig besser):

Ein Konfidenzbereich $[T_1(X), T_2(X)]$ für $\varrho(\theta)$ zum Konfidenzniveau γ heißt unverfälscht, wenn seine Kennfunktion folgende Eigenschaften besitzt:

- $K(\theta, \theta) \geq \gamma \quad \forall \theta \in \Theta$.
- $K(\theta, \theta) \geq K(\theta, \theta') \quad \forall (\theta, \theta') \in \Theta \times \Theta$ mit $\theta \neq \theta'$.

Sind die beiden Intervalle $[T_1(X), T_2(X)]$ und $[T'_1(X), T'_2(X)]$ zwei Konfidenzbereiche für $\varrho(\theta)$ zum gleichen Konfidenzniveau γ jeweils mit den zugehörigen Kennfunktionen K und K' , so heißt $[T_1(X), T_2(X)]$ gleichmäßig besser als $[T'_1(X), T'_2(X)]$, falls $K(\theta, \theta') \leq K'(\theta, \theta')$ für alle $(\theta, \theta') \in \Theta \times \Theta$ mit $\theta \neq \theta'$ gilt.

15.30 Bemerkung (Konfidenzintervall für den Erwartungswert):

Es sei X eine reelle Zufallsvariable mit $\sigma^2 := \text{Var}[X] < \infty$, $\mu := \mathbf{E}[X]$ und $X := (X_1, \dots, X_n)$ eine Stichprobe bzgl. P_X . Es gilt:

$$T_n(X) := \frac{1}{n}(X_1 + \dots + X_n) \xrightarrow[n \rightarrow \infty]{\text{f.s.}} \mu$$

und

$$\frac{nT_n(X) - n\mu}{\sqrt{n} \cdot \sigma} \xrightarrow[n \rightarrow \infty]{\text{d}} \mathcal{N}(0, 1).$$

Also gilt:

$$\begin{aligned} P(|T_n(X) - \mu| \leq \varepsilon) &= P(\mu - \varepsilon \leq T_n(X) \leq \mu + \varepsilon) \\ &= F_{T_n(X)}(\mu + \varepsilon) - F_{T_n(X)}(\mu - \varepsilon) \end{aligned}$$

und

$$P\left(\frac{nT_n(X) - n\mu}{\sqrt{n}\sigma} \leq x\right) \approx \Phi(x).$$

Mit $x_1 := -\frac{n\varepsilon}{\sigma\sqrt{n}}$ und $x_2 := \frac{n\varepsilon}{\sigma\sqrt{n}}$ ist $P(|T_n(X) - \mu| \leq \varepsilon) = P\left(x_1 \leq \frac{nT_n(X) - n\mu}{\sqrt{n}\sigma} \leq x_2\right)$. Damit ergibt sich:

$$P(|T_n(X) - \mu| \leq \varepsilon) = \Phi\left(\frac{n\varepsilon}{\sigma\sqrt{n}}\right) - \underbrace{\Phi\left(-\frac{n\varepsilon}{\sigma\sqrt{n}}\right)}_{1 - \Phi\left(\frac{n\varepsilon}{\sigma\sqrt{n}}\right)} = 2\Phi\left(\frac{n\varepsilon}{\sigma\sqrt{n}}\right) - 1,$$

d.h. gesucht ist ein $\varepsilon > 0$, so dass zu gegebenem $\gamma > 0$ folgende Bedingung erfüllt ist:

$$2\Phi\left(\frac{n\varepsilon}{\sigma\sqrt{n}}\right) - 1 \geq \gamma \quad \text{bzw.} \quad \Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) \geq \frac{1 + \gamma}{2}.$$

15.31 Beispiel:

An einer Klausur haben 200 Studierende teilgenommen. Nach der Korrektur von 16 Klausuren hat sich als Mittelwert für den Prozentsatz der erreichten Punkte $T_{16}(X) = 63.7\%$ ergeben. Die Varianz σ^2 werde durch die verzerrungsfreie Stichprobenvarianz, die sich zu $\tilde{S}_{16}^2 = 0.04$ ergeben hat, geschätzt.

Gesucht ist das Intervall $[T_{16}(X) - \varepsilon, T_{16}(X) + \varepsilon]$, in dem sich der Durchschnittswert nach Korrektur aller Klausuren mit einer Wahrscheinlichkeit von 95% befinden wird.

Es sind also $n := 16$, $\sigma^2 := 0.04$ und $\gamma := 0.95$. Nach Bemerkung 15.30 ist ein möglichst kleines $\varepsilon > 0$ gesucht, welches

$$\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) \geq \frac{1+\gamma}{2} \quad \text{bzw.} \quad \Phi(20\varepsilon) \geq \frac{39}{40}$$

erfüllt. Es ist $\Phi(1.96) \approx 0.97500$. Damit ist $\varepsilon \approx \frac{1.96}{20} = 0.098$ und es liegt der Durchschnittswert nach Korrektur aller Klausuren mit einer Wahrscheinlichkeit von 95% im Intervall $[63.7\% - 9.8\%, 63.7\% + 9.8\%] = [53.9\%, 73.5\%]$.

15.32 Beispiel (Konfidenzintervall für den Erwartungswert):

Ein Zufallsexperiment werde n mal durchgeführt. Die Varianz $\sigma^2 > 0$ sei bekannt. Für den unbekannten Erwartungswert $\mu = \mathbf{E}[X]$ verwenden wir als Schätzer das Stichprobenmittel $T_n(X) := \frac{1}{n} \sum_{i=1}^n X_i$. Asymptotisch gelten dann folgende Aussagen:

- Falls ein Konfidenzintervall für μ der Länge 2ε benötigt wird, so ist

$$\gamma = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right) - 1$$

die Wahrscheinlichkeit dafür, dass μ in dem Intervall $[T_n(X) - \varepsilon, T_n(X) + \varepsilon]$ liegt.

- Falls ein Konfidenzintervall, in dem μ mit der Wahrscheinlichkeit $\gamma \in (0, 1)$ liegt, gesucht ist, so ergibt sich die Intervalllänge 2ε aus

$$\varepsilon = \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}}$$

mit $\Phi\left(z_{\frac{1+\gamma}{2}}\right) = \frac{1+\gamma}{2}$.

- Zu vorgegebenen $\varepsilon > 0$ und $\gamma \in (0, 1)$ ist eine Stichprobe vom Umfang

$$n = \left(\frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\varepsilon}\right)^2$$

erforderlich, wenn μ mit der Wahrscheinlichkeit γ in dem Intervall $[T_n(X) - \varepsilon, T_n(x) + \varepsilon]$ liegen soll.

15.33 Beispiel:

- Ein Versuch werde $n := 100$ mal durchgeführt. Die Varianz $\sigma^2 := 5$ sei bekannt. Gesucht ist ein Konfidenzintervall, in dem der Erwartungswert mit einer Wahrscheinlichkeit von $\gamma := 99\%$ liegt.

Es ist $\frac{1+\gamma}{2} = 0.995$ und $\Phi(2.57583) \approx 0.9950$. Damit ist $z_{\frac{1+\gamma}{2}} \approx 2.57583$. Mit Satz 15.32 ergibt sich dann das Konfidenzintervall für $\gamma := 0.99$ zu $[T_n(X) - \varepsilon, T_n(X) + \varepsilon]$ mit

$$\varepsilon = \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}} \approx \frac{2.57583 \cdot \sqrt{5}}{10} \approx 0.57597.$$

D.h. die Abweichung von $T_n(X)$ vom tatsächlichen Mittelwert beträgt mit einer Wahrscheinlichkeit von $\gamma = 99\%$ höchstens $\varepsilon \approx 0.57597$.

- Bei einem Versuch mit der bekannten Varianz $\sigma^2 := 2$ ist gefragt, wie oft er durchgeführt werden muss, damit die Wahrscheinlichkeit, dass $T_n(X)$ um mehr als $\varepsilon := 5\%$ vom Mittelwert abweicht kleiner als $1 - \gamma := 1\%$ ist.

Mit $\frac{1+\gamma}{2} = 0.995$ ergibt sich zunächst $\Phi(2.57583) \approx \frac{1+\gamma}{2}$ bzw. $z_{\frac{1+\gamma}{2}} \approx 2.57583$. Nach Satz 15.32 gilt dann

$$n = \left(\frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\varepsilon} \right)^2 \approx \left(\frac{2.57583 \cdot \sqrt{2} \cdot 20}{0.05} \right)^2 \approx 5307.92$$

bzw. $n \geq 5308$.

(Siehe auch Mathematica-Notebook zu diesem Beispiel.)

Mit ähnlichen Überlegungen wie in Bemerkung 15.30 können auch Konfidenzintervalle für die Varianz aufgestellt werden, wie der folgende Satz zeigt.

15.34 Beispiel (Konfidenzintervall für die Varianz):

Bei bekanntem Erwartungswert μ und mit $Q_n(X) := \sum_{i=1}^n (X_i - \mu)^2$ liegt die Varianz eines Zufallsexperiments mit einer Wahrscheinlichkeit von $\gamma \in (0, 1)$ in

$$\left[\frac{Q_n(X)}{z_{n, \frac{1+\gamma}{2}}}, \frac{Q_n(X)}{z_{n, \frac{1-\gamma}{2}}} \right].$$

Dabei sei $\chi_n^2 \left(z_{n, \frac{1+\gamma}{2}} \right) = \frac{1+\gamma}{2}$ und es sei $\chi_n^2(z)$ die χ^2 -Verteilung mit n Freiheitsgraden.

Es sei $\sigma^2 > 0$ die Varianz des Zufallsexperiments. Die Zufallsvariable

$$\frac{Q_n(X)}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \mu_0}{\sigma} \right)^2$$

ist eine Summe von n Quadraten von $\mathcal{N}(0, 1)$ -verteilten Zufallsvariablen. Damit ist $\frac{Q_n(X)}{\sigma^2}$ χ_n^2 -verteilt und es gilt:

$$P \left(\frac{Q_n(X)}{z_{n, \frac{1+\gamma}{2}}} \leq \sigma^2 \leq \frac{Q_n(X)}{z_{n, \frac{1-\gamma}{2}}} \right) = P \left(z_{n, \frac{1-\gamma}{2}} \leq \frac{Q_n(X)}{\sigma^2} \leq z_{n, \frac{1+\gamma}{2}} \right) = \frac{1+\gamma}{2} - \frac{1-\gamma}{2} = \gamma.$$

15.35 Beispiel:

Eine Maschine bohrt Löcher der Größe $\mu = 20\text{mm}$ in Bleche. Der Produktion wurden $n := 10$ Bleche entnommen und die Größen der gebohrten Löcher wurden nachgemessen:

20.1mm, 19.8mm, 19.9mm, 20.5mm, 20.0mm, 21.0mm, 20.2mm, 19.4mm, 20.0mm, 19.7mm

Um Aussagen über die Qualität der Bohrungen zu treffen, interessiert man sich für die Varianz, mit der die Maschine arbeitet.

Als Stichprobenvarianz ergibt sich

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1.8}{9} = 0.2$$

und als Standardabweichung $\sqrt{0.2}\text{mm} = 0.45\text{mm}$. Das Konfidenzintervall zum Niveau 95% berechnet sich wie folgt:

Es ist $Q_n(X) = 1.8$, $\frac{1+\gamma}{2} = 0.975$ und $\frac{1-\gamma}{2} = 0.025$ und damit

$$\chi_n^2(20.4832) \approx 0.975 \quad \text{und} \quad \chi_n^2(3.24697) \approx 0.025.$$

Damit ergibt sich folgendes Konfidenzintervall:

$$\left[\frac{Q_n(X)}{z_{n, \frac{1+\gamma}{2}}}, \frac{Q_n(X)}{z_{n, \frac{1-\gamma}{2}}} \right] \approx \left[\frac{1.8}{20.4832}, \frac{1.8}{3.24697} \right] \approx [0.08788, 0.55436].$$

15.5 Hypothesentests

In der Praxis interessiert man sich oftmals nicht nur für den Wert eines Schätzers, sondern auch dafür, mit welcher Wahrscheinlichkeit der Schätzer den korrekten Wert geliefert hat. Eine Möglichkeit, solche Fragen zu beantworten, stellen die bereits behandelten Konfidenzintervalle dar. Eine andere Möglichkeit bieten die Hypothesentests.

Bei einem Hypothesentest wird gefragt, ob basierend auf den beobachteten Ausgängen eines Zufallsexperimentes, eine bestimmte Annahme abzulehnen ist oder nicht.

15.36 Definition ((Null-)Hypothese, Alternative):

Gegeben sei eine Stichprobe $x = (x_1, \dots, x_n)$, die eine Realisierung unabhängiger, \mathbf{P}_ϑ -verteilter Zufallsvariablen $X = (X_1, \dots, X_n)$ sind. \mathbf{P}_ϑ stellt dabei die parametrische Familie von Wahrscheinlichkeitsverteilungen auf dem Merkmalraum \mathbb{R} mit $\vartheta \in \Theta$ dar.

$H_0 \subset \Theta$ heißt (Null-)Hypothese, $H_1 := \Theta - H_0$ heißt Alternative und $H_0 : H_1$ beschreibe das Testproblem H_0 gegen H_1 . Sei $\varrho : \Theta \rightarrow \mathbb{R}$ ein Parameter und $k_0 \in \mathbb{R}$ fest, dann heißt

$$H_0 := \{\vartheta | \varrho(\vartheta) \leq k_0\} \text{ gegen } H_1 := \{\vartheta | \varrho(\vartheta) > k_0\}$$

bzw.

$$H_0 := \{\vartheta | \varrho(\vartheta) \geq k_0\} \text{ gegen } H_1 := \{\vartheta | \varrho(\vartheta) < k_0\}$$

einseitiges Testproblem und

$$H_0 := \{\vartheta | \varrho(\vartheta) = k_0\} \text{ gegen } H_1 := \{\vartheta | \varrho(\vartheta) \neq k_0\}$$

zweiseitiges Testproblem.

Hat man ein Testproblem $H_0 : H_1$ gegeben, so sind folgende Entscheidungen möglich:

- $a_{H_0} :=$ "Annahme von H_0 "
- $a_{H_1} :=$ "Ablehnung von H_0 "

Ein Hypothesentest d liefert nun zu einer Stichprobe $x = (x_1, \dots, x_n)$ die Antwort aus der Menge $\{a_{H_0}, a_{H_1}\}$, d.h. ist eine Abbildung $d : \mathbb{R}^n \rightarrow \{a_{H_0}, a_{H_1}\}$. Die Menge $\{x | d(x) = a_{H_0}\}$ heißt Annahmebereich und somit stellt die Menge $\{x | d(x) = a_{H_1}\}$ den Ablehnungsbereich bzw. kritischen Bereich dar.

Aufgrund der jeweils betrachteten Stichprobe x können Fehlentscheidungen auftreten: Liegt eine Grundgesamtheit vor, bei der $\vartheta \in H_0$ gilt und betrachtet man eine Stichprobe, die bei dem verwendeten Test d zur Ablehnung führt, so begeht man einen Fehler 1. Art (H_0 wird zu Unrecht abgelehnt). Liegt eine Grundgesamtheit mit $\vartheta \in H_1$ vor, und man entscheidet sich für die Annahme der Hypothese, so macht man einen Fehler 2. Art (H_0 wird zu Unrecht angenommen).

Die Wahrscheinlichkeit für die Fehler 1. und 2. Art sollen möglichst klein sein. Dabei hilft die im Folgenden definierte Funktion.

15.37 Definition (Gütefunktion):

Gegeben sei ein Testproblem $H_0 : H_1$ und ein Hypothesentest d . Dann heißt die Funktion $\beta : \Theta \rightarrow (0, 1)$ mit

$$\beta(\vartheta) = P_{\vartheta}(d(x) = a_{H_1}) \quad , \quad \forall \vartheta \in \Theta$$

die Gütefunktion von d .

Damit beschreibt $\beta(\vartheta)$ für $\vartheta \in H_0$ die Wahrscheinlichkeit für einen Fehler 1. Art und die Funktion $1 - \beta(\vartheta)$, auch Operationscharakteristik genannt, die Wahrscheinlichkeit für einen Fehler 2. Art. Gesucht sind demnach Hypothesentests, für die β möglichst klein auf H_0 und möglichst groß auf H_1 ist. Diese scheinbar widersprüchlichen Forderungen lassen sich wie folgt umsetzen:

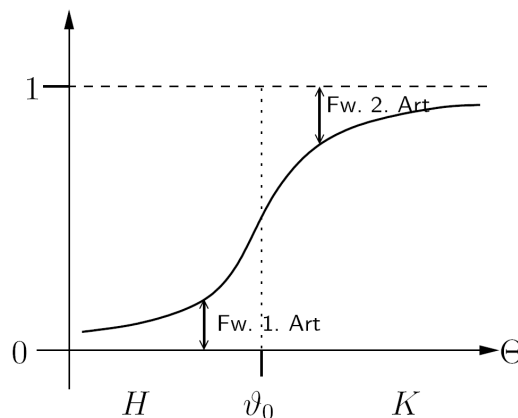


Abbildung 15.3: Der typische Verlauf einer Gütefunktion für ein einseitiges Testproblem.

15.38 Definition:

Es seien $d: \mathbb{R}^n \rightarrow \{a_{H_0}, a_{H_1}\}$ ein Hypothesentest für $H_0 : H_1$ und $\alpha \in (0, 1)$. Dann heißt d Test zum (Signifikanz-) Niveau α , falls gilt

$$P_{\vartheta}(d(x) = a_{H_1}) \leq \alpha, \quad \forall \vartheta \in H_0.$$

Betrachtet man eine Teilmenge D der Menge aller Tests zum Niveau α für $H_0 : H_1$. Ein Test d' aus dieser Teilmenge D heißt trennscharf, falls gilt

$$P_{\vartheta}(d'(x) = a_{H_1}) = \max_{d \in D} P_{\vartheta}(d(x) = a_{H_1}), \quad \forall \vartheta \in H_1.$$

Liefert ein Test zum Niveau α also $d(x) = a_{H_1}$, so wird H_0 abgelehnt und die Wahrscheinlichkeit, dass dies zu Unrecht geschieht ist $\leq \alpha$. Man sagt dann, die Ablehnung von H_0 zum Niveau α sei statistisch gesichert.

Die Aufgabe der mathematischen Statistik ist es nun, für verschiedene Testfragestellungen trennscharfe Tests zum Niveau α herzuleiten. Wie eine solche Fragestellung und der zugehörige Test aussehen, soll an folgendem Beispiel verdeutlicht werden.

15.39 Beispiel:

Es seien x_1, \dots, x_n Realisierungen unabhängiger, normalverteilter Zufallsvariablen mit dem unbekannten Erwartungswert μ und der bekannten Varianz σ_0^2 . Es wird angenommen, der Erwartungswert sei μ_0 , d.h. es wird die Hypothese $\mu = \mu_0$ aufgestellt.

Wenn alle X_i mit μ_0 und σ_0^2 normalverteilt sind, so ist das arithmetische Mittel $\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$ normalverteilt mit dem Erwartungswert μ_0 und der Varianz $\frac{\sigma_0^2}{n}$. Folglich ist

$$Y_n := \frac{\sqrt{n}}{\sigma_0}(\bar{X}_n - \mu_0)$$

standardnormalverteilt.

Die Hypothese $\mu = \mu_0$ ist dann abzulehnen, wenn der Betrag von Y_n „zu groß“ wird, d.h.:

$$P_{\mu_0}(|Y_n| > t^*) \stackrel{!}{=} \alpha$$

Da Y_n standardnormalverteilt ist, ergibt sich $t^* = u_{1-\frac{\alpha}{2}}$, also das $1 - \frac{\alpha}{2}$ - Quantil der Standardnormalverteilung. Die Hypothese $\mu = \mu_0$ wird also abgelehnt, falls

$$\left| \frac{\sqrt{n}}{\sigma_0}(\bar{X}_n - \mu_0) \right| > u_{1-\frac{\alpha}{2}}$$

ist.

Weitere Hypothesentests, die sich auf Fragestellungen bzgl. Stichproben x , die Realisierungen unabhängiger normalverteilter Zufallsvariablen sind, beziehen, gibt nachstehende Tabelle an.

Zusammenfassung der Einstichproben - Tests bei Normalverteilungsannahme			
Testgröße T	Nullhypothese H_0	Alternative H_1	Ablehnungsbereich
Gauß - Test (σ_0 bekannt) $T(x) = \frac{\sqrt{n}}{\sigma_0} (\bar{x}_n - \mu_0)$	$\mu = \mu_0$ $\mu \leq \mu_0$ $\mu \geq \mu_0$	$\mu \neq \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$ T > u_{1-\frac{\alpha}{2}}$ $T > u_{1-\alpha}$ $T < u_\alpha$
t - Test (σ_0 unbekannt) $T(x) = \sqrt{\frac{n}{s_x^2}} (\bar{x}_n - \mu_0)$	$\mu = \mu_0$ $\mu \leq \mu_0$ $\mu \geq \mu_0$	$\mu \neq \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$ T > t_{n-1; 1-\frac{\alpha}{2}}$ $T > t_{n-1; 1-\alpha}$ $T < t_{n-1; \alpha}$
χ^2 - Test $T(x) = (n-1) \frac{s_x^2}{\sigma_0^2}$	$\sigma^2 = \sigma_0^2$ $\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$	$T < \chi_{n-1; \frac{\alpha}{2}}^2$ oder $T > \chi_{n-1; 1-\frac{\alpha}{2}}^2$ $T > \chi_{n-1; 1-\alpha}^2$ $T < \chi_{n-1; \alpha}^2$

Zusammenfassung der Zweistichproben - Tests bei Normalverteilungsannahme			
Testgröße T	Nullhypothese H_0	Alternative H_1	Ablehnungsbereich
Gauß - Test (σ_1, σ_2 bekannt) $T(x, y) = \frac{\bar{y} - \bar{x}}{\sqrt{\frac{1}{m}\sigma_1^2 + \frac{1}{n}\sigma_2^2}}$	$\mu_1 = \mu_2$ $\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$	$\mu_1 \neq \mu_2$ $\mu_1 > \mu_2$ $\mu_1 < \mu_2$	$ T > u_{1-\frac{\alpha}{2}}$ $T < u_\alpha$ $T > u_{1-\alpha}$
t - Test (σ_1, σ_2 unbekannt, aber gleich!) $T(x, y) = \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\bar{y} - \bar{x}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2}}$	$\mu_1 = \mu_2$ $\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$	$\mu_1 \neq \mu_2$ $\mu_1 > \mu_2$ $\mu_1 < \mu_2$	$ T > t_{m+n-2; 1-\frac{\alpha}{2}}$ $T < t_{m+n-2; \alpha}$ $T > t_{m+n-2; 1-\alpha}$
F - Test $T(x, y) = \frac{s_x^2}{s_y^2}$	$\sigma_1^2 = \sigma_2^2$ $\sigma_1^2 \leq \sigma_2^2$ $\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$	$T < F_{m-1, n-1; \frac{\alpha}{2}}$ oder $T > F_{m-1, n-1; 1-\frac{\alpha}{2}}$ $T > F_{m-1, n-1; 1-\alpha}$ $T < F_{m-1, n-1; \alpha}$

15.40 Beispiel:

Bei der Herstellung von Büroklammern sollen etwa 100g pro Beutel abgepackt werden. Die Standardabweichung sei bekannt und betrage $\sigma_0 := 0.4$. Das Abwiegen von $n := 10$ Beuteln ergab:

101g, 98g, 97g, 102g, 99g, 99g, 101g, 98g, 100g, 102g

Getestet wird die Hypothese $\mu = \mu_0 := 100$ g zum Signifikanz-Niveau $\alpha = 5\%$.

Es ist $\bar{X}_n = \frac{997}{10}$. Damit folgt zunächst

$$Y_{(n)} := \frac{\sqrt{n}}{\sigma_0} (\bar{X}_n - \mu_0) = \frac{10}{4} \cdot \sqrt{10} \cdot \left(-\frac{3}{10}\right) \approx -2.372.$$

Für $u_{1-\frac{\alpha}{2}} = 2$ gilt $\Phi(-u_{1-\frac{\alpha}{2}}) + 1 - \Phi(u_{1-\frac{\alpha}{2}}) = 0.046 < \alpha$ und es ist $|Y_n| \approx 2.372 > u_{1-\frac{\alpha}{2}} = 2$. D.h. die Hypothese ist abzulehnen. (Siehe auch Mathematica-Notebook zu diesem Beispiel.)

15.41 Beispiel (Gut-Schlecht-Prüfung):

Bei der statistischen Qualitätskontrolle wird überprüft, ob die Ausschußwahrscheinlichkeit einer Produktion vereinbarungsgemäß nicht größer als p_0 sei, d.h. es wird die Hypothese

$$H_0 : p \leq p_0$$

aufgestellt, wobei p die tatsächliche (unbekannte) Ausschußwahrscheinlichkeit der Produktion ist. Die Hypothese H_0 wird mittels einer der Produktion entnommenen Stichprobe vom Umfang n und der darin enthaltenen Anzahl X defekter Stücke überprüft. Diese Anzahl X hängt von der Wahl der Stichprobe ab, als ihre Verteilung wurde in Stochastik I die Binomial-, die Hypergeometrische und die Poissonverteilung betrachtet.

Die Hypothese H_0 wird abgelehnt, wenn die Anzahl X der defekten Stücke in der Stichprobe "zu groß" ist, d.h. größer als die sog. Annahmезahl c . Ist X "hinreichend klein", d.h. $X < c$ so wird H_0 nicht verworfen. Beide Schlüsse können natürlich falsch sein, was auf die Begriffe der Fehler 1. und 2. Art führt.

Zur Minimierung dieser Fehler wird in der Qualitätssicherung in der Regel die Operationscharakteristik an Stelle der Gütefunktion betrachtet, sie lautet

$$L(p) := 1 - \beta(p) = P(X \leq c).$$

Um die Fehler 1. und 2. Art zu minimieren, muß die Operationscharakteristik für $p \leq p_0$ möglichst groß sein, für $p > p_0$ hingegen möglichst klein.

Bei der Konstruktion von $(n - c)$ - Stichprobenplänen wird dies erreicht, indem die Operationscharakteristik durch zwei vorgegebene Punkte, die die Zielsetzungen von Lieferanten und Abnehmer abbilden, gelegt wird (siehe Abbildung 15.4).

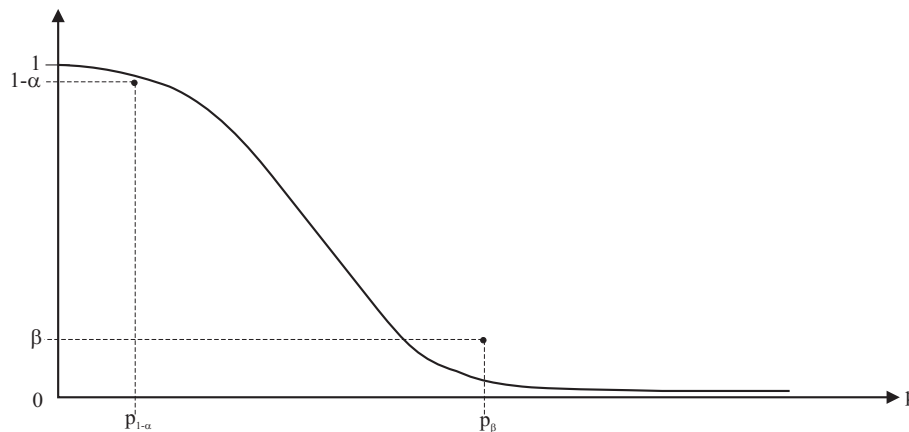


Abbildung 15.4: Beispiel einer Operationscharakteristik

Literatur zu Kapitel 15

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- A.O. ALLEN:
Probability, Statistics and Queueing Theory,

Academic Press, London, 2nd. Edition, 1990.

ISBN: 0120510510

- O. BEYER, H. HACKEL, V. PIEPER, J. TIEDGE:
Wahrscheinlichkeitsrechnung und mathematische Statistik,
7. Auflage, Teubner-Verlag, Stuttgart, 1995.
ISBN: 3-8154-2075-X
- M. FISZ:
Wahrscheinlichkeitsrechnung und mathematische Statistik,
VEB, Deutscher Verlag der Wissenschaften 1989.
ISBN: 3326000790
- U. KRENGEL:
Einführung in die Wahrscheinlichkeitstheorie und Statistik,
6. Auflage, Vieweg, 2002.
ISBN: 3528672595
- J. LEHN/H. WEGMANN:
Einführung in die Statistik,
4. Auflage, Teubner, 2004.
ISBN: 3519320711

Kapitel 16

Bedingte Erwartung und bedingte Verteilung

Dieses Kapitel stellt in gewisser Weise eine Fortsetzung des Kapitels über die bedingten Wahrscheinlichkeiten aus der Stochastik I dar. Im Folgenden wird der Begriff der bedingten Erwartung eingeführt und die bedingte Verteilung für den Fall kontinuierlicher Wahrscheinlichkeitsverteilungen definiert.

Schlüsselwörter: Bedingte Erwartung, Waldsche Gleichung, bedingte Verteilung, bedingte Dichte, Randverteilungsdichte, Randverteilung.

16.1 Bedingte Erwartung

In Stochastik I in Kapitel 3 wurde der Begriff der elementaren bedingten Wahrscheinlichkeit behandelt. Für einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ und für $A, B \in \mathfrak{F}$ mit $P(B) > 0$ wurde

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad (16.1)$$

elementare bedingte Wahrscheinlichkeit von A unter der Bedingung bzw. unter der Hypothese B genannt und $P(\cdot|B): \mathfrak{F} \rightarrow [0, 1]$ stellte sich als ein Wahrscheinlichkeitsmaß auf (Ω, \mathfrak{F}) heraus.

Der Begriff der elementaren bedingten Wahrscheinlichkeit lässt sich sofort auf reellwertige Zufallsvariablen X, Y auf $(\Omega, \mathfrak{F}, P)$ ausdehnen, wenn man Y zunächst als diskret voraussetzt und analog zu Gleichung (16.1)

$$P(X \in B|Y = y) := P(\{X \in B\}|\{Y = y\}) := \frac{P(X \in B, Y = y)}{P(Y = y)} \quad (16.2)$$

für alle $B \in \mathfrak{B}$ und $y \in \mathbb{R}$ mit $P(Y = y) > 0$ festgelegt. Für jedes dieser y definiert $P(X \in \cdot | Y = y)$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathfrak{B})$, das später bedingte Verteilung von X gegeben $Y = y$ genannt wird und mit $P_{X|Y=y}$ bezeichnet wird. Der zugehörige Erwartungswert

$$\mathbf{E}[X|Y = y] = \int_{\mathbb{R}} x dP_{X|Y=y}(x) = \frac{1}{P(Y = y)} \int_{\{\omega|Y(\omega)=y\}} X(\omega) dP(\omega) \quad (16.3)$$

heißt, sofern er existiert, bedingter Erwartungswert von X gegeben $Y = y$.

Zu beachten ist hier, dass unter der Annahme eines diskreten Y die bedingte Verteilung und der bedingte Erwartungswert gegeben $Y = y$ nur für P -fast alle y definiert ist.

Offensichtlich versagt dieses Konzept und verliert auch seine Anschaulichkeit, wenn Y keine diskrete Zufallsvariable ist. Denn in diesem Fall kann $P(Y = y) = 0 \ \forall y \in \mathbb{R}$ gelten, so dass die Festlegungen (16.2) und (16.3) keinen Sinn mehr machen würden.

Wir betrachten zunächst weiter den Fall der diskreten Zufallsvariable Y , setzen $\mathcal{D} := \{y | P(Y = y) > 0\}$ und nehmen an, dass X beschränkt ist, d.h. $|X| \leq k$. Damit gilt auch $|\mathbf{E}[X|Y = y]| \leq k$ und die Funktion

$$f(y) := \begin{cases} \mathbf{E}[X|Y = y] & , \ y \in \mathcal{D} \\ 0 & , \ \text{sonst} \end{cases}$$

bildet eine durch k beschränkte, messbare Funktion. Folglich ist $\mathbf{E}[X|Y] := f(Y)$ eine beschränkte, \mathfrak{F} - $\sigma(Y)$ -messbare Zufallsgröße, wobei $\sigma(Y)$ die von Y erzeugte σ -Algebra bezeichnet. $\mathbf{E}[X|Y]$ nennt man (als Funktion von y) bedingte Erwartung.

Wir gehen weiter davon aus, dass X eine Zufallsvariable ist, deren Realisierung nicht beobachtet werden kann, der Wert von Y sei jedoch bekannt. Gesucht ist nun eine Approximation in Abhängigkeit von Y , d.h. eine \mathfrak{F} - $\sigma(Y)$ -messbare Zufallsgröße, die X möglichst gut abbildet. Es wird sich zeigen, dass dies genau die bedingte Erwartung ist.

Dazu wird zunächst festgestellt, dass für jedes $A := \{\omega \mid Y(\omega) \in B\}$

$$\begin{aligned}
 \int_{\{Y \in B\}} \mathbf{E}[X|Y] \, dP &= \int_{B \cap \mathcal{D}} \mathbf{E}[X|Y = y] \, dP_Y(y) \\
 &= \sum_{y \in B \cap \mathcal{D}} \mathbf{E}[X|Y = y] P(Y = y) \\
 &= \sum_{y \in B \cap \mathcal{D}} \left(\frac{1}{P(Y = y)} \int_{\{Y=y\}} X \, dP \right) P(Y = y) \\
 &= \int_{\{Y \in B\}} X \, dP.
 \end{aligned}$$

Damit hat man folgende Eigenschaften, durch die die bedingte Erwartung P - fast sicher eindeutig bestimmt wird:

- 1) Die Zufallsgröße $\mathbf{E}[X|Y]$ ist \mathfrak{F} - $\sigma(Y)$ -messbar.
- 2) $\int_A \mathbf{E}[X|Y] \, dP = \int_A X \, dP$ für alle $A \in \sigma(Y)$.

Die Eindeutigkeit lässt sich wie folgt zeigen:

Sei Z eine weitere \mathfrak{F} - $\sigma(Y)$ -messbare Zufallsvariable mit

$$\int_A Z \, dP = \int_A X \, dP \quad \forall A \in \sigma(Y),$$

dann folgt für alle $A \in \sigma(Y)$:

$$\int_A (\mathbf{E}[X|Y] - Z) \, dP = 0.$$

Daraus ergibt sich mit Stochastik I Satz 9.40 und der \mathfrak{F} - $\sigma(Y)$ -Messbarkeit von $\mathbf{E}[X|Y] - Z$:

$$\mathbf{E}[X|Y] - Z = 0 \quad P - f.s.$$

Betrachtet man obige Eigenschaften, so fällt auf, dass sie von Y nur über $\sigma(Y)$ abhängen, sie können demnach auch wie folgt für eine beliebige σ -Algebra $\mathfrak{G} \subset \mathfrak{F}$ formuliert werden:

- 1') Die Zufallsgröße $\mathbf{E}[X|\mathfrak{G}]$ ist \mathfrak{G} -messbar.
- 2') $\int_A \mathbf{E}[X|\mathfrak{G}] \, dP = \int_A X \, dP$ für alle $A \in \mathfrak{G}$.

Es bleibt die Frage, bezüglich welchem Abstandsbegriff $\mathbf{E}[X|\mathfrak{G}]$ tatsächlich die beste Approximation von X bildet. Eine Antwort darauf liefert die folgende Bemerkung und der anschließende Satz.

16.1 Bemerkung:

- a) Aus der Eigenschaft 2') und der Beschränktheit von X und $\mathbf{E}[X|\mathfrak{G}]$ erhält man für alle \mathfrak{G} -messbare Zufallsvariablen Y mit $\mathbf{E}[|Y|] < \infty$:

$$\mathbf{E}[XY] = \mathbf{E}[\mathbf{E}[X|\mathfrak{G}]Y] \quad , \quad \text{d.h.} \quad \mathbf{E}[Y \cdot (X - \mathbf{E}[X|\mathfrak{G}])] = 0.$$

Damit ergibt sich für jede \mathfrak{G} -messbare Zufallsvariablen Y mit $\mathbf{E}[|Y|^2] < \infty$:

$$\begin{aligned} \mathbf{E}[(X - Y)^2] &= \mathbf{E}[(X - \mathbf{E}[X|\mathfrak{G}])^2] + \mathbf{E}[(\mathbf{E}[X|\mathfrak{G}] - Y)^2] \\ &\quad + 2\mathbf{E}[(\mathbf{E}[X|\mathfrak{G}] - Y)(X - \mathbf{E}[X|\mathfrak{G}])] \\ &= \mathbf{E}[(X - \mathbf{E}[X|\mathfrak{G}])^2] + \mathbf{E}[(\mathbf{E}[X|\mathfrak{G}] - Y)^2] \\ &\geq \mathbf{E}[(X - \mathbf{E}[X|\mathfrak{G}])^2]. \end{aligned}$$

Die gesuchte Metrik, in der $\mathbf{E}[X|\mathfrak{G}]$ tatsächlich die beste Approximation von X bildet, ist demnach die Abstandsfunktion auf dem Raum $L_2(\Omega, \mathfrak{F}, P)$:

$$d_2(X, Y) := \left(\mathbf{E}[|X - Y|^2] \right)^{\frac{1}{2}}.$$

- b) Mit $L_p(\Omega, \mathfrak{F}, P)$ wird der Vektorraum der reellen, p -fach P -integrierbaren Funktionen auf (Ω, \mathfrak{F}) bezeichnet. Dieser Raum ist unter der L_p - (Halb-)Norm

$$\|X\|_p := (\mathbf{E}[|X|^p])^{\frac{1}{p}}$$

vollständig, d.h. ein (Pseudo-)Banachraum. Es wird hier der Begriff Halbnorm benutzt, da die Eigenschaft einer Norm, aus $\|X\|_p = 0$ $X = 0$ zu folgern, nur P -f.s. gilt.

Die aus dieser Halbnorm induzierte Abstandsfunktion

$$d_p(X, Y) := \|X - Y\|_p = (\mathbf{E}[|X - Y|^p])^{\frac{1}{p}}$$

heißt L_p - (Pseudo-)Metrik.

Für den Fall $p = 2$ wird durch

$$\langle X, Y \rangle := \mathbf{E}[XY]$$

ein Skalarprodukt definiert, welches obige Halbnorm wie folgt induziert:

$$\|X\|_2^2 = \langle X, X \rangle.$$

Damit ist $L_2(\Omega, \mathfrak{F}, P)$ ein (Pseudo-)Hilbertraum, auf welchem orthogonale Projektionen existieren.

Eine orthogonale Projektion $\mathfrak{P}_{\mathfrak{G}, X}$ von $X \in L_2(\Omega, \mathfrak{F}, P)$ auf den $L_2(\Omega, \mathfrak{G}, P)$ mit $\mathfrak{G} \subseteq \mathfrak{F}$, kann als eindeutige Lösung des Minimierungsproblems

$$\|X - \mathfrak{P}_{\mathfrak{G}, X}\|_2 = \min_{Y \in L_2(\Omega, \mathfrak{G}, P)} \|X - Y\|_2$$

angesehen werden.

Mit diesen Vorüberlegungen lässt sich nun die bedingte Erwartung konstruieren:

16.2 Satz (Konstruktion der bedingten Erwartung):

Seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum, \mathfrak{G} eine beliebige σ -Algebra mit $\mathfrak{G} \subseteq \mathfrak{F}$ und $X \in L_2(\Omega, \mathfrak{F}, P)$. Dann erfüllt die P -f.s. eindeutige orthogonale Projektion $\mathfrak{P}_{\mathfrak{G}, X}$ von X auf den $L_2(\Omega, \mathfrak{G}, P)$ die Eigenschaften 1') und 2').

Beweis:

Die Existenz, die Eindeutigkeit und die \mathfrak{G} -Messbarkeit der orthogonalen Projektion $\mathfrak{P}_{\mathfrak{G}, X}$ ergeben sich aus Bemerkung 16.1 b). Damit bleibt nur die Eigenschaft 2') zu zeigen.

Für orthogonale Projektionen gilt selbstverständlich:

$$\langle X - \mathfrak{P}_{\mathfrak{G}, X}, Y \rangle = \mathbf{E}[(X - \mathfrak{P}_{\mathfrak{G}, X}) \cdot Y] = 0 \quad \forall Y \in L_2(\Omega, \mathfrak{G}, P).$$

Daraus ergibt sich für alle $A \in \mathfrak{G}$:

$$\int_A (X - \mathbf{E}[X|\mathfrak{G}]) dP = \langle X - \mathbf{E}[X|\mathfrak{G}], 1_A \rangle = 0, \quad \text{da } 1_A \in L_2(\Omega, \mathfrak{G}, P).$$

16.3 Definition (bedingte Erwartung):

Es seien X eine Zufallsvariable auf $(\Omega, \mathfrak{F}, P)$ mit $\mathbf{E}[|X|] < \infty$ und \mathfrak{G} eine beliebige σ -Algebra mit $\mathfrak{G} \subseteq \mathfrak{F}$. Die in Satz 16.2 beschriebene P -fast sicher eindeutige orthogonale Projektion $\mathfrak{P}_{\mathfrak{G}, X}$ von X auf den $L_2(\Omega, \mathfrak{G}, P)$, die die Eigenschaften

(i) $\mathbf{E}[X|\mathfrak{G}]$ ist \mathfrak{G} -messbar,

(ii) $\int_A \mathbf{E}[X|\mathfrak{G}] dP = \int_A X dP$ für alle $A \in \mathfrak{G}$.

erfüllt, heißt bedingte Erwartung von X gegeben \mathfrak{G} und wird mit $\mathbf{E}[X|\mathfrak{G}]$ bezeichnet.

Die bedingte Erwartung $\mathbf{E}[X|\mathfrak{G}]$ ist demnach die beste Approximation von X im $L_2(\Omega, \mathfrak{G}, P)$. Da er P -f.s. eindeutig bestimmt ist, bezeichnet man jede zulässige Wahl auch als **Version** der bedingten Erwartung von X gegeben \mathfrak{G} .

Die bedingte Erwartung kann alternativ auch über den Satz von Radon - Nikodym konstruiert werden, dabei geht jedoch die Anschaulichkeit, dass $\mathbf{E}[X|\mathfrak{G}]$ die beste Approximation von X ist, verloren:

16.4 Definition (μ -absolut-stetig):

Wenn aus $\mu(A) = 0$ stets folgt, dass $\nu(A) = 0$ gilt, so heißt ν μ -absolut-stetig, in Zeichen $\nu \ll \mu$.

16.5 Satz (Radon-Nikodym):

Es seien μ und ν Maße auf (Ω, \mathfrak{F}) .

a) Ist μ σ -endlich, so sind folgende Aussagen äquivalent:

(i) ν besitzt eine Dichte f bezüglich μ .

(ii) ν ist μ -stetig und σ -endlich.

b) f ist dann μ -fast-überall eindeutig bestimmt.

Beweis:

Siehe H. Bauer: „Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie“, 5. Auflage, Walter de Gruyter, Berlin, 2002; Kapitel 17 „Maße und Dichten“.

16.6 Satz:

Es sei X Zufallsvariable auf $(\Omega, \mathfrak{F}, P)$ mit $\mathbf{E}[|X|] < \infty$. \mathfrak{G} sei Unter- σ -Algebra von \mathfrak{F} . Dann existiert $\mathbf{E}[X|\mathfrak{G}]$ und ist P -fast-überall eindeutig bestimmt.

Beweis:

Es wird zunächst angenommen, dass $X \geq 0$ P -fast-sicher gilt und es wird für alle $G \in \mathfrak{G}$ definiert:

$$Q(G) := \int_G X \, dP.$$

Damit ist Q nach Definition 16.3 ein Maß auf \mathfrak{G} mit $Q(\Omega) = \mathbf{E}[X] < \infty$. Sei nun $P_{\mathfrak{G}}$ die Einschränkung von P auf \mathfrak{G} . Es gilt $Q \ll P_{\mathfrak{G}}$. Jede Nullmenge bezüglich P ist auch Nullmenge bezüglich Q . Damit folgt aus dem Satz von Radon–Nikodym, dass Q eine $P_{\mathfrak{G}}$ -fast-überall eindeutige \mathfrak{G} -messbare Dichte f bezüglich $P_{\mathfrak{G}}$ hat, d.h. es ist

$$Q(G) = \int_G X \, dP = \int_G f \, dP_{\mathfrak{G}}.$$

Es ist dann $\mathbf{E}[X|\mathfrak{G}](\omega) := f(\omega)$ die fast überall eindeutig bestimmte bedingte Erwartung von X gegeben \mathfrak{G} . Da Q endlich ist, ist auch f endlich.

Sei nun X beliebig mit $\mathbf{E}[|X|] < \infty$. Dann ist $\mathbf{E}[X|\mathfrak{G}] := \mathbf{E}[X^+|\mathfrak{G}] - \mathbf{E}[X^-|\mathfrak{G}]$ die fast überall eindeutig bestimmte bedingte Erwartung von X gegeben \mathfrak{G} . Die Differenz ist \mathfrak{G} -messbar, da beide Funktionen dies sind. Die Integrale sind, wegen der σ -Additivität, gleich. Dies liefert (siehe z.B. H. Bauer: „Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie“, Kapitel 54) die Eindeutigkeit. ■

16.7 Beispiel:

Es seien $\Omega := \{\square, \bigcirc, \triangle, \times, \heartsuit\}$, $\mathfrak{F} := \mathfrak{P}(\Omega)$, $P := (\frac{1}{3}, \frac{1}{6}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$, $\mathfrak{G} := \sigma(\{\square, \bigcirc\}, \{\triangle, \times\} \mid \{\heartsuit\})$ und $X := (1, 2, \frac{1}{2}, -4, 15)$. Dann folgt:

$\mathbf{E}[X] = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} + (-4) \cdot \frac{1}{8} + 15 \cdot \frac{1}{8} = \frac{13}{8}$ sowie

$$\begin{aligned} \mathbf{E}[X|\mathfrak{G}](\square) &= \frac{1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6}}{\frac{1}{3} + \frac{1}{6}} = \frac{4}{3} = \mathbf{E}[X|\mathfrak{G}](\bigcirc), \\ \mathbf{E}[X|\mathfrak{G}](\times) &= \frac{\frac{1}{2} \cdot \frac{1}{4} - 4 \cdot \frac{1}{8}}{\frac{1}{4} + \frac{1}{8}} = -1 = \mathbf{E}[X|\mathfrak{G}](\triangle), \\ \mathbf{E}[X|\mathfrak{G}](\heartsuit) &= \frac{15 \cdot \frac{1}{8}}{\frac{1}{8}} = 15 \end{aligned}$$

Damit folgt insgesamt: $\mathbf{E}[X|\mathfrak{G}] = (\frac{4}{3}, \frac{4}{3}, -1, -1, 15)$ und $\mathbf{E}[\mathbf{E}[X|\mathfrak{G}]] = \frac{4}{3} \cdot (\frac{1}{3} + \frac{1}{6}) - 1 \cdot (\frac{1}{4} + \frac{1}{8}) + 15 \cdot \frac{1}{8} = \frac{13}{8} = \mathbf{E}[X]$.

16.8 Satz (Eigenschaften der bedingten Erwartung):

Es seien X, Y Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit $\mathbf{E}[|X|] < \infty$ und $\mathbf{E}[|Y|] < \infty$. Weiter sei $\mathfrak{G} \subseteq \mathfrak{F}$ eine σ -Unteralgebra von \mathfrak{F} .

- (i) Aus $X = c$ mit $c \in \mathbb{R}$ folgt $\mathbf{E}[X|\mathfrak{G}] = c$ P -f.s.
- (ii) Aus $X \leq Y$ P -f.s. folgt $\mathbf{E}[X|\mathfrak{G}] \leq \mathbf{E}[Y|\mathfrak{G}]$ P -f.s.
- (iii) Für $a, b \in \mathbb{R}$ gilt $\mathbf{E}[aX + bY|\mathfrak{G}] = a\mathbf{E}[X|\mathfrak{G}] + b\mathbf{E}[Y|\mathfrak{G}]$ P -f.s.
- (iv) Ist $\sigma(X)$ unabhängig von \mathfrak{G} , so folgt $\mathbf{E}[X|\mathfrak{G}] = \mathbf{E}[X]$.
- (v) Es gilt $\mathbf{E}[\mathbf{E}[X|\mathfrak{G}]] = \mathbf{E}[X]$.

Beweis:

- (i) Die Behauptung folgt aus der Tatsache, dass die konstante Funktion $X = c$ \mathfrak{G} -messbar ist.
- (ii) $X \leq Y$ P -f.s. impliziert

$$\int_A X dP \leq \int_A Y dP \quad \forall A \in \mathfrak{G}.$$

Hieraus ergibt sich aufgrund der Definition der bedingten Erwartung

$$\int_A \mathbf{E}[X|\mathfrak{G}] dP \leq \int_A \mathbf{E}[Y|\mathfrak{G}] dP \quad \forall A \in \mathfrak{G}.$$

Mit Hilfe von Stochastik I Satz 9.41 schließt man $\mathbf{E}[X|\mathfrak{G}] \leq \mathbf{E}[Y|\mathfrak{G}]$ P -f.s.

- (iii) Für alle $A \in \mathfrak{G}$ gilt:

$$\begin{aligned} \int_A [aX + bY] dP &= \int_A aX dP + \int_A bY dP = a \int_A X dP + b \int_A Y dP \\ &= a \int_A \mathbf{E}[X|\mathfrak{G}] dP + b \int_A \mathbf{E}[Y|\mathfrak{G}] dP \\ &= \int_A a\mathbf{E}[X|\mathfrak{G}] dP + \int_A b\mathbf{E}[Y|\mathfrak{G}] dP \\ &= \int_A [a\mathbf{E}[X|\mathfrak{G}] + b\mathbf{E}[Y|\mathfrak{G}]] dP. \end{aligned}$$

- (iv) Die Abbildung $Y: \Omega \rightarrow \mathbb{R}$ sei definiert durch $Y(\omega) := \mathbf{E}[X]$ für alle $\omega \in \Omega$. Nach (i) ist Y \mathfrak{G} -messbar und es gilt

$$\int_B Y dP = \mathbf{E}[X] \cdot P(B) \quad \forall B \in \mathfrak{G}.$$

Also:

$$\begin{aligned} \int_B \mathbf{E}[X|\mathfrak{G}] dP &= \int_B X dP = \int_{\Omega} X I_B dP = \mathbf{E}[X I_B] \\ &= \mathbf{E}[X] \cdot \mathbf{E}[I_B] = \mathbf{E}[X] \cdot P(B). \end{aligned}$$

Folglich gilt

$$\int_B \mathbf{E}[X|\mathfrak{G}] dP = \int_B Y dP$$

und deshalb $\mathbf{E}[X|\mathfrak{G}] = Y$ P -f.s.

(v) Es gilt

$$\mathbf{E}[\mathbf{E}[X|\mathfrak{G}]] = \int_{\Omega} \mathbf{E}[X|\mathfrak{G}] dP.$$

Wegen $\Omega \in \mathfrak{G}$ ist aber

$$\int_{\Omega} \mathbf{E}[X|\mathfrak{G}] dP = \int_{\Omega} X dP = \mathbf{E}[X].$$

Damit wird

$$\mathbf{E}[\mathbf{E}[X|\mathfrak{G}]] = \mathbf{E}[X].$$

■

16.9 Satz:

$(X_n)_{n \in \mathbb{N}}$ sei eine Folge von Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$ und $\mathfrak{G} \subseteq \mathfrak{F}$ sei eine Unter- σ -Algebra von \mathfrak{F} .

(i) Ist $(X_n)_{n \in \mathbb{N}}$ eine monoton wachsende Folge und gilt $X_n \geq 0$ für alle $n \in \mathbb{N}$, so gilt:

$$\sup_{n \in \mathbb{N}} \mathbf{E}[X_n|\mathfrak{G}] = \mathbf{E}\left[\sup_{n \in \mathbb{N}} X_n|\mathfrak{G}\right].$$

(ii) Ist $X_n \geq 0$ für alle $n \in \mathbb{N}$, so gilt:

$$\mathbf{E}\left[\sum_{n \in \mathbb{N}} X_n|\mathfrak{G}\right] = \sum_{n \in \mathbb{N}} \mathbf{E}[X_n|\mathfrak{G}].$$

(iii) Es sei Y eine auf Ω definierte reellwertige und P -integrierbare Zufallsvariable. Außerdem gelte $|X_n| \leq Y$ für alle $n \in \mathbb{N}$ und $X_n \xrightarrow[n \rightarrow \infty]{P\text{-f.s.}} X$. Dann gilt:

$$\lim_{n \rightarrow \infty} \mathbf{E}[X_n|\mathfrak{G}] = \mathbf{E}[X|\mathfrak{G}].$$

Beweis zu Satz 16.9:

Siehe G. Alsmeyer: „Wahrscheinlichkeitstheorie“, 3. Auflage, Skripten zur Mathematischen Statistik Nr. 30 der Westfälischen Wilhelms Universität Münster, Münster, 2003; Kapitel VIII „Bedingte Erwartungswerte und Verteilungen“.

16.10 Satz:

Es seien X, Y Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit $\mathbf{E}[|X|] < \infty$ und $\mathbf{E}[|Y|] < \infty$. Außerdem sei $\mathfrak{G} \subseteq \mathfrak{F}$ eine Unter- σ -Algebra von \mathfrak{F} und X sei \mathfrak{G} -messbar. Dann gilt:

$$\mathbf{E}[XY|\mathfrak{G}] = X \cdot \mathbf{E}[Y|\mathfrak{G}] \quad P - f.s.$$

Beweis:

- Wir stellen zunächst fest, dass $\mathbf{E}[Y|\mathfrak{G}]$ \mathfrak{G} -messbar ist. Der Beweis erfolgt so dann mit algebraischer Induktion. Für $X = I_B$, $B \in \mathfrak{G}$, gilt:

$$\begin{aligned} \int_A X \mathbf{E}[Y|\mathfrak{G}] dP &= \int_A I_B \mathbf{E}[Y|\mathfrak{G}] dP = \int_{A \cap B} \mathbf{E}[Y|\mathfrak{G}] dP \\ &= \int_{A \cap B} Y dP \quad (\text{Definition der bedingten Erwartung}) \\ &= \int_A I_B Y dP = \int_A XY dP \end{aligned}$$

- Wir setzen nun $X = \sum_{i=1}^n c_i I_{B_i}$ mit $n \in \mathbb{N}$, $c_i \in \mathbb{R}$ und $B_i \in \mathfrak{G}$, $i = 1, \dots, n$.

$$\begin{aligned} \int_A X \mathbf{E}[Y|\mathfrak{G}] dP &= \int_A \left(\sum_{i=1}^n c_i I_{B_i} \right) \mathbf{E}[Y|\mathfrak{G}] dP = \sum_{i=1}^n c_i \int_A I_{B_i} \mathbf{E}[Y|\mathfrak{G}] dP \\ &= \sum_{i=1}^n c_i \int_A I_{B_i} Y dP \quad (\text{Teil 1}) \\ &= \int_A \sum_{i=1}^n c_i I_{B_i} Y dP = \int_A XY dP. \end{aligned}$$

Für nichtnegative und allgemeine messbare Abbildungen X folgert man die Behauptung aus Satz 16.9 (i) und (iii). ■

16.11 Satz (Waldsche Gleichung):

Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und N und $(X_n)_{n \in \mathbb{N}}$ seien Zufallsvariablen über $(\Omega, \mathfrak{F}, P)$. N habe Werte in \mathbb{N} und $(X_n)_{n \in \mathbb{N}}$ sei eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit Werten in \mathbb{R} . Außerdem seien N und X_j , $j \in \mathbb{N}$, unabhängig. Es sei

$$S(\omega) = \sum_{i=1}^{N(\omega)} X_i(\omega) \quad \forall \omega \in \Omega.$$

Dann gilt:

- $\mathbf{E}[S] = \mathbf{E}[N] \cdot \mathbf{E}[X_1]$, sofern $\mathbf{E}[N] < \infty$ und $\mathbf{E}[X_1] < \infty$.
- $\mathbf{Var}[S] = \mathbf{E}[N] \cdot \mathbf{Var}[X_1] + \mathbf{Var}[N] (\mathbf{E}[X_1])^2$, sofern $\mathbf{Var}[N] < \infty$ und $\mathbf{Var}[X_1] < \infty$.

Beweis:

(i) Wir wenden Satz 16.8 an, wonach

$$\mathbf{E}[S] = \mathbf{E}[\mathbf{E}[S|N]] = \mathbf{E}[\mathbf{E}[S|N^{-1}(\mathfrak{B})]]$$

gilt. $N^{-1}(\mathfrak{B})$ wird durch die abzählbar vielen Mengen $A_i := \{\omega | N(\omega) = i\}$, $i \in \mathbb{N}$, mit $\cup_{i \in \mathbb{N}} A_i = \Omega$ erzeugt.

$$\begin{aligned} \mathbf{E}[S] &= \mathbf{E}[\mathbf{E}[S|N]] = \mathbf{E}[\mathbf{E}[S|N^{-1}(\mathfrak{B})]] = \mathbf{E}\left[\sum_{\substack{i \in \mathbb{N} \\ P(A_i) > 0}} \mathbf{E}[S|A_i] I_{A_i}\right] \\ &= \mathbf{E}\left[\sum_{i \in \mathbb{N}} \sum_{j=1}^i \mathbf{E}[X_j|A_i] I_{A_i}\right] = \mathbf{E}\left[\sum_{i \in \mathbb{N}} i \mathbf{E}[X_1] I_{A_i}\right] = \mathbf{E}[X_1 \cdot N] \\ &= \mathbf{E}[X_1] \cdot \mathbf{E}[N]. \end{aligned}$$

(ii) Der Beweis verläuft ähnlich wie Teil (i). ■

16.2 (Regulär) bedingte Verteilung

Ausgehend von Definition 16.3 der bedingten Erwartung lässt sich nun auch der Begriff der bedingten Wahrscheinlichkeit verallgemeinern.

16.12 Definition (bedingte Wahrscheinlichkeit):

Es sei $\mathfrak{G} \subset \mathfrak{F}$ eine Unter- σ -Algebra. Die bedingte Wahrscheinlichkeit von $A \in \mathfrak{F}$ gegeben \mathfrak{G} ist definiert als

$$P(A|\mathfrak{G}) := \mathbf{E}[I_A|\mathfrak{G}].$$

Sie existiert in verschiedenen Versionen und ist eine \mathfrak{G} -messbare Zufallsgröße.

Es drängt sich die Frage auf, ob $P(\cdot|\mathfrak{G})(\omega)$ für alle ω ein Wahrscheinlichkeitsmaß darstellt. Ein Antwort darauf liefert der folgende Satz:

16.13 Satz:

Die bedingten Wahrscheinlichkeiten haben die folgenden Eigenschaften *P-fast-sicher*:

$$(i) \quad P(A|\mathfrak{G}) \geq 0,$$

$$(ii) \quad P(\Omega|\mathfrak{G}) = 1,$$

$$(iii) \quad (A_i)_{i=1}^n \text{ paarweise disjunkt} \implies P\left(\bigcup_{i=1}^{\infty} A_i \middle| \mathfrak{G}\right) = \sum_{i=1}^{\infty} P(A_i|\mathfrak{G}).$$

Beweis:

Aussage (i) ergibt sich unmittelbar aus Satz 16.8, während Aussage (ii) bereits der Definition 16.3 zu entnehmen ist. Es bleibt also die σ -Additivität zu zeigen:

Es seien A_1, A_2, \dots disjunkte Elemente aus \mathfrak{F} . Es wird

$$Y_n := \sum_{i=1}^n I_{A_i} \quad \text{und} \quad Y := \sum_{i=1}^{\infty} I_{A_i}$$

definiert. Somit gilt $Y_n \uparrow Y$ und es folgt fast-sicher

$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| \mathfrak{G}\right) = \mathbf{E}\left[\sum_{i=1}^{\infty} I_{A_i} \middle| \mathfrak{G}\right] = \sum_{i=1}^{\infty} \mathbf{E}[I_{A_i} | \mathfrak{G}] = \sum_{i=1}^{\infty} P(A_i | \mathfrak{G}).$$

■

Es ergibt sich jetzt das Problem, ob (iii) für beliebige, d.h. im Allgemeinen für überabzählbar viele Folgen, erfüllt ist. Es könnte aber sein, dass die Vereinigung über alle Ausnahmevereinigungen in (iii) keine Nullmenge mehr ist. Es gibt Fälle, in denen man $P(\cdot | \mathfrak{G})$ nicht so definieren kann, dass $P(\cdot | \mathfrak{G})(\omega)$ für fast alle $\omega \in \Omega$ ein Wahrscheinlichkeitsmaß ist. So kommt man zu dem Begriff der regulär bedingten Verteilungen.

16.14 Definition (regulär bedingte Wahrscheinlichkeit, regulär bedingte Verteilung):

a) Eine Version von $P(\cdot | \mathfrak{G})(\omega)$ heißt regulär bedingte Wahrscheinlichkeit, wenn sie für fast alle $\omega \in \Omega$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathfrak{G}, P_{\mathfrak{G}})$ ist.

b) Es seien $Y: (\Omega, \mathfrak{F}, P) \rightarrow (\Omega', \mathfrak{F}')$ messbar und $\mathfrak{G} \subset \mathfrak{F}$ Unter- σ -Algebra. Dann heißt

$$\hat{P}(\cdot | \mathfrak{G})(\cdot): \mathfrak{F}' \times \Omega \rightarrow \mathbb{R}$$

regulär bedingte Verteilung von Y unter \mathfrak{G} , wenn

(i) für jedes $F' \in \mathfrak{F}'$ $\hat{P}(F' | \mathfrak{G})$ eine Version von $P(Y \in F' | \mathfrak{G})$ ist.

(ii) für jedes $\omega \in \Omega$ $\hat{P}(\cdot | \mathfrak{G})(\omega)$ ein Wahrscheinlichkeitsmaß auf (Ω', \mathfrak{F}') ist.

Die Existenz regulär bedingter Verteilungen ergibt sich aus dem folgenden Satz:

16.15 Satz:

a) Es sei (Ω', \mathfrak{F}') ein Borelraum, d.h. ein vollständiger, separabler, metrischer Raum, dann existieren reguläre bedingte Verteilungen.

b) $(\mathbb{R}^n, \mathfrak{B}^n)$ und $(\mathbb{R}^\infty, \mathfrak{B}^\infty)$ sind Borelräume.

Beweis des Satzes 16.15:

Siehe H. Bauer: „Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie“, 5. Auflage, Walter de Gruyter, Berlin, 2002; Kapitel 17 „Maße und Dichten“.

Im Folgenden werden stets die Zufallsvariablen $X, Y: \Omega \rightarrow \overline{\mathbb{R}}$ mit $\mathbf{E}[|X|] < \infty$ und $\mathbf{E}[|Y|] < \infty$ betrachtet. Für diesen Fall lassen sich die Begriffe des bedingten Erwartungswertes und der bedingten Verteilung vereinfacht darstellen.

16.16 Beispiel:

1. Für die Produktion von speziellen Baugruppen werden verschiedene mit $j = 1, \dots, k$ durchnummerierte Typen von Bauteilen benötigt. Es werden zwei Zufallsvariablen eingeführt:

- Für die Wahrscheinlichkeit, dass ein Bauteil vom Typ j benötigt wird, wird $P(Y = j)$, $j \in E := \{1, \dots, k\}$ geschrieben.
- Die technische Eigenschaft der Bauteile werde durch die reellwertige Zufallsvariable X beschrieben.

Insbesondere hat man für jeden Typ j von Bauteil Kenntnis von

$$P(X \in A \mid Y = j) \quad \forall A \in \mathfrak{B}, j \in E.$$

Von jedem Typ j von Bauteil wird die Einhaltung eines Toleranzbereiches C_j gefordert. Es soll die Wahrscheinlichkeit $P((X, Y) \in C)$ berechnen werden, dass ein beliebiges für die Produktion angefordertes Bauteil die geforderte Spezifikation erfüllt. Dazu wird das Ereignis $\{(X, Y) \in C\}$ in der Form

$$\{(X, Y) \in C\} = \bigcup_{j \in E} \{X \in C_j, Y = j\}$$

zerlegt. Da die Mengen $\{X \in C_j, Y = j\}$, $j \in E$, paarweise disjunkt sind, gilt

$$P((X, Y) \in C) = P\left(\bigcup_{j \in E} \{X \in C_j, Y = j\}\right) = \sum_{j \in E} P(X \in C_j, Y = j),$$

wofür man auch

$$P((X, Y) \in C) = \sum_{j \in E} P(X \in C_j \mid Y = j) \cdot P(Y = j)$$

schreiben kann. Bezeichnet P_Y das Bildmaß von P unter Y , so kann man den letzten Ausdruck auch durch die Gleichung

$$P((X, Y) \in C) = \int_E P(X \in C_y \mid Y = y) dP_Y(y) \quad (16.4)$$

ersetzen. Diese Gleichung lässt sich in zwei Richtungen interpretieren:

Die zweidimensionale Verteilung $P((X, Y) \in \cdot)$ lässt sich mit Hilfe von $P(X \in C_y \mid Y = y)$ und P_Y erzeugen. Umgekehrt existiert zu $P((X, Y) \in C)$ und P_Y eine Funktion $P(X \in C_y \mid Y = y)$, so dass (16.4) gilt.

2. Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum. Besitzt der Zufallsvektor $(X, Y): \Omega \rightarrow \mathbb{R}^2$ eine Dichte f und gilt für die Randverteilungsdichte f_Y von f bzgl. Y $f_Y(y) > 0 \quad \forall y \in \mathbb{R}$, so kann die Funktion $P(X \in C_y \mid Y = y)$ mit

$$P((X, Y) \in C) := \int_{\mathbb{R}} P(X \in C_y \mid Y = y) dP_Y(y) \quad (16.5)$$

für jede Menge $C \in \mathfrak{B}^2$ und ihren korrespondierenden y -Schnitt $C_y := \{x \in \mathbb{R} \mid (x, y) \in C\}$ ebenso leicht konstruiert werden. Als Beispiel wird

$$C := \{(x, y) \in \mathbb{R}^2 \mid x \leq a, y \leq b\}$$

gewählt. Der zugehörige y -Schnitt C_y ist dann

$$C_y := \{x \mid (x, y) \in C\} = \begin{cases} \{x \mid x \leq a\} , & y \leq b \\ \emptyset & , \quad y > b. \end{cases}$$

Deshalb gilt:

$$P(X \in C_y \mid Y = y) = \begin{cases} P(X \leq a \mid Y = y) , & y \leq b \\ 0 & , \quad y > b. \end{cases}$$

Damit wird

$$\begin{aligned} \int_{\mathbb{R}} P(X \in C_y \mid Y = y) dP_Y(y) &= \int_{-\infty}^b P(X \in C_y \mid Y = y) dP_Y(y) \\ &\quad + \int_b^{\infty} P(X \in C_y \mid Y = y) dP_Y(y) \\ &= \int_{-\infty}^b P(X \leq a \mid Y = y) dP_Y(y) \end{aligned}$$

Es wird versuchsweise

$$P(X \leq x \mid Y = y) := \frac{1}{f_Y(y)} \int_{-\infty}^x f(s, y) ds$$

gesetzt. Damit ergibt sich

$$\begin{aligned} \int_{\mathbb{R}} P(X \in C_y \mid Y = y) dP_Y(y) &= \int_{-\infty}^b \left(\frac{1}{f_Y(y)} \int_{-\infty}^a f(s, y) ds \right) dP_Y(y) \\ &= \int_{-\infty}^b \left(\frac{1}{f_Y(y)} \int_{-\infty}^a f(s, y) ds \right) f_Y(y) dy \\ &= \int_{-\infty}^b \left(\int_{-\infty}^a f(s, y) ds \right) dy \\ &= \int_{-\infty}^a \int_{-\infty}^b f(s, y) dy ds \\ &= P(X \leq a, Y \leq b). \end{aligned}$$

Der Ausdruck

$$f_{X|Y}(x|y) := \frac{f(x, y)}{f_Y(y)}$$

macht zunächst nur für $f_Y(y) > 0$ einen Sinn. Dies ist jedoch keine wesentliche Einschränkung, da die Ausnahmemenge $N := \{(x, y) \in \mathbb{R} \mid f_Y(y) = 0\}$ eine P -Nullmenge ist. Denn es gilt:

$$\begin{aligned} P(N) &= \iint_N f(x, y) \, dx \, dy = \int_{\{y \mid f_Y(y)=0\}} \int_{\mathbb{R}} f(x, y) \, dx \, dy \\ &= \int_{\{y \mid f_Y(y)=0\}} f_Y(y) \, dy = 0. \end{aligned}$$

Auf dieser Ausnahmemenge N kann der Wert von $f_{X|Y}(x|y)$ also beliebig abgeändert werden, ohne dass sich dadurch der Ausdruck (16.5) verändern würde. Es wird weiter gezeigt, dass $f_{X|Y}(x|y)$ eine Dichte auf \mathbb{R} definiert. Da offensichtlich $f_{X|Y}(x|y) \geq 0$ gilt, bleibt nur noch

$$\int_{\mathbb{R}} f_{X|Y}(x|y) \, dx = 1$$

zu zeigen. Es gilt:

$$\begin{aligned} \int_{\mathbb{R}} f_{X|Y}(x|y) \, dx &= \int_{\mathbb{R} \setminus N} f_{X|Y}(x|y) \, dx + \int_N f_{X|Y}(x|y) \, dx \\ &= \int_{\mathbb{R} \setminus N} \frac{f(x, y)}{f_Y(y)} \, dx = \frac{1}{f_Y(y)} \int_{\mathbb{R} \setminus N} f(x, y) \, dx \\ &= \frac{1}{f_Y(y)} \cdot f_Y(y) = 1. \end{aligned}$$

Deswegen definiert man:

16.17 Definition (bedingte Dichte, bedingte Verteilung):

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X, Y): \Omega \rightarrow \mathbb{R}^2$ ein Zufallsvektor. (X, Y) besitze die Dichte $f(x, y)$ und die Randverteilungsdichte von f bzgl. Y sei f_Y .

a) Die Funktion

$$f_{X|Y}(x|y) := \begin{cases} \frac{f(x, y)}{f_Y(y)}, & \text{für } (x, y) \in \mathbb{R}^2 \text{ mit } f_Y(y) > 0 \\ 0, & \text{für } (x, y) \in \mathbb{R}^2 \text{ mit } f_Y(y) = 0 \end{cases}$$

heißt eine bedingte Dichte von X gegeben $Y = y$.

b) Für $y \in \mathbb{R}$ und $A \in \mathfrak{B}$ heißt die Abbildung

$$P(X \in A \mid Y = y) := \begin{cases} \int_A f_{X|Y}(x|y) \, dx, & f_Y(y) > 0 \\ 0, & f_Y(y) = 0 \end{cases}$$

die bedingte Verteilung von X gegeben $Y = y$. Anstelle von $P(X \in A \mid Y = y)$ schreibt man auch $P_{X|Y=y}(A)$.

16.18 Beispiel (bedingte Dichte):

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X, Y): \Omega \rightarrow \mathbb{R}^2$ ein stetiger Zufallsvektor mit einer Dichte

$$f(x, y) := \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right), \quad (x, y) \in \mathbb{R}^2.$$

1. Es wird zunächst gezeigt, dass f eine Dichte definiert. Es gilt $f(x, y) \geq 0$ für alle $(x, y) \in \mathbb{R}^2$ und es ist

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \, dx \, dy &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right) \, dx \, dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x^2 - 2xy + y^2)\right) \, dx \exp\left(-\frac{1}{2}y^2\right) \, dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x - y)^2\right) \, dx \exp\left(-\frac{1}{2}y^2\right) \, dy \\ &\stackrel{(*)}{=} \frac{1}{2\pi} \int_{\mathbb{R}} \sqrt{2\pi} \exp\left(-\frac{1}{2}y^2\right) \, dy \\ &\stackrel{(*)}{=} \frac{1}{\sqrt{2\pi}} \cdot \int_{\mathbb{R}} \exp\left(-\frac{1}{2}y^2\right) \, dy \\ &= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1. \end{aligned}$$

(*): Es ist $\int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x - y)^2\right) \, dx = \sqrt{2\pi}$ bzw. $\int_{\mathbb{R}} \exp\left(-\frac{1}{2}y^2\right) \, dy = \sqrt{2\pi}$, denn es gilt:

$$\begin{aligned} \left(\int_{\mathbb{R}} \exp\left(-\frac{t^2}{2}\right) \, dt\right)^2 &= \int_{\mathbb{R}} \int_{\mathbb{R}} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right) \, dx_1 \, dx_2 \\ &= \int_0^{2\pi} \int_0^\infty \exp\left(-\frac{r^2}{2}\right) \cdot r \, dr \, d\varphi \\ &= 2\pi \left[-\exp\left(-\frac{r^2}{2}\right)\right]_0^\infty = 2\pi. \end{aligned}$$

2. Die Randverteilungsdichte f_X von f bzgl. X lautet:

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right) \, dy \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \frac{x^2}{2}\right) \cdot \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \left(\frac{x^2}{2} - 2xy + 2y^2\right)\right) \, dy \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \frac{x^2}{2}\right) \cdot \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \left(\sqrt{2}y - \frac{x}{\sqrt{2}}\right)^2\right) \, dy. \end{aligned}$$

Mit $z := \sqrt{2}y - \frac{x}{\sqrt{2}}$ ergibt sich:

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \frac{x^2}{2}\right) \cdot \underbrace{\frac{1}{\sqrt{2}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2} z^2\right) dz}_{=\sqrt{2\pi}, \text{ s.o.}} \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \frac{x^2}{2}\right) \cdot \frac{\sqrt{2\pi}}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{x^2}{2}\right), \quad x \in \mathbb{R}. \end{aligned}$$

3. Die Randverteilungsdichte f_Y von f bzgl. Y ist:

$$\begin{aligned} f_Y(y) &= \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right) dx \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x^2 - 2xy + y^2)\right) \exp\left(-\frac{1}{2}y^2\right) dx \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}y^2\right) \cdot \underbrace{\int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x - y)^2\right) dx}_{=\sqrt{2\pi}, \text{ s.o.}} \\ &= \frac{\sqrt{2\pi}}{2\pi} \cdot \exp\left(-\frac{1}{2}y^2\right) \\ &= \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}y^2\right), \quad y \in \mathbb{R}. \end{aligned}$$

4. X und Y sind nicht unabhängig, da $f(x, y) \neq f_X(x) \cdot f_Y(y)$ ist.

5. Die bedingte Dichte $f_{X|Y}(x|y)$ ergibt sich zu

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x, y)}{f_Y(y)} = \frac{(2\pi)^{-1} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right)}{(\sqrt{2\pi})^{-1} \exp\left(-\frac{1}{2}y^2\right)} \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x - y)^2\right), \quad x \in \mathbb{R}. \end{aligned}$$

16.19 Satz:

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X, Y): \Omega \rightarrow \mathbb{R}^2$ ein Zufallsvektor mit der Dichte f . Die Randverteilungsdichte von f bzgl. Y sei f_Y . Dann gilt für alle $A, B \in \mathfrak{F}$ und $C \in \mathfrak{B}^2$:

$$(i) \quad P(X \in A, Y \in B) = \int_B P(X \in A \mid Y = y) f_Y(y) dy.$$

$$(ii) \quad P(X \in A) = \int_{\mathbb{R}} P(X \in A \mid Y = y) f_Y(y) dy.$$

- (iii) $P((X, Y) \in C) = \int P(X \in C_y \mid Y = y) f_Y(y) dy$,
 wobei $C_y := \{x \in \mathbb{R} \mid (x, y) \in C\}$ der y -Schnitt von C bedeutet.

Beweis:

- (i) Es sei wieder $N := \{(x, y) \in \mathbb{R}^2 \mid f_Y(y) = 0\}$. Unter Beachtung der Beziehung $P(Y \in B \setminus N) = P(Y \in B) - P(Y \in N) = P(Y \in B)$ erhält man

$$\begin{aligned} \int_B P(X \in A \mid Y = y) f_Y(y) dy &= \int_{B \setminus N} \left(\int_A f_{X|Y}(x|y) dx \right) f_Y(y) dy \\ &= \int_{B \setminus N} \left(\int_A \frac{f(x, y)}{f_Y(y)} dx \right) f_Y(y) dy \\ &= \int_{A \times B \setminus N} f(x, y) dx dy \\ &= P_{(X, Y)}(A \times B \setminus N) \\ &= P(X \in A, Y \in B \setminus N). \end{aligned}$$

- (ii) Man ersetze in (i) B durch \mathbb{R} .

- (iii) Die Menge C lässt sich als disjunkte Vereinigung von Mengen $C_j := A_j \times B_j$ darstellen, da \mathfrak{B}^2 von den Rechtecken erzeugt wird. Für $y \in \mathbb{R}$ gilt dann

$$\begin{aligned} C_y &= \left\{ x \in \mathbb{R} \mid (x, y) \in \bigcup_{j \in \mathbb{N}} C_j \right\} \\ &= \bigcup_{j \in \mathbb{N}} \{x \in \mathbb{R} \mid (x, y) \in C_j\} \\ &= \bigcup_{j \in \mathbb{N}} (C_j)_y \\ &= \bigcup_{\substack{j \in \mathbb{N} \\ y \in B_j}} A_j \\ &= \bigcup_{\substack{j \in \mathbb{N} \\ y \in B_j}} (A_j \times B_j) \end{aligned}$$

und das zeigt

$$P((X, Y) \in C) = \sum_{j \in \mathbb{N}} P((X, Y) \in C_j) = \sum_{j \in \mathbb{N}} P(X \in A_j, Y \in B_j)$$

sowie

$$\begin{aligned}
 \int_{\mathbb{R}} P(X \in C_y | Y = y) f_y(y) dy &= \int_{\mathbb{R}} \int_{C_y} f_{X|Y}(x|y) dx f_y(y) dy \\
 &= \int_{\mathbb{R}} \sum_{\substack{j \in \mathbb{N} \\ y \in B_j}} \int_{(C_j)_y} f_{X|Y}(x|y) dx f_y(y) dy \\
 &= \sum_{j \in \mathbb{N}} \int_{B_j} \int_{A_j} f_{X|Y}(x|y) dx f_y(y) dy \\
 &= \sum_{j \in \mathbb{N}} \int_{B_j} P(X \in A_j | Y = y) f_y(y) dy \\
 &= \sum_{j \in \mathbb{N}} P(X \in A_j, Y \in B_j).
 \end{aligned}$$

■

16.20 Bemerkung:

Ebenso wie im diskreten Fall lassen sich auch im stetigen Fall die vorangegangenen Betrachtungen auf Ausdrücke wie z.B. $P(X \in A | Y = y, Z = z)$ oder $P(X_{n+1} \in A_1, \dots, X_{n+k} \in A_k | X_0 = x_0, \dots, X_n = x_n)$ ausdehnen. Die entsprechenden Aussagen ergeben sich analog zum zweidimensionalen Fall.

16.21 Satz:

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X, Y): \Omega \rightarrow \mathbb{R}^2$ ein Zufallsvektor mit der Dichte f . $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ sei $(\mathfrak{B}^2, \mathfrak{B})$ -messbar.

a) Für alle $A \in \mathfrak{B}$ gilt:

$$P(h(X, Y) \in A | Y = y) = P(h(X, y) \in A | Y = y)$$

für P_Y -fast alle $y \in \mathbb{R}$, wobei P_Y die Randverteilung von (X, Y) bzgl. Y bedeutet.

b) Sind $h(X, Y)$ und Y stochastisch unabhängig, dann gilt

$$P(h(X, Y) \in A | Y = y) = P(h(X, y) \in A) \quad P_Y\text{-f.s.}$$

Beweis:

a) Es wird zunächst gezeigt, dass $Z := (X, Y)$ und Y eine gemeinsame Dichte $g: \mathbb{R}^3 \rightarrow \mathbb{R}^+$ mit

$$g(x, y', y) := \begin{cases} f(x, y) & , \quad y = y' \\ 0 & , \quad \text{sonst} \end{cases}$$

besitzen. Für $C := C_1 \times C_2 \in \mathfrak{B}^2$ und $B \in \mathfrak{B}$ gilt:

$$\begin{aligned}
 \int_B \left(\int_C g(x, y', y) d(x, y') \right) dy &= \int_B \left(\int_{C_y} f(x, y) dx \right) dy \\
 &= \int_{B \cap C_2} \left(\int_{C_1} f(x, y) dx \right) dy \\
 &= P((X, Y) \in C_1 \times (B \cap C_2)) \\
 &= P((X, Y) \in C_1 \times C_2, Y \in B) \\
 &= P(Z \in C, Y \in B).
 \end{aligned}$$

Die bedingte Dichte von Z bzgl. Y lautet:

$$\begin{aligned}
 g_{Z|Y}((x, y')|y) &= \frac{g(x, y', y)}{f_Y(y)} \\
 &= \begin{cases} \frac{f(x, y)}{f_Y(y)}, & y = y' \\ 0, & \text{sonst} \end{cases} \\
 &= \begin{cases} f_{X|Y}(x|y), & y = y' \\ 0, & \text{sonst.} \end{cases}
 \end{aligned}$$

Damit erhält man für $C \in \mathfrak{B}^2$:

$$\begin{aligned}
 P((X, Y) \in C \mid Y = y) &= \int_C g_{Z|Y}((x, y')|y) d(x, y') \\
 &= \int_{C_y} f_{X|Y}(x|y) dx \\
 &= P(X \in C_y \mid Y = y).
 \end{aligned}$$

Da aber

$$X \in C_y = \{x \in \mathbb{R} \mid (x, y) \in C\} \Leftrightarrow (x, y) \in C$$

ist, folgt

$$P(X \in C_y \mid Y = y) = P((X, y) \in C \mid Y = y).$$

Indem man $C := h^{-1}(A)$ wählt, erhält man die Behauptung.

- b) Mit a) bleibt zu zeigen, dass $P(h(X, y) \in A \mid Y = y) = P(h(X, y) \in A)$ P_Y -f.s. gilt wenn $h(X, y)$ und Y stochastisch unabhängig sind.

Es sei zunächst $f_Y(y) > 0$. Dann gilt:

$$\begin{aligned}
 P(h(X, y) \in A \mid Y = y) &= \int_A f_{X,Y}(h(x, y)|y) dx = \int_A \frac{f(h(x, y), y)}{f_Y(y)} dx \\
 &= \int_A \frac{f_X(h(x, y)) \cdot f_Y(y)}{f_Y(y)} dx = \int_A f_X(h(x, y)) dx \\
 &= \int_A \int_{\mathbb{R}} f(h(x, y), \tilde{y}) d\tilde{y} dx = P(h(X, y) \in A, Y \in \mathbb{R}) \\
 &= P(h(X, y) \in A).
 \end{aligned}$$

Der Fall $f_Y(y) = 0$ tritt aber P_Y -f.s. nicht auf, da $\{y | f_Y(y) = 0\}$ eine P_Y -Nullmenge ist, denn:

$$P_Y(\{y | f_Y(y) = 0\}) = \int_{\{y | f_Y(y) = 0\}} f_Y(y) dy = 0.$$

■

16.22 Beispiel:

In einen Produktionsprozess seien zwei gleichartige Maschinen integriert, deren Ausfallzeiten (Zeitspanne zwischen zwei aufeinanderfolgenden Ausfällen) X_1, X_2 stochastisch unabhängige und exponentiell verteilte Zufallsgrößen mit den Parametern $\alpha_1 = \alpha_2 = \alpha > 0$ sind. Nachdem eine dieser Maschinen ausgefallen ist und repariert wird, übernimmt die jeweils andere Maschine die anstehende Arbeit. Aufgrund einer statistischen Erhebung ist bekannt, dass die Reparaturzeit Y exponentiell verteilt ist mit dem Parameter $\beta > 0$, wobei unterstellt wird, dass die Zufallsgrößen X_1, X_2 und Y stochastisch unabhängig sind.

O.B.d.A. wird angenommen, dass die erste Maschine zu einem Zeitpunkt $t = 0$ nicht mehr in Betrieb genommen werden kann und es wird nach der Wahrscheinlichkeit, dass während der Reparatur dieser Maschine nicht auch noch die andere ausfällt, gefragt. Da die Restlaufzeit der zweiten Maschine ebenfalls exponentiell verteilt ist mit dem Parameter $\alpha > 0$, ist die gesuchte Wahrscheinlichkeit

$$P(X_2 > Y) = \int_{\mathbb{R}} P(X_2 > Y | Y = y) f_Y(y) dy.$$

Unter Berücksichtigung des Satzes 16.21 erhält man mit $h: \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, y) := x - y$

$$\begin{aligned} P(X_2 > Y) &= P(X_2 - Y > 0) \\ &= \int_{\mathbb{R}^+} P(X_2 - Y > 0 | Y = y) \cdot \beta \cdot e^{-\beta y} dy \\ &= \int_{\mathbb{R}^+} P(X_2 - y > 0) \cdot \beta \cdot e^{-\beta y} dy \\ &= \int_{\mathbb{R}^+} P(X_2 > y) \cdot \beta \cdot e^{-\beta y} dy = \int_{\mathbb{R}^+} e^{-\alpha y} \cdot \beta \cdot e^{-\beta y} dy \\ &= \beta \cdot \int_{\mathbb{R}^+} e^{-(\alpha+\beta)y} dy = \beta \cdot \left(-\frac{1}{\alpha + \beta} \right) e^{-(\alpha+\beta)y} \Big|_0^\infty \\ &= \frac{\beta}{\alpha + \beta}. \end{aligned}$$

16.23 Beispiel (Warum steht man so oft in der falschen Warteschlange?):

Stehen für die Abfertigung wie im Supermarkt oder bei der Bahn mehrere Kassen zur Auswahl, so neigt man dazu, diejenige mit der kürzesten Warteschlange auszuwählen. Doch offensichtlich garantiert diese Strategie nicht, dass man auch schneller abgefertigt wird. So entsteht der Eindruck, in der „falschen Warteschlange“ zu stehen. Da die Arbeitsaufträge der einzelnen Kunden zufällig schwanken, könnte es sein, dass in der langen Warteschlange

zufällig viele kleine Aufträge akkumuliert sind, während in der kurzen Warteschlangen große Aufträge vorherrschen.

Als Beispiel werden zwei gleichartige Bediensysteme betrachtet. In dem einen warten zum Zeitpunkt der Ankunft m , in dem anderen n Kunden, wobei angenommen werden soll, dass m kleiner oder gleich n ist. Stehen m Kunden in der Schlange und sind die aufeinander folgenden Bedienzeiten der Kunden stochastisch unabhängige und mit dem Parameter $\mu > 0$ exponentiell verteilte Zufallsvariablen, dann ist die mittlere Wartezeit $W_{(m)}$ Erlang-verteilt mit den Parametern μ und m :

$$P(W_{(m)} > t) = e^{-\mu t} \cdot \sum_{k=0}^{m-1} \frac{(\mu t)^k}{k!} \quad (t \geq 0)$$

(siehe Stochastik I, Satz 11.35).

Damit läuft die Fragestellung auf den Vergleich zweier Erlang-verteilter Zufallsgrößen hinaus, wobei die eine Erlang-verteilt ist mit den Parametern μ und m und die andere Erlang-verteilt ist mit den Parametern μ und n . Gefragt ist also nach der Wahrscheinlichkeit, dass $W_{(m)}$ größer ist als $W_{(n)}$, in Zeichen: $P(W_{(m)} > W_{(n)})$. Diese Wahrscheinlichkeit kann mit dem Instrumentarium der bedingten Verteilung ausgerechnet werden. Indem man die Gesetzmäßigkeiten der Erlang-Verteilung ausnutzt, und berücksichtigt, dass für eine mit dem Parameter μ exponentiell verteilte Zufallsgröße Z gerade $\mathbf{E}[Z^\ell] = \ell!/\mu^\ell$ gilt, erhält man

$$\begin{aligned} P(W_{(m)} > W_{(n)}) &= \int_{\mathbb{R}^+} P(W_{(m)} > W_{(n)} \mid W_{(n)} = t) \cdot f_{W_{(n)}}(t) dt \\ &= \int_{\mathbb{R}^+} e^{-\mu t} \sum_{k=0}^{m-1} \frac{(\mu t)^k}{k!} \cdot \frac{\mu^n \cdot t^{n-1}}{(n-1)!} e^{-\mu t} dt \\ &= \frac{1}{2} \frac{\mu^{n-1}}{(n-1)!} \int_{\mathbb{R}^+} \sum_{k=0}^{m-1} \frac{(\mu t)^k}{k!} \cdot t^{n-1} \cdot 2\mu \cdot e^{-2\mu t} dt \\ &= \frac{1}{2} \frac{\mu^{n-1}}{(n-1)!} \left[\sum_{k=0}^{m-1} \frac{\mu^k}{k!} \int_{\mathbb{R}^+} t^{n+k-1} \cdot 2\mu \cdot e^{-2\mu t} dt \right] \\ &= \frac{1}{2} \frac{\mu^{n-1}}{(n-1)!} \sum_{k=0}^{m-1} \frac{\mu^k}{k!} \cdot \mathbf{E}[Z^{n+k-1}], \end{aligned}$$

wobei Z eine mit den Parametern 2μ exponentiell verteilte Zufallsgröße darstellt.

$$\begin{aligned} P(W_{(m)} > W_{(n)}) &= \frac{1}{2} \frac{\mu^{n-1}}{(n-1)!} \sum_{k=0}^{m-1} \frac{\mu^k}{k!} \cdot \mathbf{E}[Z^{n+k-1}] \\ &= \frac{1}{2} \frac{\mu^{n-1}}{(n-1)!} \sum_{k=0}^{m-1} \frac{\mu^k}{k!} \cdot \frac{(n+k-1)!}{(2\mu)^{n+k-1}} \\ &= \frac{1}{2} \sum_{k=0}^{m-1} \frac{(n+k-1)!}{(n-1)!k!} \left(\frac{1}{2}\right)^{n+k-1} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \sum_{k=0}^{m-1} \binom{n+k-1}{n-1} \left(\frac{1}{2}\right)^{n+k-1} \\
 &= \left(\frac{1}{2}\right)^n \sum_{k=0}^{m-1} \binom{n+k-1}{k} \left(\frac{1}{2}\right)^k.
 \end{aligned}$$

Wertet man diese Formel für verschiedene m und n aus, so ergibt sich Abbildung 16.1.

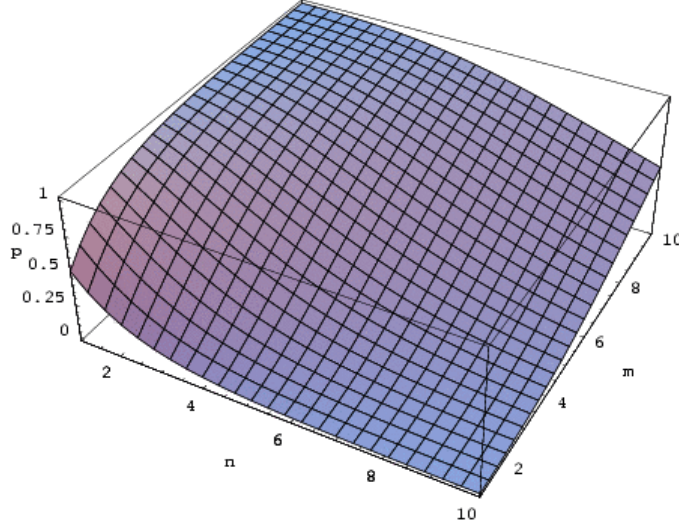


Abbildung 16.1: Darstellung von $P(W_{(m)} > W_{(n)})$ in Abhängigkeit von m und n

In der Abbildung ist deutlich zu sehen, dass die Wahrscheinlichkeit $P(W_{(m)} > W_{(n)})$ für $m > n$ relativ groß ist. Es ist aber auch zu sehen, dass sie für $m < n$ sich immer noch deutlich von 0 unterscheidet.

Der bedingte Erwartungswert lässt sich nun vereinfacht darstellen als:

16.24 Definition (bedingter Erwartungswert):

Es seien $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $(X, Y): \Omega \rightarrow \mathbb{R}^2$ ein Zufallsvektor mit Dichte $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ und Randverteilungsdichte $f_Y: \mathbb{R} \rightarrow \mathbb{R}$ bzgl. Y . Der zugehörige Erwartungswert

$$\mathbf{E}[X|Y = y] := \frac{1}{f_Y(y)} \int_{\mathbb{R}} x f(x, y) dx$$

heißt, sofern er existiert, bedingter Erwartungswert von X gegeben $Y = y$.

16.25 Beispiel:

In Beispiel 16.18 wurde der Zufallsvektor $(X, Y): \Omega \rightarrow \mathbb{R}^2$ mit der Dichte $f: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x, y) := \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 - 2xy + 2y^2)\right)$$

zugrunde gelegt. Dies führte für f_Y zu dem Ergebnis

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right), \quad y \in \mathbb{R}.$$

Damit wird

$$\mathbf{E}[X|Y = y] = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x - y)^2\right) dx = y, \quad y \in \mathbb{R}.$$

Literatur zu Kapitel 16

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- G. ALSMEYER:
Wahrscheinlichkeitstheorie,
Skripten zur Mathematischen Statistik, Nr. 30, Institut für Mathematische Statistik der
Westfälischen Wilhelms-Universität Münster, 3. Auflage 2003.
- H. BAUER:
Wahrscheinlichkeitstheorie,
5. Auflage, Walter de Gruyter, Berlin, 2002.
ISBN: 3110172364
- K.-W. GAEDE:
Zuverlässigkeit — Mathematische Modelle,
Hanser, München, 1977.
ISBN: 3446123709
- A. KLENKE:
Wahrscheinlichkeitstheorie,
Springer, Berlin, 2006.
ISBN: 3540255451
- M. KOLONKO:
Stochastik I, Skriptum zur Vorlesung,
TU Clausthal, 2005.
- P. P. SPIES:
Grundlagen stochastischer Modelle,
Hanser, München, 1982.
ISBN: 3446137114

Kapitel 17

Erneuerungstheorie

Mit Hilfe der Erneuerungstheorie lassen sich eine Vielzahl von realen Problemen, wie zum Beispiel das Ausfallverhalten von Maschinen, beschreiben und analysieren. Erneuerungsprozesse treten aber auch eingebettet in anderen Prozessen auf und bieten deshalb auch eine Handhabe zur Untersuchung allgemeinerer Prozesse.

Schlüsselwörter: Erneuerungsprozess, Erneuerungsfolge, Laplace-Transformierte, Faltung, Zählprozess, Poissonprozess, Erneuerungsfunktion, Erneuerungsgleichung, Vorwärtsrekurrenzzeit, Rückwärtsrekurrenzzeit, direkte Riemann-Integrierbarkeit, stationärer Zustand, arithmetisch verteilt, Blackwell'sches Erneuerungstheorem, Fundamentalsatz der Erneuerungstheorie, modifizierter Erneuerungsprozess, stationärer Erneuerungsprozess, Überlagerungsprozess, abbrechender Erneuerungsprozess, rekurrent, transient, periodisch.

17.1 Erneuerungsfunktion und Erneuerungsgleichung

Wir betrachten eine Beleuchtungseinrichtung mit einer Glühlampe, die nach Ausfall unverzüglich durch eine neue und gleichartige ersetzt wird. Die Zufallsvariablen S_n , $n \in \mathbb{N}$, beschreiben die aufeinanderfolgenden Zeitpunkte, an denen die Glühlampen ausgetauscht werden. X_n bezeichne die Lebensdauer der n -ten Glühlampe, $n = 1, 2, \dots$. Offensichtlich gilt: $S_n = S_{n-1} + X_n$, $n = 1, 2, \dots$ mit $S_0 := 0$. Man kann nun z.B. nach der Anzahl der Erneuerungen innerhalb eines Intervalls $[0, t]$ oder nach der zum Zeitpunkt t noch verbleibenden Zeit bis zur nächsten Erneuerung fragen. Diese und ähnliche Fragen können mit Hilfe der Erneuerungstheorie beantwortet werden.

17.1 Definition (Erneuerungsprozess, Erneuerungsfolge, Lebensdauer):

Es bezeichne $(X_n)_{n \in \mathbb{N}}$ eine Folge von stochastisch unabhängigen, identisch verteilten, nicht-negativen Zufallsvariablen über einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit $P(X_n = 0) < 1$ für alle $n \in \mathbb{N}$. Der zur Folge $(X_n)_{n \in \mathbb{N}}$ gehörende Partialsummenprozess $(S_n)_{n \in \mathbb{N}_0}$ mit

$$S_0 := 0 \quad \text{und} \quad S_n := \sum_{k=1}^n X_k$$

heißt gewöhnlicher oder auch einfacher Erneuerungsprozess (zuweilen auch Erneuerungsfolge genannt). X_n wird Lebensdauer und S_n Zeitpunkt der n -ten Erneuerung genannt. F sei die Verteilungsfunktion von X_n .

Die Verteilungsfunktion $F_n(t) := F_{S_n}(t)$ der Zufallsgröße S_n kann man rekursiv berechnen:

$$F_n(t) = F_{S_n}(t) = P(S_n \leq t) = P(S_{n-1} + X_n \leq t) = \int_0^t F_{n-1}(t-s) dF(s) \quad (n = 1, 2, \dots),$$

wobei

$$F_0(t) := \begin{cases} 1, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

Zur rechentechnischen Vereinfachung der Faltungsoperation bietet es sich an, charakteristische Funktionen zu benutzen. Da im vorliegenden Fall allerdings ausschließlich mit nichtnegativen Zufallsvariablen operiert wird, kann man ebensogut die etwas einfacher zu handhabende Laplace-Transformation verwenden.

17.2 Definition (Laplace-Transformierte):

Es sei X eine nichtnegative Zufallsvariable mit der Verteilung P_X bzw. der Verteilungsfunktion F . Dann heißt

$$\Psi_X(s) := \mathbf{E}[e^{-sX}] = \int_{\mathbb{R}^+} e^{-sx} dP_X = \int_{\mathbb{R}^+} e^{-sx} dF(x), \quad s > 0,$$

die Laplace-Transformierte der Zufallsvariable X bzw. der Verteilung P_X bzw. der Verteilungsfunktion F .

Wie im Fall der charakteristischen Funktion (siehe Definition 13.5) ist die Verteilungsfunktion F durch ihre zugehörige Laplace-Transformierte eindeutig bestimmt und umgekehrt. Da

$$\Psi_{S_0}(s) = \int_{\mathbb{R}^+} e^{-sx} dF_0(x) = e^{-s \cdot 0} \equiv 1$$

gilt, folgt aus

$$\Psi_{S_n}(s) = \mathbf{E}[e^{-sS_n}] = \mathbf{E}[e^{-s(S_{n-1}+X_n)}] = \mathbf{E}[e^{-sS_{n-1}}] \cdot \mathbf{E}[e^{-sX_n}] = \Psi_{S_{n-1}}(s) \cdot \Psi_{X_n}(s) \quad (n = 1, 2, \dots)$$

sofort

$$\Psi_{S_n}(s) = \prod_{i=1}^n \Psi_{X_i}(s) = (\Psi_{X_1}(s))^n \quad (n = 1, 2, \dots).$$

Wir können deshalb auch allgemein

$$F_n(t) = P(S_n \leq t) = F^{n*}(t), \quad t \in \mathbb{R},$$

schreiben.

17.3 Definition (Zählprozess, Erneuerungsprozess):

Die Folge $(N_t)_{t \in \mathbb{R}^+}$ mit $N_t := \max\{k \geq 1 | S_k \leq t\}$, wobei $\max \emptyset := 0$ sei, heißt der zum Erneuerungsprozess $(S_n)_{n \in \mathbb{N}_0}$ gehörende Zählprozess (zuweilen auch Erneuerungsprozess). N_t zählt die Anzahl der Erneuerungen bis zum Zeitpunkt t , wobei der Zeitpunkt $t = 0$ **nicht** als Erneuerungszeitpunkt gezählt wird.

17.4 Satz:

Es gilt für alle $t \in \mathbb{R}^+$:

$$P(N_t = n) = F_n(t) - F_{n+1}(t).$$

Beweis:

Zur Berechnung der Wahrscheinlichkeitsverteilung von N_t wird der folgende Zusammenhang benutzt:

$$\{N_t \geq n\} \iff \{S_n \leq t\}.$$

Folglich gilt:

$$\begin{aligned} P(N_t \geq n) &= P(\{N_t = n\} \cup \{N_t \geq n+1\}) = P(N_t = n) + P(N_t \geq n+1) \\ \implies P(S_n \leq t) &= P(N_t = n) + P(S_{n+1} \leq t) \\ \implies P(N_t = n) &= P(S_n \leq t) - P(S_{n+1} \leq t) = F_n(t) - F_{n+1}(t), \quad t \in \mathbb{R}^+. \end{aligned}$$

17.5 Beispiel (Poissonprozess):

Es sei $X_1 \stackrel{d}{=} \text{Exp}(\lambda)$. Dann gilt (Stochastik I, Satz 11.35):

$$S_n \stackrel{d}{=} \text{Erlang}(n, \lambda) \quad \text{bzw.} \quad P(S_n \leq t) = 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!}, \quad t \in \mathbb{R}^+.$$

Folglich ist

$$\begin{aligned} P(N_t = n) &= F_n(t) - F_{n+1}(t) \\ &= 1 - e^{-\lambda t} \sum_{i=0}^{n-1} \frac{(\lambda t)^i}{i!} - \left(1 - e^{-\lambda t} \sum_{i=0}^n \frac{(\lambda t)^i}{i!} \right) \\ &= \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad (n = 0, 1, 2, \dots), \quad t \in \mathbb{R}^+. \end{aligned}$$

N_t ist also Poisson-verteilt mit dem Parameter λt .

17.6 Bemerkung:

Alternativ zu Definition 17.3 kann man den Zählprozess $(N_t)_{t \in \mathbb{R}^+}$ auch über die Beziehung

$$N_t = \sum_{n=1}^{\infty} I_{[0,t]}(S_n)$$

entwickeln.

17.7 Definition (Erneuerungsfunktion):

Die Funktion $R: \mathbb{R}^+ \ni t \mapsto R(t) := \mathbf{E}[N_t] \in \mathbb{R}^+$ heißt Erneuerungsfunktion (renewal function) und gibt die mittlere Anzahl von Erneuerungen im Zeitraum $[0, t]$ an.

17.8 Satz:

Es gilt für alle $t \in \mathbb{R}^+$:

- a) $R(t) := \mathbf{E}[N_t] = \sum_{n=1}^{\infty} F_n(t),$
- b) $R(t) < \infty.$
- c) $R(t)$ ist monoton nichtfallend und rechtsseitig stetig.

Beweis:

a) Es gilt:

$$\mathbf{E}[N_t] = \sum_{n=1}^{\infty} n \cdot P(N_t = n) = \sum_{n=1}^{\infty} n(F_n(t) - F_{n+1}(t)) = \sum_{n=1}^{\infty} F_n(t), \quad t \in \mathbb{R}^+.$$

b) Da $F(0) = P(X_i = 0) < 1$ ist und F rechtsseitig stetig ist, existiert ein $b > 0$, so dass $F(b) < 1$ ist.

1.) Es wird $k \in \mathbb{N}$ so gewählt, dass $t \leq kb$ ist. Dann gilt

$$P(S_k \leq t) \leq P(S_k \leq kb) \leq 1 - P(X_1 > b, \dots, X_k > b) = 1 - (1 - F(b))^k =: 1 - a$$

mit $a > 0$, da $F(b) < 1$ ist.

2.) Es gilt für $m, k \in \mathbb{N}_0$:

$$\begin{aligned} P(S_{mk} \leq t) &\leq P(S_k - S_0 \leq t, \dots, S_{mk} - S_{mk-k} \leq t) = (P(S_k - \underbrace{S_0}_{=0} \leq t))^m \\ &= (P(S_k \leq t))^m \stackrel{1.)}{\leq} (1-a)^m. \end{aligned}$$

3.) Da $P(S_{mk+j} \leq t) \leq P(S_{mk} \leq t)$ für $j \in \mathbb{N}_0$ gilt, folgt:

$$\sum_{n=mk}^{mk+k-1} P(S_n \leq t) \leq k \cdot P(S_{mk} \leq t).$$

Damit ergibt sich insgesamt:

$$\begin{aligned} R(t) &:= \sum_{n=0}^{\infty} P(S_n \leq t) = \sum_{m=0}^{\infty} \sum_{n=mk}^{mk+k-1} P(S_n \leq t) \\ &\stackrel{3.)}{\leq} \sum_{m=0}^{\infty} k(P_{mk} \leq t) \stackrel{2.)}{\leq} \sum_{m=0}^{\infty} k(1-a)^m \\ &= \frac{k}{a} < \infty \quad \text{da } a > 0. \end{aligned}$$

c) Wir zeigen nur die rechtsseitige Stetigkeit. Es seien $t \geq 0$ fest, $N_t = k$ und $(t_n)_{n \in \mathbb{N}_0}$ Zeitpunkte mit $S_k \leq t_n < S_{k+1}$ und $\lim_{n \rightarrow \infty} t_n = t$. Dann gilt:

$$\begin{aligned} R(t_0) - R(t) &= \mathbf{E}[N_{t_0}] - \mathbf{E}[N_t] \\ &= \mathbf{E}[N_{t_0} - N_t] \\ &= \mathbf{E}\left[N_{t_0} - \lim_{n \rightarrow \infty} N_{t_n}\right] \\ &= \mathbf{E}\left[\lim_{n \rightarrow \infty} (N_{t_0} - N_{t_n})\right] \\ &= \lim_{n \rightarrow \infty} \mathbf{E}[N_{t_0} - N_{t_n}] \quad (\text{Satz 9.42 von der monotonen Konvergenz}) \\ &= \lim_{n \rightarrow \infty} (\mathbf{E}[N_{t_0}] - \mathbf{E}[N_{t_n}]) \\ &= \mathbf{E}[N_{t_0}] - \lim_{n \rightarrow \infty} \mathbf{E}[N_{t_n}] \\ &= R(t_0) - \lim_{n \rightarrow \infty} R(t_n) \quad (\text{da } R(t) < \infty \forall t \in \mathbb{R}^+), \end{aligned}$$

woraus

$$R(t) = \lim_{n \rightarrow \infty} R(t_n)$$

folgt. ■

17.9 Bemerkung:

Setzt man $R(t) = 0$ für $t \in \mathbb{R}^-$, dann definiert $R(t)$ eine maßdefinierende Funktion auf \mathbb{R} . Für das von R induzierte (endliche) Maß auf $(\mathbb{R}, \mathfrak{B})$ gilt:

$$\mu_R((a, b]) = R(b) - R(a) \quad \forall a, b \in \mathbb{R} \text{ mit } a \leq b.$$

17.10 Beispiel:

Es sei $X_1 \stackrel{d}{=} \text{Exp}(\beta)$. Dann ist S_n Erlang(n, β)-verteilt. Die Erlangverteilung ist ein Spezialfall der Gammaverteilung:

$$F_{\Gamma(b,p)}(x) := \begin{cases} \int_{-\infty}^x \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} dx, & b, p \in \mathbb{R}^+, \quad t \geq 0 \\ 0, & t < 0. \end{cases}$$

Mit $b := \beta$ und $p := n$ erhält man

$$F_n(t) = \begin{cases} \int_0^t \beta(\beta s)^{n-1} \cdot \frac{e^{-\beta s}}{(n-1)!} ds, & t \geq 0 \\ 0, & t < 0. \end{cases}$$

Mit Hilfe von $F_n(t)$ lässt sich die Erneuerungsfunktion wie folgt berechnen:

$$\begin{aligned} R(t) &= \sum_{n=1}^{\infty} F_n(t) = \sum_{n=1}^{\infty} \int_0^t \beta(\beta s)^{n-1} \frac{e^{-\beta s}}{(n-1)!} ds \\ &= \int_0^t \beta \sum_{n=1}^{\infty} \frac{(\beta s)^{n-1}}{(n-1)!} e^{-\beta s} ds \quad (\text{Satz 9.42 von der monotonen Konvergenz}) \\ &= \int_0^t \beta ds = \beta t, \quad t \in \mathbb{R}^+. \end{aligned}$$

17.11 Satz:

Es sei $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ messbar, dann gilt:

$$Rf = \int_{\mathbb{R}^+} f(t) dR(t) = \mathbf{E} \left[\sum_{n=1}^{\infty} f(S_n) \right].$$

Beweis:

Der Satz wird mit algebraischer Induktion bewiesen:

a) Es sei zunächst $f = I_{(s,t]}$. Dann gilt:

$$\begin{aligned} Rf &= \int_{\mathbb{R}^+} I_{(s,t]}(x) dR(x) = \mu_R((s, t]) \\ &= R(t) - R(s) = \mathbf{E}[N_t] - \mathbf{E}[N_s] = \mathbf{E}[N_t - N_s] \\ &= \mathbf{E} \left[\sum_{n=1}^{\infty} I_{[0,t]}(S_n) - \sum_{n=1}^{\infty} I_{[0,s]}(S_n) \right] \\ &= \mathbf{E} \left[\sum_{n=1}^{\infty} I_{(s,t]}(S_n) \right] \\ &= \mathbf{E} \left[\sum_{n=1}^{\infty} f(S_n) \right]. \end{aligned}$$

b) Gilt die Behauptung für $f_i \geq 0$, $i = 1, \dots, k$, so folgt

$$\begin{aligned} \int_{\mathbb{R}^+} \sum_{i=1}^k \alpha_i f_i(x) dR(x) &= \sum_{i=1}^k \alpha_i \int_{\mathbb{R}^+} f_i(x) dR(x) = \\ &= \sum_{i=1}^k \alpha_i \mathbf{E} \left[\sum_{n=1}^{\infty} f_i(S_n) \right] = \mathbf{E} \left[\sum_{n=1}^{\infty} \sum_{i=1}^k \alpha_i f_i(x) \right] = \sum_{n=1}^{\infty} \mathbf{E} \left[\sum_{i=1}^k \alpha_i f_i(x) \right]. \end{aligned}$$

Für nichtnegative messbare Abbildungen wendet man den Satz von der monotonen Konvergenz an. Allgemeine messbare Abbildungen behandelt man, in dem man ihren Positiv- und Negativteil separat behandelt. ■

Faltungen

Im Weiteren erweist es sich als zweckmäßig, den Begriff der Laplace-Transformierten und der Faltung auf andere Funktionenklassen auszudehnen.

Im Folgenden bezeichne \mathbb{M} die Menge der maßdefinierenden Funktionen auf \mathbb{R} , die in $(-\infty, 0)$ verschwinden. \mathbb{B} stehe für die Menge der auf \mathbb{R} definierten messbaren reellen Funktionen, die auf jedem Intervall der Form $[0, t]$ beschränkt sind (sogenannte lokalbeschränkte messbare reelle Funktionen).

Folgerung: $\mathbb{M} \subset \mathbb{B}$.

17.12 Definition (Faltung):

Ist $F \in \mathbb{M}$ und $g \in \mathbb{B}$, so heißt die durch

$$(F * g)(t) := \begin{cases} \int_{[0,t]} g(t-s) dF(s) & , t \geq 0 \\ 0 & , t < 0 \end{cases}$$

erklärte Funktion die Faltung von F und g .

17.13 Satz:

Es gilt: $(F * g) \in \mathbb{B}$.

Beweis:

$$|(F * g)(t)| \leq \int_0^t |g(t-s)| dF(s) \leq \sup_{0 \leq s \leq t} |g(s)| \int_0^t dF(s) = \sup_{0 \leq s \leq t} |g(s)| F(t) < \infty, \quad \forall t \in \mathbb{R}.$$

■

17.14 Satz (Eigenschaften der Faltung):

Es seien $F, F_1, F_2, F_3 \in \mathbb{M}$, $g, g_1, g_2 \in \mathbb{B}$ und $c \in \mathbb{R}$. Dann gilt:

- a) $(F_1 * F_2) * F_3 = F_1 * (F_2 * F_3),$
- b) $(F_1 + F_2) * g = F_1 * g + F_2 * g,$
- c) $F * (g_1 + g_2) = F * g_1 + F * g_2,$
- d) $F * g_1 \leq F * g_2, \text{ falls } g_1 \leq g_2,$
- e) $F * (cg) = c(F * g),$
- f) $F_1 * F_2 = F_2 * F_1,$
- g) $\overline{F * g} = \overline{F} \cdot \overline{g}.$

Beweis:

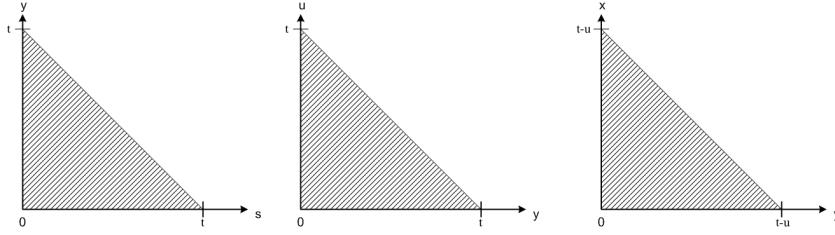
a) Es gilt

$$(F_i * F_j)(0-) = \int_0^{0-} F_j(0-s) dF_i(s) = 0 \quad \text{für } i, j = 1, 2, 3, i \neq j.$$

Damit folgt:

$$\begin{aligned}
 ((F_1 * F_2) * F_3)(t) &= \int_0^t F_3(t-s) d(F_1 * F_2)(s) \\
 &= \int_0^t \int_0^{t-s} dF_3(y) d(F_1 * F_2)(s) \\
 &\stackrel{(*)}{=} \int_0^t \int_0^{t-y} d(F_1 * F_2)(s) dF_3(y) \\
 &= \int_0^t (F_1 * F_2)(t-y) dF_3(y) \\
 &= \int_0^t \int_0^{t-y} F_2(t-y-u) dF_1(u) dF_3(y) \\
 &\stackrel{(**)}{=} \int_0^t \int_0^{t-u} F_2(t-y-u) dF_3(y) dF_1(u) \\
 &= \int_0^t \int_0^{t-u} \int_0^{t-y-u} dF_2(x) dF_3(y) dF_1(u) \\
 &\stackrel{(***)}{=} \int_0^t \int_0^{t-u} \int_0^{t-u-x} dF_3(y) dF_2(x) dF_1(u) \\
 &= \int_0^t \int_0^{t-u} F_3(t-u-x) dF_2(x) dF_1(u) \\
 &= \int_0^t (F_2 * F_3)(t-u) dF_1(u) \\
 &= (F_1 * (F_2 * F_3))(t), \quad t \in \mathbb{R}.
 \end{aligned}$$

(*), (**) und (**): Die Integration verläuft jeweils über folgende Bereiche:



b) Wegen der Linearität des Integrals gilt:

$$\begin{aligned}
 ((F_1 + F_2) * g)(t) &= \int_0^t g(t-s) d(F_1(s) + F_2(s)) \\
 &= \int_0^t g(t-s) dF_1(s) + \int_0^t g(t-s) dF_2(s) \\
 &= (F_1 * g)(t) + (F_2 * g)(t), \quad t \in \mathbb{R}.
 \end{aligned}$$

c) Gilt wegen der Linearität des Integrals.

d) Folgt mit der Monotonie des Integrals.

e) Gilt wegen der Linearität des Integrals.

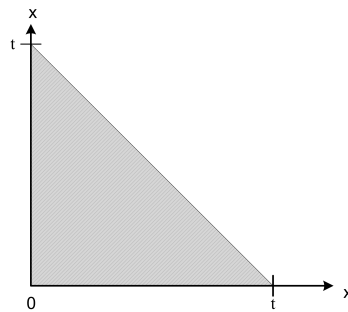
f) Es gilt

$$\int_{0-}^x dF_i(y) = \int_0^x dF_i(y) = F_i(x) \quad \text{für } i = 1, 2.$$

Damit folgt:

$$\begin{aligned}
 (F_1 * F_2)(t) &= \int_0^t F_2(t-x) dF_1(x) = \int_0^t \int_0^{t-x} dF_2(y) dF_1(x) \\
 &\stackrel{(*)}{=} \int_0^t \int_0^{t-y} dF_1(x) dF_2(y) = \int_0^t F_1(t-y) dF_2(y) \\
 &= (F_2 * F_1)(t), \quad t \in \mathbb{R}.
 \end{aligned}$$

(*): Die Integration verläuft über folgenden Bereich:



g) Es seien F_1 die Verteilung der Zufallsvariable X_1 und F_2 die Verteilung der Zufallsvariable X_2 . Dann ist $F_1 * F_2$ die Verteilung von $X_1 + X_2$, siehe Definition 17.1. Damit gilt:

$$\overline{F_1 * F_2}(s) = \mathbf{E}[e^{-s(X_1+X_2)}] = \mathbf{E}[e^{-sX_1}] \cdot \mathbf{E}[e^{-sX_2}] = \overline{F_1}(s) \cdot \overline{F_2}(s), \quad s \in \mathbb{R}.$$

■

Im Folgenden mögen diese Schreibweisen gelten:

- $F^{2*} := F * F$,
- $F^{3*} := F * F^{2*} = F * F * F$,
- $F^{m*} * F^{n*} = F^{(n+m)*}$.

17.15 Satz:

Die Erneuerungsfunktion $R(t)$ genügt der Integralgleichung

$$R(t) = F(t) + (F * R)(t) \quad (t \in \mathbb{R}^+).$$

Beweis:

X_1 ist der Zeitpunkt der ersten Erneuerung. Wir behaupten:

$$\mathbf{E}[N_t | X_1 = x] = \begin{cases} 1 + R(t - x) & , x \leq t, \\ 0 & , x > t. \end{cases}$$

Der zweite Fall ist klar: Da die erste Erneuerung erst zum Zeitpunkt $x > t$ stattfindet, kann vorher keine Erneuerung eingetreten sein. Ist t jetzt jedoch größer x , so ist bekannt, dass mindestens eine Erneuerung stattgefunden hat. Da die Zuwächse X_i jedoch stochastisch unabhängig sind, regeneriert sich der Erneuerungsprozess zum Zeitpunkt $X_1 = x$. Deswegen entspricht die mittlere Anzahl an Erneuerungen im Intervall $[x, t]$ gerade der mittleren Anzahl an Erneuerungen im Intervall $[0, t - x]$. Diese Argumentation wird als Erneuerungsargument bezeichnet.

Somit folgt:

$$\begin{aligned} R(t) &= \mathbf{E}[\mathbf{E}[N_t | X_1 = x]] \\ &= \int_{\mathbb{R}} \mathbf{E}[N_t | X_1 = x] dF(x) \\ &= \int_{x \leq t} (1 + R(t - x)) dF(x) + \int_{x > t} 0 \cdot dF(x) \\ &= \int_{x \leq t} dF(x) + \int_{x \leq t} R(t - x) dF(x) \\ &= F(t) + (F * R)(t), \quad t \in \mathbb{R}^+. \end{aligned}$$

■

Bemerkung:

Die Gleichung $h = g + F * h$ mit $F \in \mathbb{M}$ und $h, g \in \mathbb{B}$ heißt Erneuerungsgleichung.

17.16 Satz:

Die Erneuerungsgleichung $h = g + F * h$ mit $F \in \mathbb{M}$ und $h, g \in \mathbb{B}$ hat genau eine Lösung in \mathbb{B} , nämlich

$$h := g + R * g.$$

Beweis:

1. Schritt: Es wird zunächst gezeigt, dass h eine Lösung der Erneuerungsgleichung ist.

$$\begin{aligned}
 h &= g + R * g \\
 &= g + (F + F * R) * g \quad (\text{Satz 17.15}) \\
 &= g + F * g + F * (R * g) \\
 &= g + F * (g + R * g) \\
 &= g + F * h.
 \end{aligned}$$

2. Schritt: Für den Beweis der Eindeutigkeit werden zwei Lösungen $h := g + F * h$ und $h' := g + F * h'$ betrachtet. Es folgt

$$d := h - h' = F * h - F * h' = F * (h - h').$$

Die Differenz d genügt also der Gleichung $d = F * d$. Iteriert man diese Gleichung, erhält man

$$d = F * d = F^{2*} * d = \dots = F^{n*} * d = \dots$$

Aus $R(t) = \sum_{n=1}^{\infty} F^{n*}(t) < \infty$ folgt aber $\lim_{n \rightarrow \infty} F^{n*}(t) = 0 \quad \forall t \in \mathbb{R}$, so dass gilt

$$\begin{aligned}
 |d(t)| &= \lim_{n \rightarrow \infty} \left| \int_0^t d(t-x) dF^{n*}(x) \right| \\
 &\leq \sup_{0 \leq x \leq t} |d(t-x)| \lim_{n \rightarrow \infty} F^{n*}(t) = 0.
 \end{aligned}$$

3. Schritt: Zu zeigen bleibt noch, dass h lokal beschränkt ist, d.h. dass $h \in \mathbb{B}$ gilt.

$$\begin{aligned}
 \sup_{0 \leq s \leq t} |h(s)| &\leq \sup_{0 \leq s \leq t} |g(s)| + \int_0^t \left\{ \sup_{0 \leq x \leq t} |g(x)| \right\} dR(x) \\
 &\leq \sup_{0 \leq s \leq t} |g(s)| \cdot \{1 + R(t)\} < \infty.
 \end{aligned}$$

■

17.17 Bemerkung:

- Viele Größen in der Erneuerungstheorie können auf die Lösung einer Erneuerungsgleichung zurückgeführt werden.
- Mit den Definitionen

$$\bar{h}(s) := \int_{\mathbb{R}^+} e^{-sx} dh(x), \quad \bar{g}(s) := \int_{\mathbb{R}^+} e^{-sx} dg(x), \quad \bar{F}(s) := \int_{\mathbb{R}^+} e^{-sx} dF(x)$$

folgt aus $h = g + F * h$ (siehe Satz 17.14 g):

$$\bar{h}(s) = \bar{g}(s) + \bar{F}(s)\bar{h}(s) \iff \bar{h}(s) = \frac{\bar{g}(s)}{1 - \bar{F}(s)}.$$

17.18 Beispiel:

Für die Exponentialverteilung $F(x) := 1 - e^{-\lambda x}$, $x \geq 0$, gilt:

$$\begin{aligned}\bar{F}(s) &= \int_{\mathbb{R}_+} e^{-sx} dF(x) = \int_{\mathbb{R}_+} e^{-sx} \lambda e^{-\lambda x} dx \\ &= \lambda \cdot \int_{\mathbb{R}_+} e^{-x(\lambda+s)} dx = -\frac{\lambda}{\lambda+s} e^{-x(\lambda+s)} \Big|_0^\infty \\ &= \frac{\lambda}{\lambda+s}.\end{aligned}$$

Aus der Integralgleichung $R(t) = F(t) + (F * R)(t)$ aus Satz 17.15 folgt $\bar{R}(s) = \bar{F}(s) + \bar{F}(s) \cdot \bar{R}(s)$ bzw. $\bar{R}(s) = \frac{\bar{F}(s)}{1 - \bar{F}(s)}$. Damit lässt sich die Erneuerungsfunktion $R(t)$ berechnen:

$$\bar{R}(s) = \frac{\bar{F}(s)}{1 - \bar{F}(s)} = \frac{\frac{\lambda}{\lambda+s}}{\frac{s}{\lambda+s}} = \frac{\lambda}{s} \implies R(t) = \lambda \cdot t,$$

denn es gilt:

$$\bar{R}(s) = \int_{\mathbb{R}_+} e^{-st} \lambda dt = -\frac{\lambda}{s} e^{-st} \Big|_0^\infty = \frac{\lambda}{s}.$$

17.2 Vorwärts- und Rückwärtsrekurrenzzeit

Wir betrachten einen Erneuerungsprozess im zeitlichen Verlauf und fixieren einen Zeitpunkt $t \in \mathbb{R}^+$ (siehe Abbildung 17.1).

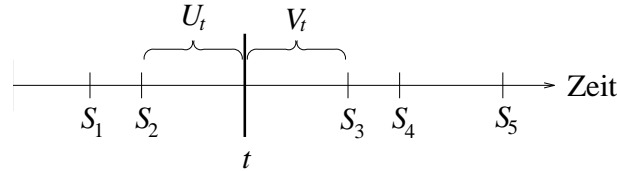


Abbildung 17.1: Darstellung der Erneuerungszeitpunkte auf einem Zeitstrahl

Die Zeitspanne von t bis zur nächsten Erneuerung wird die Restlebenszeit oder auch die Vorwärtsrekurrenzzeit genannt und mit V_t bezeichnet, und die Zeitspanne von der letzten Erneuerung bis t wird Alter oder auch Rückwärtsrekurrenzzeit genannt und mit U_t bezeichnet. Diese Zeiten können durch folgende Gleichungen mathematisch beschrieben werden:

$$V_t = S_{N_t+1} - t \quad \text{und} \quad U_t = t - S_{N_t}.$$

17.19 Satz:

Für den Erwartungswert der Vorwärtsrekurrenzzeit gilt:

$$\mathbf{E}[V_t] = \mathbf{E}[X_1] \cdot (1 + R(t)) - t \quad (t \in \mathbb{R}^+).$$

Beweis:

Aus der Definition von V_t folgt der Ansatz

$$\mathbf{E}[V_t] = \mathbf{E}[S_{N_t+1} - t] = \mathbf{E}[S_{N_t+1}] - t.$$

Es wird nun $\mathbf{E}[S_{N_t+1}] =: A(t)$ gesetzt. Um den Erwartungswert von $A(t)$ zu berechnen, wird das Erneuerungsargument benutzt.

$$\mathbf{E}[S_{N_t+1}|X_1 = x] = \begin{cases} x + A(t - x) & , x \leq t \\ x & , x > t. \end{cases}$$

Damit wird

$$\begin{aligned} A(t) &= \mathbf{E}[\mathbf{E}[S_{N_t+1}|X_1 = x]] \\ &= \int_{\mathbb{R}^+} \mathbf{E}[S_{N_t+1}|X_1 = x] dF(x) \\ &= \int_{[0,t]} (x + A(t - x)) dF(x) + \int_{(t,\infty)} x dF(x) \\ &= \int_{\mathbb{R}^+} x dF(x) + \int_{[0,t]} A(t - x) dF(x) \\ &= \mathbf{E}[X_1] + \int_{[0,t]} A(t - x) dF(x) \\ &= \mathbf{E}[X_1] + (A + F)(t) \\ &= \mathbf{E}[X_1] + \int_{[0,t]} \mathbf{E}[X_1] dR(x) \quad (\text{Erneuerungsgleichung, Satz 17.16}) \\ &= \mathbf{E}[X_1] \left(1 + \int_{[0,t]} dR(x) \right) \\ &= \mathbf{E}[X_1] (1 + R(t)). \end{aligned}$$

■

17.20 Beispiel:

Es sei $X_1 \stackrel{d}{=} \text{Exp}(\lambda)$. Dann folgt $\mathbf{E}[X_1] = \frac{1}{\lambda}$ und nach Beispiel 17.10 gilt $R(t) = \lambda \cdot t$. Es folgt mit Satz 17.19:

$$\mathbf{E}[V_t] = \mathbf{E}[X_1] (1 + R(t)) - t = \frac{1}{\lambda} (1 + \lambda \cdot t) - t = \frac{1}{\lambda}.$$

Das Ergebnis verwundert angesichts der Gedächtnislosigkeit der Exponentialverteilung nicht.

17.21 Satz:

Für die Verteilung der Vorwärtsrekurrenz gilt:

$$P(V_t \leq v) = \begin{cases} \int_t^{t+v} (1 - F(t + v - y)) dR(y) & , v \geq 0 \\ 0 & , v < 0. \end{cases}$$

Beweis:

Es wird zunächst das Ereignis $\{V_t \leq v\}$ betrachtet. Es gilt:

$$\{V_t \leq v\} \iff \{S_{N_t+1} \leq t+v\} \iff \{S_{N_t} \leq t; t \leq S_{N_t+1} \leq t+v\}.$$

Es lässt sich also schreiben

$$\begin{aligned} P(V_t \leq v) &= P\left(\bigcup_{n=0}^{\infty} \{S_n \leq t; t < S_{n+1} \leq t+v\}\right) \\ &= \sum_{n=0}^{\infty} P(S_n \leq t, t < S_{n+1} \leq t+v) \\ &= F(t+v) - F(t) + \sum_{n=1}^{\infty} P(S_n \leq t, t < S_{n+1} \leq t+v) \\ &= F(t+v) - F(t) + \sum_{n=1}^{\infty} \int_0^t P(t < S_{n+1} \leq t+v | S_n = y) dF_n(y) \\ &= F(t+v) - F(t) + \sum_{n=1}^{\infty} \int_0^t P(t-y < X_{n+1} \leq t+v-y) dF_n(y) \end{aligned}$$

Mit Hilfe von Satz 17.11 und der Beziehung $\mathbf{E}[\sum_{n=1}^{\infty} X_n] = \sum_{n=1}^{\infty} \mathbf{E}[X_n]$ für $X_n \geq 0$ schließen wir weiter:

$$\begin{aligned} P(V_t \leq v) &= F(t+v) - F(t) + \int_0^t P(t-y < X_{n+1} \leq t+v-y) dR(y) \\ &= F(t+v) - F(t) + \int_0^t (F(t+v-y) - F(t-y)) dR(y) \\ &= F(t+v) + \int_0^{t+v} F(t+v-y) dR(y) - F(t) - \int_0^t F(t-y) dR(y) \\ &\quad - \int_t^{t+v} F(t+v-y) dR(y). \end{aligned}$$

Mit $R = F + F * R$ bzw. $F = R - F * R$ folgt nun:

$$\begin{aligned} P(V_t \leq v) &= R(t+v) - (F * R)(t+v) + \int_0^{t+v} F(t+v-y) dR(y) - R(t) + (F * R)(t) \\ &\quad - \int_0^t F(t-y) dR(y) - \int_t^{t+v} F(t+v-y) dR(y) \\ &= R(t+v) - R(t) - \int_t^{t+v} F(t+v-y) dR(y) \end{aligned}$$

$$\begin{aligned}
 &= \int_t^{t+v} dR(y) - \int_t^{t+v} F(t+v-y) dR(y) \\
 &= \int_t^{t+v} (1 - F(t+v-y)) dR(y).
 \end{aligned}$$

■

17.22 Beispiel:

Es sei $X_1 \stackrel{d}{=} \text{Exp}(\lambda)$. Beispiel 17.10 zufolge ist dann $R(t) = \lambda \cdot t$, $t \in \mathbb{R}_0^+$. Damit gilt nach Satz 17.21:

$$P(V_t \leq v) = \int_t^{t+v} e^{-\lambda(t+v-y)} \lambda dy = e^{-\lambda(t+v-y)} \Big|_t^{t+v} = 1 - e^{-\lambda v},$$

womit die Gedächtnislosigkeit der Exponentialverteilung ein weiteres Mal gezeigt wäre.

17.23 Satz:

Für die Verteilung der Rückwärtsrekurrenzzeit gilt:

$$P(U_t \leq u) = \begin{cases} \int_{t-u}^t (1 - F(t-\tau)) dR(\tau) & , u \geq 0 \\ 0 & , u < 0. \end{cases}$$

Beweis:

Analog zum Beweis von Satz 17.21 gilt:

$$\begin{aligned}
 P(U_t \leq u) &= P(t - S_{N_t} \leq u) \\
 &= P(S_{N_t} \geq t - u) \\
 &= P\left(\bigcup_{n=0}^{\infty} \{t - u \leq S_n \leq t; S_{n+1} > t\}\right) \\
 &= \sum_{n=0}^{\infty} P(t - u \leq S_n \leq t; S_{n+1} > t) \\
 &= \sum_{n=0}^{\infty} \int_{t-u}^t P(S_{n+1} > t | S_n = \tau) dF_n(\tau) \\
 &= \sum_{n=0}^{\infty} \int_{t-u}^t P(X_{n+1} > t - \tau) dF_n(\tau) \\
 &= \sum_{n=0}^{\infty} \int_{t-u}^t (1 - F(t - \tau)) dF_n(\tau).
 \end{aligned}$$

Die Behauptung folgt nun mit Satz 17.11.

■

Bemerkung:

Eine alternative Herleitung für Satz 17.23 wäre:

$$\{U_t > x\} \iff \{\text{keine Erneuerung in } [t-x, t]\} \iff \{V_{t-x} > x\}.$$

17.24 Beispiel:

Es sei $X_1 \stackrel{d}{=} \text{Exp}(\lambda)$. Nach Beispiel 17.10 folgt dann $R(t) = \lambda \cdot t$. Damit gilt nach Satz 17.23:

$$F(x) := P(U_t \leq u) = \int_{t-u}^t e^{-\lambda(t-x)} dR(x) = \lambda \cdot \int_{t-u}^t e^{-\lambda(t-x)} dx = e^{-\lambda(t-x)} \Big|_{t-u}^t = 1 - e^{-\lambda u}.$$

Dies gilt für alle $0 \leq u \leq t$. Es folgt:

$$\mathbf{E}[U_t] = \int_{\mathbb{R}^+} (1 - F(x)) dx = \int_0^t e^{-\lambda x} dx = \frac{1}{\lambda} (1 - e^{-\lambda t}).$$

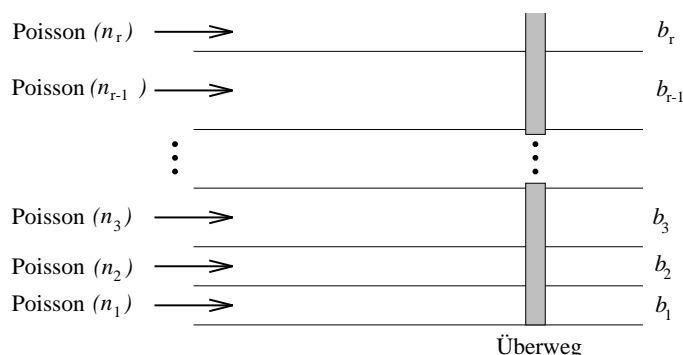
17.3 Anwendungsbeispiel 1

Problem:

Eine Einbahnstraße bestehe aus r Spuren, deren Breiten variabel und die in Höhe des Fußgängerüberweges durch $(r-1)$ Straßeninseln voneinander getrennt sind. In den r Spuren sei der Verkehr unabhängig mit einer Poisson-Verteilung verteilt; die Parameter n_i für die i -te Spur ($i = 1, \dots, r$) seien im Folgenden fest. Ein Fußgänger überquert die Straße in r Etappen, wobei sich jede seiner r Einzel-Wartezeiten nur nach dem Verkehr in der jeweils vor ihm liegenden Spur richtet; es sei Z_r die Gesamt-wartezeit bei r Spuren. Es geht darum, durch geschickte Wahl der Spurbreiten diese Gesamt-Wartezeit zu minimieren.

Lösung:

Gegeben sind r Spuren mit den Nummern $1, \dots, r$, deren jeweilige Breite durch die Zahl b_i ($i = 1, \dots, r$) beschrieben sei. Die Breite der Straße ist somit $b_1 + \dots + b_r =: b$.



Annahmen:

- (1) Die Ankunftszeiten der Fahrzeuge am Überweg bilden Poissonprozesse mit Parametern $n_i :=$ „mittlere Anzahl von ankommenden Fahrzeugen in Spur i pro Minute“. Einzelne Spuren sind unabhängig.
- (2) Der Fußgänger überquere die Straße in r Etappen. Der Fußgänger überquert die Spur sofort, wenn ein Auto vorbeigefahren ist und der Abstand zum nächsten größer ist, als die Überquerungszeit.

Weitere Parameter:

- T bezeichne die Überquerungszeit für die ganze Straße.
- $T_i := T \cdot \frac{b_i}{b}$ bezeichne die Überquerungszeit für die Spur i . Dabei wird angenommen, dass der Fußgänger die Fahrbahn mit konstanter Geschwindigkeit überquert und zwar jeder Fußgänger mit der gleichen Geschwindigkeit.
- W_i sei die Wartezeit vor der Überquerung der i -ten Spur.
- $Z_r := W_1 + \dots + W_r$ sei die Gesamtwartezeit.

Lösung für eine Fahrspur

Wegen der Unabhängigkeitsannahme in (1) ist die Verteilung von Z_r gleich der Faltung der Verteilungen der W_i ($i = 1, \dots, r$). Es wird daher zunächst folgende Situation betrachtet:



Gegeben sei eine Spur der Breite b mit der Überquerungszeit T . Der Verkehr ist ein Poissonprozess mit Parameter n . Die Wartezeit wird durch W beschrieben.

Es sei I_k die Zeit zwischen der Ankunft des $(k-1)$ -ten und k -ten Autos. Da es sich beim Verkehr um einen Poissonprozess handelt, bilden die I_k eine i.i.d. Folge von exponentialverteilten Zufallsgrößen, d.h. zu einem beliebigen Zeitpunkt t ist die Zeit bis zur Ankunft des nächsten Autos verteilt nach $\text{Exp}(n)$.

Es sei $S_k := I_1 + \dots + I_k$ die Ankunftszeit des k -ten Autos. S_k ist Erlangverteilt mit den Parametern n und k .

Wegen Annahme (2) ist W fast sicher gleich S_k für ein $k \in \mathbb{N}_0$. Folglich gilt für alle Borelmengen $A \subset \mathbb{R}$:

$$\begin{aligned} P(W \in A) &= \sum_{k=0}^{\infty} P(W \in A, W = S_k) = \sum_{k=0}^{\infty} P(W \in A | W = S_k) \cdot P(W = S_k) \\ &= \sum_{k=0}^{\infty} P(S_k \in A) \cdot P(W = S_k). \end{aligned}$$

Definiert man $Q_k := P(S_k \in A)$ und $p_k := P(W = S_k)$ für $k \in \mathbb{N}_0$ so gilt:

$$P(W \in A) = \sum_{k=0}^{\infty} Q_k \cdot p_k \quad \text{mit} \quad \sum_{k=0}^{\infty} p_k = 1.$$

Folglich hat W eine gewichtete Erlangverteilung mit Gewicht p_k bei $\text{Erlang}(n, k)$.

$$p_k := P(W = S_k) = P(I_1 \leq T, \dots, I_k \leq T, I_{k+1} > T) \stackrel{\text{iid}}{=} (P(I_1 \leq T))^k (1 - P(I_1 \leq T)).$$

Mit der Definition $p := p_T := P(I_1 \leq T) = 1 - e^{-nT}$ gilt also:

$$p_k = p^k \cdot (1 - p) \quad (k \in \mathbb{N}_0).$$

Da $P(W \in A) = \sum_{k=0}^{\infty} P(S_k \in A) \cdot p_k$ gilt, folgt für die charakteristische Funktion der Verteilung von W :

$$\psi = \sum_{k=0}^{\infty} p_k \varphi_k,$$

wobei φ_k die charakteristische Funktion der Verteilung von S_k ist, d.h. die charakteristische Funktion der Erlang-Verteilung mit den Parametern n und k .

Laut Beispiel 13.6.7 besitzt die $\text{Erlang}(\alpha, k)$ -Verteilung die charakteristische Funktion

$$\varphi_k(\zeta) = \frac{1}{(1 - i\frac{\zeta}{\alpha})^k}.$$

Damit folgt:

$$\begin{aligned} \psi(\zeta) &= (1 - p) \cdot \sum_{k=0}^{\infty} \left(\frac{p}{1 - i\frac{\zeta}{n}} \right)^k \\ &= (1 - p) \left(1 - \frac{p}{1 - i\frac{\zeta}{n}} \right)^{-1} \\ &= (1 - p) \frac{1 - i\frac{\zeta}{n}}{(1 - p) - i\frac{\zeta}{n}} \\ &= (1 - p) \frac{(1 - p) - i\frac{\zeta}{n} + p}{(1 - p) - i\frac{\zeta}{n}} \\ &= (1 - p) + \frac{p(1 - p)}{(1 - p) - i\frac{\zeta}{n}} \quad \text{für } \left| \frac{p}{1 - i\frac{\zeta}{n}} \right| < 1. \end{aligned}$$

Nach Satz 13.16 c) gilt $\psi'(0) = i\mathbf{E}[W]$. Es ist

$$\begin{aligned} \psi'(\zeta) &= \frac{-p(1 - p) \left(-\frac{i}{n}\right)}{\left((1 - p) - i\frac{\zeta}{n}\right)^2} = \frac{p(1 - p)i}{n \left((1 - p)^2 - \frac{2(1 - p)i\zeta}{n} - \frac{\zeta^2}{n^2}\right)} \\ &= \frac{p(1 - p)i}{n(1 - p)^2 - 2(1 - p)i\zeta - \frac{\zeta^2}{n}}. \end{aligned}$$

Also gilt $\psi'(0) = \frac{p(1 - p)i}{n(1 - p)^2} = \frac{ip}{n(1 - p)}$, was

$$\mathbf{E}[W] = \frac{p}{n(1 - p)}$$

nach sich zieht.

Lösung für alle Fahrspuren

Mit $n = n_i$ und $p = 1 - e^{-n_i T_i}$ gilt

$$\mathbf{E}[W_i] = \frac{1 - e^{-n_i T_i}}{n_i e^{-n_i T_i}} = \frac{e^{n_i T_i} - 1}{n_i}.$$

Damit ergibt sich als Gesamtergebnis:

$$\mathbf{E}[Z_r] = \sum_{i=1}^r \frac{e^{n_i T_i} - 1}{n_i} = \sum_{i=1}^r \frac{e^{n_i b_i T/b} - 1}{n_i} - \sum_{i=1}^r \frac{1}{n_i}.$$

Bei gegebenen n_1, \dots, n_r , b und T ist das Minimum von $\mathbf{E}[Z_r](b_1, \dots, b_r)$ unter der Nebenbedingung $b_1 + \dots + b_r = b$ gesucht. Dazu äquivalent ist

$$\mathbf{E}[Z_r] = \sum_{i=1}^{r-1} \frac{e^{n_i b_i T/b} - 1}{n_i} + \frac{e^{n_r(b-b_1-\dots-b_{r-1})T/b} - 1}{n_r} - \sum_{i=1}^r \frac{1}{n_i},$$

was bezüglich b_1, \dots, b_{r-1} ohne Nebenbedingungen zu minimieren ist. Es ist

$$\begin{aligned} \frac{\partial \mathbf{E}[Z_r]}{\partial b_j} &= \frac{1}{n_j} \cdot n_j \cdot \frac{T}{b} \cdot e^{n_j b_j T/b} + \frac{1}{n_r} \cdot n_r \cdot (-1) \cdot \frac{T}{b} \cdot e^{n_r(b-b_1-\dots-b_{r-1})T/b} \\ &= \frac{T}{b} \cdot \left(e^{n_j b_j T/b} - e^{n_r(b-b_1-\dots-b_{r-1})T/b} \right). \end{aligned}$$

Mit der notwendigen Bedingung für ein Minimum folgt für $j = 1, \dots, r-1$:

$$\begin{aligned} \frac{\partial \mathbf{E}[Z_r]}{\partial b_j} &= 0 \\ \iff e^{n_j b_j T/b} &= e^{n_r(b-b_1-\dots-b_{r-1})T/b} \\ \iff n_j b_j &= n_r(b-b_1-\dots-b_{r-1}). \end{aligned}$$

Die b_1, b_2, \dots, b_r ergeben sich aus dem Gleichungssystem

$$\begin{aligned} b &= b_1 + \dots + b_r \\ n_r b_r &= n_j b_j \quad j = 1, \dots, r-1. \end{aligned}$$

Es ergibt sich $b_j = \frac{n_r}{n_j} b_r$. Eingesetzt in die 1. Gleichung liefert dies

$$b = \left(\frac{n_r}{n_1} + \dots + \frac{n_r}{n_r} \right) b_r = n_r b_r \sum_{j=1}^r \frac{1}{n_j}.$$

Folglich gilt:

$$b_r = b \cdot \frac{\frac{1}{n_r}}{\sum_{j=1}^r \frac{1}{n_j}} \quad \text{und} \quad b_i = b \cdot \frac{\frac{1}{n_i}}{\sum_{j=1}^r \frac{1}{n_j}}.$$

Da n_1, \dots, n_r gegeben sind, sind die b_i die optimalen Spurbreiten. Es gilt

$$n_j b_j \frac{T}{b} = \frac{T}{\sum_{k=1}^r \frac{1}{n_k}}$$

und somit ist unsere Gesamtwarezeit gegeben durch

$$\mathbf{E}[Z_r] = \left(\exp \left\{ \frac{T}{\sum_{j=1}^r \frac{1}{n_j}} \right\} - 1 \right) \cdot \sum_{j=1}^r \frac{1}{n_j}.$$

17.4 Stationärer Zustand

Bisher wurden die Kenngrößen von Erneuerungsprozessen jeweils zu einem festen Zeitpunkt $t \in \mathbb{R}^+$ betrachtet. Es stellt sich deshalb die Frage, wie sich diese Kenngrößen für $t \rightarrow \infty$ verhalten.

17.25 Definition (arithmetisch verteilt, Gitterkonstante):

Eine diskrete Zufallsvariable X , die nur Werte der Form $x_m := m \cdot d$ mit $d > 0$ und $m \in \mathbb{Z}$ annehmen kann, heißt arithmetisch verteilt. Das größte d mit dieser Eigenschaft heißt Gitterkonstante.

17.26 Satz (Blackwell'sches Erneuerungstheorem (1949)):

Es sei $(X_n)_{n \in \mathbb{N}_0}$ ein (einfacher) Erneuerungsprozess mit der Verteilungsfunktion F , Erwartungswert $\mu := \mathbf{E}[X_1]$ und Erneuerungsfunktion $R(t)$. Ist F nicht arithmetisch, dann gilt für ein festes $h > 0$:

$$\lim_{t \rightarrow \infty} (R(t) - R(t-h)) = \begin{cases} \frac{h}{\mu} & , \mu < \infty \\ 0 & , \mu = \infty. \end{cases}$$

Zum Beweis des Blackwell'schen Erneuerungstheorems werden zunächst einige Hilfsergebnisse zur Verfügung gestellt, die im Wesentlichen die charakteristische Funktionen bzw. die Fourier-Transformierte betreffen. Dabei soll der Begriff der Fourier-Transformierten hier auf Lebesgue-integrierbare (kurz: λ -integrierbare) Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ durch die Definition

$$\hat{f}(t) := \int_{\mathbb{R}} e^{itx} f(x) dx$$

ausgedehnt werden. Darüberhinaus sei im Folgenden für $A \subseteq \mathbb{R}$

$$1_A(x) := \begin{cases} 1 & , x \in A \\ 0 & , \text{sonst} \end{cases}$$

17.27 Lemma:

Sei $h_\alpha(x) = (2\alpha - |x|)1_{(-2\alpha, 2\alpha)}(x)$ für $\alpha > 0$. Dann ist $h_\alpha \in C_0$ und $\hat{h}_\alpha \in C_b$ mit

$$\hat{h}_\alpha(t) = \frac{\sin^2(\alpha t)}{\alpha^2 t^2} \quad (t \neq 0) \quad \text{und} \quad \hat{h}_\alpha(0) = 1.$$

Außerdem ist $(h_\alpha + h_\beta)^\wedge = \hat{h}_\alpha + \hat{h}_\beta > 0$ für alle $\alpha, \beta \in \mathbb{R}$ mit $\frac{\alpha}{\beta} \notin \mathbb{Q}$.

Beweis:

Die Darstellung der Fourier-Transformierten kann direkt unter Verwendung der üblichen Additionstheoreme für trigonometrische Funktionen nachgerechnet werden. Offensichtlich ist dann $\hat{h}_\alpha + \hat{h}_\beta \geq 0$. Gleichheit gilt genau für $\sin \alpha t = \sin \beta t = 0$. Dies ist genau dann der Fall, wenn $t \in \frac{\pi}{\alpha}\mathbb{Z} \cap \frac{\pi}{\beta}\mathbb{Z}$ ist. Für $\frac{\alpha}{\beta} \notin \mathbb{Q}$ ist diese Menge aber leer. ■

Das nachfolgende Lemma wird in einer etwas allgemeineren Version in der Literatur als Parseval-Gleichung bezeichnet. Hier soll allerdings nur ein Spezialfall aufgeführt werden.

17.28 Lemma (Parseval-Gleichung):

Es seien μ ein endliches Maß mit charakteristischer Funktion φ und h eine nichtnegative λ -integrierbare Funktion mit Fourier-Transformierter \hat{h} . Dann gilt

$$\int_{\mathbb{R}} \hat{h}(x) d\mu(x) = \int_{\mathbb{R}} \varphi(x) h(x) dx.$$

Beweis:

Nach Definition sind

$$\varphi(t) = \int_{\mathbb{R}} e^{itx} d\mu(x) \quad \text{und} \quad \hat{h}(t) = \int_{\mathbb{R}} e^{itx} f(x) dx$$

Die Behauptung kommt nun unmittelbar aus dem Satz von Fubini. ■

17.29 Lemma (Riemann-Lebesgue-Lemma):

Ist f λ -integrierbar mit Fourier-Transformierter \hat{f} , so ist

$$\lim_{|t| \rightarrow \infty} \hat{f}(t) = 0.$$

Beweis:

Für $f = 1_{(a,b)}$ für beliebige $a, b \in \mathbb{R}$ folgt die Behauptung unmittelbar aus

$$\hat{f}(t) = \frac{e^{ibt} - e^{iat}}{it}.$$

Sei nun f eine beliebige λ -integrierbare Funktion und $\varepsilon > 0$. Dann existieren paarweise disjunkte Intervalle I_1, \dots, I_n und $c_1, \dots, c_n \in \mathbb{C}$, so dass

$$\int_{\mathbb{R}} |f(x) - g(x)| dx < \varepsilon \quad \text{für} \quad g = \sum_{j=1}^n c_j 1_{I_j}$$

gilt. Damit folgt für alle $t \in \mathbb{R}$

$$\left| \hat{f}(t) - \hat{g}(t) \right| = \left| \int_{\mathbb{R}} e^{itx} (f(x) - g(x)) dx \right| \leq \int_{\mathbb{R}} |f(x) - g(x)| dx < \varepsilon.$$

Da g nach Definition eine Linearkombination der Funktionen 1_{I_j} ist, gilt

$$\lim_{|t| \rightarrow \infty} \hat{g}(t) = 0$$

und es folgt die Behauptung. ■

Dass folgende Lemma verallgemeinert den Stetigkeitssatz von Lévy, indem statt für die schwache Konvergenz endlicher Maße ein Kriterium für die vage Konvergenz lokal endlicher Maße angegeben wird.

17.30 Lemma:

Es sei $(\mu_n)_{n \in \mathbb{N}}$ eine Folge lokal endlicher Maße auf \mathbb{R} . Gilt für eine in 0 stetige Funktion g und eine strikt positive Funktion $h \in C_b$

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} e^{itx} h(x) d\mu_n(x) = g(t)$$

für alle $t \in \mathbb{R}$, so existiert ein lokal endliches Maß μ , so dass $\mu_n \xrightarrow{v} \mu$ und

$$g(t) = \int_{\mathbb{R}} e^{itx} h(x) d\mu(x) \quad (17.1)$$

für alle $t \in \mathbb{R}$ gilt.

Beweis:

Definiere die Folge von Maßen $(\hat{\mu}_n)_{n \in \mathbb{N}}$ durch $d\hat{\mu}_n = h d\mu_n$, d.h.

$$\hat{\mu}_n(A) = \int_A h d\mu_n \quad (A \in \mathfrak{B}).$$

Dann sind alle $\hat{\mu}_n$ endlich, da $\hat{\mu}_n(\mathbb{R}) = g(0) < \infty$ ist. Ihre charakteristischen Funktionen sind

$$\int_{\mathbb{R}} e^{itx} h(x) d\mu_n(x),$$

die nach Voraussetzung gegen die in 0 stetige Funktion g konvergieren. Nach dem Stetigkeitssatz von Lévy 14.13 existiert ein endliches Maß $\hat{\mu}$ mit charakteristischer Funktion g und $\hat{\mu}_n \xrightarrow{w} \hat{\mu}$. Wird nun $d\mu = \frac{1}{h} d\hat{\mu}$, d.h.

$$\mu(A) = \int_A \frac{1}{h} d\hat{\mu} \quad (A \in \mathfrak{B})$$

definiert, so folgt wegen der Positivität von h für jedes beschränkte Intervall I

$$\mu(I) \leq \sup_{x \in I} \left\{ \frac{1}{h(x)} \right\} \hat{\mu}(I) < \infty,$$

d.h. μ ist lokal endlich.

Für eine beliebige Funktion $f \in C_0$ ist auch $\frac{f}{h} \in C_0$ und es folgt

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f d\mu_n = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{f}{h} d\hat{\mu}_n = \int_{\mathbb{R}} \frac{f}{h} d\hat{\mu} = \int_{\mathbb{R}} f d\mu,$$

also $\mu_n \xrightarrow{v} \mu$.

Die Darstellung (17.1) folgt unmittelbar aus der Definition von μ und daraus, dass g die charakteristische Funktion von $\hat{\mu}$ ist. ■

Beweis des Blackwell'schen Erneuerungstheorems:

Die Erneuerungsfunktion R ist monoton nichtfallend und rechtsseitig stetig, definiert also ein Maß $\tilde{\mu}$ auf \mathbb{R} (auch Erneuerungsmaß genannt), für $A \in \mathfrak{B}$ gilt damit

$$\tilde{\mu}(A) = \sum_{n=0}^{\infty} P(S_n \in A).$$

Es wird nun eine Familie von Maßen $\tilde{\mu}^a$ auf \mathbb{R} durch $\tilde{\mu}^a(A) := \tilde{\mu}(a + A)$ definiert ($a + A := \{a + x : x \in A\}$) und gezeigt, dass

$$\tilde{\mu}^a \xrightarrow{v} \frac{1}{\mu} \lambda \quad (a \rightarrow \infty)$$

gilt. Speziell folgt dann

$$R(t) - R(t - h) = \tilde{\mu}((t - h, t]) = \tilde{\mu}^t((-h, 0]) \xrightarrow{t \rightarrow \infty} \frac{1}{\mu} \lambda((-h, 0]) = \frac{h}{\mu}.$$

(Für $\mu = \infty$ erhält man jeweils die Konvergenz gegen 0.)

Zum Nachweis der vagen Konvergenz der $(\tilde{\mu}^a)$ wird im Wesentlichen Lemma 17.30 angewendet - allerdings auf die leicht modifizierten Maße $\tilde{\nu}, \tilde{\nu}^a$, die durch

$$\tilde{\nu}(A) := \tilde{\mu}(A) \tilde{\mu}(-A) \quad \text{und} \quad \tilde{\nu}^a(A) = \tilde{\mu}^a(A) + \tilde{\mu}^a(-A) \quad (A \in \mathfrak{B})$$

definiert sind. Da $\tilde{\mu}(A) = 0$ für $A \subseteq \mathbb{R}^-$ gilt, ist das Grenzverhalten von $\tilde{\mu}^a$ und $\tilde{\nu}^a$ gleich.

Eine geeignete Funktion h zur Anwendung von Lemma 17.30 ist, wie noch zu zeigen ist, eine Funktion $\hat{h} = \hat{h}_\alpha + \hat{h}_\beta$ mit irrationalem Quotienten $\frac{\alpha}{\beta}$ aus Lemma 17.27, denn es gilt dann

$$\lim_{a \rightarrow \infty} \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) = \frac{1}{\mu} \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\lambda(x). \quad (17.2)$$

Es verbleibt also noch, diese Identität nachzuweisen.

Zu diesem Zweck werden einige Hilfsmaße definiert. In Verallgemeinerung von $\tilde{\mu} =: \tilde{\mu}_1$ bzw. $\tilde{\nu} =: \tilde{\nu}_1$ werden Maße $\tilde{\mu}_s$ und $\tilde{\nu}_s$ für $s \in (0, 1)$ durch

$$\tilde{\mu}_s(A) := \sum_{n=0}^{\infty} s^n \cdot P(S_n \in A) \quad \text{und} \quad \tilde{\nu}_s(A) = \tilde{\mu}_s(A) + \tilde{\mu}_s(-A), \quad A \in \mathfrak{B}$$

definiert. Für $s \rightarrow 1-$ konvergieren die $\tilde{\mu}_s$ vage gegen $\tilde{\nu}$ (und damit auch $\tilde{\nu}_s \xrightarrow{v} \tilde{\nu}$), da $\tilde{\mu}_s(B) \xrightarrow{s \rightarrow 1-} \tilde{\mu}(B)$ für alle beschränkten $B \in \mathfrak{B}$ (vergleiche Satz 14.11).

Anders als $\tilde{\mu}$ sind alle $\tilde{\mu}_s$ (und damit auch alle $\tilde{\nu}_s$), $s \in (0, 1)$, endliche Maße mit $\tilde{\mu}_s(\mathbb{R}) = \frac{1}{1-s}$. Ist φ die charakteristische Funktion von X_1 , so erhält man φ^n als charakteristische Funktion von S_n und für $s \in (0, 1)$

$$\psi_s := \sum_{n=0}^{\infty} s^n \cdot \varphi^n = \frac{1}{1 - s \cdot \varphi}$$

als charakteristische Funktion von $\tilde{\mu}_s$ und

$$\psi_s(-t) + \psi_s(t) = \overline{\psi_s(t)} + \psi_s(t) = 2 \operatorname{Re}(\psi_s(t))$$

als charakteristische Funktion von $\tilde{\nu}_s$.

Der Grenzwert

$$\psi(t) := \lim_{s \rightarrow 1-} \psi_s(t) = \frac{1}{1 - \varphi(t)}$$

existiert zwar für $t \neq 0$, ist aber nicht die charakteristische Funktion von $\tilde{\mu}$, da zu $\tilde{\mu}$ als unendlichem Maß keine charakteristische Funktion existiert.

Mit diesen Bezeichnungen gilt

17.31 Lemma:

Sei $b > 0$ beliebig und $\mu = E[X_1] < \infty$. Dann ist $\operatorname{Re}(\psi)$ λ -integrierbar auf $[-b, b]$ und für alle $f \in C_0$ gilt

$$\lim_{s \rightarrow 1-} \int_{-b}^b f(t) \operatorname{Re}(\psi_s)(t) dt = \frac{\pi f(0)}{\mu} + \int_{-b}^b f(t) \operatorname{Re}(\psi)(t) dt.$$

Der Beweis dieses Lemmas wird zurückgestellt und es wird zunächst der Nachweis der Identität (17.2) fortgesetzt. Analog zu $\tilde{\nu}^a$ werden die Maße $\tilde{\nu}_s^a$ durch

$$\tilde{\nu}_s^a(A) := \tilde{\nu}_s(a + A) = \sum_{n=0}^{\infty} s^n (P(S_n - a \in A) + P(-S_n - a \in A)) \quad (A \in \mathfrak{B})$$

definiert. Die charakteristische Funktion von $\tilde{\nu}_s^a$ ist nach Satz 13.16

$$2e^{-iat} \operatorname{Re}(\psi_s)(t).$$

Mit der Bezeichnung $h_t(x) := h(x - t)$ ist $\hat{h}_t(x) = e^{itx} \hat{h}(x)$. Unter Verwendung von Lemma 17.28 folgt damit

$$\int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) = \int_{\mathbb{R}} \hat{h}_t(x) d\tilde{\nu}_s^a(x) = \int_{-\infty}^{\infty} 2e^{-iax} \operatorname{Re}(\psi_s)(x) h_t(x) dx. \quad (17.3)$$

Nun folgt (17.2) durch drei Schritte aus (17.3):

- (i): Grenzübergang $s \rightarrow 1-$ auf der linken Seite von (17.3).
- (ii): Grenzübergang $s \rightarrow 1-$ auf der rechten Seite von (17.3).
- (iii): Grenzübergang $a \rightarrow \infty$ auf der rechten Seite.

Zu (i): Es gilt

$$\lim_{s \rightarrow 1-} \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) = \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x). \quad (17.4)$$

Zum Nachweis sei $\varepsilon > 0$ beliebig. Wähle nun $N \in \mathbb{N}$ so groß, dass

$$\sum_{n \geq N} \frac{1}{n^2} < \varepsilon$$

ist und $s < 1$ so groß, dass

$$\int_{-N}^N e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) - \int_{-N}^N e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) < \varepsilon$$

(Existenz von N und s folgt aus der Konvergenz der Reihe $\sum \frac{1}{n^2}$ sowie aus $\tilde{\nu}_s^a \xrightarrow{v} \tilde{\nu}^a$). Außerdem sind die Differenzen $R(t+a) - R(t)$ für festes $a > 0$ durch eine Konstante $C_a \in \mathbb{R}$ beschränkt, wie man unter Verwendung des Erneuerungsarguments schließen kann. Dazu sei $t+y$ der erste Erneuerungszeitpunkt in $(t, t+a)$. Es folgt

$$R(t+a) - R(t) = 1 + R(t+a) - R(t+y) = 1 + R(a-y) \leq 1 + R(a) =: C_a,$$

speziell ist $R(t+1) - R(t) \leq C_1$ für alle $t \in \mathbb{R}^+$. Damit folgt

$$\int_n^{n+1} d\tilde{\nu}_s^a(x) \leq \int_n^{n+1} d\tilde{\nu}^a(x) \leq C$$

mit einer von $n \in \mathbb{Z}$ unabhängigen Konstanten C . Da $\hat{h}(x) \leq \frac{C'}{x^2}$ mit einer weiteren Konstanten C' gilt, folgt

$$\begin{aligned} & \left| \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) - \int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) \right| \\ & \leq \left| \int_{-N}^N e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) - \int_{-N}^N e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) \right| \\ & + \left| \int_{\mathbb{R} \setminus [-N, N]} e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) \right| + \left| \int_{\mathbb{R} \setminus [-N, N]} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) \right| \\ & < \varepsilon + \sum_{n=N}^{\infty} \left(\left| \int_n^{n+1} e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) \right| + \left| \int_n^{n+1} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) \right| \right) \\ & + \sum_{n=-\infty}^N \left(\left| \int_{n-1}^n e^{itx} \hat{h}(x) d\tilde{\nu}_s^a(x) \right| + \left| \int_{n-1}^n e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) \right| \right) \\ & \leq \varepsilon + C' \sum_{n=N}^{\infty} \frac{1}{n^2} \left(\left| \int_n^{n+1} d\tilde{\nu}_s^a(x) \right| + \left| \int_n^{n+1} d\tilde{\nu}^a(x) \right| \right) \\ & + C' \sum_{n=-\infty}^{-N} \frac{1}{n^2} \left(\left| \int_{n-1}^n d\tilde{\nu}_s^a(x) \right| + \left| \int_{n-1}^n d\tilde{\nu}^a(x) \right| \right) \\ & < \varepsilon + 4C'C \sum_{n=N}^{\infty} \frac{1}{n^2} < (1 + 4C'C)\varepsilon. \end{aligned}$$

Zu (ii): Es wird $b > 0$ so gewählt, dass $[-b, b]$ den Träger von h_t umfasst ($h_t \in C_0$, da $h \in C_0$). Wird nun in Lemma 17.31 $f(x) = 2e^{-iax}h(x-t)$ gesetzt, so folgt

$$\begin{aligned} \lim_{s \rightarrow 1-} \int_{-\infty}^{\infty} 2e^{-iax} \operatorname{Re}(\psi_s)(x) h_t(x) dx &= \lim_{s \rightarrow 1-} \int_{-b}^b 2e^{-iax} \operatorname{Re}(\psi_s)(x) h(x-t) dx \\ &= \frac{2\pi}{\mu} h(-t) + \int_{-b}^b 2e^{-iax} \operatorname{Re}(\psi)(x) h(x-t) dx. \end{aligned}$$

Einsetzen dieser Identität sowie von (17.4) in (17.3) liefert

$$\int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) = \frac{2\pi}{\mu} h(-t) + \int_{-b}^b 2e^{-iax} \operatorname{Re}(\psi)(x) h(x-t) dx.$$

Formuliert man Satz 13.13 für allgemeine Fourier-Transformierte, so erhält man

$$h(-t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \hat{h}(x) dx$$

und damit

$$\int_{\mathbb{R}} e^{itx} \hat{h}(x) d\tilde{\nu}^a(x) = \frac{1}{\mu} \int_{-\infty}^{\infty} e^{itx} \hat{h}(x) dx + \int_{-b}^b 2e^{-iax} \operatorname{Re}(\psi)(x) h(x-t) dx. \quad (17.5)$$

Zu (iii): Es wird nun der Grenzübergang $a \rightarrow \infty$ betrachtet. Das letzte Integral entspricht der Fourier-Transformation einer nach Lemma 17.31 λ -integrierbaren Funktion und nach Lemma 17.29 verschwindet es beim Grenzübergang.

Damit ist (17.2) gezeigt und zum vollständigen Beweis des Blackwell'schen Erneuerungstheorems bleibt nun noch der Beweis von Lemma 17.31 nachzutragen.

Beweis (von Lemma 17.31):

Im ersten Schritt wird die λ -Integrierbarkeit von $\operatorname{Re}(\psi)$ auf $[-b, b]$ gezeigt. Wegen $\psi = \frac{1}{1-\varphi}$ und $\varphi(t) \neq 0$ für $t \in [-b, b] \setminus \{0\}$ ist $\operatorname{Re}(\psi)$ auf $[-b, b] \setminus \{0\}$ stetig und es genügt die Integrierbarkeit auf einem Intervall $[-\varepsilon, \varepsilon]$ für ein beliebig kleines $\varepsilon > 0$ zu zeigen. Hier wird ε so klein gewählt, dass $|\varphi(t) - 1| \geq \frac{\mu|t|}{2}$ für alle $t \in [-\varepsilon, \varepsilon]$ ist. Dies ist möglich, da $\lim_{t \rightarrow 0} \frac{\varphi(t)-1}{t} = \varphi'(0) = i\mu$ (vergleiche Satz 13.16(c)) gilt. Dann ist

$$\int_{-\varepsilon}^{\varepsilon} \operatorname{Re}(\psi)(t) dt = \int_{-\varepsilon}^{\varepsilon} \operatorname{Re}\left(\frac{1}{1-\varphi(t)}\right) dt = \int_{-\varepsilon}^{\varepsilon} \frac{\operatorname{Re}(1-\bar{\varphi})(t)}{|1-\varphi(t)|^2} dt \geq 0$$

und

$$\int_{-\varepsilon}^{\varepsilon} \operatorname{Re}(\psi)(t) dt \leq \frac{2}{\mu} \int_{-\varepsilon}^{\varepsilon} \frac{\operatorname{Re}(1-\bar{\varphi})(t)}{t^2} dt = \frac{4}{\mu} \int_0^{\varepsilon} \frac{E[1 - \cos(tX_1)]}{t^2} dt.$$

Nach dem Satz von Fubini ist

$$\begin{aligned} \int_0^{\varepsilon} \frac{E[1 - \cos(tX_1)]}{t^2} dt &= E \left[\int_0^{\varepsilon} \frac{1 - \cos(tX_1)}{t^2} dt \right] = E \left[|X_1| \int_0^{\varepsilon X_1} \frac{1 - \cos t}{t^2} dt \right] \\ &\leq E[|X_1|] \int_0^{\infty} \frac{1 - \cos t}{t^2} dt < \infty. \end{aligned}$$

Im zweiten Teil des Beweises muss noch die Integralformel nachgewiesen werden. Dazu wird für beliebiges $f \in C_0$

$$\begin{aligned} I(s, b) &:= \int_{-b}^b f(t) \operatorname{Re}(\psi_s - \psi)(t) dt = \int_{-b}^b f(t) \operatorname{Re} \left(\frac{1}{1-s\varphi} - \frac{1}{1-\varphi} \right) dt \\ &= \int_{-b}^b f(t) \frac{s-1}{|1-s\varphi(t)|^2} \operatorname{Re} \left(\frac{\varphi(1-s\bar{\varphi})}{1-\varphi} \right) (t) dt \end{aligned}$$

definiert. Zu zeigen ist

$$\lim_{s \rightarrow 1-} I(s, b) = 0.$$

Der Integrant von $I(s, b)$ konvergiert auf $[-b, b] \setminus [-\varepsilon, \varepsilon]$ für $s \rightarrow 1-$ gegen 0, also gilt

$$I(s, b) - I(s, \varepsilon) \xrightarrow{s \rightarrow 1-} 0.$$

Daher bleibt nur $I(s, \varepsilon)$ für ein beliebig kleines $\varepsilon > 0$ zu untersuchen. Wegen

$$\operatorname{Re} \left(\frac{\varphi(1 - s\bar{\varphi})}{1 - \varphi} \right) (t) = \operatorname{Re} \left(\frac{\varphi(1 - \bar{\varphi})}{1 - \varphi} \right) (t) - (s - 1) \operatorname{Re} \left(\frac{\varphi\bar{\varphi}}{1 - \varphi} \right) (t)$$

ist $I(s, \varepsilon) = I_1(s, \varepsilon) + I_2(s, \varepsilon)$ mit

$$I_1(s, \varepsilon) = \int_{-\varepsilon}^{\varepsilon} f(t) \frac{s - 1}{|1 - s\varphi(t)|^2} \operatorname{Re} \left(\frac{\varphi(1 - \bar{\varphi})}{1 - \varphi} \right) (t) dt$$

und

$$I_2(s, \varepsilon) = - \int_{-\varepsilon}^{\varepsilon} f(t) \frac{(s - 1)^2}{|1 - s\varphi(t)|^2} |\varphi(t)|^2 \operatorname{Re}(\psi)(t) dt.$$

Wegen $|1 - s\varphi(t)| \geq 1 - s$ ist der Integrand von $I_2(s, \varepsilon)$ unabhängig von s beschränkt und mit dem Satz von der majorisierten Konvergenz folgt $\lim_{s \rightarrow 1} I_2(s, \varepsilon) = 0$ für alle $\varepsilon \in (0, b]$, so dass nur noch der Term $I_1(s, \varepsilon)$ zu untersuchen ist.

Wähle dazu $\eta \in (0, 1)$ beliebig und definiere die Funktion g

$$g(s, t) := \frac{|1 - s\varphi(t)|^2}{(1 - s)^2 + \mu^2 t^2}$$

und die auf $[-\varepsilon, \varepsilon]$ stetige Funktion h durch

$$h(t) := -f(t) \operatorname{Re} \left(\frac{\varphi(1 - \bar{\varphi})}{1 - \varphi} \right) (t), \quad t \neq 0 \quad \text{und} \quad h(0) = f(0).$$

Die Stetigkeit von h kommt dabei aus

$$\lim_{t \rightarrow 0} \frac{1 - \overline{\varphi(t)}}{1 - \varphi(t)} = \frac{\overline{\varphi'(0)}}{\varphi'(0)} = -1.$$

Damit gilt dann

$$I_1(s, \varepsilon) = \int_{-\varepsilon}^{\varepsilon} \frac{h(t)}{g(s, t)} \frac{1 - s}{(1 - s)^2 + \mu^2 t^2} dt = \int_{\frac{-\varepsilon}{1-s}}^{\frac{\varepsilon}{1-s}} \frac{h(t(1-s))}{g(s, t(1-s))} \frac{1}{1 + \mu^2 t^2} dt.$$

Erneute Benutzung von $\varphi'(0) = i\mu$ liefert

$$\lim_{t \rightarrow 0} \lim_{s \rightarrow 1-} g(s, t) = \lim_{t \rightarrow 0} \frac{|1 - \varphi(t)|^2}{\mu^2 t^2} = \frac{|\varphi'(0)|^2}{\mu^2} = 1,$$

was die Existenz eines $\varepsilon \in (0, b)$ zeigt, so dass für alle $s \in [1 - \varepsilon, 1)$ und $t \in [-\varepsilon, \varepsilon]$ die Ungleichungen

$$1 - \eta \leq g(s, t) \leq 1 + \eta \quad \text{und} \quad \frac{1 - \eta}{1 + \eta} h(0) \leq \frac{h(t)}{g(s, t)} \leq \frac{1 + \eta}{1 - \eta} h(0)$$

gelten. Für $s \in [1 - \varepsilon, 1)$ folgen

$$I_1(s, \varepsilon) \leq \frac{1 + \eta}{1 - \eta} h(0) \int_{-\infty}^{\infty} \frac{dt}{1 + \mu^2 t^2} = \frac{1 + \eta}{1 - \eta} \frac{\pi h(0)}{\mu} \quad \text{und} \quad I_1(s, \varepsilon) \geq \frac{1 - \eta}{1 + \eta} \frac{\pi h(0)}{\mu}.$$

Da $\eta > 0$ beliebig vorgegeben war und $h(0) = f(0)$ ist, folgt

$$\lim_{s \rightarrow 1-} I(s, b) = \frac{\pi f(0)}{\mu}$$

und damit die Behauptung. ■

17.32 Bemerkung:

Das Blackwell'sche Erneuerungstheorem kann auch für den arithmetischen Fall formuliert werden. Mit $\mu := E[X_1]$ und Gitterkonstante d gilt hier

$$\lim_{n \rightarrow \infty} (R(nd) - R(nd - 0)) = \frac{d}{\mu}.$$

In Vorbereitung auf den Fundamentalsatz der Erneuerungstheorie erläutern wir zunächst den Begriff der direkten Riemann-Integrierbarkeit.

17.33 Definition (direkte Riemann-Integrierbarkeit):

Es sei g eine reellwertige Funktion auf \mathbb{R}^+ . Für ein $\delta > 0$ und $n \in \mathbb{N}_0$ werden

$$\begin{aligned} I_n^\delta &:= (\delta \cdot n, \delta \cdot (n + 1)], \\ m_n^\delta &:= \inf\{g(t) | t \in I_n^\delta\} \quad \text{und} \quad M_n^\delta := \sup\{g(t) | t \in I_n^\delta\}, \\ \underline{\sigma}(\delta) &:= \delta \cdot \sum_{n \in \mathbb{N}_0} m_n^\delta \quad \text{und} \quad \bar{\sigma}(\delta) := \delta \cdot \sum_{n \in \mathbb{N}_0} M_n^\delta \end{aligned}$$

gesetzt. g heißt direkt Riemann-integrierbar, falls $\underline{\sigma}(\delta)$ und $\bar{\sigma}(\delta)$ beide für alle $\delta > 0$ absolut konvergieren und

$$\lim_{\delta \rightarrow 0} (\bar{\sigma}(\delta) - \underline{\sigma}(\delta)) = 0$$

ist. Im Folgenden bezeichne \mathcal{R} die Menge der direkt Riemann-integrierbaren Funktionen auf \mathbb{R}^+ .

Bemerkung:

Ersetzt man den Definitionsbereich von g durch ein Kompaktum, erhält man den Begriff der gewöhnlichen Riemann-Integrierbarkeit.

17.34 Beispiel:

- Es sei $g(x) := e^{-x}$, $x \in \mathbb{R}^+$. Dann ist

$$\begin{aligned} m_n^\delta &= \inf\{g(t) | t \in I_n^\delta\} = g(\delta \cdot (n + 1)) = e^{-\delta(n+1)}, \\ M_n^\delta &= \sup\{g(t) | t \in I_n^\delta\} = g(\delta \cdot n) = e^{-\delta n}. \end{aligned}$$

Die beiden unendlichen Reihen $\underline{\sigma}(\delta)$ und $\overline{\sigma}(\delta)$ sind absolut konvergent:

$$\begin{aligned}\underline{\sigma}(\delta) &= \delta \sum_{n \in \mathbb{N}_0} e^{-\delta(n+1)} = \delta e^{-\delta} \sum_{n \in \mathbb{N}_0} (e^{-\delta})^n = \delta e^{-\delta} \cdot \frac{1}{1 - e^{-\delta}}, \\ \overline{\sigma}(\delta) &= \delta \sum_{n \in \mathbb{N}_0} e^{-\delta n} = \delta \sum_{n \in \mathbb{N}_0} (e^{-\delta})^n = \delta \frac{1}{1 - e^{-\delta}}.\end{aligned}$$

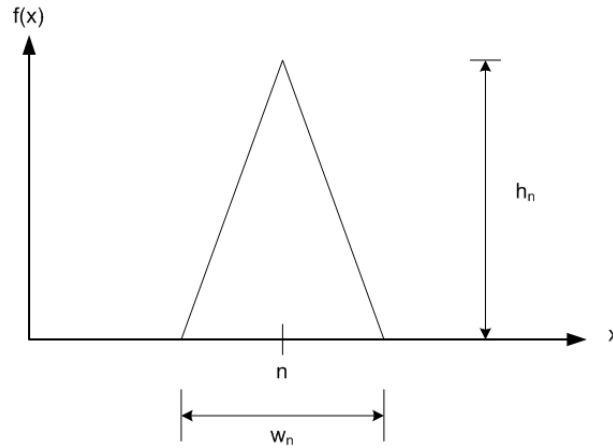
Es gilt weiter:

$$\lim_{\delta \rightarrow 0} \left(\frac{\delta}{1 - e^{-\delta}} - \frac{\delta e^{-\delta}}{1 - e^{-\delta}} \right) = \lim_{\delta \rightarrow 0} \left(\delta \frac{1 - e^{-\delta}}{1 - e^{-\delta}} \right) = 0,$$

d.h. $g(x)$ ist direkt Riemann-integrierbar (vgl. auch Satz 17.36).

- Es sei

$$f(x) := \sum_{n=1}^{\infty} I_{[-\frac{w_n}{2}, \frac{w_n}{2}]}(x - n) \cdot \begin{cases} h_n + \frac{2h_n}{w_n}(x - n), & \text{für } -\frac{w_n}{2} \leq x - n < 0 \\ h_n - \frac{2h_n}{w_n}(x - n), & \text{für } 0 < x - n \leq \frac{w_n}{2}. \end{cases}$$



Der Graph der Funktion f_n besteht aus lauter Dreiecken. Das n -te Dreieck beginnt dabei bei $n - \frac{w_n}{2}$, erreicht bei n seinen höchsten Punkt (nämlich (n, h_n)) und endet bei $n + \frac{w_n}{2}$. Es sei $w_n < 1$, d.h. die Dreiecke überlappen sich nicht. Es sei $w_n \xrightarrow{n \rightarrow \infty} 0$ und $h_n \xrightarrow{n \rightarrow \infty} \infty$ mit der Bedingung $\sum_{n=1}^{\infty} h_n w_n < \infty$. Dann gilt:

$$\int_0^{\infty} f(x) dx = \sum_{k=1}^{\infty} \frac{1}{2} h_n w_n < \infty,$$

d.h. f ist Riemann-integrierbar. f ist aber nicht direkt Riemann-integrierbar, da:

$$\overline{\sigma}(1) = 1 \cdot \sum_{n \in \mathbb{N}_0} M_n^1 = \sum_{n \in \mathbb{N}_0} \sup\{f(t) | t \in (n, n+1]\} = \sum_{n \in \mathbb{N}_0} h_n = \infty.$$

17.35 Satz:

Es sei g eine reellwertige Funktion auf \mathbb{R}^+ . Ist g direkt-Riemann-integrierbar, dann ist g auch uneigentlich Riemann-integrierbar auf \mathbb{R}^+ und es gilt:

$$\lim_{\delta \rightarrow 0} \underline{\sigma}(\delta) = \lim_{\delta \rightarrow 0} \bar{\sigma}(\delta) = \int_{\mathbb{R}^+} g(t) dt,$$

wobei $\int_{\mathbb{R}^+} g(t) dt$ das uneigentliche Riemann-Integral von g bezeichnet.

Beweis:

Aufgrund der Voraussetzung gilt $\bar{\sigma}(\delta) < \infty$, $\underline{\sigma}(\delta) < \infty$ sowie

$$\begin{aligned} 0 &= \lim_{\delta \rightarrow 0} [\bar{\sigma}(\delta) - \underline{\sigma}(\delta)] \\ &= \lim_{\delta \rightarrow 0} \sum_{n=0}^{\infty} M_n^\delta \cdot \delta - \sum_{n=0}^{\infty} m_n^\delta \cdot \delta \\ &= \lim_{\delta \rightarrow 0} \sum_{n=0}^{\infty} \delta \cdot [M_n^\delta - m_n^\delta] \\ &\geq \lim_{\delta \rightarrow 0} \sum_{n|n\delta \leq a}^{\infty} \delta \cdot [M_n^\delta - m_n^\delta]. \end{aligned}$$

Womit gezeigt ist, dass g zumindest über $[0, a]$ Riemann-integrierbar ist.

Da $\bar{\sigma}(\delta) < \infty \forall \delta > 0$ ist, kann geschlossen werden, dass zu jedem $\varepsilon > 0$ ein $a := a(\varepsilon)$ existiert mit $\sum_{n>a} M_n^1 < \varepsilon$. Für alle $\delta > 0$ gilt deshalb:

$$\bar{\sigma}(\delta) - \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta = \sum_{n|n\delta > a} M_n^\delta \cdot \delta \leq \sum_{n>a} M_n^1 \cdot 1 < \varepsilon.$$

(i) Da g direkt-Riemann-integrierbar ist, gilt:

$$\lim_{\delta \rightarrow 0} \bar{\sigma}(\delta) = \sigma_0 < \infty$$

bzw.

$$|\sigma_0 - \bar{\sigma}(\delta)| < \varepsilon \forall \delta \leq \delta_0.$$

(ii) Da g Riemann-integrierbar ist auf $[0, a]$ gilt:

$$\left| \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta - \int_0^a g(t) dt \right| < \varepsilon \forall \delta < \delta_1.$$

(iii) Zusammenfassung:

$$\begin{aligned} \left| \sigma_0 - \int_0^a g(t) dt \right| &= \left| \sigma_0 - \bar{\sigma}(\delta) + \bar{\sigma}(\delta) - \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta + \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta - \int_0^a g(t) dt \right| \\ &\leq |\sigma_0 - \bar{\sigma}(\delta)| + \left| \bar{\sigma}(\delta) - \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta \right| + \left| \sum_{n|n\delta \leq a} M_n^\delta \cdot \delta - \int_0^a g(t) dt \right| \\ &\leq 3\varepsilon \forall \delta \leq \min(\delta_0, \delta_1, 1). \end{aligned}$$

■

17.36 Satz:

Ist $g \geq 0$ nichtwachsend, dann gilt $g \in \mathcal{R}$ genau dann, wenn g Riemann-integrierbar ist.

Beweis:

Es wird angenommen, dass g nicht wachsend und Riemann-integrierbar ist.

$$\begin{aligned} \infty &> \int_0^\infty g(t) dt = \sum_{n=0}^\infty \int_{n\delta}^{(n+1)\delta} g(t) dt \geq \sum_{n=0}^\infty g((n+1)\delta) \cdot \delta \\ &= \delta \cdot \sum_{n=0}^\infty g((n+1)\delta) = \delta \cdot \sum_{n=0}^\infty m_n^\delta = \underline{\sigma}(\delta). \end{aligned}$$

Außerdem gilt:

$$\begin{aligned} \sum_{n=0}^N M_n^\delta \cdot \delta - \sum_{n=0}^N m_n^\delta \cdot \delta &= \delta \cdot \sum_{n=0}^N [g(n \cdot \delta) - g((n+1) \cdot \delta)] \\ &= \delta \cdot [g(0) - g(N+1)] \xrightarrow{N \rightarrow \infty} \delta \cdot [g(0) - g(\infty)] < \infty. \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \bar{\sigma}(\delta) < \infty &\iff \underline{\sigma}(\delta) < \infty, \\ \bar{\sigma}(\delta) - \underline{\sigma}(\delta) &\leq \delta \cdot [g(0) - g(\infty)]. \end{aligned}$$

Schlussfolgerungen:

- (i) $\underline{\sigma}(\delta) < \infty$,
- (ii) $\bar{\sigma}(\delta) < \infty$,
- (iii) $\bar{\sigma}(\delta) - \underline{\sigma}(\delta) \rightarrow 0$ für $\delta \rightarrow 0$.

Die Umkehrung folgt aus Satz 17.35.

■

17.37 Satz:

Ist g Riemann-integrierbar auf $[0, \infty)$ und $g \leq z$, $z \in \mathcal{R}$, dann ist auch $g \in \mathcal{R}$.

Beweis:

Siehe S. I. Resnick: „Adventures in Stochastic Processes“, Birkhäuser, Boston, 1992 in Kapitel 3.10.

17.38 Satz (Fundamentalsatz der Erneuerungstheorie (Smith, 1955)):

Es sei $(X_n)_{n \in \mathbb{N}_0}$ ein (einfacher) Erneuerungsprozess mit der Verteilungsfunktion F , Erwartungswert $\mu := \mathbf{E}[X_1]$ und Erneuerungsfunktion $R(t)$. Außerdem sei $g \in \mathcal{R}$.

a) Ist F nicht arithmetisch, dann gilt

$$\lim_{t \rightarrow \infty} (R * g)(t) = \begin{cases} \frac{1}{\mu} \int_0^\infty g(x) dx & , \mu < \infty \\ 0 & , \mu = \infty. \end{cases}$$

b) Ist F arithmetisch mit der Gitterkonstanten d , dann gilt für alle $c > 0$

$$\lim_{n \rightarrow \infty} (R * g)(c + nd) = \begin{cases} \frac{d}{\mu} \sum_{n \in \mathbb{N}_0} g(c + nd) & , \mu < \infty \\ 0 & , \mu = \infty. \end{cases}$$

Beweis:

Der Beweis soll hier nur für den nichtarithmetischen Fall geführt werden. Dazu seien g eine direkt Riemann-integrierbare Funktion und (für $\delta > 0$) I_n^δ , m_n^δ und M_n^δ wie in Definition 17.33. Darüberhinaus seien

$$g_\delta(x) := \sum_{n=0}^{\infty} m_n^\delta 1_{I_n^\delta}(x) \quad \text{und} \quad g^\delta(x) := \sum_{n=0}^{\infty} M_n^\delta 1_{I_n^\delta}(x).$$

Mit dieser Notation ist

$$\underline{\sigma}(\delta) = \int_0^\infty g_\delta(x) dx, \quad \bar{\sigma}(\delta) = \int_0^\infty g^\delta(x) dx$$

und $g_\delta(x) \leq g(x) \leq g^\delta(x)$ für alle $x \in \mathbb{R}^+$. Wähle nun $\delta \in (0, 1)$ und $N \in \mathbb{N}$ so groß, dass

$$\sum_{n > N} |M_n^\delta| < \delta$$

ist.

Die Differenzen $R(t + a) - R(t)$ sind für festes $a > 0$ unabhängig von $t \in \mathbb{R}^+$ durch eine Konstante C_a beschränkt, wie schon im Beweis des Blackwell'schen Erneuerungstheorems gezeigt wurde. Damit folgt

$$\begin{aligned} g^\delta * R(t) &= \sum_{n=0}^{\infty} M_n^\delta (R(t - n\delta) - R(t - (n+1)\delta)) \\ &\leq \sum_{n \leq N} M_n^\delta (R(t - n\delta) - R(t - (n+1)\delta)) + C_a \delta. \end{aligned}$$

Unter Verwendung des Blackwell'schen Erneuerungstheorems ergibt sich damit

$$\begin{aligned} \limsup_{t \rightarrow \infty} g^\delta * R(t) &\leq \sum_{n \leq N} M_n^\delta \lim_{t \rightarrow \infty} (R(t - n\delta) - R(t - (n+1)\delta)) + C_a \delta \\ &= \frac{\delta}{\mu} \sum_{n \leq N} M_n^\delta + C_a \delta \leq \frac{1}{\mu} \int_0^\infty g^\delta(x) dx + \frac{\delta^2}{\mu} + C_a \delta = \frac{\bar{\sigma}(\delta)}{\mu} + \frac{\delta^2}{\mu} + C_a \delta. \end{aligned}$$

Daraus folgt

$$\limsup_{t \rightarrow \infty} g * U(t) \leq \lim_{\delta \rightarrow 0+} \limsup_{t \rightarrow \infty} g^\delta * U(t) \leq \lim_{\delta \rightarrow 0+} \left(\frac{\bar{\sigma}(\delta)}{\mu} + \frac{\delta^2}{\mu} + C_a \delta \right) = \frac{1}{\mu} \int_0^\infty g(x) dx.$$

Entsprechend erhält man auch

$$\liminf_{t \rightarrow \infty} g * U(t) \geq \frac{1}{\mu} \int_0^\infty g(x) dx,$$

was zusammen den Fundamentalsatz im nichtarithmetischen Fall liefert. ■

17.39 Lemma (Cesaro's Lemma):

Es sei $(\beta_n)_{n \in \mathbb{N}}$ eine reellwertige Zahlenfolge mit

$$\lim_{n \rightarrow \infty} \beta_n = 0.$$

Dann gilt:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{k=1}^n \beta_k = 0.$$

Beweis:

Aufgrund der Voraussetzung existiert eine Konstante K mit $|\beta_n| \leq K \forall n \in \mathbb{N}$. Außerdem existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $|\beta_n| < \varepsilon \forall n \geq n_0$. Damit wird

$$\begin{aligned} \left| \frac{\beta_1 + \dots + \beta_n}{n} \right| &\leq \left| \frac{\beta_1 + \dots + \beta_{n_0}}{n} \right| + \left| \frac{\beta_{n_0+1} + \dots + \beta_n}{n} \right| \\ &\leq \frac{n_0 \cdot K}{n} + \frac{n - n_0}{n} \cdot \varepsilon \leq 2\varepsilon \end{aligned}$$

für fast alle n . ■

17.40 Satz:

Es bezeichne $(X_n)_{n \in \mathbb{N}}$ eine (einfache) Erneuerungsfolge mit Lebensdauerverteilung F . Ist F nicht arithmetisch und $\mathbf{E}[X_1] < \infty$, dann gilt:

$$\lim_{t \rightarrow \infty} \frac{R(t)}{t} = \frac{1}{\mathbf{E}[X_1]}.$$

Beweis:

Es sei $\beta_n := R(n) - R(n-1)$, $n \in \mathbb{N}_0$. Dem Blackwell'schen Erneuerungstheorem entnimmt man

$$\beta_n \longrightarrow \frac{1}{\mathbf{E}[X_1]}.$$

Mit Hilfssatz 17.39 folgt dann

$$\frac{1}{\mathbf{E}[X_1]} = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{k=1}^n \beta_k = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{k=1}^n R(k) - R(k-1) = \lim_{n \rightarrow \infty} \frac{R(n)}{n}.$$

Es bezeichne $\lfloor x \rfloor$ die größte ganze Zahl $\leq x$. Da R monoton nicht fallen ist, gilt wegen $\lfloor t \rfloor \leq t \leq \lfloor t \rfloor + 1$

$$R(\lfloor t \rfloor) \leq R(t) \leq R(\lfloor t \rfloor + 1)$$

bzw.

$$\frac{\lfloor t \rfloor \cdot R(\lfloor t \rfloor)}{t \cdot \lfloor t \rfloor} \leq \frac{R(t)}{t} \leq \frac{R(\lfloor t \rfloor + 1)}{\lfloor t \rfloor + 1} \cdot \frac{\lfloor t \rfloor + 1}{t}.$$

Indem man den Grenzübergang $t \rightarrow \infty$ durchführt, erhält man:

$$\frac{1}{\mathbf{E}[X_1]} \leq \lim_{t \rightarrow \infty} \frac{R(t)}{t} \leq \frac{1}{\mathbf{E}[X_1]}.$$

■

17.41 Satz:

Ist N_t die Anzahl der in $[0, t]$ gelegenen Erneuerungen eines Erneuerungsprozesses S_n und ist $\mu = \mathbf{E}[X_j] = \mathbf{E}[S_j - S_{j-1}] < \infty$, so gilt

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} = \frac{1}{\mu} \quad P\text{-f.s.}$$

Beweis:

Die Aussage bezieht sich direkt auf N_t , **nicht** auf $\mathbf{E}[N_t]$!

Es ist nach Definition $S_{N_t} \leq t < S_{N_t+1}$ und damit

$$\frac{S_{N_t}}{N_t} \leq \frac{t}{N_t} < \frac{S_{N_t+1}}{N_t} = \frac{S_{N_t+1}}{N_t+1} \cdot \frac{N_t+1}{N_t}. \quad (17.6)$$

Ferner gilt für alle $n \in \mathbb{N}$

$$\lim_{t \rightarrow \infty} P(N_t \geq n) = \lim_{t \rightarrow \infty} P(S_n \leq t) = 1$$

also $N_t \rightarrow \infty$ für $t \rightarrow \infty$ fast sicher. Daher gilt

$$\frac{N_t+1}{N_t} \xrightarrow[t \rightarrow \infty]{} 1 \quad P\text{-f.s.}$$

in Gleichung (17.6). Mit $S_{N_t} = X_1 + \dots + X_{N_t}$ folgt

$$\frac{S_{N_t}}{N_t} = \frac{X_1 + \dots + X_{N_t}}{N_t}. \quad (17.7)$$

Alle Lebensdauern X_i haben nach Voraussetzung dieselbe Verteilung mit $\mu = \mathbf{E}[X_1] < \infty$. Wende daher auf die rechte Seite in Gleichung (17.7) das starke Gesetz der großen Zahlen an:

$$\lim_{t \rightarrow \infty} \frac{S_{N_t}}{N_t} = \lim_{t \rightarrow \infty} \frac{X_1 + \dots + X_{N_t}}{N_t} = \lim_{n \rightarrow \infty} \frac{X_1 + \dots + X_n}{n} = \mu \quad P\text{-f.s.}$$

Für $t \rightarrow \infty$ liefert nun (17.6) fast sicher

$$\mu \leq \lim_{t \rightarrow \infty} \frac{t}{N_t} \leq \mu \cdot 1$$

d.h.

$$\lim_{t \rightarrow \infty} \frac{t}{N_t} = \mu \quad P\text{-f.s.}$$

In einem Erneuerungsprozeß gilt nach Voraussetzung $P(X_1 = 0) = F(0) < 1$, d.h. es ist $\mu > 0$ und mit Kehrwertbildung folgert die Behauptung. ■

17.42 Satz (Grenzverteilung der Vorwärtsrekurrenzzeit):

Ist X_1 nicht arithmetisch verteilt und $\mu := \mathbf{E}[X_1] < \infty$, dann gilt für die Grenzverteilung der Vorwärtsrekurrenzzeit

$$\lim_{t \rightarrow \infty} P(V_t \leq v) = \begin{cases} \frac{1}{\mu} \int_0^v (1 - F(y)) dy & , v \geq 0 \\ 0 & , v < 0. \end{cases}$$

Beweis:

Die folgende Gleichung stammt aus dem Beweis zu Satz 17.21:

$$\begin{aligned} P(V_t \leq v) &= F(t+v) - F(t) + \int_0^t (F(t+v-y) - F(t-y)) dR(y) \\ &= F(t+v) - F(t) + \int_0^t (1-F(t-y)) dR(y) - \int_0^t (1-F(t-v-y)) dR(y). \end{aligned}$$

Da $F(t+v) - F(t) \rightarrow 0$ für $t \rightarrow \infty$ folgt aus dem Fundamentalsatz der Erneuerungstheorie:

$$\begin{aligned} \lim_{t \rightarrow \infty} P(V_t \leq v) &= \frac{1}{\mu} \int_0^\infty (1-F(y)) dy - \frac{1}{\mu} \int_0^\infty (1-F(v+y)) dy \\ &= \frac{1}{\mu} \int_0^v (1-F(y)) dy. \end{aligned}$$

■

17.43 Satz (Erwartungswert Grenzverteilung der Vorwärtsrekurrenzzzeit):

Ist X_1 nicht arithmetisch verteilt mit $\mu := \mathbf{E}[X_1] < \infty$ und $\sigma^2 := \mathbf{Var}[X_1] < \infty$, dann gilt für den Erwartungswert $\mathbf{E}[V_\infty]$ der Grenzverteilung der Vorwärtsrekurrenzzzeit

$$\mathbf{E}[V_\infty] = \lim_{t \rightarrow \infty} \mathbf{E}[V_t] = \frac{\sigma^2 + \mu^2}{2\mu}.$$

Beweis:

Satz 17.42 gibt die Verteilungsfunktion von V_∞ an. Deren zugehörige Dichte existiert und ist gegeben durch

$$\frac{d}{dv} P(V_\infty \leq v) = \frac{1}{\mu} (1 - F(y)).$$

Hieraus folgt für den zugehörigen Erwartungswert:

$$\mathbf{E}[V_\infty] = \int_0^\infty x \frac{1}{\mu} (1 - F(x)) dx.$$

Für eine nichtnegative Zufallsgröße Y gilt bekanntlich

$$\mathbf{E}[Y^r] = r \int_0^\infty x^{r-1} (1 - F(x)) dx.$$

Folglich ist

$$\mathbf{E}[V_\infty] = \frac{1}{\mu} \int_0^\infty x (1 - F(x)) dx = \frac{1}{2\mu} \mathbf{E}[X_1^2] = \frac{\sigma^2 + \mu^2}{2\mu}.$$

■

17.44 Satz (Grenzverteilung der Rückwärtsrekurrenzzzeit):

Es sei X_1 nicht arithmetisch-verteilt und sei $\mathbf{E}[X_1] < \infty$, dann gilt für die Grenzverteilung der Rückwärtsrekurrenzzzeit

$$\lim_{t \rightarrow \infty} P(U_t \leq u) = \begin{cases} \frac{1}{\mu} \int_0^u (1 - F(x)) dx & , u \geq 0 \\ 0 & , u < 0. \end{cases}$$

Beweis:

Für die Verteilungsfunktion der Rückwärtsrekurrenzzzeit gilt

$$P(U_t \leq u) = \int_{t-u}^t (1 - F(t-x)) dR(x) \quad 0 \leq u < t.$$

Es wird

$$g(z) := \begin{cases} 1 - F(z) & , 0 \leq z < u \\ 0 & , z \geq u \end{cases}$$

gesetzt und der Fundamentalsatz der Erneuerungstheorie angewandt:

$$\lim_{t \rightarrow \infty} P(U_t \leq u) = \lim_{t \rightarrow \infty} \int_0^t g(t-x) dR(x) = \lim_{t \rightarrow \infty} \frac{1}{\mu} \int_0^\infty g(x) dx = \frac{1}{\mu} \int_0^u (1 - F(x)) dx.$$

■

17.45 Satz:

Es seien $(X_n)_{n \in \mathbb{N}}$ ein (einfacher) Erneuerungsprozess mit $\mu := \mathbf{E}[X_1] < \infty$ und $\sigma^2 := \mathbf{Var}[X_1] < \infty$. Dann gilt für den zugehörigen Zählprozess $(N_t)_{t \in \mathbb{R}^+}$:

$$\lim_{t \rightarrow \infty} P\left(\frac{N_t - t \cdot \mu^{-1}}{\sqrt{t \cdot \sigma^2 \cdot \mu^{-3}}} \leq x\right) = \Phi(x), \quad x \in \mathbb{R},$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung bezeichnet, d.h. N_t ist asymptotisch normalverteilt mit dem Erwartungswert $\mathbf{E}[N_t] = t/\mu$ und der Varianz $\mathbf{Var}[N_t] = t \cdot \sigma^2 / \mu^3$.

Beweis:

Es bezeichne $(S_n)_{n \in \mathbb{N}_0}$ den mit $(X_n)_{n \in \mathbb{N}}$ korrespondierenden Erneuerungsprozess. Da die aufeinanderfolgenden Lebensdauern X_i eine i.i.d. Folge bilden, genügt S_n dem zentralen Grenzwertsatz:

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq x\right) = \Phi(x), \quad x \in \mathbb{R}.$$

Mit der Festlegung

$$k(t) := x \cdot \sqrt{t \cdot \sigma^2 \cdot \mu^3} + t\mu^{-1}$$

gilt

$$P\left(\frac{N_t - t\mu^{-1}}{\sqrt{t \cdot \sigma^2 \cdot \mu^3}} \leq x\right) = P(N_t \leq k(t)).$$

Aus der Gleichwertigkeit der Ereignisse $\{N_t \leq n\}$ und $\{S_{n+1} > t\}$ kann gefolgert werden

$$P(N_t \leq k(t)) = P(S_{[k(t)+1]} > t) = P\left(\frac{S_{[k(t)+1]} - \mu \cdot (k(t) + 1)}{\sigma \cdot \sqrt{k(t) + 1}} > \xi(t)\right),$$

wobei

$$\xi(t) := \frac{t - \mu \cdot (k(t) + 1)}{\sigma \cdot \sqrt{k(t) + 1}}$$

ist. Da aber

$$\begin{aligned} \lim_{t \rightarrow \infty} \xi(t) &= \lim_{t \rightarrow \infty} \frac{t - \mu \cdot [x \cdot \sqrt{t \cdot \sigma^2 \cdot \mu^{-3}} + t \cdot \mu^{-1} + 1]}{\sigma \cdot \sqrt{k(t) + 1}} \\ &= \lim_{t \rightarrow \infty} \frac{-x \cdot \sqrt{\sigma^2 \cdot t \mu^{-1}} - \mu}{\sigma \cdot \sqrt{k(t) + 1}} = \lim_{t \rightarrow \infty} \frac{-x(\sigma^2 \cdot t \cdot \mu^{-1})^{-1/2} \cdot \sigma \cdot \mu^{-1}}{(k(t) + 1)^{-1/2} \cdot k'(t)} \\ &= -x \cdot \frac{\sigma}{\mu} \cdot (\sigma^2)^{-1/2} \cdot \mu = -x \end{aligned}$$

gilt, folgt schließlich

$$\begin{aligned} \lim_{t \rightarrow \infty} P\left(\frac{N_t - t \cdot \mu^{-1}}{\sqrt{t \cdot \sigma^2 \cdot \mu^{-3}}} \leq x\right) &= \lim_{t \rightarrow \infty} P\left(\frac{S_{[k(t)+1]} - \mu \cdot (k(t) + 1)}{\sigma \cdot \sqrt{k(t) + 1}} > \xi(t)\right) \\ &= 1 - \Phi(-x) = \Phi(x), \quad x \in \mathbb{R}^+. \end{aligned}$$

■

17.5 Anwendungsbeispiel 2

Problem:

Für eine Fabrikation benötigt man 200 gleichartige Vorrichtungen. Aus vorangegangenen Produktionen weiß man, dass die Vorrichtungen eine mittlere Lebensdauer von 1000 Stunden haben. Eine statistische Analyse ergab außerdem, dass die Lebensdauern einer 2-Erlang-Verteilung genügen. Die Vorrichtungen sind in jedem Monat 500 Stunden lang in Betrieb. Man möchte nun gerne wissen, wie viele Vorrichtungen man im ersten, zweiten, dritten usw. Betriebsmonat ersetzen muss.

Lösung:

Modellannahmen: Die Lebensdauern der einzelner Vorrichtungen sind unabhängig voneinander verteilt nach $\text{Erlang}(\alpha, 2)$ mit Erwartungswert 1000 Stunden. Die Zeit wird in Einheiten von 1 Monat gemessen, was einer Betriebsdauer der Vorrichtungen von 500 Stunden entspricht. Dann ist 2 die erwartete Lebensdauer, $\frac{2}{\alpha} = 2$, also $\alpha = 1$. Fortan sei N_t die Anzahl der Erneuerungen einer festen Maschine im Zeitintervall $(0, t]$ mit einem $t \geq 0$. T_k mit $k \in \mathbb{N}$ bezeichnet den Zeitpunkt der k -ten Erneuerung dieser Maschine. Dann ist

$$T_k \stackrel{d}{=} \text{Erlang}(\alpha, 2)^{k*} = \text{Erlang}(\alpha, 2k) \quad \text{und} \quad \mathbf{E}[N_t] = \sum_{k=1}^{\infty} F_k(t), \quad t \in \mathbb{R}^+,$$

wobei F_k die Verteilungsfunktion von T_k ist. Bezeichnet man die Dichte von F_k mit f_k , so gilt

$$\mathbf{E}[N_t] = \sum_{k=1}^{\infty} \int_0^t f_k(s) ds = \int_0^t \sum_{k=1}^{\infty} f_k(s) ds, \quad t \in \mathbb{R}^+.$$

Hierbei ist

$$\sum_{k=1}^{\infty} f_k(s) = \sum_{k=1}^{\infty} \frac{\alpha^{2k}}{(2k-1)!} s^{2k-1} e^{-\alpha s} = \alpha e^{-\alpha s} \sum_{k=1}^{\infty} \frac{(\alpha s)^{2k-1}}{(2k-1)!}, \quad s \in \mathbb{R}^+.$$

Beachtet man, dass $e^x - e^{-x} = \sum_{k=0}^{\infty} \frac{x^k}{k!} (1 - (-1)^k) = 2 \cdot \sum_{k \text{ ungerade}} \frac{x^k}{k!}$ für k ungerade ist, so folgt

$$\sum_{k=1}^{\infty} \frac{x^{2k-1}}{(2k-1)!} = \frac{1}{2} (e^x - e^{-x}) = \sinh x,$$

also

$$\begin{aligned} \mathbf{E}[N_t] &= \int_0^t \alpha e^{-\alpha s} \cdot \frac{1}{2} (e^{\alpha s} - e^{-\alpha s}) ds = \frac{\alpha}{2} \int_0^t (1 - e^{-2\alpha s}) ds = \frac{\alpha}{2} \left(s + \frac{1}{2\alpha} e^{-2\alpha s} \right) \Big|_0^t \\ &= \frac{\alpha}{2} \left(t + \frac{1}{2\alpha} (e^{-2\alpha t} - 1) \right), \quad t \in \mathbb{R}^+. \end{aligned}$$

Hier war $\alpha = 1$, also gilt mit $t > 0$:

$$\mathbf{E}[N_t] = \frac{1}{2} \left(t + \frac{1}{2} (e^{-2t} - 1) \right), \quad t \in \mathbb{R}^+.$$

Für $t \rightarrow \infty$ erhält man mit dem Satz von Blackwell:

$$\lim_{t \rightarrow \infty} (\mathbf{E}[N_t] - \mathbf{E}[N_{t-1}]) = \frac{1}{2}.$$

Die erwartete Gesamtzahl von Erneuerungen in $(0, t]$ ist dann (wegen der vorausgesetzten Unabhängigkeit)

$$200 \cdot \mathbf{E}[N_t] = 100 \cdot \left(t + \frac{1}{2} (e^{-2t} - 1) \right), \quad t \in \mathbb{R}^+.$$

Die asymptotische mittlere Anzahl von Erneuerungen in $(t-1, t]$ ist

$$200 \cdot \mathbf{E}[N_t - N_{t-1}] = 100.$$

t (Monate)	# Erneuerungen in $(0, t]$	# Erneuerungen in $(t-1, t]$
1	56.77	56.77
2	150.92	94.15
3	250.12	99.20
4	350.02	99.90
5	450.00	99.98
6	550.00	100.00
∞		100.00

17.6 Modifizierte Erneuerungsprozesse

Aus verschiedenen Gründen erscheint es zweckmäßig, für die erste Erneuerungszeit eine andere Verteilung als für die übrigen Erneuerungszeiten zuzulassen. Denn nicht immer kann man die Beobachtungszeitpunkte so wählen, dass der Ursprung der Zeitrechnung mit der ersten Erneuerung zusammenfällt.

17.46 Definition (modifizierter Erneuerungsprozess):

Sind X_1, X_2, \dots unabhängige nichtnegative Zufallsgrößen jeweils mit den Verteilungsfunktionen $F_1(t) := P(X_1 \leq t)$ und $F(t) = P(X_i \leq t)$ für $i = 2, 3, \dots$ und $F(0) < 1$, so heißt die gemäß

$$S'_0 := 0 \quad \text{und} \quad S'_n := \sum_{k=1}^n X_k$$

gebildete Folge $(S'_n)_{n \in \mathbb{N}_0}$ ein modifizierter Erneuerungsprozess.

Für die modifizierten Erneuerungsprozesse lässt sich nun eine analoge Theorie aufbauen, wobei sich zeigt, dass die Verteilung der ersten Erneuerung gar keine so große Rolle spielt. Daher sollen im Folgenden nur Ergebnisse angegeben werden:

Für die Erneuerungsfunktion eines modifizierten Erneuerungsprozesses gilt

$$R'(t) = \mathbf{E}[N'_t] = \sum_{k=1}^{\infty} F_1 * F^{(k-1)*}(t), \quad t \in \mathbb{R}^+, \quad \text{wobei} \quad F^{0*}(t) := \begin{cases} 1 & , t \geq 0 \\ 0 & , t < 0. \end{cases}$$

$R'(t)$ genügt der Erneuerungsgleichung

$$R'(t) = F_1(t) + (F * R')(t), \quad t \in \mathbb{R}^+.$$

Die Verteilung der Vorwärtsrekurrenzzeit eines modifizierten Erneuerungsprozesses lautet

$$P(V'_t \leq v) = \begin{cases} \int_t^{t+v} (1 - F(t+v-y)) dR'(y) & , v \geq 0 \\ 0 & , v < 0. \end{cases}$$

Somit gilt für den stationären Zustand

$$\lim_{t \rightarrow \infty} P(V'_t \leq v) = \begin{cases} \frac{1}{\mathbf{E}[X_2]} \int_0^v (1 - F(x)) dx & , v \geq 0 \wedge \mathbf{E}[X_2] < \infty \\ 0 & , v < 0 \vee \mathbf{E}[X_2] = \infty. \end{cases}$$

Ein analoges Resultat gilt für die Rückwärtsrekurrenzzeit:

$$P(U'_t \leq u) = \begin{cases} 0 & , u \leq 0 \\ \int_{t-u}^t (1 - F(t-y)) dR'(y) & , 0 < u < t \\ 1 & , u \geq t. \end{cases}$$

Der stationäre Zustand der Rückwärtsrekurrenzzeit ist

$$\lim_{t \rightarrow \infty} P(U'_t \leq u) = \begin{cases} \frac{1}{\mathbf{E}[X_2]} \int_0^u (1 - F(x)) dx & , u \geq 0 \wedge \mathbf{E}[X_2] < \infty \\ 0 & , u < 0 \vee \mathbf{E}[X_2] = \infty. \end{cases}$$

Entsprechend gilt:

$$\lim_{t \rightarrow \infty} R'(t) - R'(t - h) = \frac{h}{\mathbf{E}[X_2]} \quad \text{und} \quad \lim_{t \rightarrow \infty} \frac{R'(t)}{t} = \frac{1}{\mathbf{E}[X_2]}.$$

Stationäre Erneuerungsprozesse

Ein spezieller modifizierter Erneuerungsprozess ist der stationäre Erneuerungsprozess. Wählt man $F_1(x) := P(V_\infty \leq x)$, so entspricht das der Betrachtungsweise, dass zur Zeit $t = 0$ der Erneuerungsprozess schon „unendlich“ lange gelaufen ist. Man vermutet, dass in dem so modifizierten Erneuerungsprozess die asymptotischen Aussagen von Anfang an gelten, dass also der Prozess in diesem Sinne „stationär“ ist.

17.47 Definition (stationär):

Ein modifizierter Erneuerungsprozess heißt stationär, wenn X_1 und V_∞ dieselbe Verteilung besitzen.

17.48 Satz:

Es seien $F(0) = 0$, F keine Einpunktverteilung und $\mathbf{E}[X_2] < \infty$. Dann gilt: $(S'_n)_{n \in \mathbb{N}_0}$ ist genau dann stationär, wenn

$$R'(t) = \frac{t}{\mathbf{E}[X_2]}$$

für alle $t \geq 0$ gilt.

Beweis:

Es gilt für die Laplace-Transformierte von R' :

$$\overline{R'}(\theta) = \frac{\overline{F_1}(\theta)}{1 - \overline{F}(\theta)}.$$

Ebenso gilt nach Voraussetzung

$$F_1(t) = \frac{1}{\mathbf{E}[X_2]} \int_0^t (1 - F(y)) dy.$$

Es folgt

$$\overline{F_1}(\theta) = \int_0^\infty e^{it\theta} dF_1(t) = \frac{1}{\mathbf{E}[X_2]} \int_0^\infty e^{it\theta} (1 - F(t)) dt$$

und mit partieller Integration

$$\begin{aligned}
 \overline{F}_1(\theta) &= \frac{1}{\mathbf{E}[X_2]} \left(- \int_0^\infty \frac{1}{i\theta} e^{it\theta} d(1 - F(t)) - \frac{1}{i\theta} \right) \\
 &= \frac{1}{\mathbf{E}[X_2]} \left(\int_0^\infty \frac{1}{i\theta} e^{it\theta} dF(t) - \frac{1}{i\theta} \right) \\
 &= \frac{1 - \overline{F}(\theta)}{i\theta \mathbf{E}[X_2]}.
 \end{aligned}$$

Somit folgt insgesamt

$$\overline{R}'(\theta) = -\frac{1}{i\theta \mathbf{E}[X_2]} \implies R'(t) = \frac{t}{\mathbf{E}[X_2]}.$$

■

17.49 Satz:

Für einen stationären Erneuerungsprozess gilt

$$P(V'_t \leq v) = \begin{cases} \frac{1}{\mathbf{E}[X_2]} \int_0^v (1 - F(x)) dx & , v \geq 0 \\ 0 & , v < 0. \end{cases}$$

Beweis:

$$\begin{aligned}
 P(V'_t \leq v) &= \int_t^{t+v} (1 - F(t + v - y)) dR'(y) \\
 &= \frac{1}{\mathbf{E}[X_2]} \int_t^{t+v} (1 - F(t + v - y)) dy \\
 &= \frac{1}{\mathbf{E}[X_2]} \int_0^v (1 - F(x)) dx.
 \end{aligned}$$

■

17.7 Die Überlagerung von Erneuerungsprozessen

Eine Überlagerung oder Superposition von Erneuerungsprozessen liegt vor, wenn $n \geq 2$ unabhängige Erneuerungsprozesse $(N_t^{(i)})_{t \in T}$, $i = 1, \dots, n$ gleichzeitig laufen. Der Prozess $(N_t^\Sigma)_{t \in T}$ mit

$$N_t^\Sigma := \sum_{i=1}^n N_t^{(i)} \tag{17.8}$$

heißt der zugehörige Überlagerungsprozess. Verteilungen und Parameter des Überlagerungsprozesses werden fortan mit einem großen Sigma (Σ) gekennzeichnet. Die Überlagerung von Erneuerungsprozessen ist deshalb von Interesse, weil man den Prozess der Anrufwiederholung als Überlagerungsprozess von einfachen Erneuerungsprozessen auffassen kann. Eine besondere Rolle spielt dabei der Poissonprozess.

17.50 Definition (Poissonprozess):

Ein Poissonprozess $(Y_t)_{t \in T}$ ist ein Prozess mit unabhängigen Zuwächsen für den $Y_0 = 0$ fast sicher gilt und dessen Zuwächse $(Y_t - Y_s)$ ($0 \leq s < t$) eine Poissonsche Verteilung mit dem Parameter $\Lambda_t - \Lambda_s$ besitzen,

$$P((Y_t - Y_s) = k) = (\Lambda_t - \Lambda_s) \cdot \frac{(\Lambda_t - \Lambda_s)^k}{k!} \quad (k \in \mathbb{N}_0; 0 \leq s < t), \quad (17.9)$$

wobei Λ_t eine nichtnegative, nichtabnehmende Funktion von t ist. $(Y_t)_{t \in T}$ heißt homogen, wenn $\Lambda_t = \lambda t$ ($\lambda > 0$) gilt, andernfalls inhomogen.

17.51 Satz:

Es sei $(X_n)_{n \in \mathbb{N}}$ eine einfache Erneuerungsfolge mit Verteilungsfunktion $F(x) := 1 - e^{-\lambda x}$ für $x \geq 0$ und $F(x) = 0$ für $x < 0$. Der mit $(X_n)_{n \in \mathbb{N}}$ assoziierte Erneuerungsprozess ist ein homogener Poissonprozess mit Parameter λ .

Einen Beweis für diesen Satz findet man in der Monographie von J.W. Cohen (1969).

Die Bedeutung des Poissonprozesses für die Erneuerungs- und Warteschlangentheorie wird durch den nachfolgenden Satz herausgestellt. Er besagt, dass bei der Überlagerung von unabhängigen, identischen einfachen Erneuerungsprozessen im Allgemeinen die Eigenschaft der Einzelprozesse (z.B. die Unabhängigkeit und die Verteilungsgleichheit der X_n) verlorengehen. Die mathematische Analyse, aber auch die Simulation der aus der Superposition beliebiger Erneuerungsprozesse hervorgehenden Punktprozesse erfordern einen großen Aufwand. Außerdem ist dieses Thema in der Literatur noch nicht erschöpfend behandelt worden, so dass der Gestaltung der Wiederholungsprozesse gewisse Grenzen gesetzt sind.

17.52 Satz:

Gegeben seien $n \geq 2$ unabhängige, identische einfache Erneuerungsprozesse mit der Verteilungsfunktion $F(x)$. Es wird vorausgesetzt, dass $F(x)$ keine Gitterverteilung ist. Dann gilt: Der zugehörige Überlagerungsprozess ist genau dann ein Erneuerungsprozess, wenn die Einzelprozesse (und damit auch der Überlagerungsprozess) homogene Poissonprozesse sind.

Beweis:

„ \Rightarrow “: Die Aussage ist in einem Satz enthalten, der bei J.L. Doob (1967) aufgeschrieben ist. Danach ist ein Erneuerungsprozess (allgemeiner noch ein Punktprozess) genau dann ein homogener Poissonprozess, wenn für jede endliche Menge von punktfremden Intervallen $(u_i, t_i]$, $i = 1, \dots, k$ gilt:

$$P\left(\bigcap_{i=1}^k (N_{t_i} - N_{u_i} = m_i)\right) = \prod_{i=1}^k \frac{(\lambda(t_i - u_i))^{m_i}}{m_i!} \cdot e^{-\lambda(t_i - u_i)} \quad (m_i \in \mathbb{N}_0; \lambda > 0). \quad (17.10)$$

Die zu (17.10) analoge Formel für den Summenprozess $(N_t^\Sigma)_{t \in T}$ kann wegen der Unabhängigkeit der Einzelprozesse leicht berechnet werden. Man stellt fest, dass man in

(17.10) lediglich den Parameter λ durch $n\lambda$ ersetzen muss. Also ist $(N_t^\Sigma)_{t \in T}$ ein Poissonprozess mit dem Parameter $\lambda^\Sigma = n\lambda$.

„ \Leftarrow “: Da die Einzelprozesse identisch sind, werden ihre Parameter und Verteilungen im weiteren Text nicht indiziert. Es ist

$$V_t^\Sigma = \min_{1 \leq k \leq n} V_t^{(k)},$$

woraus

$$P(V_t^\Sigma > v) = (P(V_t > v))^n, \quad v \geq 0,$$

folgt. Es ist dann aber auch

$$\frac{1}{\mathbf{E}[X_1^\Sigma]} \cdot \int_v^\infty (1 - F^\Sigma(y)) dy = \frac{1}{(\mathbf{E}[X_1])^n} \cdot \left(\int_v^\infty (1 - F(y)) dy \right)^n. \quad (17.11)$$

In Gleichung (17.11) darf $1 - F^\Sigma(y)$ durch $(1 - F(y))^n$ ersetzt werden, denn es gilt:

$$P(X_1^\Sigma > y) = P(N_y^{(1)} = 0, \dots, N_y^{(n)} = 0) = (P(X_1 > y))^n = (1 - F(y))^n$$

Differenziert man (17.11) nach v und potenziert beide Seiten mit $1/(n-1)$, so erhält man:

$$1 - F(v) = \omega \cdot \int_v^\infty (1 - F(y)) dy, \quad \omega := \sqrt[n-1]{n \cdot \mathbf{E}[X_1^\Sigma] / (\mathbf{E}[X_1])^n}. \quad (17.12)$$

Aus (17.12) gewinnt man das Anfangswertproblem

$$\frac{d}{dv} G(v) = -\omega \cdot G(v), \quad G(0) = 1, \quad (17.13)$$

wobei $G(v) := 1 - F(v)$ gesetzt wurde. Die Lösung von (17.13) aber ist $G(v) = e^{-\omega v}$. Die Behauptung folgt jetzt aus Satz 17.51. ■

17.8 Abbrechende Erneuerungsprozesse

In den zurückliegenden Abschnitten wurde stets vorausgesetzt, dass die Lebensdauerverteilung F eine eigentliche Verteilungsfunktion darstellt (d.h. $\lim_{t \rightarrow \infty} F(t) = 1$). Für weitere Zwecke erweist es sich als zweckmäßig, den Begriff des Erneuerungsprozesses etwas allgemeiner zu fassen und $\lim_{t \rightarrow \infty} F(t) < 1$ zuzulassen. In diesem Fall kann es vorkommen, dass der Erneuerungsprozess nach einer zufälligen Zeit abbricht, d.h. das

$$\sup_{t \in \mathbb{R}^+} N_t < \infty$$

gilt.

17.53 Definition (rekurrent, transient, periodisch):

Es sei $(X_n)_{n \in \mathbb{N}}$ eine (einfache) Erneuerungsfolge mit Lebensdauerverteilung F . $(X_n)_{n \in \mathbb{N}}$ heißt

rekurrent, wenn $\lim_{t \rightarrow \infty} F(t) = 1$ ist,

transient, wenn $\lim_{t \rightarrow \infty} F(t) < 1$ ist,

periodisch, wenn F arithmetisch ist.

aperiodisch, wenn F arithmetisch ist mit $d = 1$.

Es wird

$$N := \sup_{t \in \mathbb{R}} N_t$$

gesetzt, dann gilt

$$P(N = k) = [1 - F(\infty)] \cdot F(\infty)^k \quad (k = 0, 1, 2, \dots).$$

Die Größe

$$L := \sup\{S_n \mid S_n < \infty\}$$

wird als Lebensdauer eines transienten Erneuerungsprozesses bezeichnet.

17.54 Satz:

Es bezeichne $(S_n)_{n \in \mathbb{N}_0}$ eine (einfache) Erneuerungsfolge mit Lebensdauerverteilung F und Erneuerungsfunktion R . Dann gilt:

$$P(L \leq t) = 1 - F(\infty) + R(t) \cdot [1 - F(\infty)], \quad t \in \mathbb{R}^+.$$

Beweis:

$$\begin{aligned} P(L > t) &= P(L > t \mid t < X_1 < \infty) + P(L > t \mid X_1 \leq t) \\ &= \int_t^\infty P(L > t \mid X_1 = x) dF(x) + \int_0^t P(L > t \mid X_1 = x) dF(x) \\ &= \int_t^\infty 1 dF(x) + \int_0^t P(L > t - x) dF(x) \\ &= F(\infty) - F(t) + \int_0^t P(L > t - x) dF(x), \quad t \in \mathbb{R}^+. \end{aligned}$$

Die Lösung dieser Erneuerungsgleichung ist

$$\begin{aligned} P(L > T) &= F(\infty) - F(t) + \int_0^t [F(\infty) - F(t - x)] dR(x) \\ &= F(\infty) - F(t) + F(\infty) \cdot R(t) - \int_0^t F(t - x) dR(x), \quad t \in \mathbb{R}^+. \end{aligned}$$

Da die Erneuerungsfunktion der Gleichung

$$R(t) = F(t) + (F * R)(t), \quad t \in \mathbb{R}^+,$$

bzw.

$$(F * R)(t) = R(t) - F(t), \quad t \in \mathbb{R}^+,$$

genügt, erhält man

$$\begin{aligned} P(L > t) &= F(\infty) - F(t) + F(\infty) \cdot R(t) - R(t) + F(t) \\ &= F(\infty) + [F(\infty) - 1] \cdot R(t), \quad t \in \mathbb{R}^+, \end{aligned}$$

bzw.

$$P(L \leq t) = 1 - P(L > t) = 1 - F(\infty) + [1 - F(\infty)] \cdot R(t), \quad t \in \mathbb{R}^+.$$

■

Literatur zu Kapitel 17

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- G. ALSMEYER:
Erneuerungstheorie,
Teubner-Verlag, Stuttgart, 1991.
ISBN: 3519027305
- E. CINLAR:
Introduction to stochastic processes,
Prentice-Hall, 1975.
- J. W. COHEN:
The single server queue,
North Holland Publishing Company, Amsterdam, 1969.
- K.-W. GAEDE:
Zuverlässigkeit — Mathematische Modelle,
Hanser, München, 1977.
ISBN: 3446123709
- S. KARLIN/H.M. TAYLOR:
A first course in stochastic processes,
Academic Press, 1975.
- J. KOHLAS:
Stochastische Methoden des Operations Research,
Teubner, 1977.
ISBN: 3519023423

- M. KOLONKO:
Stochastik II, Skriptum zur Vorlesung,
TU Clausthal, 2006.
- S. I. RESNICK:
Adventures in Stochastic Processes,
Birkhäuser, Boston, 1992.
ISBN:0817635912

Kapitel 18

Regenerative stochastische Prozesse

Beobachtet man ein Warteschlangensystem (z.B. die Warteschlangen vor den Check-In-Schaltern auf einem Flughafen oder vor den Fahrkartenschaltern in den Service-Zentren der Bahn), so stellt man fest, dass die Systeme von Zeit zu Zeit immer wieder leerlaufen und die Warteschlangen sich danach nach demselben Prinzip wie in den vorangegangenen Zyklen entwickeln. Dieses Prinzip der steten Erneuerung wird in der Stochastik durch das Prinzip der regenerativen Prozesse nachgebildet und ist der Schlüssel zu vielen praktischen Problemstellungen.

Schlüsselwörter: Filtration, Stoppzeit, regenerativer stochastischer Prozess, Regenerationspunkte, Regenerationszyklen.
--

18.1 Einleitung

In diesem Kapitel beschäftigen wir uns mit einer speziellen Klasse stochastischer Prozesse. Unter einem stochastischen Prozess versteht man eine Familie $Z = (Z_t)_{t \in T}$ von Zufallsvariablen über einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ mit Werten in einem Messraum (E, \mathcal{E}) , wobei T eine nichtleere Indexmenge bezeichnet. In der Regel setzen wir $T = \mathbb{N}_0$ bzw. $T = \mathbb{R}^+$ und bezeichnen Z entsprechend entweder als zeitdiskreten oder zeitstetigen Prozess. Für jedes $t \in T$ ist Z_t als Funktion von ω $(\mathfrak{F}, \mathcal{E})$ -messbar. Für festes $\omega \in \Omega$ heißt die Funktion $t \rightarrow Z_t(\omega)$, $t \in T$, eine Realisierung oder Trajektorie von Z .

18.1 Definition (Filtration):

Es sei (Ω, \mathfrak{F}) ein Messraum. Eine Folge $(\mathfrak{F}_t)_{t \in T}$ von σ -Algebren mit $\mathfrak{F}_s \subset \mathfrak{F}_t \subset \mathfrak{F}$ für alle $s, t \in T$ mit $s < t$ heißt Filtration auf (Ω, \mathfrak{F}) .

18.2 Definition (Stoppzeit):

Es sei $(\Omega, \mathfrak{F}, P, (Z_t)_{t \in T})$ ein stochastischer Prozess und $(\mathfrak{F}_t)_{t \in T}$ mit $\mathfrak{F}_t := \sigma(Z_s, s \in T \cap [0, t])$, $t \in T$, seine kanonische Filtration. Eine Zufallsgröße $\tau: \Omega \rightarrow T \cup \{\infty\}$ heißt Stoppzeit bzgl. $(F_t)_{t \in T}$ bzw. $(Z_t)_{t \in T}$, falls $\{\tau \leq t\} \in \mathfrak{F}_t \forall t \in T$.

18.3 Definition (regenerativer Prozess):

Ein stochastischer Prozess $Z = (\Omega, \mathfrak{F}, P, (Z_t)_{t \in T})$ mit Werten in einem Messraum (E, \mathcal{E}) heißt regenerativer Prozess, falls es einen Erneuerungsprozess $(S_n)_{n \in \mathbb{N}_0}$ mit den Lebensdauern (Zuwächsen) X_1, X_2, \dots gibt, so dass folgende Bedingungen erfüllt sind:

- a) $((Z_{t+S_n})_{t \in T}, (X_k)_{k > n})$ und $((Z_t)_{t \in T \cap [0, S_n]}, S_0, S_1, \dots, S_n)$ sind stochastisch unabhängig für alle $n \in \mathbb{N}_0$.
- b) Die Prozesse $((Z_{t+S_n})_{t \in T}, (X_k)_{k > n})$, $n \in \mathbb{N}_0$ haben alle dieselbe Verteilung.

18.4 Bemerkung:

Die Definition 18.3 besagt, dass ein regenerativer stochastischer Prozess in aufeinanderfolgende Zyklen $(Z_t)_{t \in T \cap [S_{n-1}, S_n)}$, $n \in \mathbb{N}$, zerfällt, die nicht nur identisch verteilt sondern auch stochastisch unabhängig sind.

Insbesondere kann zu jedem Zeitpunkt $t \in T$ und für jedes $n \in \mathbb{N}$ allein anhand der Kenntnis über \mathfrak{F}_t entschieden werden, ob das Ereignis $\{S_n \leq t\}$ eingetreten ist oder nicht. Die Zeitpunkte $(S_n)_{n \in \mathbb{N}_0}$ stellen also Stoppzeiten bzgl. Z dar und werden Regenerationspunkte genannt. Das Intervall $[S_{n-1}, S_n)$ bezeichnet man als n -ten Regenerationszyklus. Aufgrund der besonderen Struktur eines regenerativen stochastischen Prozesses ist auch sofort klar, dass mit $(Z_t)_{t \in T}$ auch $(f(Z_t))_{t \in T}$ ein regenerativer Prozess ist, sofern $f: (E, \mathcal{E}) \rightarrow (E', \mathcal{E}')$ messbar ist.

18.5 Satz:

Es sei $Z = (\Omega, \mathfrak{F}, P, (Z_t)_{t \in \mathbb{R}^+})$ ein regenerativer Prozess mit Zustandsraum $(\mathbb{R}, \mathfrak{B})$ und eingebettetem Erneuerungsprozess $(S_n)_{n \in \mathbb{N}_0}$. F bezeichne die Verteilungsfunktion von S_1 und

$R = (R(t))_{t \in \mathbb{R}^+}$ die zu $(S_n)_{n \in \mathbb{N}_0}$ gehörende Erneuerungsfunktion. Ferner sei $\psi_A(t) := P(Z_t \in A, S_1 > t) \forall A \in \mathfrak{B}$ und $\forall t \in \mathbb{R}^+$. Dann gilt:

$$P(Z_t \in A) = \Psi_A(t) + \int_{[0,t]} \psi_A(t-s) dR(s) \quad \forall A \in \mathfrak{B} \text{ und } \forall t \in \mathbb{R}^+.$$

Beweis:

$$\begin{aligned} P(Z_t \in A) &= P(Z_t \in A, S_1 > t) + P(Z_t \in A, S_1 \leq t) \\ &= P(Z_t \in A, S_1 > t) + P(Z_{t-S_1} \in A) \\ &= \psi_A(t) + \int_0^t P(Z_{t-s} \in A) dF(s) \quad \forall A \in \mathfrak{B}, t \in \mathbb{R}^+. \end{aligned}$$

Die eindeutige Lösung dieser Erneuerungsgleichung liefert

$$P(Z_t \in A) = \psi_A(t) + \int_{[0,t]} \psi_A(t-s) dR(s), \quad A \in \mathfrak{B}, t \in \mathbb{R}^+.$$

■

18.2 Zeitstetige regenerative stochastische Prozesse

18.6 Satz:

Es sei Z ein zeitstetiger regenerativer Prozess mit Zustandsraum $(\mathbb{R}, \mathfrak{B})$ und eingebettetem Erneuerungsprozess $(S_n)_{n \in \mathbb{N}}$. Ist S_1 nicht arithmetisch verteilt mit Erwartungswert $\mathbf{E}[S_1] < \infty$ und ist Ψ_A uneigentlich Riemann-integrierbar, dann gilt:

$$Z_t \xrightarrow{i.V.} Z^*.$$

Beweis:

Satz 18.5 und der Fundamentalsatz der Erneuerungstheorie (Satz 17.38) besagen, dass

$$P(Z^* \in A) = \lim_{t \rightarrow \infty} P(Z_t \in A) = \frac{1}{\mathbf{E}[S_1]} \int_0^\infty \Psi_A(s) ds \quad \forall A \in \mathfrak{B}$$

gilt, sofern Ψ_A direkt-Riemann-integrierbar ist. Zunächst gilt

$$\Psi_A(t) = P(Z_t \in A, S_1 > t) \leq P(S_1 > t) = 1 - F(t), \quad A \in \mathfrak{B}, t \in \mathbb{R}^+,$$

wobei F wieder die Verteilungsfunktion von S_1 ist. Da $1 - F(t)$ nichtnegativ, nicht wachsend und wegen

$$\mathbf{E}[S_1] = \int_0^\infty (1 - F(t)) dt < \infty$$

auch uneigentlich Riemann-integrierbar ist, ist $1 - F(t)$ nach Satz 17.36 direkt-Riemann-integrierbar. Die direkte Riemann-Integrierbarkeit von Ψ_A folgt nun aus Satz 17.37. Es bleibt zu zeigen, dass $P(Z^* \in A)$ eine Wahrscheinlichkeitsverteilung bildet. Dazu wird gezeigt, dass $P(Z^* \in \mathbb{R}) = 1$ gilt:

$$\begin{aligned} P(Z^* \in \mathbb{R}) &= \frac{1}{\mathbf{E}[S_1]} \int_0^\infty \psi_{\mathbb{R}}(s) ds = \frac{1}{\mathbf{E}[S_1]} \int_0^\infty P(Z_s \in \mathbb{R}, S_1 > s) ds \\ &= \frac{1}{\mathbf{E}[S_1]} \int_0^\infty (1 - P(S_1 \leq s)) ds = \frac{1}{\mathbf{E}[S_1]} \cdot \mathbf{E}[S_1] = 1. \end{aligned}$$

■

Im Folgenden geht es um Aussagen über das Integral

$$\int_0^t Z_s(\omega) ds, \quad \omega \in \Omega,$$

längs eines Pfades $t \rightarrow Z_t(\omega)$ von $(Z_t)_{t \in \mathbb{R}^+}$ und den zugehörigen Erwartungswert

$$\mathbf{E} \left[\int_0^t Z_s ds \right] = \int_{\Omega} \left[\int_0^t Z_s(\omega) ds \right] dP(\omega).$$

Wir fordern deshalb die Messbarkeit der Abbildung

$$Z_t: \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}^+, \quad (t, \omega) \rightarrow Z_t(\omega)$$

bzgl. $\mathfrak{B}^+ \otimes \mathfrak{F}$ und \mathfrak{B}^+ . Erfüllt $(Z_t)_{t \in \mathbb{R}^+}$ diese Bedingung, nennt man $(Z_t)_{t \in \mathbb{R}^+}$ messbar.

18.7 Lemma:

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von stochastisch unabhängigen und identisch verteilten Zufallsgrößen über $(\Omega, \mathfrak{F}, P)$ mit $\mathbf{E}[|X_n|] < \infty$ für alle $n \in \mathbb{N}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0 \quad P\text{-f.s.}$$

Beweis:

Es sei $\tilde{F}(x)$ die Verteilungsfunktion von $|X_n|$. Dann gilt für alle $\varepsilon > 0$:

$$\begin{aligned} \mathbf{E}[|X_n|] < \infty &\implies \int_0^\infty (1 - \tilde{F}(x)) dx < \infty \implies \sum_{n=1}^\infty (1 - \tilde{F}(n\varepsilon)) < \infty \\ &\iff \sum_{n=1}^\infty P(|X_n| > n\varepsilon) < \infty \implies P(|X_n| > n\varepsilon) \xrightarrow{n \rightarrow \infty} 0 \\ &\iff P(|X_n| \leq n\varepsilon) \xrightarrow{n \rightarrow \infty} 1 \iff \frac{|X_n|}{n} \xrightarrow{n \rightarrow \infty} 0 \quad P\text{-f.s.} \end{aligned}$$

■

Mit Hilfe von Lemma 18.7 kann man nun zeigen:

18.8 Satz:

Es sei $Z = (\Omega, \mathfrak{F}, P, (Z_t)_{t \in \mathbb{R}^+})$ ein messbarer regenerativer Prozess mit dem zugehörigen Zustandsraum $(\mathbb{R}^+, \mathfrak{B}^+)$. Der eingebettete Erneuerungsprozess sei $(S_n)_{n \in \mathbb{N}_0}$. S_1 sei nicht arithmetisch verteilt mit $\mathbf{E}[S_1] < \infty$ und $\mathbf{E}\left[\int_0^{S_1} Z_s ds\right] < \infty$. Dann gilt:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t Z_s ds = \frac{\mathbf{E}\left[\int_0^{S_1} Z_s ds\right]}{\mathbf{E}[S_1]} \quad P\text{-f.s.}$$

Beweis:

Wir definieren

$$Y_k(\omega) := \int_{[S_{k-1}(\omega), S_k(\omega))} Z_s(\omega) ds \quad \forall \omega \in \Omega \text{ und } \forall k \in \mathbb{N}.$$

Der zu $(S_n)_{n \in \mathbb{N}_0}$ gehörende Zählprozess sei $(N_t)_{t \in \mathbb{R}^+}$. Dann gilt:

$$\begin{aligned} \frac{1}{t} \int_0^t Z_s ds &= \frac{1}{t} \int_0^{S_{N_t+1}} Z_s ds - \frac{1}{t} \int_t^{S_{N_t+1}} Z_s ds \\ &= \frac{1}{t} \sum_{k=1}^{N_t+1} Y_k - \frac{1}{t} \int_t^{S_{N_t+1}} Z_s ds \\ &= \frac{N_t+1}{t} \cdot \frac{1}{N_t+1} \sum_{k=1}^{N_t+1} Y_k - \frac{1}{t} \int_t^{S_{N_t+1}} Z_s ds, \quad t \in \mathbb{R}^+. \end{aligned}$$

Satz 17.41 besagt

$$\frac{N_t}{t} \rightarrow \frac{1}{\mathbf{E}[S_1]} \quad P\text{-f.s.}$$

und aufgrund des starken Gesetzes der großen Zahlen gilt

$$\frac{1}{N_t+1} \sum_{k=1}^{N_t+1} Y_k \rightarrow \mathbf{E}[Y_1] \quad P\text{-f.s.}$$

Es bleibt folglich zu zeigen, dass der Ausdruck $t^{-1} \int_t^{S_{N_t+1}} Z_s ds$ P -f.s. gegen 0 konvergiert. Es ist aber

$$\int_t^{S_{N_t+1}} Z_s ds \leq \int_{S_{N_t}}^{S_{N_t+1}} Z_s ds = Y_{N_t+1}$$

Nach Voraussetzung ist $\mathbf{E}[Y_{N_t+1}] = \mathbf{E}[Y_1] < \infty$. Lemma 18.7 besagt, dass $\frac{Y_{N_t+1}}{N_t+1} \rightarrow 0$ P.-f.s. gilt. Aufgrund der Beziehung

$$\frac{N_t + 1}{t} \cdot \frac{Y_{N_t+1}}{N_t + 1} = \frac{Y_{N_t+1}}{t}$$

und der Voraussetzung $\mathbf{E}[S_1] < \infty$ muss deshalb dasselbe auch für $\frac{Y_{N_t+1}}{t}$ gelten. ■

18.9 Satz:

Es sei $(Z_t)_{t \in \mathbb{R}^+}$ ein messbarer regenerativer Prozess mit Zustandsraum $(\mathbb{R}, \mathfrak{B})$ und eingebettetem Erneuerungsprozess $(S_n)_{n \in \mathbb{N}_0}$. S_1 sei nicht arithmetisch verteilt mit $\mathbf{E}[S_1] < \infty$.

Außerdem sei $\mathbf{E} \left[\int_0^{S_1} f(Z_s) ds \right] < \infty$ und $f: \mathbb{R} \rightarrow \mathbb{R}$ sei $(\mathfrak{B}, \mathfrak{B})$ -messbar. Dann gilt

$$\lim_{t \rightarrow \infty} \mathbf{E}[f(Z_t)] = \mathbf{E} \left[\int_0^{S_1} f(Z_s) ds \right] / \mathbf{E}[S_1].$$

Beweis:

Es sei zunächst $f \geq 0$ angenommen. Für ein beliebiges $c > 0$ sei $f^c := \min\{f, c\}$. Mit Hilfe des Erneuerungsargumentes folgt

$$\begin{aligned} \mathbf{E}[f^c(Z_t)] &= \mathbf{E}[f^c(Z_t), S_1 > t] + \mathbf{E}[f^c(Z_t), S_1 \leq t] \\ &= \mathbf{E}[f^c(Z_t), S_1 > t] + \int_0^t \mathbf{E}[f^c(Z_{t-s})] dF(s), \quad t \in \mathbb{R}^+, \end{aligned}$$

wobei F die Verteilungsfunktion von S_1 bezeichnet. Da f^c lokal beschränkt und S_1 mit Wahrscheinlichkeit Eins endlich ist, existiert nicht nur $\mathbf{E}[f^c(Z_t), S_1 > t]$ für alle $t \in \mathbb{R}^+$, sondern außerdem ist $\mathbf{E}[f^c(Z_t), S_1 > t]$ als Funktion von t auch direkt Riemann-integrierbar. Es folgt deshalb aus dem Fundamentalsatz der Erneuerungstheorie, dass

$$\lim_{t \rightarrow \infty} \mathbf{E}[f^c(Z_t)] = \frac{1}{\mathbf{E}[S_1]} \int_0^\infty \mathbf{E}[f^c(Z_s), S_1 > s] ds < \infty \quad (18.1)$$

gilt. Aufgrund des Satzes von Fubini kann man das Integral in (18.1) auch in der Form

$$\int_0^\infty \mathbf{E}[f^c(Z_t), S_1 > t] dt = \int_0^\infty \mathbf{E}[f^c(Z_t) \cdot I_{\{S_1 > t\}}] dt = \mathbf{E} \left[\int_0^{S_1} f^c(Z_s) ds \right] \quad (18.2)$$

schreiben. Die Gleichungen (18.1) und (18.2) ergeben zusammen

$$\lim_{t \rightarrow \infty} \mathbf{E}[f^c(Z_t)] = \mathbf{E} \left[\int_0^{S_1} f^c(Z_s) ds \right] / \mathbf{E}[S_1].$$

Durch den Grenzübergang $c \rightarrow \infty$ und dem Satz von der monotonen Konvergenz erhält man die Aussage des Satzes 18.9 für beliebige nichtnegative messbare Abbildungen f . Mit $f = f^+ - f^-$ folgert man die Behauptung für beliebige messbare Abbildungen. ■

18.10 Bemerkung (Zeitmittel — Raummittel):

Die Kombination von Satz 18.8 und Satz 18.9 ergibt

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t Z_s \, ds = \mathbf{E}[Z^*].$$

Dieses wichtige Ergebnis wird in der Literatur häufig unter dem Stichwort „Zeitmittel gleich Raummittel“ zitiert.

18.3 Zeitdiskrete regenerative stochastische Prozesse

Grundsätzlich lassen sich alle Ergebnisse über zeitdiskrete regenerative stochastische Prozesse aus den Sätzen über zeitstetige regenerative stochastische Prozesse herleiten, indem man für alle $t \in \mathbb{R}^+$ und $k \in \mathbb{N}_0$

$$Z_t = W_k \quad \text{für } t \in [k, k+1)$$

vereinbart. Allerdings ist bei zeitdiskreten regenerativen stochastischen Prozessen zu beachten, dass die Lebensdauerverteilung $F(x) := P(X_1 \leq x)$ arithmetisch ist. Im Hinblick auf Grenzwertbetrachtungen muss man insbesondere zwischen den Fällen $d = 1$ (aperiodischer Fall) und $d \neq 1$ (periodischer Fall) unterscheiden.

18.11 Satz:

Es sei $W := (\Omega, \mathfrak{F}, P, (W_k)_{k \in \mathbb{N}_0})$ ein zeitdiskreter regenerativer stochastischer Prozess mit Zustandsraum $(\mathbb{R}, \mathfrak{B})$, und eingebettetem Erneuerungsprozess $(S_k)_{k \in \mathbb{N}_0}$. Die zugehörige Erneuerungsfunktion sei $(R(k))_{k \in \mathbb{N}_0}$. Ferner sei $\psi_A(k) := P(W_k \in A, S_1 > k)$ für $k \in \mathbb{N}_0$ und $A \in \mathfrak{B}$. Dann gilt

$$P(W_k \in A) = \Psi_A(k) + \sum_{n=1}^k \Psi_A(k-n) \cdot [R(n) - R(n-1)], \quad k \in \mathbb{N}_0, \quad A \in \mathfrak{B}.$$

Beweis:

Es gilt

$$\begin{aligned} P(W_k \in A) &= P(W_k \in A, S_1 > k) + P(W_k \in A, S_1 \leq k) \\ &= P(W_k \in A, S_1 > k) + P(W_{k-S_1} \in A) \\ &= \Psi_A(k) + \sum_{n=1}^k P(W_{k-n} \in A) \cdot P(S_1 = n), \quad A \in \mathfrak{B}, \quad k \in \mathbb{N}_0. \end{aligned}$$

Die Lösung dieser diskreten Erneuerungsgleichung ist

$$P(W_k \in A) = \Psi_A(k) + \sum_{n=1}^k \Psi_A(k-n) \cdot [R(n) - R(n-1)] \quad \forall A \in \mathfrak{B}.$$

■

18.12 Satz:

Es sei $W := (\Omega, \mathfrak{F}, P, (W_k)_{k \in \mathbb{N}_0})$ ein zeitdiskreter regenerativer stochastischer Prozess mit Zustandsraum $(\mathbb{R}, \mathfrak{B})$ und eingebettetem Erneuerungsprozess $(S_k)_{k \in \mathbb{N}_0}$. Ist S_1 arithmetisch mit $d = 1$ und $\mathbf{E}[S_1] < \infty$, dann gilt:

$$W_k \xrightarrow{i.V.} W^*.$$

Beweis:

Aufgrund der Voraussetzung gilt

$$\begin{aligned} \sum_{n=0}^{\infty} \Psi_A(n) &= \sum_{n=0}^{\infty} P(W_n \in A, S_1 > n) \leq \sum_{n=0}^{\infty} P(S_1 > n) \\ &= \mathbf{E}[S_1] \quad \forall A \in \mathfrak{B}. \end{aligned}$$

Damit lässt sich der Fundamentalsatz der Erneuerungstheorie für den arithmetischen Fall mit $d := 1$ und $c := 0$ anwenden und man erhält

$$P(W^* \in A) = \lim_{k \rightarrow \infty} P(W_k \in A) = \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} \Psi_A(n), \quad A \in \mathfrak{B}.$$

Außerdem gilt

$$\begin{aligned} P(W^* \in \mathbb{R}) &= \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} \Psi_{\mathbb{R}}(n) = \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} P(W_n \in \mathbb{R}, S_1 > n) \\ &= \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} P(S_1 > n) = \frac{1}{\mathbf{E}[S_1]} \cdot \mathbf{E}[S_1] = 1. \end{aligned} \quad \blacksquare$$

18.13 Satz:

Es sei $W := (\Omega, \mathfrak{F}, P, (W_k)_{k \in \mathbb{N}_0})$ ein zeitdiskreter regenerativer stochastischer Prozess mit Zustandsraum $(\mathbb{R}^+, \mathfrak{B}^+)$ und eingebettetem Erneuerungsprozess $(S_k)_{k \in \mathbb{N}_0}$. Ist S_1 arithmetisch mit $d = 1$, $\mathbf{E}[S_1] < \infty$ und ist $\mathbf{E}\left[\sum_{k=0}^{S_1-1} W_k\right] < \infty$, dann gilt:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{k=0}^{n-1} W_k = \frac{\mathbf{E}\left[\sum_{k=0}^{S_1-1} W_k\right]}{\mathbf{E}[S_1]} \quad P\text{-f.s.}$$

Beweis:

$(N_n)_{n \in \mathbb{N}_0}$ sei der mit $(S_n)_{n \in \mathbb{N}_0}$ assoziierte Zählprozess. Mit derselben Argumentation wie im Beweis von Satz 18.8 ergibt sich dann:

$$\frac{1}{n} \cdot \sum_{k=0}^{n-1} W_k = \frac{1}{n} \cdot \left[\sum_{k=0}^{S_{N_{n-1}+1}-1} W_k - \sum_{k=n}^{S_{N_{n-1}+1}-1} W_k \right].$$

Das elementare Erneuerungstheorem und das starke Gesetz der großen Zahlen besagen:

$$\frac{1}{n} \cdot \sum_{k=0}^{S_{N_{n-1}+1}-1} W_k = \frac{N_{n-1} + 1}{n} \cdot \frac{1}{N_{n-1} + 1} \cdot \sum_{k=1}^{N_{n-1}+1} W_k \longrightarrow \frac{\mathbf{E}\left[\sum_{k=0}^{S_1-1} W_k\right]}{\mathbf{E}[S_1]} \quad P\text{-f.s.}$$

und $\frac{1}{n} \sum_{k=S_{N_{n-1}}}^{n-1} W_k$ strebt aufgrund der Abschätzung

$$\sum_{k=n}^{S_{N_{n-1}}} W_k \leq \sum_{k=S_{N_{n-1}}}^{S_{N_{n-1}+1}-1} W_k$$

und Lemma 18.7 P -f.s. gegen 0. ■

18.14 Satz:

Es sei $W := (\Omega, \mathfrak{F}, P, (W_k)_{k \in \mathbb{N}_0})$ regenerativer stochastischer Prozess mit dem Zustandsraum $(\mathbb{R}, \mathfrak{B})$ und eingebettetem Erneuerungsprozess $(S_k)_{k \in \mathbb{N}_0}$. Ist S_1 arithmetisch mit $d = 1$, $\mathbf{E}[S_1] < \infty$ und $\mathbf{E}\left[\sum_{n=0}^{S_1-1} f(W_n)\right] < \infty$. Dann gilt

$$\lim_{k \rightarrow \infty} \mathbf{E}[f(W_k)] = \frac{\mathbf{E}\left[\sum_{n=0}^{S_1-1} f(W_n)\right]}{\mathbf{E}[S_1]} \quad P\text{-f.s.}$$

Beweis:

Mit Hilfe des Erneuerungsarguments und des Fundamentalsatzes der Erneuerungstheorie deduziert man wie im zeitstetien Fall:

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathbf{E}[f^c(W_k)] &= \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} \mathbf{E}[f^c(W_n, S_1 > n)]. \\ &= \frac{1}{\mathbf{E}[S_1]} \cdot \sum_{n=0}^{\infty} \mathbf{E}[f^c(W_n) \cdot I_{\{S_1 > n\}}] \\ &= \frac{1}{\mathbf{E}[S_1]} \cdot \mathbf{E}\left[\sum_{n=0}^{S_1-1} f^c(W_n)\right]. \end{aligned}$$

Die weiteren Beweisschritte sind mit dem kontinuierlichen Fall identisch. ■

Literatur zu Kapitel 18

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- A.O. ALLEN:
Probability, Statistics and Queueing Theory,
Academic Press, London, 2nd. Edition, 1990.
ISBN: 0120510510
- E. CINLAR:
Introduction to stochastic processes,
Prentice-Hall, 1975.
- D. L. IGLEHART/G. S. SHEDLER:
An Introduction to the Regenerative Method for Simulation Analysis,
Springer-Verlag, Berlin, 1977.
ISBN: 3540084088

- D. L. IGLEHART/G. S. SHEDLER:
Regenerative Simulation of Response Times in Networks of Queues,
Springer-Verlag, Berlin, 1980.
ISBN: 3540099425
- J. KOHLAS:
Stochastische Methoden des Operations Research,
Teubner, 1977.
ISBN: 3519023423
- M. KOLONKO:
Stochastik II, Skriptum zur Vorlesung,
TU Clausthal, 2006.
- S. I. RESNICK:
Adventures in Stochastic Processes,
Birkhäuser, Boston, 1992.
ISBN:0817635912

Kapitel 19

Instandhaltungsstrategien

Die Instandhaltung ist ein Teilgebiet der in Stochastik I eingeführten Zuverlässigkeitstheorie. Sie beschäftigt sich mit der Untersuchung und Optimierung von Maßnahmen zur Erhaltung bzw. Wiederherstellung der Funktionstüchtigkeit von Systemen. Es werden unterschiedliche Strategien zur optimalen zeitlichen Planung von Instandhaltungsmaßnahmen betrachtet, deren Effektivität anhand von Kriterien wie Kostenrate oder Verfügbarkeit verglichen werden.

Schlüsselwörter: vollständige Erneuerung, minimale Reparatur, Kostenrate, Verfügbarkeit, altersunabhängige Erneuerung, altersabhängige Erneuerung, Blockerneuerung, prophylaktische Erneuerung, Havarieerneuerung.

19.1 Einleitung

Zu den Aufgabenstellungen der Zuverlässigkeitstheorie gehören u.a. die Modellierung des Ausfallverhaltens und der Abnutzung von Systemen, sowie die Untersuchung und Optimierung von Maßnahmen zur Erhaltung bzw. Wiederherstellung der Funktionstüchtigkeit von Systemen. Der erste Punkt wurde bereits in Stochastik I behandelt. Der Zweite führt auf die sog. Instandhaltungstheorie, welche folgende Maßnahmen umfasst:

- Pflege und Wartung
- Inspektion, z.B. Überprüfung und Diagnose
- Instandsetzung

Der Schwerpunkt liegt in diesem Kapitel auf den Maßnahmen zur Instandsetzung einfacher binärer Systeme, d.h. von Systemen mit den Zuständen *intakt* und *ausgefallen*. Man unterscheidet dabei zwischen prophylaktischer Instandsetzung (vorbeugende Maßnahmen) und Havarieinstandsetzung (wiederherstellende Maßnahmen). Gesucht ist die optimale zeitliche Planung solcher Vorgehensweisen, die durch unterschiedliche Strategien beschrieben und hinsichtlich der Größen Kostenrate und Verfügbarkeit auf ihre Effektivität untersucht werden.

Zunächst werden dafür einige Grundbegriffe aus der Zuverlässigkeitstheorie wiederholt:

19.2 Wiederholung der Zuverlässigkeit

19.1 Definition (Lebensdauer, Lebensdauerfunktion, Zuverlässigkeitsfunktion):

Als Lebensdauer einer Maschine oder Komponente wird die Zeit zwischen Inbetriebnahme und Ausfall bezeichnet. Sie wird durch eine nichtnegative Zufallsgröße T beschrieben.

Die Lebensdauerfunktion sei die Verteilungsfunktion der Lebensdauer T , d.h.

$$F(t) := P(T \leq t), \quad t \in \mathbb{R}^+.$$

Mit Hilfe der Lebensdauerfunktion lässt sich die Überlebens- bzw. Zuverlässigkeitsfunktion $\bar{F}(t)$ der Komponente als

$$\bar{F}(t) := 1 - F(t) = P(T > t), \quad t \in \mathbb{R}^+$$

definieren.

19.2 Definition (Ausfallrate):

Es sei F eine stetige Verteilungsfunktion. Dann heißt

$$a(t) := \lim_{h \rightarrow 0+} \frac{1}{h} P(t < T \leq t + h \mid T > t)$$

die Ausfallrate eines Bauteils mit der Lebensdauer T .

$\Delta h \cdot a(t)$ ist die Wahrscheinlichkeit dafür, dass eine Komponente nach Erreichen des Lebensalters t innerhalb der Zeitspanne Δh ausfällt.

19.3 Satz:

Es sei $a(t)$ die Ausfallrate eines Bauteils mit stetiger Lebensdauerverteilung. Dann gilt für die Zuverlässigkeitsfunktion:

$$\bar{F}(t) = \exp\left(-\int_0^t a(u) \, du\right).$$

Neben der Ausfallrate ist die bedingte Restlebensdauer, d.h. die Lebensdauerverteilung unter der Bedingung, dass das Bauteil bereits bis zum Zeitpunkt t_0 überlebt hat, von Interesse.

19.4 Definition (Bedingte Restlebensdauer, bedingte Überlebenswahrscheinlichkeit):

Es sei

$$\bar{F}_{t_0}(t) := P(T > t + t_0 | T > t_0) = \frac{\bar{F}(t + t_0)}{\bar{F}(t_0)} = \frac{1 - F(t + t_0)}{1 - F(t_0)}$$

die bedingte Überlebenswahrscheinlichkeit ab dem Zeitpunkt $t_0 \in \mathbb{R}^+$, d.h. die Verteilung der Restlebensdauer unter der Annahme, dass das Bauteil bereits bis zum Zeitpunkt $t_0 \in \mathbb{R}^+$ überlebt hat. Demgegenüber stellt

$$F_{t_0}(t) := 1 - \bar{F}_{t_0}(t) = P(T < t + t_0 | T > t_0) = \frac{F(t + t_0) - F(t_0)}{\bar{F}(t_0)}$$

die bedingte Restlebensdauer dar.

Die betrachteten Lebensdauerfunktionen $F(t)$ lassen sich in die parametrische und die nichtparametrische Klasse unterteilen. In der ersten Klasse werden die Lebensdauern durch bekannte Verteilungen modellieren, wie z.B. Exponentialverteilung, Weibullverteilung oder logarithmische Normalverteilung. Die Lebensdauerverteilung ist dann durch die entsprechenden Parameter der gewählten Verteilungsfunktion charakterisiert. In der nichtparametrischen Klasse wird die Verteilungsfunktion der Lebensdauern durch die Eigenschaften der Ausfallrate bzw. der Überlebenswahrscheinlichkeit beschrieben.

19.5 Definition (IFR, DFR):

Eine Verteilungsfunktion $F(t)$ heißt *Increasing Failure Rate (IFR) Verteilung* bzw.

Decreasing Failure Rate (DFR) Verteilung, wenn die bedingte Überlebenswahrscheinlichkeit $\bar{F}_{t_0}(t)$ bei beliebigem, aber festem $t > 0$ monoton in t_0 fällt bzw. in t_0 wächst.

19.3 Einführung

Bei den weiteren Untersuchungen zu den verschiedenen Instandsetzungsmaßnahmen und Strategien wird davon ausgegangen, dass die folgenden allgemeinen Voraussetzungen erfüllt sind:

1. Der Übergang vom Arbeitszustand in den Ausfallzustand des Systems erfolgt sprunghaft.
2. Es wird immer ein hinreichend langer Zeitraum untersucht, so dass Aussagen über Grenzwerte getroffen werden können.
3. Falls nicht anders angegeben, wird davon ausgegangen, dass die Instandsetzungsmaßnahme für das System in vernachlässigbarer Zeit erfolgt.

4. Nach einer Instandsetzungsmaßnahme geht das System sofort wieder in den Betriebszustand über.
5. Bei den Lebensdauern wird eine IFR(Increasing Failure Rate)–Verteilung vorausgesetzt.

Die Maßnahmen zur Instandsetzung lassen sich wie folgt unterteilen:

19.6 Definition (totale Erneuerung, Havarieerneuerung, prophylaktische Erneuerung):

- Eine vollständige oder auch totale Erneuerung versetzt das System bezüglich seines Ausfallverhaltens in den Neuzustand zurück. Wird bei Ausfall des Systems eine totale Erneuerung durchgeführt, spricht man auch von Havarieerneuerung. Eine totale Erneuerung nach einer bestimmten Betriebsdauer τ , um einem Ausfall vorzubeugen, heißt prophylaktisch.
- Eine minimale Reparatur bei einem Ausfall zum Zeitpunkt t versetzt das System bezüglich seines Ausfallverhaltens in den Zustand unmittelbar vor dem Ausfall. Die Lebensdauerverteilung nach einer minimalen Reparatur lautet

$$F_t(x) = \frac{F(t+x) - F(t)}{\bar{F}(t)}.$$

F_t ist die Verteilungsfunktion der restlichen Lebensdauer eines Systems, dass bereits t Zeiteinheiten gearbeitet hat.

Die in diesem Kapitel betrachteten Instandhaltungsstrategien basieren darauf, dass die betreffenden Systeme nach zufälligen oder vorher festgelegten Zeiträumen vollständig erneuert werden. Die Betriebszeit kann demnach in Zyklen unterteilt werden. Im Folgenden bezeichne $(T_i)_{i \in \mathbb{N}_0}$ diejenigen Zeitpunkte, zu denen das System vollständig erneuert wird. Die Länge des i -ten Zyklus sei $Y_i := T_i - T_{i-1}$, $i \in \mathbb{N}$. Diese Zufallsgrößen sind unabhängig und identisch verteilt.

Weiter sei $(Z_t)_{t \in \mathbb{R}^+}$ der Zustandsprozess des Systems. In diesem Kapitel werden nur binäre Systeme betrachten, daher kann Z_t lediglich zwei Werte annehmen. Es gilt $Z_t = 1$, wenn das System zum Zeitpunkt t in Betrieb, kurz up, ist und $Z_t = 0$, wenn das System zum Zeitpunkt t außer Betrieb, kurz down, ist.

Es lässt sich folgendes feststellen:

- $(Z_t)_{t \in \mathbb{R}^+}$ ist ein regenerativer Prozess mit den Regenerationspunkten $(T_i)_{i \in \mathbb{N}_0}$.
- Die Punkte $(T_i)_{i \in \mathbb{N}_0}$ definieren einen gewöhnlichen Erneuerungsprozess.

Damit werden nun Funktionen für die Kostenrate, d.h. die Instandsetzungskosten pro Zeiteinheit, und die Verfügbarkeit, also die Wahrscheinlichkeit für ausfallfreies Arbeiten zu einem festen Zeitpunkt, bestimmt.

Es sei f_k die Kostenfunktion des Systems in Abhängigkeit des Betriebszustandes, d.h. up oder down. Somit sind die kumulativen Kosten C_i des i -ten Zyklus gegeben durch

$$C_i = \int_{T_{i-1}}^{T_i} f_k(Z_t) dt.$$

Die zeitabhängige Kostenrate des Gesamtsystems lautet daher:

$$K(t)(\omega) = \frac{1}{t} \int_0^t f_k(Z_t(\omega)) dt.$$

Zur Erinnerung:

19.7 Satz (vgl. Satz 18.6):

Es sei Z ein regenerativer Prozess mit eingebettetem Erneuerungsprozess $(T_i)_{i \in \mathbb{N}_0}$, dessen Verteilungsfunktion nicht arithmetisch ist. Gilt $\mathbf{E}[T_1] < \infty$ und $P(Z_t \in A, T_1 > t)$ ist Riemann-integrierbar, so folgt $Z_t \xrightarrow{i.V.} Z^$.*

Die Voraussetzungen des Satzes sind hier erfüllt: Es liegt eine Riemann-integrierbare IFR-Lebensdauerverteilung vor, somit ist $\mathbf{E}[T_1] < \infty$. Man kann hier also von einer Konvergenz in Verteilung gegen einen stationären Prozess Z^* ausgehen. Damit gilt für die stationäre Kostenrate K :

$$\frac{1}{t} \int_0^t f_k(Z_t(\hat{\omega})) dt \xrightarrow{i.V.} \int_{\Omega} f_k(Z^*(\omega)) dP(\omega) =: K$$

für fast alle $\hat{\omega} \in \Omega$. Für $t = T_n$ folgt

$$\frac{1}{t} \int_0^t f_k(Z_t) dt = \frac{\sum_{i=1}^n C_i}{\sum_{i=1}^n Y_i},$$

und es gilt fast sicher

$$\lim_{n \rightarrow \infty} \frac{\frac{1}{n} \sum_{i=1}^n C_i}{\frac{1}{n} \sum_{i=1}^n Y_i} = K.$$

Falls also $\mathbf{E}[C_1] < \infty$ und $\mathbf{E}[Y_1] < \infty$ gilt, so gilt fast sicher

$$\mathbf{E}[C_1] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n C_i, \quad \mathbf{E}[Y_1] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i.$$

Damit kann die stationäre Kostenrate K wie folgt berechnet werden:

$$K = \frac{\mathbf{E}[C_1]}{\mathbf{E}[Y_1]}. \quad (19.1)$$

Analog zur Kostenrate lässt sich eine Formel für die stationäre Verfügbarkeit herleiten:

Die Betriebszeit, d.h. up-time, im i -ten Regenerationszyklus B_i ist gegeben durch

$$B_i = \int_{T_{i-1}}^{T_i} Z_t dt.$$

Die zeitabhängige Verfügbarkeit sei

$$V(t)(\omega) = \frac{1}{t} \int_0^t Z_t(\omega) dt.$$

Folgt man nun dem oben aufgezeichneten Rechenweg, so gelangt man zum Ergebnis für die stationäre Verfügbarkeit:

$$V = \frac{\mathbf{E}[B_1]}{\mathbf{E}[Y_1]}. \quad (19.2)$$

Mit den bisher erzielten Ergebnissen sollen nun spezielle Strategien zur Instandhaltung untersucht werden. Man unterteilt die hier vorgestellten Strategien in drei Bereiche:

1. Altersunabhängige Erneuerung, d.h. Instandsetzungsmaßnahmen werden nur bei Ausfall ergriffen.
2. Altersabhängige Erneuerung, auch streng periodische Erneuerung genannt. Bei diesen Strategien werden Instandsetzungsmaßnahmen sowohl bei Ausfall, als auch vorbeugend nach einer gewissen Betriebsdauer τ durchgeführt.
3. Blockerneuerung, d.h. bei Ausfall und zu festen Zeiten werden Maßnahmen zur Instandsetzung verlangt.

Die nachfolgende Darstellung ist stark an die Lehrbücher von F. Beichelt „Zuverlässigkeits- und Instandhaltungstheorie“, Teubner, Stuttgart, 1993 und F. Beichelt/P. Franken „Zuverlässigkeit und Instandhaltung: Mathematische Methoden“, Hanser-Verlag, München, 1984 angelehnt.

19.4 Altersunabhängige Erneuerungen

Strategie 0 — Bei Ausfall Totalerneuerung

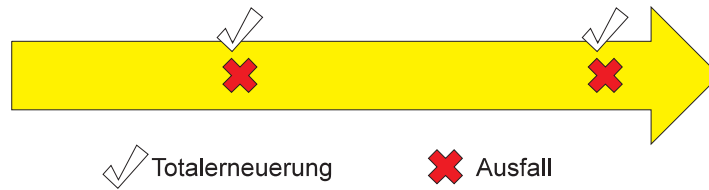


Abbildung 19.1:

Die erste Strategie, die hier untersucht werden soll, ist die totale Erneuerung beim Ausfall des Systems, sie wird auch als Ausfallstrategie bezeichnet. In der Praxis kann dies wie folgt aussehen: Steht ein Reservesystem zur Verfügung, so kann beim Ausfall des Systems ohne nennenswerte Zeiteinbußen auf dieses umgeschaltet werden.

Das hier genutzte mathematische Modell ist der Erneuerungsprozess. k_v stehe für die mittleren Kosten für eine totale Erneuerung. $\mathbf{E}[X]$ bezeichne die mittlere Lebensdauer des Systems. Nach Gleichung (19.1) gilt somit für die stationäre Kostenrate:

$$K = \frac{k_v}{\mathbf{E}[X]}$$

Für die stationäre Verfügbarkeit gilt bei Vorhandensein eines Reservesystems $V = 1$, da das Umschalten auf das Reservesystem in vernachlässigbarer Zeit erfolgt und somit das System ständig in Betrieb ist.

Wird davon ausgegangen, dass kein Reservesystem zur Verfügung steht, und die totale Erneuerung somit eine mittlere Zeit d_v beansprucht, so ist dies nur eine leichte Abwandlung. Es gilt dann:

$$K = \frac{k_v}{\mathbf{E}[X] + d_v}, \quad V = \frac{\mathbf{E}[X]}{\mathbf{E}[X] + d_v}.$$

Strategie 0' — Bei Ausfall minimale Reparatur

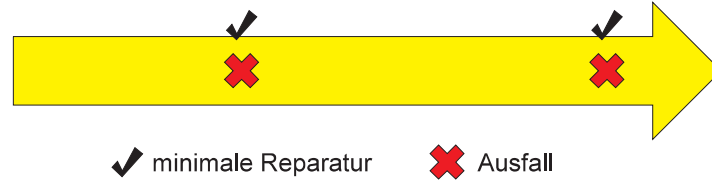


Abbildung 19.2:

Bei dieser Strategie führt man bei jedem Ausfall eine minimale Reparatur durch. Totale Erneuerungen treten hier überhaupt nicht auf.

Es gibt hier im Allgemeinen keine Erneuerungspunkte und keine Zyklen. Für die Kostenrate gilt allgemein

$$K = \lim_{t \rightarrow \infty} \frac{1}{t} (\text{„mittlere Kosten im Intervall } [0, t]\text{“}).$$

Zur Zeit t gibt für kleines Δt folgender Wert die Wahrscheinlichkeit für einen Ausfall an:

$$a(t) \cdot \Delta t + o(\Delta t).$$

Es entstehen Kosten k_m im Intervall $[t, t + \Delta t]$ mit Wahrscheinlichkeit $a(t)\Delta t + o(\Delta t)$. Damit ist $a(t)$ die Dichtefunktion der Kosten und die mittleren Kosten im Intervall $[0, t]$ ergeben sich aus

$$\int_0^t k_m a(s) ds = k_m \int_0^t a(s) ds := k_m A(t).$$

Dabei heißt $A(t)$ kumulative Ausfallrate. Es folgt:

$$K = \lim_{t \rightarrow \infty} \frac{1}{t} k_m A(t) = k_m \lim_{t \rightarrow \infty} \frac{A(t)}{t}.$$

Für eine unbeschränkte Ausfallrate $a(t)$ ist also auch K unbeschränkt. Bei im Alter quasi “alterungsfreien” Systemen, d.h. für beschränkte $a := \lim_{t \rightarrow \infty} a(t)$, gilt $K = k_m a$.

19.5 Altersabhängige Erneuerung

Strategie 1 — Bei Ausfall und nach Betriebsdauer \tilde{t} erfolgt Totalerneuerung

Bei dieser Strategie werden nicht nur sogenannte Havarieerneuerungen durchgeführt, d.h. totale Erneuerungen bei Ausfall des Systems, sondern auch prophylaktische Erneuerungen nach einer bestimmten Betriebsdauer \tilde{t} , um so einem Ausfall vorzubeugen. Der Vorteil in der Praxis liegt dabei darin, dass sich eine prophylaktische Erneuerung wesentlich besser planen lässt, wenn man zum Beispiel Service-Teams anderer Firmen zur Instandsetzung benötigt. Bei der prophylaktischen Erneuerung sind im Allgemeinen sowohl mittlere Kosten als auch mittlere Erneuerungszeit niedriger als bei der Havarieerneuerung.

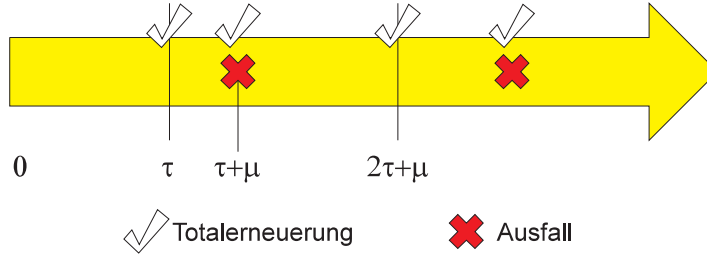


Abbildung 19.3:

Es sei X_h die zufällige Zeit zwischen zwei benachbarten Havarieerneuerungen und X_p die zufällige Zeit zwischen zwei prophylaktischen Erneuerungen. Y bleibt die zufällige Zeit zwischen zwei Erneuerungen beliebigen Typs.

X_h , X_p und Y definieren jeweils einfache Erneuerungsprozesse.

Für einen Zeitpunkt t existiert ein $n \in \mathbb{N}_0$ mit $n\tilde{t} < t \leq (n+1)\tilde{t}$. Damit ergibt sich die Wahrscheinlichkeit, dass der Abstand zwischen zwei Havarieerneuerungen größer als t ist aus der Wahrscheinlichkeit, dass das System n mal prophylaktisch erneuert wurde und dann die Zeit $t - n\tilde{t}$ überlebt hat:

$$P(X_h > t) = P(\{X_1 > \tilde{t}\} \cap \dots \cap \{X_n > \tilde{t}\} \cap \{X_{n+1} > t - n\tilde{t}\}) = (\bar{F}(\tilde{t}))^n \bar{F}(t - n\tilde{t}).$$

Somit folgt

$$P(X_h \leq t) = 1 - (\bar{F}(\tilde{t}))^n \bar{F}(t - n\tilde{t}).$$

Für die zufällige Zeit bis zur nächsten Totalerneuerung $Y = \min(X, \tilde{t})$ gilt:

$$P(Y \leq t) = \begin{cases} F(t) & , 0 \leq t < \tilde{t} \\ 1 & , t \geq \tilde{t} \end{cases}.$$

Damit ergeben sich:

$$\begin{aligned} \mathbf{E}[X_h] &= \int_0^\infty P(X_h > t) dt \\ &= \sum_{n=0}^\infty (\bar{F}(\tilde{t}))^n \int_{n\tilde{t}}^{(n+1)\tilde{t}} \bar{F}(t - n\tilde{t}) dt \\ &= \sum_{n=0}^\infty (\bar{F}(\tilde{t}))^n \int_0^{\tilde{t}} \bar{F}(t) dt \\ &= \frac{1}{F(\tilde{t})} \int_0^{\tilde{t}} \bar{F}(t) dt \end{aligned} \tag{19.3}$$

und

$$\mathbf{E}[Y] = \int_0^{\tilde{t}} \bar{F}(t) dt \tag{19.4}$$

Es bezeichne $N(\tilde{t}) = 1/\mathbf{E}[Y]$ die mittlere Anzahl von Erneuerungen je Zeiteinheit, d.h. die Erneuerungsrate. Analog sei $N_h(\tilde{t}) = 1/\mathbf{E}[X_h]$ die Rate der Havarieerneuerungen und $N_p(\tilde{t}) = 1/\mathbf{E}[X_p]$ die Rate der prophylaktischen Erneuerungen. Es gilt:

$$N(\tilde{t}) = N_h(\tilde{t}) + N_p(\tilde{t}). \quad (19.5)$$

Setzt man (19.3) und (19.4) in (19.5) ein, so folgt

$$\mathbf{E}[X_p] = \frac{1}{\overline{F}(\tilde{t})} \int_0^{\tilde{t}} \overline{F}(t) dt. \quad (19.6)$$

Für das Monotonieverhalten der Raten $N_h(\tilde{t})$ und $N_p(\tilde{t})$ folgt aufgrund der angenommenen IFR-Verteilung für die Lebensdauer:

19.8 Satz:

Für stetige Lebensdauerverteilungen $F(t)$ vom Typ IFR gilt mit wachsendem $\tilde{t} > 0$:

1. $N_h(\tilde{t})$ ist monoton steigend,
2. $N_p(\tilde{t})$ ist monoton fallend.

Beweis:

Vgl. Beichelt/Franken, S. 103f. ■

Optimierung der Kostenrate

Es seien k_h die mittleren Kosten für eine Havarieerneuerung und k_p die mittleren Kosten für eine prophylaktische Erneuerung. Dann gilt für die Kostenrate:

$$\begin{aligned} K(\tilde{t}) &= k_h \cdot N_h(\tilde{t}) + k_p \cdot N_p(\tilde{t}) \\ &= \frac{k_h \cdot F(\tilde{t}) + k_p \cdot \overline{F}(\tilde{t})}{\int_0^{\tilde{t}} \overline{F}(t) dt}. \end{aligned} \quad (19.7)$$

19.9 Bemerkung:

Zur Gleichung (19.7) gelangt man auch, wenn man von $K = \mathbf{E}[K_1] / \mathbf{E}[Y_1]$ ausgeht. Dabei ist $\mathbf{E}[K_1] = k_h F(\tilde{t}) + k_p \overline{F}(\tilde{t})$.

Aus Gleichung (19.7) und obigem Satz sieht man, dass bei zunehmendem \tilde{t} die durch Ausfälle bzw. durch Havarieerneuerungen verursachten Kosten zunehmen, die Kosten für prophylaktische Erneuerungen jedoch zurückgehen. Gesucht ist nun ein optimales Erneuerungsintervall \tilde{t}^* um diese gegenläufigen Tendenzen auszugleichen.

Für das gesuchte optimale \tilde{t}^* gilt $K(\tilde{t}^*) \leq K(\tilde{t})$ für alle $\tilde{t} \in \mathbb{R}^+$. Das Minimum von K wird nun durch Differentiation von (19.7) bestimmt:

$$\frac{dK}{d\tilde{t}} = \frac{k_h f(\tilde{t}) - k_p f(\tilde{t})}{\int_0^{\tilde{t}} \overline{F}(t) dt} - \frac{k_h F(\tilde{t}) + k_p \overline{F}(\tilde{t})}{\left(\int_0^{\tilde{t}} \overline{F}(t) dt\right)^2} \overline{F}(\tilde{t}).$$

Setzt man $K' = 0$ und beachtet, dass der Nenner immer größer ist als Null, so folgt:

$$\begin{aligned} (k_h - k_p)f(\tilde{t}) \int_0^{\tilde{t}} \bar{F}(t) dt &= \bar{F}(\tilde{t})(F(\tilde{t})(k_h - k_p) + k_p) \\ \iff a(\tilde{t}) \int_0^{\tilde{t}} \bar{F}(t) dt &= F(\tilde{t}) + \frac{k_p}{k_h - k_p}. \end{aligned}$$

Man setzt $k := k_p/k_h < 1$, da die Kosten für eine prophylaktische Erneuerung niedriger sind als für eine Havarieerneuerung. Es ergibt sich die folgende zu lösende Gleichung:

$$a(\tilde{t}) \int_0^{\tilde{t}} \bar{F}(t) dt - F(\tilde{t}) = \frac{k}{1 - k}. \quad (19.8)$$

Diese Gleichung ist nicht einfach aufzulösen. Mit Hilfe des folgenden Satzes können jedoch in Satz 19.11 Aussagen über die Existenz einer Lösung getroffen werden.

19.10 Satz:

$a(\tilde{t}) \int_0^{\tilde{t}} \bar{F}(t) dt - F(\tilde{t})$ ist monoton wachsend.

Beweis:

Es sei $0 \leq x \leq t_1 < t_2$. Weil $a(x)\bar{F}(x) = f(x)$, und $a(x)$ monoton wachsend ist, folgt

$$0 \leq a(t_1)\bar{F}(x) - f(x) \leq a(t_2)\bar{F}(x) - f(x).$$

Durch Integrieren ergibt sich

$$a(t_1) \int_0^{t_1} \bar{F}(x) - f(x) dx \leq a(t_2) \int_0^{t_1} \bar{F}(x) - f(x) dx.$$

Durch Ausdehnen des Integrationsbereiches erhält man aufgrund der Nichtnegativität des Integranden

$$a(t_1) \int_0^{t_1} \bar{F}(x) - f(x) dx \leq a(t_2) \int_0^{t_2} \bar{F}(x) - f(x) dx.$$

■

19.11 Satz:

Es sei

$$\lim_{\tilde{t} \rightarrow \infty} a(\tilde{t}) > \frac{1}{\mathbf{E}[X](1 - k)}, \quad (19.9)$$

dann gibt es genau ein \tilde{t}^* mit $K(\tilde{t}^*) \leq K(\tilde{t})$ für alle $\tilde{t} > 0$.

Beweis:

Betrachte man die Gleichung (19.8) für $\tilde{t} \rightarrow \infty$ so ergibt sich:

$$\lim_{\tilde{t} \rightarrow \infty} a(\tilde{t}) \int_0^{\tilde{t}} \bar{F}(t) dt - F(\tilde{t}) = \lim_{\tilde{t} \rightarrow \infty} a(\tilde{t})\mathbf{E}[X] - 1,$$

wenn also

$$\lim_{\tilde{t} \rightarrow \infty} a(\tilde{t}) \cdot \mathbf{E}[X] - 1 > \frac{k}{1 - k}$$

gilt, so folgt die Behauptung. ■

19.12 Bemerkung:

Die Voraussetzung aus Satz 19.11 ist natürlich erfüllt, falls $a(t) \xrightarrow{t \rightarrow \infty} \infty$ gilt, die Ausfallrate des Systems also immer größer wird, d.h. das System altert.

Die Kostenrate für das optimale \tilde{t}^* lautet:

$$\begin{aligned}
 K(\tilde{t}^*) &\stackrel{(19.7)}{=} \frac{(k_h - k_p)F(\tilde{t}^*) + k_p}{\int_0^{\tilde{t}^*} \bar{F}(t)dt} \\
 &\stackrel{(19.8)}{=} \frac{(k_h - k_p) \left[a(\tilde{t}^*) \int_0^{\tilde{t}^*} \bar{F}(t)dt - \frac{k}{1-k} \right] + k_p}{\int_0^{\tilde{t}^*} \bar{F}(t)dt} \\
 &= (k_h - k_p)a(\tilde{t}^*) + \overbrace{\frac{-(k_h - k_p)\frac{k}{1-k} + k_p}{\int_0^{\tilde{t}^*} \bar{F}(t)dt}}^{=0} \\
 &= (k_h - k_p)a(\tilde{t}^*)
 \end{aligned} \tag{19.10}$$

Optimierung der Verfügbarkeit

Bei einer Optimierung bezüglich der Verfügbarkeit des Systems wird wie folgt vorgegangen: Es sei d_h die mittlere Zeit, die für eine Havarieerneuerung benötigt wird und d_p die mittlere Zeit für eine prophylaktische Erneuerung. Wie zu Beginn dieses Teilkapitels angedeutet, kann $0 < d_p < d_h < \infty$ vorausgesetzt werden.

Die Betriebszeit des Systems in einem Zyklus ist $Y := \min(X, \tilde{t})$. Die Zykluslänge L beträgt

$$L = \begin{cases} X + d_h & , \min(X, \tilde{t}) = X \\ \tilde{t} + d_p & , \min(X, \tilde{t}) = \tilde{t} \end{cases} \quad \begin{array}{l} \text{(mit Wahrscheinlichkeit } F(\tilde{t})) \\ \text{(mit Wahrscheinlichkeit } \bar{F}(\tilde{t})) \end{array}$$

Somit berechnet sich der Erwartungswert

$$\begin{aligned}
 \mathbf{E}[L] &= \mathbf{E}[L \mid X \leq \tilde{t}] \cdot F(\tilde{t}) + \mathbf{E}[L \mid X > \tilde{t}] \cdot \bar{F}(\tilde{t}) \\
 &= \left(\frac{\int_0^{\tilde{t}} x dF(x)}{F(\tilde{t})} + d_h \right) \cdot F(\tilde{t}) + (\tilde{t} + d_p) \cdot \bar{F}(\tilde{t}) \\
 &= \int_0^{\tilde{t}} \bar{F}(x) dx + d_h F(\tilde{t}) + d_p \bar{F}(\tilde{t}).
 \end{aligned} \tag{19.11}$$

Damit ergibt sich die stationäre Verfügbarkeit aus (19.2), (19.4) und (19.11):

$$V(\tilde{t}) = \frac{\int_0^{\tilde{t}} \bar{F}(x) dx}{\int_0^{\tilde{t}} \bar{F}(x) dx + d_h F(\tilde{t}) + d_p \bar{F}(\tilde{t})}. \tag{19.12}$$

Das Problem besteht in der Maximierung von $V(\tilde{t})$. Hierzu nutzt man (19.7) und (19.8), da $1/V(\tilde{t}) - 1$ und $K(\tilde{t})$ sich nur durch die Konstanten k_h , k_p und d_h , d_p unterscheiden. Das Problem der Maximierung von $V(\tilde{t})$ ist demnach äquivalent zum Problem der Minimierung von $K(\tilde{t})$.

Das optimale \tilde{t}^* für die Maximierung von $V(\tilde{t})$ lässt sich also durch das Lösen von (19.8) mit $k := d_p/d_h$ berechnen. Die Beziehung (19.9) sichert die Existenz der Lösung. Die zugehörige maximale Verfügbarkeit beträgt somit:

$$V(\tilde{t}^*) = \frac{1}{1 + (d_h - d_p)a(\tilde{t}^*)}. \quad (19.13)$$

19.6 Blockerneuerungen

Wendet man Strategie 1 an, sind die Zeitpunkte, zu denen prophylaktische Massnahmen stattfinden, zu Beginn des Arbeitsprozesses nicht genau bekannt. Die geplanten Erneuerungen nach jeweils \tilde{t} Einheiten werden lediglich mit der Wahrscheinlichkeit $\bar{F}(\tilde{t})$ durchgeführt. Erfordern prophylaktische Erneuerungen eine sorgfältige Vorbereitung, ist diese Strategie daher nicht geeignet. Statt dessen greift man auf die nachfolgend beschriebenen Blockerneuerungen zurück, bei denen die Zeitpunkte für prophylaktische Erneuerungen zu Arbeitsbeginn des Systems bereits fest vorgegeben sind.

Im Folgenden werden dieselben Bezeichnungen wie bei Strategie 1 benutzt.

Strategie 2 — Bei Ausfall und zu festen Zeiten erfolgt Totalerneuerung

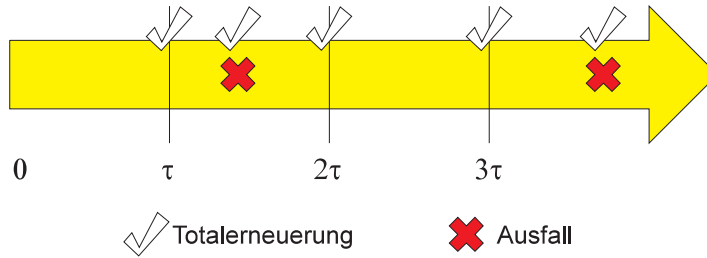


Abbildung 19.4:

Bei dieser Strategie erfolgt bei Ausfall und zu fixierten Zeiten $n\tilde{t}$, $n \in \mathbb{N}$, eine vollständige Erneuerung des Systems. Der Betriebsprozess zerfällt in statistisch äquivalente Zyklen $[n\tilde{t}, (n+1)\tilde{t}]$. In einem Zyklus läuft ein gewöhnlicher Erneuerungsprozess mit Erneuerungszeit $X \stackrel{d}{=} F$. Es sei $R(t)$ die Erneuerungsfunktion, also der Erwartungswert für eine Havarieerneuerung. Dann betragen die mittleren Kosten pro Zyklus

$$\mathbf{E}[C] = k_p + k_h R(\tilde{t}).$$

Die Kostenrate ist dementsprechend

$$K(\tilde{t}) = \frac{k_p + k_h R(\tilde{t})}{\tilde{t}}. \quad (19.14)$$

Eine optimale Zykluslänge \tilde{t}^* ergibt sich aus der Gleichung

$$\begin{aligned} \frac{dK}{dt} &= \frac{k_h R'(\tilde{t})}{\tilde{t}} - \frac{k_p + k_h R(\tilde{t})}{\tilde{t}^2} = 0 \\ \iff \tilde{t} R'(\tilde{t}) - R(\tilde{t}) &= \frac{k_p}{k_h}, \end{aligned} \quad (19.15)$$

dabei ist $R'(\tilde{t})$ die Erneuerungsichte. Existiert ein optimales Erneuerungsintervall \tilde{t}^* dann beträgt die minimale Kostenrate nach (19.14) und (19.15)

$$K(\tilde{t}^*) = k_h R'(\tilde{t}^*).$$

Der Nachteil dieser Strategie ist offensichtlich: Es werden unter Umständen recht neue Systeme erneuert.

Strategie 2' — Nur zu festen Zeiten erfolgt eine Totalerneuerung

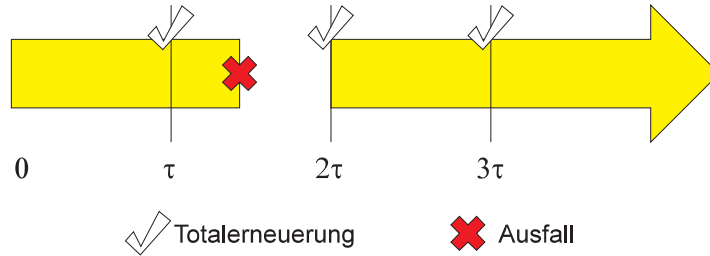


Abbildung 19.5:

Diese Strategie ist der vorherigen ähnlich, es wird jedoch auf die totale Erneuerung bei einem Ausfall verzichtet. Das System wird demnach nur zu den festen Zeiten $n\tilde{t}$, $n \in \mathbb{N}$, erneuert. Fällt das System in der Zeit zwischen diesen Erneuerungszeitpunkten aus, steht es bis zur nächsten Wiederherstellung still. Bei einem solchen Stillstand entsteht jedoch ein finanzieller Verlust, der durch eine monoton wachsende Funktion $v(t)$ beschrieben wird. Zusätzlich wird angenommen, dass v differenzierbar ist und $v(0) = 0$ gilt. Dann ergibt sich die Kostenrate aus (19.1):

$$K(\tilde{t}) = \frac{k_p + \int_0^{\tilde{t}} v(\tilde{t} - t) dF(t)}{\tilde{t}}.$$

Eine optimale Zykluslänge \tilde{t}^* erhält man durch Lösen folgender Gleichung:

$$\int_0^{\tilde{t}} \tilde{t} v'(\tilde{t} - t) - v(\tilde{t} - t) dF(t) = k_p.$$

Es existiert eine Lösung, falls $\lim_{t \rightarrow \infty} v(t)/t = \infty$. Die optimale Kostenrate ist dann

$$K(\tilde{t}^*) = \int_0^{\tilde{t}^*} v'(\tilde{t}^* - t) dF(t).$$

Es sind noch viele weitere Modifikationen möglich, die grundsätzlich auf den Strategien 2 und 2' beruhen. Zum Beispiel:

- Ausgefallene Systeme werden durch gebrauchte ersetzt, zum Beispiel Systemen aus früheren prophylaktischen Erneuerungen.
- Ist das System zum Zeitpunkt $n\tilde{t}$ noch jünger als ein vorher festgelegter Wert T , $0 < T < \tilde{t}$, so wird die prophylaktische Erneuerung auf den Zeitpunkt $(n+1)\tilde{t}$ verschoben.
- Das System wird zu festgelegten Zeitpunkten $n\tilde{t}$ vollständig erneuert. Bei einem Ausfall wird im Allgemeinen eine Totalerneuerung durchgeführt. Sollte das System jedoch im Intervall $[n\tilde{t} - T, n\tilde{t}]$, $0 < T < \tilde{t}$, ausfallen, so wird auf die Totalerneuerung verzichtet.

Die meisten der vorgestellten Strategien beschränken sich auf die Planung vollständiger Erneuerungen. Die folgenden Strategien enthalten nun auch die Option der minimalen Reparatur:

Strategie 3 — Bei Ausfall minimale Reparatur und zu festen Zeiten Totalerneuerung

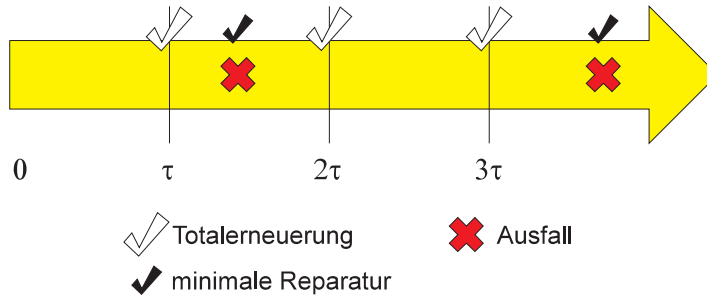


Abbildung 19.6:

In dieser Strategie wird bei einem Ausfall eine minimale Reparatur durchgeführt. Das gesamte System wird zusätzlich zu festen Zeiten $n\tilde{t}$, $n \in \mathbb{N}_0$, total erneuert.

Die vollständigen Erneuerungen erzeugen Zyklen konstanter Länge \tilde{t} . Die mittleren Kosten je Zyklus setzen sich zusammen, unter Zuhilfenahme der Ergebnisse von Strategie 0', aus den mittleren prophylaktischen Kosten je Zyklus und den Kosten einer minimalen Reparatur multipliziert mit der kumulierten Ausfallrate:

$$\mathbf{E}[C] = k_p + k_m \cdot A(\tilde{t}).$$

Es folgt die Kostenrate

$$K(\tilde{t}) = \frac{k_p + k_m A(\tilde{t})}{\tilde{t}},$$

und ein optimales \tilde{t}^* ergibt sich wieder durch Nullsetzen von $dK/d\tilde{t}$:

$$\tilde{t}a(\tilde{t}) - A(\tilde{t}) = \frac{k_p}{k_m}. \quad (19.16)$$

Die optimale Kostenrate lautet somit

$$K(\tilde{t}^*) = k_m \cdot a(\tilde{t}^*).$$

Bezüglich der Verfügbarkeit ergibt sich folgendes: Sind die Zeiten für eine vollständige Erneuerung d_p und eine minimale Reparatur d_m nicht vernachlässigbar klein, so gilt für die Verfügbarkeit des Systems:

$$V(\tilde{t}) = \frac{\tilde{t}}{\tilde{t} + d_m A(\tilde{t}) + d_p}.$$

Das optimale \tilde{t}^* ist, wie bereits in Strategie 1 beschrieben, die Lösung der Gleichung (19.16) mit d_p/d_m statt k_p/k_m . Es ergibt sich die maximale Verfügbarkeit

$$V(\tilde{t}^*) = \frac{1}{1 + d_m a(\tilde{t}^*)}.$$

Strategie 4 — Bei Ausfall minimale Reparatur und bei erstem Ausfall nach fester Zeit Totalerneuerung

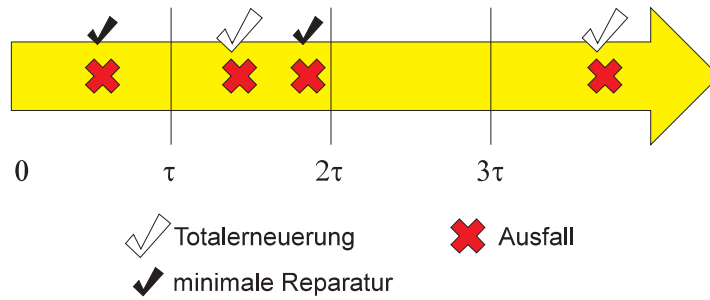


Abbildung 19.7:

Bei dieser Strategie wird bei einem Ausfall im Allgemeinen eine minimale Reparatur durchgeführt. Beim ersten Ausfall nach einer festen Zeit $n\tilde{t}$, $n \in \mathbb{N}_0$, jedoch gibt es eine Totalerneuerung.

Die Kosten für einen Zyklus sind analog zu Strategie 3:

$$\mathbf{E}[C] = k_p + k_m \cdot A(\tilde{t}).$$

Die Zykluslänge ist in diesem Fall aber

$$\mathbf{E}[Y] = \tilde{t} + \mathbf{E}[X_{\tilde{t}}]$$

mit $X_{\tilde{t}}$ als der Restlebensdauer des Systems nach \tilde{t} Zeiteinheiten. Für den Erwartungswert der Restlebensdauer gilt:

$$\mathbf{E}[X_{\tilde{t}}] = \frac{1}{\bar{F}(\tilde{t})} \int_0^\infty \bar{F}(\tilde{t} + x) dx =: r(\tilde{t}).$$

Es ergibt sich damit die Kostenrate

$$K(\tilde{t}) = \frac{k_p + k_m \cdot A(\tilde{t})}{\tilde{t} + r(\tilde{t})}.$$

Das optimale \tilde{t}^* ist die Lösung der folgenden Gleichung, die man wieder durch Differenzieren und Nullsetzen der Kostengleichung erhält:

$$\left(A(\tilde{t}) + \frac{k_p}{k_m} - 1 \right) r(\tilde{t}) = \tilde{t}.$$

Im Falle der Existenz eines \tilde{t}^* beträgt die optimale Kostenrate demzufolge:

$$K(\tilde{t}^*) = \frac{k_m}{r(\tilde{t}^*)}.$$

Neben den hier Vorgestellten sind noch weitere Strategien anwendbar. Zum Beispiel:

- Ausgefallene Systeme werden nach Strategie 4 vollständig wiederhergestellt, mit dem Zusatz, dass eine Erneuerung spätestens \tilde{t}' Zeiteinheiten nach der letzten Erneuerung stattfindet.
- Nach den ersten $n - 1$ Ausfällen wird das System jeweils durch eine minimale Reparatur instand gesetzt und erst nach dem n -ten Ausfall erfolgt eine vollständige Erneuerung.
- Ein System kann durch verschiedene Typen ausfallen. Dementsprechend unterschiedliche Erneuerungsstrategien werden zur Instandsetzung angewendet.

Literatur zu Kapitel 19

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- A.O. ALLEN:
Probability, Statistics and Queueing Theory,
Academic Press, London, 2nd. Edition, 1990.
ISBN: 0120510510
- F. BEICHELT:
Zuverlässigkeits- und Instandhaltungstheorie,
Teubner, Stuttgart, 1993.
ISBN: 3519029855
- F. BEICHELT/P. FRANKEN:
Zuverlässigkeit und Instandhaltung: Mathematische Methoden,
Hanser-Verlag, München, 1984.
ISBN: 3446139060

Kapitel 20

Monte–Carlo–Simulation

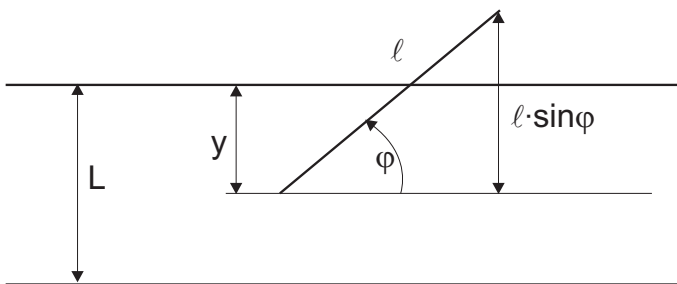
Stochastische Systeme, die aufgrund ihrer Komplexität keiner analytischen Untersuchung zugänglich sind, können experimentell mit Hilfe von Monte–Carlo–Simulationen untersucht werden. Dabei kann es sich z.B. um die Ausbreitung einer ansteckenden Krankheit, den Materialfluss durch ein Produktionssystem oder die zeitliche Entwicklung von Börsenkursen handeln. Will man solche Systeme in der Praxis erforschen, um Abhängigkeiten herauszufinden oder Vorhersagen zu treffen, muss man umfangreiche Experimente durchführen, die kostspielig und langwierig sind. Oft scheitern solche Experimente aber auch daran, dass das reale System überhaupt nicht manipuliert werden kann, oder die Zeit nicht ausreicht, um alle Untersuchungen durchzuführen. Ziel der Simulation ist es daher, die realen Abläufe durch ein stochastisches Experiment auf dem Computer zu ersetzen. Wozu die Natur Tage und Jahre braucht, kann auf diese Weise in wenigen Sekunden oder Minuten nachgestellt werden. Als Ergebnis erhält man Daten, die mit Hilfe statistischer Verfahren ausgewertet werden können.

Schlüsselwörter: Buffonsches Nadelexperiment, Zufallszahlen, linearer Kongruenzgenerator, Periodenlänge, Diskrepanz, Gitterverteilung, nicht–gleichverteilte Zufallszahlen, Inversionsmethode, Verwerfungsmethode, Monte–Carlo–Integration, „Hit or Miss“-Methode, Sample Mean-Methode, stochastische Simulation von Warteschlangen.

20.1 Einleitung

Das Buffon'sche Nadelexperiment ist möglicherweise das älteste Monte-Carlo-Experiment, das in der wissenschaftlichen Literatur dokumentiert ist. Es dient der näherungsweisen Berechnung der Zahl π .

Auf eine ebene Fläche werden Parallelen im Abstand L gezeichnet. Es wird eine Nadel der Länge $l < L$ auf die Fläche geworfen und nach der Wahrscheinlichkeit gefragt, dass die Nadel eine der Parallelen schneidet.

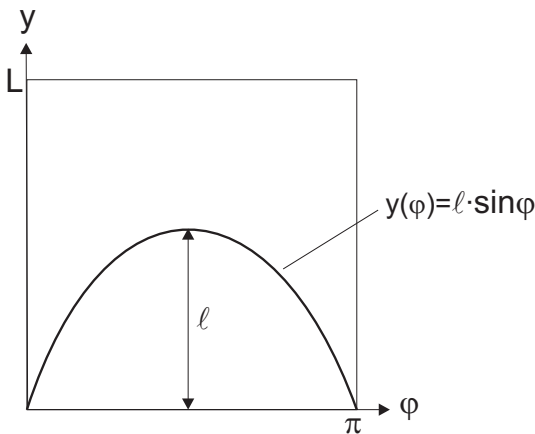


Es ist $0 \leq y \leq L$, $0 \leq \varphi \leq \pi$ (aus Symmetriegründen).

$$\begin{aligned}\Omega &= \{(y, \varphi) \in \mathbb{R}^2 \mid 0 \leq y \leq L; 0 \leq \varphi \leq \pi\} \\ I(\Omega) &= \pi \cdot L\end{aligned}$$

Die Nadel schneidet genau dann eine Parallele, wenn sie sich in den Positionen (y, φ) mit $y \leq l \cdot \sin \varphi$ befindet.

$$\Rightarrow A = \{(y, \varphi) \in \mathbb{R}^2 \mid 0 \leq y \leq L, 0 \leq \varphi \leq \pi, y \leq l \cdot \sin \varphi\}$$



$$\begin{aligned}I(A) &= \int_0^\pi l \cdot \sin \varphi \, d\varphi = [-l \cdot \cos \varphi]_0^\pi \\ &= -l \cdot \cos \pi - (-l) \cdot \cos 0 = l + l = 2l\end{aligned}$$

$$\Rightarrow P(A) = \frac{I(A)}{I(\Omega)} = \frac{2l}{\pi l}$$

entsprechend

$$rh(A) = \frac{I(A)}{I(\Omega)} \approx \frac{2l}{\pi l} \Rightarrow \pi \approx \frac{2l}{rh(E) \cdot L}.$$

Man erzeuge sich auf Ω gleichverteilte Tupel (y, φ) und frage ab, ob $y \leq l \cdot \sin \varphi$ ist.

Da man in der Praxis nicht immer Nadeln fallen lassen oder Würfel werfen kann, stellt sich die Frage, wie man auf einem Computer auf algorithmischem Weg Zahlenfolgen generieren kann, die vorgegebenen Wahrscheinlichkeitsverteilungen genügen (im Fall des Buffon'schen Nadelproblems z.B. auf einem Rechteck gleichverteilte Tupel und ein auf $[0, 2\pi]$ gleichverteilter Winkel).

Algorithmen zur Erzeugung von Zufallszahlen werden Zufallszahlengeneratoren genannt. Man unterscheidet zwei Arten von Zufallszahlengeneratoren:

- **Generatoren echter Zufallszahlen:**

Echte Zufallszahlen lassen sich mit Hilfe von Computern nur schwer erzeugen, da Computerprogramme immer deterministisch ablaufen. Zur Erzeugung echter Zufallszahlen werden oft physikalische Experimente verwendet (thermisches Rauschen bzw. Spannungsschwankungen in bestimmten Bauteilen, radioaktiver Zerfall, in Form von statischem Rauschen aufgenommene elektromagnetische Wellen, usw.).

Echte Zufallszahlen werden besonders in der Kryptographie benötigt, da für sie gilt, dass aus den ersten n Zahlen nie die $n + 1$ -te Zahl vorhergesagt werden kann.

- **Pseudo-Zufallszahlen Generatoren:**

Diese Generatoren erzeugen streng genommen keine wirklichen Zufallszahlen, sondern nur Zahlenfolgen, die gewisse Eigenschaften von Zufallszahlen aufweisen. Meist wird eine bestimmte Verteilung nachgebildet. So erfüllen beispielsweise schon die Zahlen

$$a + 0 \cdot \frac{b-a}{n}, a + 1 \cdot \frac{b-a}{n}, \dots, a + (n-1) \cdot \frac{b-a}{n}, a + n \cdot \frac{b-a}{n}$$

die Eigenschaft auf $[a, b]$ gleichmäßig verteilt zu sein, dennoch sind diese Zahlen alles andere als zufällig.

Für die meisten Simulationsanwendung genügt es, wenn die Zahlen pseudo-zufällig sind. Daher werden in diesem Kapitel nur Pseudo-Zufallszahlen Generatoren behandelt.

Die Erzeugung von Pseudo-Zufallszahlen erfolgt in zwei Schritten:

1. Zunächst werden auf $[0, 1)$ gleichverteilte Pseudo-Zufallszahlen erzeugt. Dafür wird üblicherweise die in diesem Abschnitt behandelte lineare Kongruenzmethode verwendet.
2. Danach werden die Zahlen so transformiert, dass sie als Realisierungen einer vorgegebenen Zufallsvariable und Wahrscheinlichkeitsverteilung angesehen werden können. (Diese Transformationen werden im nächsten Abschnitt behandelt.)

20.2 Erzeugung gleichverteilter Zufallszahlen

Die gebräuchlichsten Algorithmen zur Erzeugung gleichverteilter Zufallszahlen basieren auf der von D.H. Lehmer und A. Rotenberg vorgeschlagenen linearen Kongruenzmethode. Ausgehend von einem Startwert $x_0 \in \mathbb{N}_0$ werden ganze Zahlen x_n nach der Vorschrift

$$x_{n+1} \equiv a \cdot x_n + b \pmod{m} \quad (20.1)$$

berechnet. Die Brüche

$$u_n := \frac{x_n}{m} \quad (n = 0, 1, 2, \dots)$$

werden dann als Stichprobe einer auf $[0, 1)$ gleichverteilten Zufallsvariable verwendet. Der Modul m , der Multiplikator a , das Inkrement b sowie der Startwert x_0 sind hierbei vorzugebende nichtnegative ganze Zahlen mit $0 \leq a, b, x_0 < m$. Der Kongruenzgenerator heißt gemischt, wenn $b \neq 0$ ist, andernfalls heißt er multiplikativ. Die Spezifikation eines linearen Kongruenzgenerators erfolgt kurz durch das Quadrupel (m, a, b, x_0) .

20.1 Bemerkung:

Um die Zahlen u_0, u_1, \dots als Realisierung von stochastisch unabhängigen auf $[0, 1)$ gleichverteilten Zufallsvariablen U_0, U_1, \dots mit der gemeinsamen Verteilungsfunktion

$$F(x) := \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

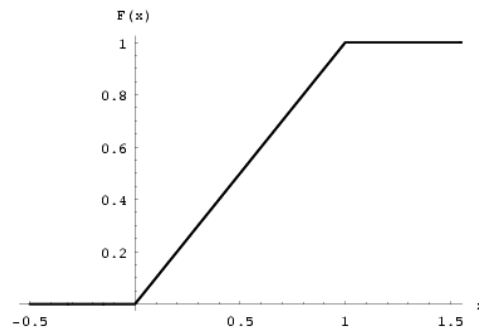


Abbildung 20.1: Verteilungsfunktion $F(x)$ der Gleichverteilung auf $[0, 1]$

dem Erwartungswert

$$\mathbf{E}[U] = \int_{-\infty}^{\infty} x \, dF(x) = \int_0^1 x \cdot 1 \, dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

und der Varianz

$$\mathbf{Var}[U] = \int_{-\infty}^{\infty} (x - \mathbf{E}[U])^2 \, dF(x) = \frac{1}{12}$$

auffassen zu können, müssen eine Reihe von Bedingungen erfüllt sein, die man durch geeignete Manipulationen der Parameter a , b und x_0 herbeizuführen versucht.

Zunächst lässt sich feststellen, dass Zufallszahlen, die auf einer Menge $\{0, \frac{1}{m}, \dots, \frac{m-1}{m}\} \subset [0, 1)$ gleichverteilt sind, für $m \rightarrow \infty$ gegen eine Gleichverteilung auf $[0, 1)$ konvergieren. Deshalb reicht es für viele Anwendungen aus, einen Zufallszahlengenerator zu verwenden, der für ein hinreichend großes m auf $\{0, \frac{1}{m}, \dots, \frac{m-1}{m}\} \subset [0, 1)$ gleichverteilte Zufallszahlen liefert.

20.2 Satz:

Es sei X_m eine auf der Menge $M := \{0, \frac{1}{m}, \dots, \frac{m-1}{m}\} \subset [0, 1)$ gleichverteilte Zufallsvariable und F_m die zugehörige Verteilungsfunktion. Dann gilt

$$\lim_{m \rightarrow \infty} F_m(x) = F(x) \quad \forall x \in \mathbb{R}$$

mit

$$F(x) := \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1. \end{cases}$$

Beweis:

Die zu X_m gehörende Verteilungsfunktion lautet

$$F_m(x) = \begin{cases} 0, & x < 0 \\ \sum_{k=0}^{m-1} I_{[0,x]} \left(\frac{k}{m} \right) \cdot \frac{1}{m}, & x \geq 0. \end{cases}$$

Mit $\frac{k}{m} \leq x \iff k \leq x \cdot m$ und den Eigenschaften der Gaußklammer (ganzzahliger Anteil) gilt für $x \in [0, 1]$:

$$\begin{aligned} F_m(x) &= \sum_{k=0}^{m-1} I_{[0,x]} \left(\frac{k}{m} \right) \cdot \frac{1}{m} \\ &= \lfloor x \cdot m \rfloor \cdot \frac{1}{m} = \frac{x \cdot m - \{x \cdot m\}}{m} \xrightarrow{m \rightarrow \infty} x. \end{aligned}$$

(Dabei ist $\{x\} := x - \lfloor x \rfloor$.) ■

Gesucht ist damit also ein Zufallszahlengenerator, der auf einer Menge obiger Form gleichverteilte Zufallszahlen liefert.

Als Nächstes wird man verlangen, dass die aufeinander folgenden Zahlen u_0, u_1, \dots als Realisierungen stochastisch unabhängiger Zufallsvariablen angesehen werden können.

20.3 Bemerkung (Stochastische Unabhängigkeit und n -dimensionale Diskrepanz):

Da für stochastisch unabhängige auf $[0, 1)$ gleichverteilte Zufallsvariablen U_0, U_1, \dots unter anderem

$$\begin{aligned} &P(a_1 < U_1 \leq b_1, \dots, a_n < U_n \leq b_n) \\ &= [F(b_1) - F(a_1)] \cdot [F(b_2) - F(a_2)] \cdot \dots \cdot [F(b_n) - F(a_n)] \\ &= \prod_{k=1}^n (b_k - a_k) \end{aligned}$$

für $0 \leq a_i \leq b_i \leq 1$ gilt, wird man verlangen, dass die n -dimensionale Diskrepanz

$$\sup_{\substack{k=1, \dots, n \\ 0 < a_k \leq b_k < 1}} |\{(u_i, \dots, u_{i+n-1}) \mid a_k \leq u_{i+k-1} < b_k\}| - m \cdot \prod_{k=1}^n (b_k - a_k)$$

für alle $n \geq 1$ möglichst klein ist.

Diese Bedingung ist gleichbedeutend damit, dass die aufeinanderfolgenden n -Tupel (u_i, \dots, u_{i+n-1}) den Einheitswürfel gleichmäßig ausfüllen. Eine (unendliche) Zahlenfolge, die diese Forderung für alle n exakt erfüllt, wird vollständig gleichverteilt genannt. Deswegen wird man nun verlangen, dass ein Zufallszahlengenerator nicht nur auf $[0, 1)$ gleichverteilte Zufallszahlen produziert, sondern auch über eine ausgewogene Gitterstruktur im $[0, 1)^n$ verfügt.

Da die Größen x_n nur Werte zwischen 0 und $m - 1$ annehmen können, liefert jede Rekursion der Form (20.1) zwangsläufig eine periodische Zahlenfolge, wobei auch noch eine gewisse Vorperiode auftreten kann.

20.4 Definition (Periodenlänge):

Es sei (m, a, b, x_0) ein linearer Kongruenzgenerator. Die kleinste natürliche Zahl $L := L(m, a, b, x_0)$, für die es ein i_0 gibt derart, dass

$$x_{i+L} = x_i \quad \text{für alle } i \geq i_0$$

gilt, heißt die Periodenlänge des Generators. Es gilt immer $L \leq m$.

Um eine möglichst feine Unterteilung des Einheitswürfels $[0, 1)^n$ zu erzielen, ist es folglich notwendig, dass der Generator eine möglichst große Periodenlänge aufweist. Die nachfolgenden Beispiele zeigen, dass die Periodenlänge ganz wesentlich von der Wahl der Parameter a und x_0 abhängt.

20.5 Beispiel:

Im Falle des Generators

$$x_{i+1} \equiv x_i + 3 \pmod{11} \tag{20.2}$$

erhält man unabhängig von der Wahl des Anfangswerts x_0 jedesmal die maximale Periodenlänge $L = 11$:

$$\begin{aligned} x_0 = 1: & \quad 1, 4, 7, 10, 2, 5, 8, 0, 3, 6, 9, \dots \\ x_0 = 2: & \quad 2, 5, 8, 0, 3, 6, 9, 1, 4, 7, 10, 2, \dots \\ x_0 = 3: & \quad 3, 6, 9, 1, 4, 7, 10, 2, 5, 8, 0, 3, \dots \\ & \quad \vdots \end{aligned}$$

Wählt man dagegen in der Rekursion (20.2) anstelle von $a := 1$ den Multiplikator $a := 10$, erhält man

$$\begin{aligned} x_0 = 0: & \quad 0, 3, 0, \dots \\ x_0 = 1: & \quad 1, 2, 1, \dots \\ x_0 = 2: & \quad 2, 1, 2, \dots \\ x_0 = 3: & \quad 3, 0, 3, \dots \\ x_0 = 4: & \quad 4, 10, 4, \dots \\ & \quad \vdots \end{aligned}$$

(Siehe auch Mathematica-Notebook zu diesem Beispiel.)

Aussagen darüber, wie die maximale Periodenlänge zu erzielen ist, trifft der Satz 20.9, für dessen Beweis aber zunächst einige Hilfsmittel aus der Algebra und der Zahlentheorie benötigt werden.

20.6 Satz:

Es sei $m = m' \cdot m''$ mit $m', m'' \in \mathbb{N}$ und $\mathbf{ggT}(m', m'') = 1$. Die Periodenlänge L des durch (m, a, b, x_0) bestimmten linearen Kongruenzgenerators ist das kleinste gemeinsame Vielfache der Periodenlänge L', L'' der durch (m', a, b, x_0) bzw. durch (m'', a, b, x_0) bestimmten linearen Kongruenzgeneratoren, d.h.

$$L = \mathbf{kgV}(L', L'').$$

Beweis:

Sei $(y_i)_{i \in \mathbb{N}}$ die (nicht modulo m reduzierte) Folge mit

$$y_i = ay_{i-1} + b, \quad y_0 = x_0.$$

Ferner seien $(x'_i)_{i \in \mathbb{N}}$ und $(x''_i)_{i \in \mathbb{N}}$ die von den linearen Kongruenzgeneratoren zu (m', a, b, x_0) und (m'', a, b, x_0) erzeugten Folgen in $\{0, 1, \dots, m'\}$ bzw. $\{0, 1, \dots, m''\}$. Nach der Definition der Periodenlängen gilt für alle natürlichen Zahlen n

$$y_{i+nL'} \equiv x'_{i+nL'} = x'_i \pmod{m'} \equiv y_i \quad \text{für alle } i \geq i'_0$$

und

$$y_{i+nL''} \equiv x''_{i+nL''} = x''_i \pmod{m''} \equiv y_i \quad \text{für alle } i \geq i''_0.$$

Damit gibt es $c'_{i,n}, c''_{i,n} \in \mathbb{Z}$ mit

$$y_{i+nL'} = y_i + c'_{i,n} \cdot m' \quad \text{und} \quad y_{i+nL''} = y_i + c''_{i,n} \cdot m''$$

für alle $i \geq i_0 := \max(i'_0, i''_0)$. Setzt man nun

$$L_0 := \mathbf{kgV}(L', L'') = n' \cdot L' = n'' \cdot L'' \quad \text{für } n', n'' \in \mathbb{N},$$

dann gilt für alle $i \geq i_0$:

$$y_i + c'_{i,n} \cdot m' = y_{i+n'L'} = y_{i+L_0} = y_{i+n''L''} = y_i + c''_{i,n''} \cdot m''.$$

Wegen $\mathbf{ggT}(m', m'') = 1$ ist m' ein Teiler von $c''_{i,n''}$, d.h. $c''_{i,n''} = c_i \cdot m'$ mit $c_i \in \mathbb{N}$. Somit folgt

$$y_{i+L_0} = y_i + c_i \cdot m' \cdot m'' = y_i + c_i \cdot m \quad \text{für alle } i \geq i_0.$$

Damit findet man $x_{i+L_0} \equiv x_i \pmod{m}$ und wegen $0 \leq x_j < m$ dann auch $x_{i+L_0} = x_i$ für alle $i \geq i_0$. Da L minimal gewählt ist, gilt $L_0 = \mathbf{kgV}(L', L'') \geq L$. Deshalb muss L ein Teiler von L_0 sein, andernfalls würde sich ein Widerspruch zur Minimaleigenschaft der Periodenlänge ergeben. Andererseits erhält man entsprechend aus

$$y_{i+L} = x_i + c_i \cdot m = x_i + c_i \cdot m' \cdot m'' \quad \text{für alle } i \geq i_0$$

mit $c_i \in \mathbb{N}$, dass $x'_{i+L} = x'_i$ und entsprechend $x''_{i+L} = x''_i$. Also sind L' und L'' und damit auch $L_0 = \mathbf{kgV}(L', L'')$ Teiler von L . Insgesamt gilt daher $L = \mathbf{kgV}(L', L'')$. ■

20.7 Bemerkung:

Durch wiederholte Anwendung der Gleichung 20.1 erhält man

$$x_i = a^i \cdot x_0 + (a^{i-1} + \dots + a + 1)b \pmod{m}, \quad (20.3)$$

d.h.

$$x_i = \begin{cases} x_0 + ib \pmod{m} & , \quad \text{für } a = 1, \\ a^i x_0 + \frac{a^i - 1}{a - 1} \cdot b \pmod{m} & , \quad \text{für } a \neq 1. \end{cases} \quad (20.4)$$

Wegen Satz 20.6 werden also im Folgenden die Ausdrücke $a^i - 1 \pmod{p^e}$, wobei p eine Primzahl und e eine natürliche Zahl ist, näher zu betrachten sein.

20.8 Satz:

Für eine Primzahl p und ein natürliche Zahl e mit $p^e > 2$ gilt:

Aus

$$x \equiv 1 \pmod{p^e} \text{ und } x \not\equiv 1 \pmod{p^{e+1}}$$

folgt

$$x^p \equiv 1 \pmod{p^{e+1}} \text{ und } x^p \not\equiv 1 \pmod{p^{e+2}}.$$

Beweis:

Wegen der Voraussetzungen ist $x = 1 + qp^e$ mit einer ganzen Zahl q , die kein Vielfaches von p ist. Damit gilt also

$$\begin{aligned} x^p &= (1 + qp^e)^p = \sum_{i=0}^p \binom{p}{i} \cdot (qp^e)^i \\ &= 1 + qp^{e+1} \cdot \left[1 + \frac{1}{p} \binom{p}{2} qp^e + \frac{1}{p} \binom{p}{3} q^2 p^{2e} + \dots + \frac{1}{p} \binom{p}{p} q^{p-1} p^{(p-1)e} \right]. \end{aligned}$$

Hierbei sind für $2 \leq i < p$ die Binomialkoeffizienten

$$\binom{p}{i} = \frac{p(p-1) \dots (p-i+1)}{i!}$$

durch p teilbar, da die Binomialkoeffizienten stets ganze Zahlen sind und hier im Zähler die Primzahl $p > i$ auftritt. Daher sind für alle i mit $2 \leq i < p$ die Terme

$$\frac{1}{p} \binom{p}{i} q^{i-1} p^{(i-1)e}$$

durch $p^{(i-1)e}$ teilbar. Der letzte Term ist $q^{p-1} p^{(p-1)e-1}$, der durch p teilbar ist, da wegen $p^e > 2$ die Ungleichung $(p-1)e > 1$ gilt. Also bekommen wir für x^p die Darstellung

$$x^p = 1 - qp^{e+1} \cdot (1 + np) \quad \text{mit } n \in \mathbb{N},$$

d.h.

$$x^p \equiv 1 \pmod{p^{e+1}} \text{ und } x^p \not\equiv 1 \pmod{p^{e+2}}.$$

■

20.9 Satz:

Der Lehmer-Generator hat genau dann die Periodenlänge m , wenn die folgenden Bedingungen erfüllt sind:

- a) b ist relativ prim zu m , d.h. $b \nmid m$,
- b) $a - 1$ ist Vielfaches von p für alle Primteiler p von m , und
- c) $a - 1$ ist Vielfaches von 4, falls 4 Teiler von m ist.

Folgerung:

Insbesondere hat damit jeder Lehmer-Generator der Form $x_{n+1} \equiv x_n + b \pmod{m}$ mit $b \nmid m$ die maximale Periodenlänge m .

Vorbemerkung zum Beweis:

Wegen Satz 20.6 hat der lineare Kongruenzgenerator mit Modul m genau dann die maximal mögliche Periodenlänge m , wenn die entsprechenden linearen Kongruenzgeneratoren mit den Primzahlpotenzen aus der Primfaktorenzerlegung $p_1^{e_1} \cdot p_2^{e_2} \cdot \dots \cdot p_k^{e_k}$ von m jeweils die maximal mögliche Periodenlänge $p_j^{e_j}$ besitzen.

Daher kann im Folgenden angenommen werden, dass $m = p^e$ gilt, wobei p eine Primzahl und e eine natürliche Zahl ist.

Der Beweis wird für beide Fälle $a = 1$ und $a > 1$ getrennt durchgeführt.

Für $a = 1$ hat der Generator wegen Gleichung (20.4) genau dann die Periodenlänge p^e , wenn b teilerfremd zu p^e ist. Also ist in diesem Fall die Behauptung des Satzes richtig.

Es bleibt der Satz für den Fall $a > 1$ zu beweisen. Es wird gezeigt:

„ \Rightarrow “: Besitzt der Generator maximale Periodenlänge, so gelten die Bedingungen a) - c).

„ \Leftarrow “: Gelten die Bedingungen a) - c), so hat der Generator maximale Periodenlänge.

Beweis:

„ \Rightarrow “: Da die Periodenlänge p^e nur dann erreicht wird, wenn jede der möglichen ganzen Zahl x mit $0 \leq x < p^e$ in der Periode auftritt, kann o.B.d.A. auch noch $x_0 = 0$ angenommen werden. Also ergibt sich aufgrund der Gleichung (20.4) für das n -te Element der Folge die Darstellung

$$x_n \equiv \frac{a^n - 1}{a - 1} \cdot b \pmod{p^e}. \quad (20.5)$$

• **Nachweis von Bedingungen a):**

An obiger Gleichung erkennt man, dass $\text{ggT}(b, p^e)$ jedes x_n teilt und daher gleich 1 sein muss. Die Bedingung aus Satz 20.9 a) ist also notwendig.

• **Nachweis von Bedingungen b):**

Weiterhin erkennt man aus (20.5), dass die Periodenlänge genau dann p^e ist, wenn der kleinste Wert von k mit $x_k = x_0 = 0$ gerade $k = p^e$ ist.

Um die Notwendigkeit der Bedingungen b) und c) aus Satz 20.9 zu zeigen, sei also $k = p^e$ vorausgesetzt. Dann gilt mit (20.5) insbesondere

$$\frac{a^{p^e} - 1}{a - 1} \cdot b \equiv 0 \pmod{p^e}. \quad (20.6)$$

Angenommen, es gelte $a \not\equiv 1 \pmod{p}$; dann ist 20.6 äquivalent zu

$$a^{p^e} - 1 \equiv 0 \pmod{p^e}.$$

Hieraus folgt insbesondere

$$a^{p^e} \equiv 1 \pmod{p}.$$

Aber nach dem Satz von Fermat gilt $a^p \equiv a \pmod{p}$, woraus

$$a^{p^e} \equiv a^{p^{e-1}} \equiv \dots \equiv a^p \equiv a \pmod{p}$$

und damit $a \equiv 1 \pmod{p}$ folgt. Dieser Widerspruch zur Annahme zeigt, dass die Bedingung b) notwendig ist.

• **Nachweis von Bedingungen c):**

Um die Notwendigkeit der Bedingung c) zu überprüfen, wird angenommen, dass $p = 2$ ist. Dann ist wegen a) das Inkrement b ungerade. Wäre jetzt der Multiplikator a gerade, so wären alle x_j für $j \in \mathbb{N}$ ungerade, was nicht sein kann. Es bleibt also noch zu zeigen, dass auch $a \equiv 3 \pmod{4}$ zum Widerspruch führt. In diesem Falle hätte man $a \equiv 1 \pmod{2}$ und $a \not\equiv 1 \pmod{2^2}$, also mit Satz 20.8

$$a^2 \equiv 1 \pmod{2^2} \quad \text{und} \quad a^2 \not\equiv 1 \pmod{2^3}.$$

Man rechnet aber sofort nach, dass $a^2 \equiv 9 \equiv 1 \pmod{2^3}$ gilt, was den gewünschten Widerspruch liefert. Also ist auch die Bedingung c) notwendig und damit die Richtung „ \Rightarrow “ des Satzes bewiesen.

„ \Leftarrow “: Gesucht ist der kleinste Wert n , so dass $x_n = x_0$ ist, d.h.

$$\frac{(a^n - 1)(x_0(a - 1) + b)}{a - 1} \equiv 0 \pmod{m}.$$

Nach den Voraussetzungen a) und b) ist $x_0(a - 1) + b$ relativ prim zu m . Es wird demnach der kleinste Wert n gesucht, so dass gilt:

$$\frac{a^n - 1}{a - 1} \equiv 0 \pmod{m}.$$

Es wird gezeigt, dass dieser kleinste Wert n gleich dem Modul $m = p^e$ ist.

Da die Bedingung b) des Satzes erfüllt ist, lässt sich schreiben:

$$a = 1 + kp^f, \quad k \not\equiv 0 \pmod{p}, \quad k \neq 0, \quad f \in \mathbb{N}.$$

p^f ist demnach die maximale Potenz von p , die in $a - 1$ enthalten ist, d.h.

$$a \equiv 1 \pmod{p^f} \quad \text{und} \quad a \not\equiv 1 \pmod{p^{f+1}}.$$

Im Fall $m = p = 2$ sind die Bedingungen a) - c) trivialerweise hinreichend.

Sei nun $p > 2$ oder $f > 1$. Dann ist auch $p^f > 2$ und damit Satz 20.8 anwendbar. Man erhält dann für alle natürlichen Zahlen g :

$$a^{p^g} \equiv 1 \pmod{p^{f+g}} \quad \text{und} \quad a^{p^g} \not\equiv 1 \pmod{p^{f+g+1}}.$$

Daher gelten:

$$\frac{a^{p^g} - 1}{a - 1} \equiv 0 \pmod{p^g} \text{ und } \frac{a^{p^g} - 1}{a - 1} \not\equiv 0 \pmod{p^{g+1}}$$

bzw.

$$\frac{a^{p^g} - 1}{a - 1} \equiv 0 \pmod{p^g} \text{ und } \frac{a^{p^{g-1}} - 1}{a - 1} \not\equiv 0 \pmod{p^g}$$

und insbesondere für $g = e$ und $g = e - 1$

$$\frac{a^{p^e} - 1}{a - 1} \equiv 0 \pmod{p^e} \text{ und } \frac{a^{p^{e-1}} - 1}{a - 1} \not\equiv 0 \pmod{p^e}.$$

Damit hat der lineare Kongruenzgenerator die maximale Periodenlänge p^e , und folglich sind die Bedingungen a) - c) in Satz 20.9 auch hinreichend. ■

20.10 Bemerkung:

In den Bedingungen von Satz 20.9 a) - c) kommt der Startwert x_0 nicht vor. Also erhält man für jeden Startwert x_0 mit $0 \leq x_0 < m$ gemäß der Rechenvorschrift (20.1) eine reinperiodische Folge mit maximaler Periodenlänge m , wenn die Bedingungen a)-c) von Satz 20.9 erfüllt sind. Daher bezeichnet man in diesem Fall die Gleichung (20.1) auch ohne Angabe von x_0 als linearen Kongruenzgenerator. Durch die Wahl von x_0 bestimmt man dabei lediglich, an welcher Stelle man in der Periode beginnt.

20.11 Beispiel:

Es sei $x_{n+1} \equiv 5x_n + 1 \pmod{16}$. Damit ergibt sich folgende Rekursion:

$$\begin{array}{ccccccccccccccc} 1 & \rightarrow & 6 & \rightarrow & 15 & \rightarrow & 12 & \rightarrow & 13 & \rightarrow & 2 & \rightarrow & 11 & \rightarrow & 8 \\ \uparrow & & & & & & & & & & & & & & \downarrow \\ 0 & \leftarrow & 3 & \leftarrow & 10 & \leftarrow & 5 & \leftarrow & 4 & \leftarrow & 7 & \leftarrow & 14 & \leftarrow & 9 \end{array}$$

Trägt man die aufeinanderfolgenden Tupel $(1, 6)$, $(6, 15)$, $(15, 12)$, ... in ein Koordinatensystem ein (siehe Abbildung 20.2), so stellt man fest, dass sie auf nur 4 Hyperebenen zu liegen kommen.

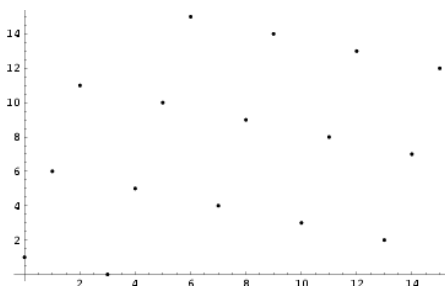


Abbildung 20.2: Darstellung der Tupel (x_n, x_{n+1})

Gitter und Gitterbasen

Um zu guten Gleichverteilungen in $[0, 1]^n$ zu gelangen, muss man deshalb die Anzahl paralleler Hyperebenen, auf denen die aufeinanderfolgenden Tupel liegen, und deren Abstand zueinander untersuchen. Dabei spielt die Menge

$$T_n := \{\vec{u}_i := (u_i, u_{i+1}, \dots, u_{i+n-1}) \mid i = 0, 1, 2, \dots, m-1\}$$

der aufeinanderfolgenden n -Tupel eine besondere Rolle. Der Einfachheit halber wird T_n zu der Menge

$$G_n := \{u + z \mid u \in T_n, z \in \mathbb{Z}^n\}$$

in alle Koordinatenrichtungen mod 1 periodisch fortgesetzt. Für die linearen Kongruenzgeneratoren stellt die Menge G_n ein Gitter bzw. die Vereinigungsmenge von endlich vielen verschiedenen Gittern dar.

Für die weitere Betrachtung soll zunächst der Begriff der Gitterbasis eingeführt werden:

20.12 Definition (aufgespanntes Gitter, Gitterbasis):

Es seien $\vec{g}_1, \dots, \vec{g}_n$ n linear unabhängige Vektoren im \mathbb{R}^n . Die Menge

$$G := \{\vec{g} \in \mathbb{R}^n : \vec{g} = \sum_{i=1}^n z_i \vec{g}_i : z_i \in \mathbb{Z}, 1 \leq i \leq n\}$$

wird das von $\vec{g}_1, \dots, \vec{g}_n$ aufgespannte Gitter genannt; $\vec{g}_1, \dots, \vec{g}_n$ wird als Gitterbasis bezeichnet.

20.13 Bemerkung:

Zu einem Gitter gibt es verschiedene Basen, die das Gitter erzeugen. Durch eine Basistransformation mit einer unimodularen Matrix, d.h. mit einer ganzzahligen Matrix U mit $\det U = \pm 1$, erhält man aus einer Gitterbasis wieder eine solche. Jedoch ist $|\det(\vec{g}_1, \dots, \vec{g}_n)|$ unabhängig von der betrachteten Gitterbasis $\vec{g}_1, \dots, \vec{g}_n$.

20.14 Beispiel (Natürliches Gitter):

Die Punkte im \mathbb{R}^n (als Vektoren ausgefasst) mit ganzzahligen Koordinaten bilden ein Gitter, das von den Einheitsvektoren (der Orthonormalbasis des \mathbb{R}^n) erzeugt wird. Dieses Gitter wird auch als natürliches Gitter bezeichnet.

Zu einem durch (m, a, b, x_0) bestimmten gemischten Kongruenzgenerator sei für eine natürliche Zahl n

$$V_n := \{\vec{x} = (x_i, \dots, x_{i+n-1})^T : x_j \equiv ax_{j-1} + b \pmod{m}, i < j < i + n, 0 \leq i < m\}$$

die Menge von allen erzeugten Punkten \vec{x} ; diese liegen im n -dimensionalen Würfel $[0, m]^n$. Es werde die Menge V_n zu der unendlichen Menge

$$G_n := \{\vec{g} = \vec{x} + m\vec{z} : \vec{x} \in V_n, z \in \mathbb{Z}^n\} \quad (20.7)$$

periodisch mit Periode m fortgesetzt. Mit dieser Bezeichnung gilt der folgende Satz.

20.15 Satz:

Der durch (m, a, b, x_0) bestimmte lineare Kongruenzgenerator habe die maximale Periodenlänge m . Dann bildet die in (20.7) definierte Menge G_n für alle x_0 mit $0 \leq x_0 < m$ ein um den Vektor

$$\vec{g}_0 := b(0, 1, 1 + a, \dots, 1 + a + \dots + a^{n-2})^T$$

verschobenes Gitter mit der Gitterbasis

$$\begin{aligned} \vec{g}_1 &:= (1, a, \dots, a^{n-1})^T, \\ \vec{g}_2 &:= (0, m, 0, \dots, 0)^T, \\ &\vdots \\ \vec{g}_n &:= (0, 0, \dots, 0, m)^T. \end{aligned}$$

Beweis:

Zu zeigen: $G_n := \{g = x + mz \mid x \in V_n, z \in \mathbb{Z}^n\}$ stellt das beschriebene Gitter dar.

1. Alle Punkte von G_n liegen auf dem im Satz definierten Gitter.
2. Alle Gitterpunkte lassen sich als Punkte aus der Menge G_n darstellen.

• Zu 1.

Da nach Voraussetzung der Generator die maximale Periodenlänge m hat, werden für jedes x_0 mit $0 \leq x_0 < m$ alle Zahlen μ mit $\mu \in \{0, 1, \dots, m-1\}$ erzeugt. Analog zur Gleichung (20.3) erhält man für den k -ten Nachfolger von x_i

$$x_{i+k} \equiv a^k \cdot x_i + b \cdot (1 + a + \dots + a^{k-1}) \pmod{m}. \quad (20.8)$$

Daher besteht G_n für alle x_0 aus den Punkten \vec{g} der Form

$$\begin{aligned} \vec{g} &= (\mu, a\mu + b, \dots, a^k\mu + b(1 + a + \dots + a^{k-1}), \dots, a^{n-1}\mu \\ &\quad + b(1 + a + \dots + a^{n-2}))^T + m \cdot (z_1, \dots, z_n)^T \\ &= (0, b, \dots, b(1 + a + \dots + a^{n-2}))^T + \mu(1, a, \dots, a^{n-1})^T + m \cdot (z_1, \dots, z_n)^T, \end{aligned}$$

wobei μ von 0 bis $m-1$ variiert und z_1, \dots, z_n alle ganzen Zahlen durchlaufen. Also lässt sich jeder Punkt von G_n in der Form

$$\vec{g} = \vec{g}_0 + \mu\vec{g}_1 + z_1(m, 0, \dots, 0)^T + z_2\vec{g}_2 + \dots + z_n\vec{g}_n \quad (20.9)$$

und wegen

$$(m, 0, \dots, 0)^T = m\vec{g}_1 - a\vec{g}_2 - \dots - a^{n-1}\vec{g}_n \quad (20.10)$$

als

$$\vec{g} = \vec{g}_0 + (\mu + z_1m)\vec{g}_1 + (z_2 - z_1a)\vec{g}_2 + \dots + (z_n - z_1a^{n-1})\vec{g}_n \quad (20.11)$$

mit $0 \leq \mu + z_1m < m$ und $z_1, \dots, z_n \in \mathbb{Z}$ darstellen. Also liegen alle Punkte von G_n auf dem in der Behauptung des Satzes beschriebenen Gitter.

• Zu 2.

Es bleibt zu zeigen, dass sich umgekehrt auch jeder Punkt

$$\vec{g}' = \vec{g}_0 + z'_1\vec{g}_1 + z'_2\vec{g}_2 + \dots + z'_n\vec{g}_n, \quad z'_1, \dots, z'_n \in \mathbb{Z},$$

in der Form (20.11) darstellen lässt. Dies ist aber mit

$$z_1 := \left\lfloor \frac{z'_1}{m} \right\rfloor, \quad \mu := z'_1 - z_1 m, \quad z_2 := z'_2 + z_1 a, \dots, \quad z_n := z'_n + z_1 a^{n-1} \quad (20.12)$$

erfüllt. Wegen $\det(\vec{g}_1, \dots, \vec{g}_n) = m^{n-1}$ sind die Vektoren $\vec{g}_1, \dots, \vec{g}_n$ linear unabhängig und erzeugen ein Gitter mit dem Gittervolumen m^{n-1} . ■

Der Nachteil dieser Beschreibung ist, dass die Basisvektoren für ein solches Gitter nicht eindeutig sind. Dies führt zu dem im Folgenden definierten Begriff der reduzierten Basis von Minkowski. Dabei beschreiben die Basisvektoren gerade das kleinste in das Gitter eingeschriebene Parallelogramm. Die Bestimmung dieser reduzierten Basis ist relativ aufwendig. Es liefert jedoch ein weiteres Mass für die Güte eines Gitters.

20.16 Definition (Minkowski-reduzierte Basis):

Eine Gitterbasis $\vec{e}_1, \dots, \vec{e}_n$ heisst Minkowski-reduziert, wenn jeder Vektor \vec{e}_k die kürzeste Länge unter allen Vektoren der Form $z_k g_k + \dots + z_n g_n$, $z_i \in \mathbb{Z}$, $\text{ggT}(z_1, \dots, z_n) = 1$ hat.

Auch die Minkowski-reduzierten Basisvektoren eines Gitters sind nicht eindeutig bestimmt. Doch lässt sich zeigen, dass die von zwei Minkowski-Basen aufgespannten Parallelepipede stets kongruent sind. Da die Vektoren einer Minkowski-Basis aufgrund der Definition der Grösse nach geordnet sind, kann man zur Beurteilung der Gitterstruktur den sogenannten Beyer-Quotienten heranziehen.

20.17 Definition (Beyer-Quotient):

Der Beyer-Quotient q_n ist definiert durch

$$q_n := \frac{|\vec{e}_1|}{|\vec{e}_n|}$$

mit \vec{e}_1 als dem kürzesten und \vec{e}_n als dem längsten Gitter-Basisvektor.

Liegt der Beyer-Quotient nahe bei 1, so bedeutet dies, dass alle Basisvektoren nahezu gleich lang sind. Wenn ausserdem alle Winkel zwischen je zwei Basisvektoren annähernd 90° sind, so kann davon ausgegangen werden, dass die Punkte des Gitters den Einheitswürfel gleichmässig ausfüllen. Von einem guten Zufallszahlengenerator wird man deshalb verlangen, dass seine Beyer-Quotienten q_n für möglichst viele n nahe bei 1 liegen.

20.3 Erzeugung nichtgleichverteilter Zufallszahlen

Nichtgleichverteilte Zufallszahlen werden durch geeignete Transformation aus auf $[0, 1)$ gleichverteilter Zufallszahlen gewonnen. Viele dieser Verfahren basieren auf der sogenannten Umkehrmethode und dem Transformationssatz für Dichten.

Um die Umkehrmethode einführen zu können, wird zunächst der Begriff der Pseudo-Inversen benötigt.

20.18 Definition (Pseudo-Inverse):

Es sei $g: \mathbb{R} \rightarrow \mathbb{R}$ eine monoton nichtfallende rechtsseitig stetige Funktion und es seien

$$\alpha := \inf\{g(x) \mid x \in \mathbb{R}\} \quad \text{und} \quad \beta := \sup\{g(x) \mid x \in \mathbb{R}\}.$$

Dann ist auf dem offenen Intervall (α, β) die Pseudo-Inverse g^{-1} von g durch

$$g^{-1}(y) := \inf\{x \in \mathbb{R} \mid g(x) \geq y\}, \quad \alpha < y < \beta,$$

erklärt.

20.19 Satz (Eigenschaften der Pseudo-Inversen):

Es sei $g: \mathbb{R} \rightarrow \mathbb{R}$ eine monoton nichtfallende rechtsseitig stetige Funktion und g^{-1} ihre Pseudo-Inverse. Dann gilt:

- a) g^{-1} ist auf (α, β) monoton nichtfallend und linksseitig stetig.
- b) Es gilt $g(g^{-1}(y)) \geq y$ für alle $\alpha < y < \beta$.
Ist g in $g^{-1}(y)$ stetig, so gilt $g(g^{-1}(y)) = y$ für alle $\alpha < y < \beta$.
- c) Es gilt $g^{-1}(g(x)) \geq x$ für alle $\alpha < g(x) < \beta$.
Ist g^{-1} in $g(x)$ stetig, so gilt $g^{-1}(g(x)) = x$ für alle $\alpha < g(x) < \beta$.
- d) Es gilt: $y \leq g(x) \Leftrightarrow g^{-1}(y) \leq x$ für $\alpha < g(x), y < \beta$.

Beweis:

Übung. ■

20.20 Satz (Umkehrmethode):

Es sei $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $U: \Omega \rightarrow \mathbb{R}$ eine auf dem Intervall $[0, 1]$ Rechteckverteilte (gleichverteilte) Zufallsvariable. Ferner sei Q ein Wahrscheinlichkeitsmaß auf \mathfrak{B} mit Verteilungsfunktion F . Die Pseudo-Inverse von F sei F^{-1} . Dann besitzt die für alle $\omega \in \Omega$ durch

$$X(\omega) := \begin{cases} F^{-1}(U(\omega)) & , \quad U(\omega) \in (0, 1) \\ 0 & , \quad \text{sonst} \end{cases}$$

definierte Zufallsvariable $X: \Omega \rightarrow \mathbb{R}$ die Verteilung $P_X = Q$.

Beweis:

Zunächst wird gezeigt, dass X messbar ist. Bezugnehmend auf Definition 20.18 ist $\alpha := 0$ und $\beta := 1$. Aus der linksseitigen Stetigkeit von F^{-1} erhält man:

$$X^{-1}(\{0\}) = U^{-1}\left((F^{-1})^{-1}(\{0\})\right) \cup U^{-1}((-\infty, 0] \cup [1, \infty)).$$

Wegen $(F^{-1})^{-1}(\{0\}) \in (0, 1) \cap \mathfrak{B}$ und $U^{-1}((-\infty, 0] \cup [1, \infty)) \in \mathfrak{B}$ ist $X^{-1}(\{0\}) \in \mathfrak{F}$. Sofern $0 \notin \mathcal{B} \subset \mathfrak{B}$ ist, ergibt sich wegen $(F^{-1})^{-1}(\mathcal{B}) \in (0, 1) \cap \mathfrak{B}$ auch

$$X^{-1}(\mathcal{B}) = U^{-1}((F^{-1})^{-1}(\mathcal{B})) \in \mathfrak{F}.$$

Mit Hilfe von Satz 20.19 d) schließt man weiter:

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x) \quad \text{für } 0 < F(x) < 1,$$

wobei $P(U \notin (0, 1)) = 0$ ist. ■

20.21 Beispiel (Rechteck-verteilte Zufallsvariablen):

Es sei U eine auf dem Intervall $[a, b]$, $a < b$, Rechteck-verteilte Zufallsvariable. Die zu U gehörende Verteilungsfunktion lautet:

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

Offensichtlich ist

$$F^{-1}(y) = y(b-a) + a \quad \text{für } 0 < y < 1.$$

Hieraus folgt: Ist U eine auf $[0, 1]$ gleichverteilte Zufallsvariable, dann ist $X := U(b-a) + a$ eine auf $[a, b]$ gleichverteilte Zufallsvariable.

20.22 Beispiel (Exponentiell verteilte Zufallsvariablen):

Für die Verteilungsfunktion F einer mit dem Parameter λ exponentiell verteilten Zufallsvariablen gilt:

$$F(x) := \begin{cases} 1 - e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad x < 0. \end{cases}$$

Die Umkehrfunktion lautet

$$F^{-1}(y) = -\frac{1}{\lambda} \ln(1-y) \quad \text{für } 0 < y < 1.$$

Hieraus folgt: Ist U eine auf $[0, 1]$ gleichverteilte Zufallsvariable, dann ist $X := -\frac{1}{\lambda} \ln(1-U)$ eine mit dem Parameter $\lambda > 0$ exponentiell-verteilte Zufallsvariable. Da wegen

$$P(1-U \leq x) = P(U \geq 1-x) = P(U > 1-x) = 1 - (1-x) = x$$

mit U auch $1-U$ auf $[0, 1]$ gleichverteilt ist, ist auch $X := -\frac{1}{\lambda} \ln(U)$ eine mit dem Parameter $\lambda > 0$ exponentiell verteilte Zufallsvariable.

20.23 Beispiel (Cauchy-verteilte Zufallsvariablen):

Für die Dichte f einer mit den Parametern $\alpha > 0$ und $\beta > 0$ Cauchy-verteilten Zufallsvariablen gilt:

$$f(x) := \frac{\beta}{\pi[\beta^2 + (x-\alpha)^2]}, \quad x \in \mathbb{R}.$$

Die zugehörige Verteilungsfunktion ist

$$F(x) := \frac{1}{2} + \pi^{-1} \arctan\left(\frac{x-\alpha}{\beta}\right), \quad x \in \mathbb{R},$$

(vgl. Stochastik I Kapitel 7.3). Die zugehörige Inverse lautet

$$F^{-1}(y) = \alpha + \beta \tan\left(\pi \left[y - \frac{1}{2}\right]\right) = \alpha - \frac{\beta}{\tan(\pi y)}, \quad 0 < y < 1.$$

Mit anderen Worten: Ist U auf $[0, 1]$ gleichverteilt, dann ist

$$X := \alpha - \frac{\beta}{\tan(\pi U)}$$

Cauchy-verteilt mit den Parametern $\alpha > 0$ und $\beta > 0$.

20.24 Beispiel (Weibull-verteilte Zufallsvariablen):

Für die Dichte f einer mit den Parametern $\lambda > 0$ und $\beta > 0$ Weibull-verteilten Zufallsvariablen gilt:

$$f(x) := \begin{cases} \lambda \cdot \beta \cdot x^{\beta-1} \cdot e^{-\lambda x^\beta} & , 0 \leq x < \infty \\ 0 & , x < 0. \end{cases}$$

Die zugehörige Verteilungsfunktion lautet

$$F(x) := \begin{cases} 1 - e^{-\lambda x^\beta} & , x \geq 0 \\ 0 & , x < 0. \end{cases}$$

Die zu F Inverse ist

$$F^{-1}(y) = \left(-\frac{1}{\lambda} \ln(1-y) \right)^{\frac{1}{\beta}} \quad \text{für } 0 < y < 1.$$

Mit anderen Worten: Ist U eine auf $[0, 1]$ gleichverteilte Zufallsvariable, dann ist

$$X := \left(-\frac{1}{\lambda} \ln(1-U) \right)^{\frac{1}{\beta}}$$

Weibull-verteilt mit den Parametern $\lambda > 0$ und $\beta > 0$.

20.25 Beispiel (Bernoulli-verteilte Zufallsvariablen):

Es sei X eine Bernoulli-verteilte Zufallsvariable mit $P(X = a) = p$ und $P(X = b) = 1-p$, $a < b$. X hat die Verteilungsfunktion

$$F(x) := \begin{cases} 0 & , x < a \\ p & , a \leq x < b \\ 1 & , x \geq b. \end{cases}$$

Die Pseudo-Inverse lautet

$$F^{-1}(y) = \begin{cases} a & , 0 < y \leq p \\ b & , p < y \leq 1. \end{cases}$$

Bernoulli-verteilte Zufallszahlen erzeugt man also, indem man zunächst auf $[0, 1)$ gleichverteilte Zufallszahlen erzeugt und dann entscheidet: Fällt u in das Intervall $(0, p]$ setzt man $x = a$, fällt u in das Intervall $(p, 1]$, setzt man $x = b$.

20.26 Beispiel (Geometrisch-verteilte Zufallsvariablen):

Gesucht ist die Realisierung einer geometrisch verteilten Zufallsgrösse. Es gilt für $i \in \mathbb{N}$:

$$P(X = i) = p^i(1-p) \implies F(k) = \sum_{i=0}^k p^i(1-p) = (1-p) \frac{1-p^{k+1}}{1-p} = 1 - p^{k+1}.$$

Gesucht ist nun das kleinste k , für das $F(k) = 1 - p^{k+1} \geq U$ gilt.

$$\begin{aligned} 1 - U \geq p^{k+1} & \iff \ln(1-U) \geq (k+1) \ln p \iff k+1 \geq \frac{\ln(1-U)}{\ln p} \\ & \iff k = \left\lfloor \frac{\ln(1-U)}{\ln p} \right\rfloor \iff k = \left\lfloor \frac{\ln U}{\ln p} \right\rfloor. \end{aligned}$$

Das von Box und Muller stammende Verfahren zur Erzeugung $\mathcal{N}(0, 1)$ -verteilter Zufallszahlen basiert auf folgendem Satz.

20.27 Satz:

Es seien U_1 und U_2 stochastisch unabhängige $\mathcal{R}([0, 1])$ -verteilte Zufallsvariablen. Dann sind die Zufallsvariablen

$$X_1 := \sqrt{-2 \ln U_1} \cdot \cos(2\pi U_2) \quad \text{und} \quad X_2 := \sqrt{-2 \ln U_1} \cdot \sin(2\pi U_2)$$

stochastisch unabhängig und jeweils $\mathcal{N}(0, 1)$ -verteilt.

Beweis:

Wir setzen $G: (0, 1)^2 \rightarrow \mathbb{R}^2$ mit

$$G(u_1, u_2) := (x_1, x_2) := \left(\sqrt{-2 \ln u_1} \cdot \cos(2\pi u_2), \sqrt{-2 \ln u_1} \cdot \sin(2\pi u_2) \right), \quad (u_1, u_2) \in (0, 1)^2$$

und wenden den Transformationssatz für Dichten an. Die zu G gehörende Funktionaldeterminante lautet

$$\begin{aligned} \left| \begin{array}{cc} \frac{\partial G_1}{\partial u_1} & \frac{\partial G_1}{\partial u_2} \\ \frac{\partial G_2}{\partial u_1} & \frac{\partial G_2}{\partial u_2} \end{array} \right| &= \left| \begin{array}{cc} -\frac{\cos(2\pi u_2)}{u_1 \sqrt{-2 \ln u_1}} & -2\pi \sqrt{-2 \ln u_1} \cdot \sin(2\pi u_2) \\ -\frac{\sin(2\pi u_2)}{u_1 \sqrt{-2 \ln u_1}} & 2\pi \sqrt{-2 \ln u_1} \cdot \cos(2\pi u_2) \end{array} \right| \\ &= \frac{2\pi \cos^2(2\pi u_2)}{u_1} + \frac{2\pi \sin^2(2\pi u_2)}{u_1} \\ &= \frac{2\pi}{u_1}. \end{aligned}$$

Es ist

$$x_1^2 + x_2^2 = -2 \ln u_1 \quad \text{bzw.} \quad u_1 = \exp\left(-\frac{x_1^2 + x_2^2}{2}\right), \quad (x_1, x_2) \in \mathbb{R}^2.$$

Damit erhält man

$$\begin{aligned} f_{(X_1, X_2)}(x_1, x_2) &= \frac{1}{\left| \det \left(\frac{\partial G_j}{\partial u_j}(G^{-1}(x_1, x_2)) \right) \right|} \\ &= \frac{u_1}{2\pi} = \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}}, \quad (x_1, x_2) \in \mathbb{R}^2, \end{aligned}$$

und

$$f_{X_1}(x_1) = \int_{\mathbb{R}} \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}} dx_2 = \frac{1}{2\pi} e^{-\frac{x_1^2}{2}} \int_{\mathbb{R}} e^{-\frac{x_2^2}{2}} dx_2 = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}}, \quad x_1 \in \mathbb{R}$$

bzw.

$$f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}}, \quad x_2 \in \mathbb{R}.$$

Also sind X_1 und X_2 jeweils $\mathcal{N}(0, 1)$ -verteilt und stochastisch unabhängig. ■

20.28 Bemerkung:

Bekannte andere Transformationsmethoden sind die Verwerfungsmethode und die Quotenmethode.

20.4 Monte-Carlo-Integration

Die Lösung mathematischer Probleme mittels stochastischer Experimente soll zunächst anhand der Monte-Carlo-Integration demonstriert werden.

Es sei $g(x)$ eine reelle auf $[a, b]$ integrierbare Funktion mit $0 \leq g(x) \leq c$ für alle $x \in [a, b]$ und ein $c \in \mathbb{R}$. Es wird das Problem der Berechnung von

$$I := \int_a^b g(x) dx$$

betrachtet. Dazu werden die beiden folgenden Mengen definiert:

$$A := \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq g(x)\} \quad \text{und} \quad D := \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq c\}.$$

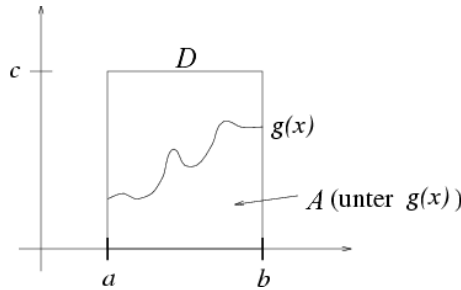


Abbildung 20.3: Die Mengen A und D bezogen auf eine Funktion $g(x)$

Sei nun $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum und $Z := (U, V): \Omega \rightarrow D$ ein auf D gleichverteilter Zufallsvektor, d.h. Z besitzt eine Dichte der Form

$$f_Z(u, v) := f(u, v) := \begin{cases} \frac{1}{c(b-a)} & , \quad (u, v) \in D \\ 0 & , \quad (u, v) \notin D. \end{cases}$$

Dann gilt:

$$\begin{aligned} P(Z \in A) &= \iint_A f(u, v) du dv = \frac{1}{c(b-a)} \iint_A du dv = \frac{\int_a^b g(x) dx}{c(b-a)} \\ &= \frac{I}{c(b-a)}. \end{aligned}$$

Sei weiter $(Z_i)_{i \in \mathbb{N}}$ eine i.i.d. Folge von auf D gleichverteilten Zufallsvektoren. Es sei

$$\theta_n^{(1)} := \frac{c(b-a)}{n} \sum_{i=1}^n I_A(Z_i).$$

Da $\mathbf{E}[I_A(Z_1)] = P(Z_1 \in A) =: \theta < \infty$ gilt, folgt mit dem starkem Gesetz der großen Zahlen

$$\theta_n^{(1)} \xrightarrow[n \rightarrow \infty]{P\text{-f.s.}} \varrho(\theta) := c(b-a) \cdot \mathbf{E}[\theta] = I = \int_a^b g(x) dx.$$

Wenn es also gelänge, mit Hilfe eines physikalischen Experiments oder eines linearen Kongruenzgenerators Realisierungen eines auf Ω gleichverteilten Zufallsvektors zu erzeugen, könnte man $\theta_n^{(1)}$ als Schätzverfahren für I verwenden. Diese Vorgehensweise zum Schätzen von I wird „Hit or Miss“ genannt.

20.29 Beispiel:

Bei der numerischen Integration des Integrals

$$\int_0^{\pi} \sin x \, dx$$

erzeugt man sich zunächst auf $D := [0, \pi] \times [0, 1]$ gleichverteilte Punkte (u, v) und setzt die Anzahl der Treffer $((u, v) \in A)$ in Beziehung zum Umfang der Stichprobe.

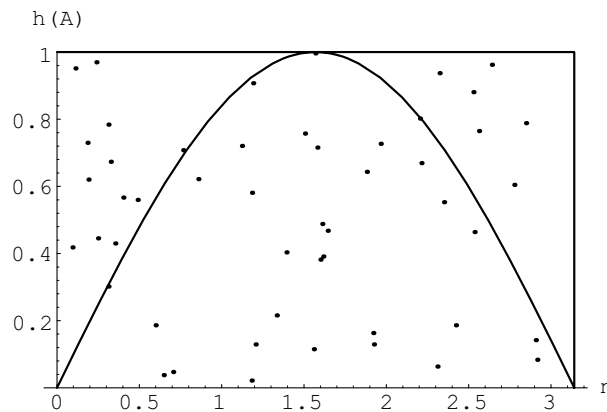


Abbildung 20.4: Im Beispiel ist $a := 0$, $b := \pi$, $c := 1$, $N := 50$ und die Anzahl der Treffer 32.

Damit folgt:

$$\int_0^{\pi} \sin x \, dx \approx 1 \cdot (\pi - 0) \cdot \frac{32}{50} \approx 2.010619298,$$

was dem tatsächlichen Wert 2 schon sehr nahe kommt.
(Siehe auch Mathematica-Notebook zu diesem Beispiel.)

Algorithmus zur „Hit or Miss“-Methode:

1. Erzeuge auf $[0, 1] \times [0, 1]$ gleichverteilte Tupel $(u_1, w_1), \dots, (u_N, w_N)$.
2. $N_H := 0$
3. Für $i = 1, \dots, N$ führe aus:
 Berechne $x_i := a + u_i(b - a)$
 Berechne $g(x_i)$
 Falls $g(x_i) > cw_i$, dann setze $N_H = N_H + 1$.
4. Führe aus $I \approx c(b - a) \frac{N_H}{N}$.

Zur „Hit or Miss“-Methode gibt es noch eine interessante Alternative, die sogenannte „Sample Mean“-Methode (Mittelwertmethode). Dazu bezeichne X eine auf dem Intervall $[a, b]$ gleichverteilte Zufallsvariable, d.h. die Verteilung von $X: \Omega \rightarrow [a, b]$ besitzt eine Dichte der Form

$$f_X(x) := f(x) := \begin{cases} \frac{1}{b-a}, & \text{für } a \leq x \leq b \text{ (} a < b \text{)}, \\ 0, & \text{sonst.} \end{cases}$$

Dann lässt sich schreiben:

$$I := \int_a^b g(x) dx = \int_a^b \frac{g(x)}{f(x)} f(x) dx = \mathbf{E}_P \left[\frac{g(X)}{f(X)} \right] = \mathbf{E}_{P_X} \left[\frac{g}{f} \right].$$

Dieser Zusammenhang legt für I den Schätzer

$$\theta_n^{(2)} := \frac{1}{n} \sum_{i=1}^n \frac{g(X_i)}{f(X_i)} = \frac{b-a}{n} \sum_{i=1}^n g(X_i)$$

nahe, wobei $(X_i)_{i \in \mathbb{N}}$ eine i.i.d. Folge von auf $[a, b]$ gleichverteilten Zufallsvariablen ist. Das starke Gesetz der großen Zahlen liefert für $\theta_n^{(2)}$ ebenfalls

$$\theta_n^{(2)} \xrightarrow[n \rightarrow \infty]{P\text{-f.s.}} \mathbf{E}_P \left[\frac{g(X_1)}{f(X_1)} \right] = I = \int_a^b g(x) dx.$$

Algorithmus zur „Sample Mean“-Methode:

1. Erzeuge eine Folge u_1, \dots, u_N stochastisch unabhängiger, auf $[0, 1)$ gleichverteilter Zufallszahlen.
2. Berechne $x_i := a + u_i(b-a)$ für $i = 1, \dots, N$.
3. Berechne $g(x_i)$ für $i = 1, \dots, N$.
4. Führe aus $I \approx (b-a) \frac{1}{N} \sum_{i=1}^N g(x_i)$.

Beide Schätzer sind erwartungstreu. Denn es gilt:

$$\begin{aligned} \mathbf{E}_\theta \left[\theta_n^{(1)} \right] &= \mathbf{E}_\theta \left[c(b-a) \cdot \frac{1}{n} \sum_{i=1}^n I_A(Z_i) \right] \\ &= \frac{c(b-a)}{n} \cdot \sum_{i=1}^n \mathbf{E}_\theta [I_A(Z_i)] \\ &= c(b-a) \cdot \mathbf{E}_\theta [I_A(Z_1)] \\ &= c \cdot (b-a) \cdot P(Z_1 \in A) \\ &= c \cdot (b-a) \cdot \theta \\ &= I = \varrho(\theta). \end{aligned}$$

$$\begin{aligned}
 \mathbf{E}_\theta \left[\theta_n^{(2)} \right] &= \mathbf{E}_\theta \left[\frac{1}{n} \sum_{i=1}^n \frac{g(X_i)}{f(X_i)} \right] \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbf{E}_\theta \left[\frac{g(X_i)}{f(X_i)} \right] \\
 &= \mathbf{E}_\theta \left[\frac{g(X_1)}{f(X_1)} \right] \\
 &= I.
 \end{aligned}$$

Für die zugehörige Varianz gilt:

$$\begin{aligned}
 \mathbf{Var}_\theta \left[\theta_n^{(1)} \right] &= \mathbf{Var}_\theta \left[c \cdot (b-a) \cdot \frac{1}{n} \sum_{i=1}^n I_A(Z_i) \right] \\
 &= \frac{c^2(b-a)^2}{n^2} \sum_{i=1}^n \mathbf{Var}_\theta [I_A(Z_i)] \\
 &= \frac{c^2(b-a)^2}{n} \mathbf{Var}_\theta [I_A(Z_1)] \\
 &= \frac{c^2(b-a)^2}{n} \cdot (\mathbf{E}_\theta [(I_A(Z_1))^2] - (\mathbf{E}_\theta [I_A(Z_1)])^2) \\
 &= \frac{c^2(b-a)^2}{n} \cdot (\mathbf{E}_\theta [I_A(Z_1)] - (\mathbf{E}_\theta [I_A(Z_1)])^2) \\
 &= \frac{c^2(b-a)^2}{n} \cdot (P(Z_1 \in A) - (P(Z_1 \in A))^2) \\
 &= \frac{c^2(b-a)^2}{n} \cdot \left(\frac{I}{c \cdot (b-a)} - \frac{I^2}{c^2 \cdot (b-a)^2} \right) \\
 &= \frac{1}{n} (c \cdot (b-a) I - I^2).
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{Var}_\theta \left[\theta_n^{(2)} \right] &= \mathbf{Var}_\theta \left[\frac{1}{n} \sum_{i=1}^n \frac{g(X_i)}{f(X_i)} \right] \\
 &= \frac{1}{n^2} \sum_{i=1}^n \mathbf{Var}_\theta \left[\frac{g(X_i)}{f(X_i)} \right] \\
 &= \frac{1}{n} \mathbf{Var}_\theta \left[\frac{g(X_1)}{f(X_1)} \right] \\
 &= \frac{1}{n} \left(\mathbf{E}_\theta \left[\left(\frac{g(X_1)}{f(X_1)} \right)^2 \right] - \left(\mathbf{E}_\theta \left[\frac{g(X_1)}{f(X_1)} \right] \right)^2 \right) \\
 &= \frac{1}{n} \left(\int_a^b \frac{g^2(x)}{f^2(x)} \cdot f(x) \, dx - I^2 \right)
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \left(\int_a^b \frac{g^2(x)}{f(x)} dx - I^2 \right) \\
&= \frac{1}{n} \left((b-a) \cdot \int_a^b g^2(x) dx - I^2 \right) \\
&\leq \frac{1}{n} \left(c(b-a) \cdot \int_a^b g(x) dx - I^2 \right) \quad (\text{da } g(x) \leq c \ \forall x \in \mathbb{R} \text{ vorausgesetzt war}) \\
&= \frac{1}{n} (c(b-a) \cdot I - I^2),
\end{aligned}$$

woran man erkennt, dass $\theta_n^{(2)}$ gleichmäßig besser ist als $\theta_n^{(1)}$.

20.5 Stochastische Simulation von Warteschlangen

Warteschlangen sind aus betriebswirtschaftlicher Sicht meist nicht wünschenswert. Denn die Zeit, die man in der Warteschlange zubringt, ist Untätigkeitszeit, die weder den Kunden noch dem Betreiber des Systems zugute kommt. Den wirtschaftlichen Einfluss von Warteschlangen kann man am besten am Beispiel der Produktion verdeutlichen. Lange Durchlaufzeiten durch die Produktion haben zur Folge, dass man neue Produkte nicht schnell genug auf den Markt bringen kann und der Konkurrenz das Feld überlassen muss. Da lange Durchlaufzeiten mit hohen Beständen korreliert sind, entstehen durch die auf Bearbeitung wartenden Halbfertigfabrikate außerdem hohe Kapitalbindungskosten, die sich negativ auf das Betriebsergebnis auswirken.

Das Phänomen des Wartens wird seit fast einem Jahrhundert wissenschaftlich erforscht. Bereits 1917 publizierte der dänische Ingenieur und Mathematiker A.K. Erlang, der bei einer Kopenhagener Telefongesellschaft beschäftigt war, eine mathematische Formel, mit deren Hilfe man Fernsprechvermittlungsstellen dimensionieren konnte. Nach Erlang waren es hauptsächlich Nachrichtentechniker, die mathematische Verfahren benutzten, um den Telefonverkehr durchgängiger und effizienter zu machen. Mitte des vergangenen Jahrhunderts hieß es, dass die Länder, die über die schlechtesten Telefonsysteme verfügten, die besten Mathematiker auf dem Gebiet der Warteschlangentheorie hervorbringen würden. Mit dem Aufkommen der Datenverarbeitung wurden diese Methoden auch zur Konzeption von Rechensystemen verwendet. Ziel der Analysen ist es, bereits im Vorfeld der Planung Engpässe und Schwachstellen zu erkennen. In letzter Zeit wird die Warteschlangentheorie auch immer häufiger auf Fragen der Produktion, des Verkehrs und der Modellierung von Geschäftsprozessen ausgedehnt. Mittlerweile sind mehr als 10000 wissenschaftliche Publikationen über Warteschlangenprobleme erschienen, die sich auf die unterschiedlichsten Bereiche des täglichen Lebens beziehen.

Beschreibung von Bediensystemen

Die Warteschlangentheorie verwendet zur Beschreibung von Bedienungssystemen ein einfaches Grundmodell. Dieses Modell ist in Abbildung 20.5 dargestellt. Es besteht aus dem sogenannten Bedienungsschalter, der über eine oder mehrere parallel arbeitende gleichartige

Maschinen bzw. Arbeitsplätze verfügt, und aus einem Warteraum. Die Kunden treffen einzeln und zu zufälligen Zeitpunkten vor dem Bedienungssystem ein. Ein neu ankommender Kunde wird bedient, sofern mindestens eine der Maschinen frei ist, andernfalls muss er sich in die Warteschlange einreihen.

Die Begriffe Kunde und Schalter können in der Praxis unterschiedliche Bedeutungen haben: Fahrzeuge, die an einer Verkehrsampel warten; Computerprogramme, die in einem Rechnernetz zirkulieren; Telefonanrufe, die an einer Vermittlungsstelle einfallen; Werkstücke, die von einer Maschine bearbeitet werden; Patienten, die in einer Arztpraxis auf ihre Behandlung warten, usw.

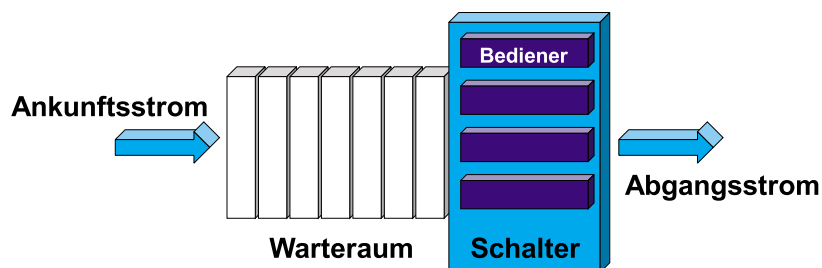


Abbildung 20.5: Grundmodell eines Bedienungssystems

Das Grundmodell kann auf vielfältige Weise variiert werden:

- Die Kunden werden nicht einzeln, sondern gruppenweise bedient (Systeme mit Gruppenbedienung)
Anwendung: Losfertigung in einem Produktionsbetrieb, Pendel-Busse.
- Einige Kunden verlassen das System, bevor sie bedient worden sind (Wartesysteme mit Zeitbeschränkungen)
Anwendung: Lagerhaltung von verderblicher Ware.
- Nicht alle Bedienungsgeräte stehen jedem Kunden zur Verfügung (Bedienungssysteme mit eingeschränkter Erreichbarkeit)
Anwendung: Fertigungsstraßen mit dedizierten Maschinen, Koppelanforderungen in einem Fernsprechnet.
- Einige Kunden scheuen sich, in das Bedienungssystem einzutreten, weil ihnen die Warteschlange zu lang erscheint (Wartesysteme mit ungeduldigen Kunden)
Anwendung: Übliches Kundenverhalten an einem Post-, Bank- oder Fahrkartenschalter.
- Ein Kunde höherer Priorität verdrängt einen Kunden niedrigerer Priorität aus dem Bedienungsprozess (Bedienungssysteme mit Prioritätensteuerung)
Anwendung: Expres-Los-Steuerung in einem Fertigungsprozess.
- Ein Kunde, der bei seiner Ankunft nicht sofort bedient werden kann, geht verloren (Verlustsysteme)
Anwendung: Telefonate in einem Fernsprechnet.

Die Kunden fallen zu den zufälligen Zeitpunkten $(T_n)_{n \in \mathbb{N}}$ in das System ein. Die Zeitspanne $I_n := T_n - T_{n-1}$ zwischen der Ankunft des $(n-1)$ -ten Kunden und des n -ten Kunden wird als Zwischenankunftszeit bezeichnet, wobei $T_0 = 0$ gesetzt wird. Von den Zufallsvariablen I_n , $n = 1, 2, \dots$, wird vorausgesetzt, dass sie stochastisch unabhängig und identisch verteilt sind mit der Verteilungsfunktion $F_I(x)$, dem Erwartungswert $\mathbf{E}[I]$ und der Varianz $\mathbf{Var}[I]$. Der Kehrwert

$$\lambda := \frac{1}{\mathbf{E}[I]}$$

heißt Ankunftsrate und gibt an, wieviele Kunden im Durchschnitt pro Zeiteinheit in das System einfallen.

Die Bedienungszeiten S_n , $n = 1, 2, \dots$, der aufeinanderfolgenden Kunden werden ebenfalls als stochastisch unabhängige und identisch verteilte Zufallsvariablen aufgefasst. Die Verteilungsfunktion der Bedienungszeiten wird mit $F_S(x)$ bezeichnet. Für den zugehörigen Erwartungswert und die zugehörige Varianz werden die Symbole $\mathbf{E}[S]$ und $\mathbf{Var}[S]$ verwendet. Der Kehrwert

$$\mu := \frac{1}{\mathbf{E}[S]}$$

heißt Bedienrate und gibt an, wieviele Kunden im Durchschnitt pro Zeiteinheit von einem Bedienungsgerät abgefertigt werden können. Sind mehrere parallele und gleichartige Bedienungsgeräte vorhanden, erhöht sich die Bedienungsrate entsprechend der Anzahl der Geräte.

Die Bedienungsregel legt fest, in welcher Reihenfolge die wartenden Kunden abgefertigt werden sollen. Folgende Regeln und Bezeichnungen sind gebräuchlich:

FIFO (FCFS) First In, First Out (First Come, First Served). Die Bedienung erfolgt in der Reihenfolge der Ankünfte.

LIFO (LCFS) Last In, First Out (Last Come, First Served). Die Bedienung erfolgt in umgekehrter Reihenfolge der Ankünfte.

SIRO Selection In Random Order. Der nächste Kunde wird zufällig ausgewählt.

Non-preemptive Priority relative Priorität. Manche Kunden werden gegenüber anderen Kunden vorrangig behandelt. Der laufende Bedienungsprozess wird jedoch nicht unterbrochen.

Preemptive Priorität absolute Priorität. Besitzt der neu ankommende Kunde gegenüber den anderen Kunden im System eine höhere Priorität, so wird der laufende Bedienungsprozess unterbrochen und mit der neuen Forderung fortgesetzt. Die alte Forderung wird zurückgestellt.

RR Round Robin. Jeder Kunde kann den Bediener jeweils nur für ein bestimmtes Zeitintervall in Anspruch nehmen. Kunden, deren Abfertigung mehr Zeit benötigt, müssen sich deshalb mehrmals hintereinander in die Warteschlange einreihen.

Zur symbolischen Kennzeichnung der Bedienungssysteme haben D.G. Kendall und B.W. Gnedenko die Notation

$$A/B/c/m$$

eingeführt. Die Buchstaben A und B markieren hierbei den Verteilungstyp der Zwischenankunfts- und Bedienungszeiten. Der Buchstabe c steht für die Anzahl der parallelen Bediener, m bezeichnet die Kapazität des Warteraums.

Für den Verteilungstyp sind folgende Abkürzungen gebräuchlich:

D	Deterministische Verteilung,
M	Exponentialverteilung, (M steht für die sogenannte Markov-Eigenschaft der Exponentialverteilung),
E_k	Erlang-Verteilung mit den Parameter k ($k = 1, 2, \dots$),
H_k	Hyperexponentialverteilung mit dem Parameter k ($k = 1, 2, \dots$),
PH	Phasen-Typ-Verteilung,
G	Allgemeine Verteilung.

20.30 Beispiel:

Die Notation $M/G/3/\infty$ z.B. kennzeichnet ein Bedienungssystem mit exponentialverteilten Zwischenankunftszeiten, beliebig verteilten Bedienungszeiten, drei parallelen Bedienern und unendlicher Warteraumkapazität.

20.31 Definition (Anzahl der Kunden im System, Verweilzeit, Durchlaufzeit):

Die Leistungsbewertung von Bedienungssystemen erfolgt auf der Basis folgender stochastischer Prozesse:

- Die Anzahl der Kunden im System $(N_t)_{t \geq 0}$.
Dieser Prozess gibt an, wie viele Kunden sich zur Zeit t im Bedienungssystem (einschließlich Schalter) aufhalten.
- Der Prozess der aufeinanderfolgenden Verweilzeiten (bzw. Durchlaufzeiten) $(V_n)_{n \in \mathbb{N}}$.
Die Zufallsvariable V_n bezeichnet die Zeit, die der n -te Kunde im Bedienungssystem verweilt.
- Der Prozess der aufeinanderfolgenden Wartezeiten $(W_n)_{n \in \mathbb{N}}$.
Die Zufallsvariable W_n bezeichnet die Zeit, die der n -te Kunde in der Warteschlange verbringt, bevor er bedient wird.

Simulation von Warteschlangensystemen

Für einfache Warteschlangenmodelle (insbesondere für $M/M/1/\infty$) existieren umfangreiche Theorien zur analytischen Bestimmung der zugehörigen Kenngrößen (siehe Stochastik IV). Aber bereits für das Warteschlangenmodell $G/G/1/\infty$ ist bis heute keine analytische Lösung bekannt, weshalb man auf stochastische Simulationen angewiesen ist. Die dafür erforderliche Technik soll exemplarisch am Prozess $(W_n)_{n \in \mathbb{N}}$ der aufeinanderfolgenden Wartezeiten erläutert werden.

Offensichtlich gilt $W_n = 0$ genau dann, wenn der n -te Kunde erst bei oder nach Verlassen des $(n-1)$ -ten Kunden in das System eintrifft. Unter Berücksichtigung der oben benutzten Bezeichnungen kann diese Situation mathematisch wie folgt charakterisiert werden:

$$\begin{aligned}
 W_n = 0 &\Leftrightarrow T_n \geq T_{n-1} + W_{n-1} + S_{n-1} \\
 &\Leftrightarrow I_n \geq W_{n-1} + S_{n-1} \quad \left(\text{da } T_n = \sum_{k=1}^n I_k \right) \\
 &\Leftrightarrow W_{n-1} \leq I_n - S_{n-1} \quad (n = 1, 2, \dots),
 \end{aligned}$$

wobei als Anfangsbedingung $W_0 := S_0 := 0$ gesetzt wird. Im Fall $T_n < T_{n-1} + W_{n-1} + S_{n-1}$ erhält man die Wartezeit des n -ten Kunden, indem man von der Verweilzeit V_{n-1} des $(n-1)$ -ten Kunden die Zwischenankunftszeit I_n , d.h. die Zeit zwischen der Ankunft des n -ten und $(n-1)$ -ten Kunden abzieht:

$$W_n = V_{n-1} - I_n = W_{n-1} + S_{n-1} - I_n \quad (n = 1, 2, \dots).$$

Fasst man beide Bedingungen zusammen, erhält man:

$$\begin{aligned}
 W_0 &= 0 \\
 W_n &= \max(W_{n-1} - I_n + S_{n-1}, 0) =: (W_{n-1} - I_n + S_{n-1})^+ \quad (n = 1, 2, \dots).
 \end{aligned}$$

Um die charakteristischen Größen von W_n wie z.B. $\mathbf{E}[W_n]$ zu ermitteln, wird das Prinzip der wiederholten Versuche angewandt, d.h. man erzeugt sich unabhängige Realisierungen

$$(I_{n,k})_{n \in \mathbb{N}, 1 \leq k \leq N} \quad \text{und} \quad (S_{n,k})_{n \in \mathbb{N}, 1 \leq k \leq N}$$

der aufeinanderfolgenden Zwischenankunfts- und Bedienzeiten und schätzt $\mathbf{E}[W_n]$ mit Hilfe des Stichprobenmittels

$$\bar{W}_n = \frac{1}{N} \sum_{k=1}^N W_{n,k},$$

wobei $W_{n,k} = (W_{n-1,k} - I_{n,k} + S_{n-1,k})^+$, $n \in \mathbb{N}, 1 \leq k \leq N$, und $W_{0,k} := S_{0,k} := 0$ für $k = 1, \dots, N$ ist. (Siehe auch MatLab-Programm zur Simulation von Warteschlangen.)

In der Praxis wird man feststellen, dass die Werte $W_{n,k}$ für große n vergleichsweise stark streuen. Für Spezialfälle (z.B. des Bediensystems $M/M/1/\infty$) kann bewiesen werden, dass $\varrho \rightarrow 1$ $\mathbf{Var}[W_n] \rightarrow \infty$ nach sich zieht, so dass man bei der Simulation Schwierigkeiten haben wird, $\mathbf{E}[W_n]$ hinreichend genau einzugrenzen. Deshalb stellt sich die Frage nach Techniken, mit denen die Varianz des Schätzers geeignet reduziert werden kann.

Varianzreduktion

Die Verfahren zur Varianzreduktion basieren auf Prinzipien der antithetischen (gegenläufigen) Variablen und der sogenannten Kontrollvariable.

Varianzreduktion mittels antithetischer Variablen

20.32 Satz:

Es seien X und Y zwei identisch verteilte Zufallsvariablen mit $\mathbf{Var}[X] < \infty$ und $\mathbf{Var}[Y] < \infty$. Dann gilt:

$$a) \mathbf{E}\left[\frac{1}{2}(X + Y)\right] = \mathbf{E}[X],$$

$$b) \mathbf{Var}\left[\frac{1}{2}(X + Y)\right] = \frac{1}{2}(\mathbf{Var}[X] + \mathbf{COV}[X, Y]).$$

Beweis:

$$a) \mathbf{E}\left[\frac{1}{2}(X + Y)\right] = \frac{1}{2}(\mathbf{E}[X + Y]) = \frac{1}{2}(\mathbf{E}[X] + \mathbf{E}[Y]) = \frac{1}{2} \cdot 2\mathbf{E}[X] = \mathbf{E}[X].$$

$$b) \mathbf{Var}\left[\frac{1}{2}(X + Y)\right] = \frac{1}{4} \mathbf{Var}[X + Y] = \frac{1}{4}(\mathbf{Var}[X] + \mathbf{Var}[Y] + 2\mathbf{COV}[X, Y]) = \frac{1}{4}(2\mathbf{Var}[X] + 2\mathbf{COV}[X, Y]) = \frac{1}{2}(\mathbf{Var}[X] + \mathbf{COV}[X, Y]). \blacksquare$$

Dieser Satz kann für die Simulation in folgender Weise genutzt werden:

Sind $T_n^{(1)}$ und $T_n^{(2)}$ zwei erwartungstreue Schätzer für $\varrho(\theta)$, die identisch verteilt und negativ korreliert sind, dann ist $T_n := \frac{1}{2}(T_n^{(1)} + T_n^{(2)})$ ebenfalls ein erwartungstreuer Schätzer für $\varrho(\theta)$, der aber eine geringere Varianz besitzt als $T_n^{(1)}$ bzw. $T_n^{(2)}$. Indem in Satz 20.32 X mit $T_n^{(1)}$ und Y mit $T_n^{(2)}$ identifiziert wird, erhält man

$$\mathbf{E}[T_n] = \mathbf{E}\left[\frac{1}{2}(T_n^{(1)} + T_n^{(2)})\right] = \mathbf{E}[T_n^{(1)}]$$

und

$$\mathbf{Var}[T_n] = \mathbf{Var}\left[\frac{1}{2}(T_n^{(1)} + T_n^{(2)})\right] = \frac{1}{2}[\mathbf{Var}[T_n^{(1)}] + \mathbf{COV}[T_n^{(1)}, T_n^{(2)}]] \leq \frac{1}{2} \mathbf{Var}[T_n^{(1)}].$$

Bei dieser Vorgehensweise wird also die Varianz um mindestens die Hälfte reduziert. Demgegenüber steht eine Verdopplung des Aufwandes bei der Stichprobenerhebung.

20.33 Satz:

Es bezeichne U eine $\mathcal{R}([0, 1])$ -verteilte Zufallsvariable. Dann gilt:

$$\mathbf{COV}[U, 1 - U] = -\frac{1}{12}.$$

Beweis:

Mit U ist auch $1 - U$ $\mathcal{R}([0, 1])$ -verteilt. Deshalb gilt:

$$\mathbf{E}[U] = \mathbf{E}[1 - U] = \int_0^1 x \, dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

und

$$\mathbf{E}[U^2] = \mathbf{E}[(1 - U)^2] = \int_0^1 x^2 \, dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}.$$

Damit wird

$$\begin{aligned} \mathbf{COV}[U, 1 - U] &= \mathbf{E}[U(1 - U)] - \mathbf{E}[U] \cdot \mathbf{E}[1 - U] \\ &= \mathbf{E}[U] - \mathbf{E}[U^2] - \mathbf{E}[U(1 - \mathbf{E}[U])] \\ &= (\mathbf{E}[U])^2 - \mathbf{E}[U^2] \\ &= \left(\frac{1}{2}\right)^2 - \frac{1}{3} = -\frac{1}{12}. \end{aligned} \blacksquare$$

Es seien nun F_I die Verteilungsfunktion der Zwischenankunftszeiten und F_S die Verteilungsfunktion der Bedienzeiten. Die Erkenntnisse aus den Sätzen 20.32 und 20.33 kann man sich im Zusammenhang mit der Simulation eines G/G/1/∞-Bediensystems in der folgenden Weise zunutze machen:

Es seien $(U_{n,k}^{(1)})_{n \in \mathbb{N}, 1 \leq k \leq N}$ und $(U_{n,k}^{(2)})_{n \in \mathbb{N}, 1 \leq k \leq N}$ Sequenzen stochastisch unabhängiger $\mathcal{R}([0, 1])$ -verteilter Zufallsvariablen. Hiermit korrespondieren die Folgen der aufeinanderfolgenden Zwischenankunftszeiten $I_{n,k} = F_I^{-1}(U_{n,k}^{(1)})$ und Bedienzeiten $S_{n,k} = F_S^{-1}(U_{n,k}^{(2)})$, $n \in \mathbb{N}$ und $1 \leq k \leq N$, wobei F_I die Verteilung der Zwischenankunftszeiten und F_S die Verteilung der Bedienzeiten bedeuten. Angesichts Satz 20.33 erzeugt man hierzu Folgen $I'_{n,k} := F_I^{-1}(1 - U_{n,k}^{(1)})$ sowie $S'_{n,k} := F_S^{-1}(1 - U_{n,k}^{(2)})$, $n \in \mathbb{N}$ und $1 \leq k \leq N$, und ersetzt das Stichprobenmittel \bar{W}_n durch den Schätzer

$$\widetilde{W}_n = \frac{1}{N} \sum_{k=1}^N \frac{W_{n,k} + W'_{n,k}}{2},$$

wobei

$$\begin{aligned} W_{n,k} &:= \left[W_{n-1,k} - F_I^{-1}(U_{n,k}^{(1)}) + F_S^{-1}(U_{n-1,k}^{(2)}) \right]^+, \\ W'_{n,k} &:= \left[W'_{n-1,k} - F_I^{-1}(1 - U_{n,k}^{(1)}) + F_S^{-1}(1 - U_{n-1,k}^{(2)}) \right]^+ \end{aligned}$$

ist. Man kann nun zeigen, dass sich die Gegenläufigkeiten der Variablen $U_{n,k}^{(1)}$ und $1 - U_{n,k}^{(1)}$ bzw. $U_{n,k}^{(2)}$ und $1 - U_{n,k}^{(2)}$ auf die Zufallsvariablen $W_{n,k}$ und $W'_{n,k}$ übertragen, so dass sich tatsächlich ein Effekt in der Form

$$\mathbf{Var}[\widetilde{W}_n] < \frac{1}{2} \mathbf{Var}[\bar{W}_n]$$

einstellt (vgl. Mitchell, B.: Various Reduction by Antithetic Variates in GI/G/1 Queueing Simulation, Oper. Res. 21, 1973, 988-997). Hierbei spielt die Monotonie von F_I und F_S bzw. F_I^{-1} und F_S^{-1} eine wesentliche Rolle.

Varianzreduktion mittels Kontrollvariablen

Das Prinzip der Kontrollvariablen nutzt folgenden Zusammenhang aus:

20.34 Satz:

Es seien X und Y Zufallsvariablen mit $\mathbf{Var}[X] < \infty$, $\mathbf{Var}[Y] < \infty$ und $c \in \mathbb{R}$. Dann gilt:

(i)

$$\begin{aligned} \mathbf{E}[X + c(Y - \mathbf{E}[Y])] &= \mathbf{E}[X] \\ \mathbf{Var}[X + c(Y - \mathbf{E}[Y])] &= \mathbf{Var}[X] + c^2 \mathbf{Var}[Y] + 2c \mathbf{COV}[X, Y]. \end{aligned}$$

(ii) Die Funktion $f: c \mapsto \mathbf{Var}[X + c(Y - \mathbf{E}[Y])]$ nimmt ihr Minimum an für

$$c = c^* := -\frac{\mathbf{COV}[X, Y]}{\mathbf{Var}[Y]}.$$

Beweis:

(i) Es gilt:

$$\begin{aligned}\mathbf{E}[X + c(Y - \mathbf{E}[Y])] &= \mathbf{E}[X] + c(\mathbf{E}[Y] - \mathbf{E}[Y]) = \mathbf{E}[X] \\ \mathbf{Var}[X + c(Y - \mathbf{E}[Y])] &= \mathbf{Var}[X] + \mathbf{Var}[c(Y - \mathbf{E}[Y])] + 2\mathbf{COV}[X, c(Y - \mathbf{E}[Y])] \\ &= \mathbf{Var}[X] + c^2 \mathbf{Var}[Y] + 2c\mathbf{COV}[X, Y].\end{aligned}$$

(ii) Es wird

$$f(c) = \mathbf{Var}[X] + c^2 \mathbf{Var}[Y] + 2c\mathbf{COV}[X, Y], \quad \forall c \in \mathbb{R}$$

gesetzt und differenziert

$$\frac{d}{dc}f(c) = 2c \cdot \mathbf{Var}[Y] + 2\mathbf{COV}[X, Y], \quad \forall c \in \mathbb{R}.$$

Die Bedingung $\frac{d}{dc}f(c) = 0$ führt auf

$$c = c^* := -\frac{\mathbf{COV}[X, Y]}{\mathbf{Var}[Y]}.$$

Und wegen $\frac{d^2}{dc^2}f(c) = 2 \mathbf{Var}[Y] \geq 0$ ist c^* Minimum von f . ■

In der Praxis kann dieses Prinzip wie folgt genutzt werden:

Ist $X := T_n$ ein erwartungstreuer Schätzer für $\varrho(\theta)$ und Y eine Zufallsvariable mit bekanntem Erwartungswert $\mathbf{E}[Y]$ und bekannter Varianz $\mathbf{Var}[Y]$, dann ist

$$\hat{T}_n := T_n + c^*(Y - \mathbf{E}[Y])$$

ebenfalls ein erwartungstreuer Schätzer für $\varrho(\theta)$, wobei

$$\mathbf{Var}[\hat{T}_n] \leq \mathbf{Var}[T_n]$$

gilt. Ein Nachteil der Methode ist, dass die Kovarianz $\mathbf{COV}[T_n, Y]$ nur empirisch anhand eines Verlaufs geschätzt werden kann.

20.6 Statistik regenerativer Prozesse

Die Theorie der regenerativen Prozesse lässt sich zur Auswertung von Simulationen nutzen. Aufgrund der Definition eines regenerativen Prozesses weiß man, dass die Zufallsvektoren

$$(Y_1, X_1), (Y_2, X_2), \dots, (Y_n, X_n), \dots$$

mit

$$Y_j := \int_{[S_{j-1}, S_j)} f(Z_s) ds \quad \text{bzw.} \quad Y_j := \sum_{n=S_{j-1}}^{S_j-1} f(Z_n)$$

und $X_j := S_j - S_{j-1}$, $j \in \mathbb{N}$, unabhängig und identisch verteilt sind.

Es soll das Problem der Schätzung von $r := \mathbf{E}[f(Z^*)] := \lim_{t \rightarrow \infty} \mathbf{E}[f(Z_t)]$ betrachtet werden. Aufgrund der Sätze 18.8 und 18.9 ist dies gleichbedeutend mit der Schätzung von $\mathbf{E}[Y_1]/\mathbf{E}[X_1]$. Ebenfalls interessant ist die Berechnung eines Konfidenzintervalles für unsere Schätzung \hat{r} , d.h. eines Intervalls, in dem der gesuchte Wert mit einer gewissen Wahrscheinlichkeit liegt.

Hierfür werden die Zufallsvariablen $V_j := Y_j - rX_j$, $j \in \mathbb{N}$, betrachtet. Die Folge $(V_j)_{j \in \mathbb{N}}$ ist eine i.i.d. Folge. Für V_j gilt insbesondere:

$$\mathbf{E}[V_j] = \mathbf{E}[Y_j] - r\mathbf{E}[X_j] = \mathbf{E}[Y_j] - \left(\frac{\mathbf{E}[Y_1]}{\mathbf{E}[X_1]} \right) \cdot \mathbf{E}[X_j] = 0.$$

Wir nehmen $0 < \sigma^2 = \mathbf{Var}[V_j] = \mathbf{E}[V_j^2] < \infty$ an und setzen

$$\bar{Y} := \frac{1}{n} \sum_{j=1}^n Y_j, \quad \bar{X} := \frac{1}{n} \sum_{j=1}^n X_j \quad \text{und} \quad \bar{V} := \frac{1}{n} \sum_{j=1}^n V_j.$$

Somit gilt

$$\bar{V} = \bar{Y} - r \cdot \bar{X} = \frac{1}{n} \sum_{j=1}^n (Y_j - r \cdot X_j).$$

Mit Hilfe des zentralen Grenzwertsatzes schließt man, dass für $x \in \mathbb{R}$ gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left(\frac{n \cdot \bar{V} - n \cdot 0}{\sigma \cdot \sqrt{n}} \leq x \right) &= \lim_{n \rightarrow \infty} P \left(\frac{\bar{V} \cdot \sqrt{n}}{\sigma} \leq x \right) \\ &= \lim_{n \rightarrow \infty} P \left(\frac{(\bar{Y} - r\bar{X}) \cdot \sqrt{n}}{\sigma} \leq x \right) = \lim_{n \rightarrow \infty} P \left(\frac{(\hat{r} - r) \cdot \sqrt{n}}{\sigma/\bar{X}} \leq x \right) = \Phi(x), \quad x \in \mathbb{R}, \end{aligned}$$

wobei Φ die Standardnormalverteilung bezeichnet und $\hat{r} := \frac{\bar{Y}}{\bar{X}}$ als Schätzer für r verwendet wurde.

Man geht nun davon aus, dass der Fehler standardnormalverteilt ist. Es sei nun $z_0^* := \Phi^{-1}(1 - \frac{\theta}{2})$ bzw. $\Phi(z_0^*) = 1 - \frac{\theta}{2}$ das Quantil der Ordnung $1 - \theta$ der Standardnormalverteilung. Damit gilt

$$\lim_{n \rightarrow \infty} P \left(-z_0^* \leq \frac{\sqrt{n}(\hat{r} - r)}{\sigma/\bar{X}} \leq z_0^* \right) = \lim_{n \rightarrow \infty} P \left(\hat{r} - \frac{z_0^* \sigma}{\bar{X} \sqrt{n}} \leq r \leq \hat{r} + \frac{z_0^* \sigma}{\bar{X} \sqrt{n}} \right) = 1 - \theta.$$

Hieraus ergibt sich das Konfidenzintervall zum Niveau $1 - \theta$

$$r := \hat{r} \pm \frac{z_0^* \cdot \sigma}{\bar{X} \sqrt{n}}.$$

Das Problem hierbei ist, dass die Streuung σ unbekannt ist. Sie lässt sich aber wie folgt schätzen:

$$\sigma^2 = \mathbf{E}[(Y_1 - rX_1)^2] = \mathbf{Var}[Y_1] - 2 \cdot r \cdot \mathbf{COV}[Y_1, X_1] + r^2 \cdot \mathbf{Var}[X_1] \approx s_{11} - 2\hat{r}s_{12} + \hat{r}^2 s_{22},$$

wobei

$$\begin{aligned}
 s_{11} &= \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y})^2 = \frac{1}{n-1} \sum_{j=1}^n Y_j^2 - \frac{1}{n(n-1)} \left(\sum_{j=1}^n Y_j \right)^2 \\
 s_{22} &= \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2 = \frac{1}{n-1} \sum_{j=1}^n X_j^2 - \frac{1}{n(n-1)} \left(\sum_{j=1}^n X_j \right)^2 \\
 s_{12} &= \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y})(X_j - \bar{X}) = \frac{1}{n-1} \sum_{j=1}^n Y_j X_j - \frac{1}{n(n-1)} \left(\sum_{j=1}^n Y_j \right) \left(\sum_{j=1}^n X_j \right).
 \end{aligned}$$

Zusammenfassung:

- Man beobachtet n Regenerationszyklen.
- Daraus lassen sich für jeden Zyklus Y_j und X_j berechnen.
- Aus Y_j und X_j können \bar{Y} , \bar{X} , \hat{r} , s_{11} , s_{12} , s_{22} und s^2 bestimmt werden.
- Man berechnet $r := \hat{r} \pm \frac{z_0^* \cdot s}{X \sqrt{n}}$.

20.35 Beispiel:

Mittlere Wartezeit in einem $G/G/1/\infty$ -System:

1. Stichprobe:

$w_1 := 0$	$w_{11} := 0$	$w_{21} := 10$	$w_{31} := 16$
$w_2 := 10$	$w_{12} := 12$	$w_{22} := 8$	$w_{32} := 4$
$w_3 := 8$	$w_{13} := 9$	$w_{23} := 14$	$w_{33} := 0$
$w_4 := 15$	$w_{14} := 18$	$w_{24} := 14$	$w_{34} := 8$
$w_5 := 17$	$w_{15} := 18$	$w_{25} := 13$	$w_{35} := 18$
$w_6 := 13$	$w_{16} := 10$	$w_{26} := 10$	$w_{36} := 12$
$w_7 := 18$	$w_{17} := 0$	$w_{27} := 8$	$w_{37} := 18$
$w_8 := 12$	$w_{18} := 16$	$w_{28} := 12$	$w_{38} := 19$
$w_9 := 2$	$w_{19} := 0$	$w_{29} := 13$	$w_{39} := 8$
$w_{10} := 5$	$w_{20} := 5$	$w_{30} := 23$	$w_{40} := 4$
			$w_{41} := 0$

2. Es gibt $n := 5$ Zyklen.

$$\begin{aligned}
 X_1 &= 10, & Y_1 &:= \sum_{i=1}^{10} w_i = 100 \\
 X_2 &= 6, & Y_2 &:= \sum_{i=11}^{16} w_i = 67 \\
 X_3 &= 2, & Y_3 &:= \sum_{i=17}^{18} w_i = 16 \\
 X_4 &= 14, & Y_4 &:= \sum_{i=19}^{32} w_i = 150 \\
 X_5 &= 8, & Y_5 &:= \sum_{i=33}^{40} w_i = 87
 \end{aligned}$$

3. Berechnung weiterer Größen:

$$\begin{aligned}
 \bar{Y} &:= \frac{1}{5} \cdot \sum_{i=1}^5 Y_i = 84 \\
 \bar{X} &:= \frac{1}{5} \cdot \sum_{i=1}^5 X_i = 8 \\
 \hat{r} &:= \bar{Y}/\bar{X} = 10.5 \\
 s_{11} &:= \frac{1}{4} \cdot \sum_{i=1}^5 Y_i^2 - \frac{1}{20} \left(\sum_{i=1}^5 Y_i \right)^2 = 2383.5 \\
 s_{12} &= 217.5 \\
 s_{22} &= 20 \\
 s^2 &= 21
 \end{aligned}$$

4. Konfidenzintervall zum Niveau 0.9:

$$I = 10.5 \pm \frac{z_0^* \cdot \sqrt{21}}{8 \cdot \sqrt{5}} = 10.5 \pm 0.4214.$$

Literatur zu Kapitel 20

Folgende Bücher werden als begleitende Literatur zu diesem Kapitel des Skriptes empfohlen:

- A.O. ALLEN:
Probability, Statistics and Queueing Theory,
 Academic Press, London, 2nd. Edition, 1990.
 ISBN: 0120510510
- U.N. BHAT:
Sixty years of queueing theory,
 Management Science 15,280-294, 1969.

- E. BROCKEMEYER, H.L. HALSTROM, A. JENSEN:
The Life and Works of A.K. Erlang,
Acta Polystechnica Scandinavia, Mathematics and Computing Machinery Series, 6,
1960.
- M. A. CRANE, A. J. LEMOINE:
An Introduction to the Regenerative Method for Simulation Analysis,
Lecture notes in Control and Information Science 4Springer-Verlag, Berlin, 1977.
ISBN: 3540084088
- G. S. FISHMAN:
Monte-Carlo: Concepts, Algorithms and Applications,
Springer, New York, 1996.
ISBN: 038794527X
- D. GROSS/C.M. HARRIS:
Fundamentals of Queueing Theory,
John Wiley and Sons, New York, 2nd. Edition, 1986.
ISBN: 0471170836
- L. KLEINROCK:
Queueing Systems, Volume 1,
John Wiley and Sons, 1975.
- L. KLEINROCK:
Queueing Systems, Volume 2,
John Wiley and Sons, 1976.
- S.S. LAVENBERG:
Computer Performance Modeling Handbook,
Academic Press, San Diego, 1983.
ISBN: 0124387209
- A. M. LAW, W.D. KELTON:
Simulation, Modelling and Analysis,
Mc Graw Hill, New York, 1982.
ISBN: 0070366969
- R. Y. RUBINSTEIN:
Simulation and The monte Carlo Method,
John Wiley & Sons, New York, 1981.
ISBN: 0471089176
- T.L. SAATY:
Elements of Queueing Theory with Applications,
Mc Graw Hill, New York, 1961.
ISBN: 0486645533
- P. TRAN-GIA:
Analytische Leistungsbewertung verteilter Systeme,

Springer-Verlag, Berlin, 1996.
ISBN: 3540606661

Anhang A

Alternative Definition der Laplace-Transformierten

Oftmals wird die Laplace-Transformation auch wie folgt definiert:

$$\hat{g}(s) := \int_{\mathbb{R}^+} e^{-st} g(t) dt.$$

Die zugehörige Faltung lautet dann:

$$(g_1 \hat{*} g_2)(t) := \begin{cases} \int_{[0,t]} g_1(t-s) g_2(s) ds & , t \geq 0 \\ 0 & , t < 0. \end{cases}$$

Die Sätze 17.13 und 17.14 gelten für diese Definition ebenfalls:

Analog zu Satz 17.13 gilt: Für $g_1, g_2 \in \mathbb{B}$ gilt $g_1 \hat{*} g_2 \in \mathbb{B}$, denn:

$$|(g_1 \hat{*} g_2)(t)| \leq \int_0^t |g(t-s)| |g(s)| ds \leq \sup_{0 \leq s \leq t} |g_1(s)| \cdot \sup_{0 \leq s \leq t} |g_2(s)| \cdot t < \infty.$$

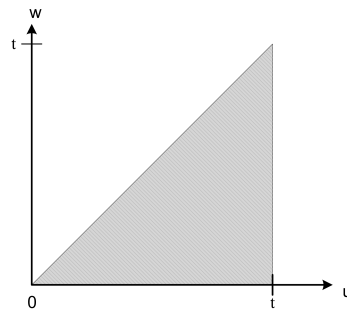
Mit den Voraussetzungen $F, F_1, F_2, F_3, g, g_1, g_2 \in \mathbb{B}$ und $c \in \mathbb{R}$ gilt Satz 17.14 ebenfalls für die Laplace-Transformation \hat{f} und die Faltung $\hat{*}$. Wie oben sind nur die Punkte a), f) und g) zu zeigen (alle anderen folgen wie oben direkt aus Eigenschaften des Integrals):

a) Es gilt:

$$\begin{aligned} ((f_1 \hat{*} f_2) \hat{*} f_3)(t) &= \int_0^t (f_1 \hat{*} f_2)(t-s) f_3(s) ds \\ &\stackrel{(*)}{=} \int_t^0 (f_1 \hat{*} f_2)(u) f_3(t-u) (-1) du \\ &= \int_0^t \int_0^u f_1(u-v) f_2(v) dv f_3(t-u) du \\ &\stackrel{(**)}{=} \int_0^t \int_u^0 f_1(w) f_2(u-w) (-1) dw f_3(t-u) du \end{aligned}$$

$$\begin{aligned}
 &= \int_0^t \int_0^u f_1(w) f_2(u-w) f_3(t-u) \, dw \, du \\
 &\stackrel{(***)}{=} \int_0^t \int_w^t f_1(w) f_2(u-w) f_3(t-u) \, du \, dw \\
 &\stackrel{****}{=} \int_0^t f_1(w) \int_{t-w}^0 f_2(t-w-x) f_3(x) (-1) \, dx \, dw \\
 &= \int_0^t f_1(w) (f_2 \hat{*} f_3)(t-w) \, dw \\
 &= (f_1 \hat{*} (f_2 \hat{*} f_3))(t).
 \end{aligned}$$

(*): Substitution von s durch $u := t - s$, (**): Substitution von v durch $w := u - v$,
 (***): Die Integration verläuft über folgenden Bereich:



(****): Substitution von u durch $x := t - u$.

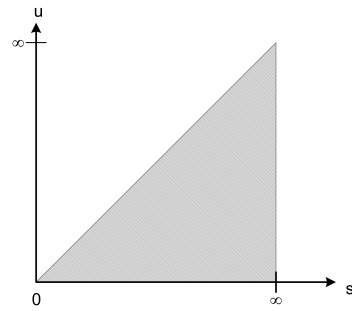
f) Mit der Substitution $u := t - s$ und $\frac{du}{ds} = -1$ gilt:

$$\begin{aligned}
 (g_1 \hat{*} g_2)(t) &= \int_0^t g_1(t-s) g_2(s) \, ds = \int_t^0 g_1(u) g_2(t-u) (-1) \, du \\
 &= \int_0^t g_1(u) g_2(t-u) \, du = (g_2 \hat{*} g_1)(t).
 \end{aligned}$$

g) Es gilt:

$$\begin{aligned}
 \widehat{f \hat{*} g} &= \int_0^\infty e^{-st} (f \hat{*} g)(s) \, ds \\
 &= \int_0^\infty \int_0^s f(s-u) g(u) e^{-st} \, du \, ds \\
 &\stackrel{(*)}{=} \int_0^\infty \int_u^\infty f(s-u) g(u) e^{-st} \, ds \, du \\
 &= \int_0^\infty \int_u^\infty f(s-u) e^{-t(s-u)} g(u) e^{-tu} \, ds \, du \\
 &\stackrel{(**)}{=} \int_0^\infty \int_0^\infty f(v) e^{-tv} g(u) e^{-tu} \, dv \, du \\
 &= \int_0^\infty f(v) e^{-tv} \, dv \cdot \int_0^\infty g(u) e^{-tu} \, du \\
 &= \hat{f}(t) \cdot \hat{g}(t).
 \end{aligned}$$

(*) : Die Integration verläuft über folgenden Bereich:



(**): Substitution von s durch $v := s - u$.

Anhang B

Tabelle der χ^2 -Verteilung

$F(x)$	Anzahl der Freiheitsgrade									
	1	2	3	4	5	6	7	8	9	10
0.001	0.00	0.00	0.02	0.09	0.21	0.38	0.60	0.86	1.15	1.48
0.005	0.00	0.01	0.07	0.21	0.41	0.68	0.99	1.34	1.73	2.16
0.01	0.00	0.02	0.11	0.30	0.55	0.87	1.24	1.65	2.09	2.56
0.025	0.00	0.05	0.22	0.48	0.83	1.24	1.69	2.18	2.70	3.25
0.05	0.00	0.10	0.35	0.71	1.15	1.64	2.17	2.73	3.33	3.94
0.1	0.02	0.21	0.58	1.06	1.61	2.20	2.83	3.49	4.17	4.87
0.25	0.10	0.58	1.21	1.92	2.67	3.45	4.25	5.07	5.90	6.74
0.5	0.45	1.39	2.37	3.36	4.35	5.35	6.35	7.34	8.34	9.34
0.75	1.32	2.77	4.11	5.39	6.63	7.84	9.04	10.22	11.39	12.55
0.9	2.71	4.61	6.25	7.78	9.24	10.64	12.02	13.36	14.68	15.99
0.95	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31
0.975	5.02	7.38	9.35	11.14	12.83	14.45	16.01	17.53	19.02	20.48
0.99	6.63	9.21	11.35	13.28	15.09	16.81	18.48	20.09	21.67	23.21
0.995	7.88	10.69	12.84	14.86	16.75	18.55	20.28	21.96	23.59	25.19
0.999	10.83	13.82	16.27	18.47	20.52	22.46	24.32	26.13	27.88	29.59

$F(x)$	Anzahl der Freiheitsgrade									
	11	12	13	14	15	16	17	18	19	20
0.001	1.83	2.21	2.62	3.04	3.48	3.94	4.42	4.90	5.41	5.92
0.005	2.60	3.07	3.57	4.07	4.60	5.14	5.70	6.26	6.84	7.43
0.01	3.05	3.57	4.11	4.66	5.23	5.81	6.41	7.01	7.63	8.26
0.025	3.82	4.40	5.01	5.63	6.26	6.91	7.56	8.23	8.91	9.59
0.05	4.57	5.23	5.89	6.57	7.26	7.96	8.67	9.39	10.12	10.85
0.1	5.58	6.30	7.04	7.79	8.55	9.31	10.09	10.86	11.65	12.44
0.25	7.58	8.44	9.30	10.17	11.04	11.91	12.79	13.68	14.56	15.45
0.5	10.34	11.34	12.34	13.34	14.34	15.34	16.34	17.34	18.34	19.34
0.75	13.70	14.85	15.98	17.12	18.25	19.37	20.49	21.60	22.72	23.83
0.9	17.28	18.55	19.81	21.06	22.31	23.54	24.77	25.99	27.20	28.41
0.95	19.68	21.03	22.36	23.68	25.00	26.30	27.59	28.87	30.14	31.41
0.975	21.92	23.34	24.74	26.12	27.49	28.85	30.19	31.53	32.85	34.17
0.99	24.73	26.22	27.69	29.14	30.58	32.00	33.41	34.81	36.19	37.57
0.995	26.76	28.30	29.82	31.32	32.80	34.27	35.72	37.16	38.58	40.00
0.999	31.26	32.91	34.53	36.12	37.70	39.25	40.79	42.31	43.82	45.32

Beispiel: Bei 3 Freiheitsgraden ist $F = 0.99$ für $x = 11.35$.

$F(x)$	Anzahl der Freiheitsgrade									
	21	22	23	24	25	26	27	28	29	30
0,001	6,4	7,0	7,5	8,1	8,7	9,2	9,8	10,4	11,0	11,6
0,005	8,0	8,6	9,3	9,9	10,5	11,2	11,8	12,5	13,1	13,8
0,01	8,9	9,5	10,2	10,9	11,5	12,2	12,9	13,6	14,3	15,0
0,025	10,3	11,0	11,7	12,4	13,1	13,8	14,6	15,3	16,0	16,8
0,05	11,6	12,3	13,1	13,8	14,6	15,4	16,2	16,9	17,7	18,5
0,1	13,2	14,0	14,8	15,7	16,5	17,3	18,1	18,9	19,8	20,6
0,25	6,3	17,2	18,1	19,0	19,9	20,8	21,7	22,7	23,6	24,5
0,5	20,3	21,3	22,3	23,3	24,3	25,3	26,3	27,3	28,3	29,3
0,75	24,9	26,0	27,1	28,2	29,3	30,4	31,5	32,6	33,7	34,8
0,9	29,6	30,8	32,0	33,2	34,4	35,6	36,7	37,9	39,1	40,3
0,95	32,7	33,9	35,2	36,4	37,7	38,9	40,1	41,3	42,6	43,8
0,975	35,5	36,8	38,1	39,4	40,6	41,9	43,2	44,5	45,7	47,0
0,99	38,9	40,3	41,6	43,0	44,3	45,6	47,0	48,3	49,6	50,9
0,995	41,4	42,8	44,2	45,6	46,9	48,3	49,6	51,0	52,3	53,7
0,999	46,8	48,3	49,7	51,2	52,6	54,1	55,5	56,9	58,3	59,7

$F(x)$	Anzahl der Freiheitsgrade							
	40	50	60	70	80	90	100	> 100 (Näherung)
0,001	17,9	24,7	31,7	39,0	46,5	54,2	61,9	$(h - 3,09)^2/2$
0,005	20,7	28,0	35,5	43,3	51,2	59,2	67,3	$(h - 2,58)^2/2$
0,01	22,2	29,7	37,5	45,4	53,5	61,8	70,1	$(h - 2,33)^2/2$
0,025	24,4	32,4	40,5	48,8	57,2	65,6	74,2	$(h - 1,96)^2/2$
0,05	26,5	34,8	43,2	51,7	60,4	69,1	77,9	$(h - 1,64)^2/2$
0,1	29,1	37,7	46,5	55,3	64,3	73,3	82,4	$(h - 1,28)^2/2$
0,25	33,7	42,9	52,3	61,7	71,1	80,6	90,1	$(h - 0,67)^2/2$
0,5	39,3	49,3	59,3	69,3	79,3	89,3	99,3	$h^2/2$
0,75	45,6	56,3	67,0	77,6	88,1	98,6	109,1	$(h + 0,67)^2/2$
0,9	51,8	63,2	74,4	85,5	96,6	107,6	118,5	$(h + 1,28)^2/2$
0,95	55,8	67,5	79,1	90,5	101,9	113,1	124,3	$(h + 1,64)^2/2$
0,975	59,3	71,4	83,3	95,0	106,6	118,1	129,6	$(h + 1,96)^2/2$
0,99	63,7	76,2	88,4	100,4	112,3	124,1	135,8	$(h + 2,33)^2/2$
0,995	66,8	79,5	92,0	104,2	116,3	128,3	140,2	$(h + 2,58)^2/2$
0,999	73,4	86,7	99,6	112,3	124,8	137,2	149,4	$(h + 3,09)^2/2$

In der letzten Spalte ist $h = \sqrt{2m - 1}$
 $(m = \text{Anzahl der Freiheitsgrade})$

Anhang C

Tabelle der Standardnormalverteilung

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{x^2}{2}} dx$$

z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9777	0.9783	0.9788	0.9793	0.9798	0.9803	0.9807	0.9812	0.9816
	499	844	083	217	248	178	007	738	372	911
2.1	0.9821	0.9825	0.9829	0.9834	0.9838	0.9842	0.9846	0.9849	0.9853	0.9857
	356	708	970	142	226	224	137	966	713	379
2.2	0.9860	0.9864	0.9867	0.9871	0.9874	0.9877	0.9880	0.9883	0.9886	0.9889
	966	474	906	263	545	755	894	962	962	893
2.3	0.9892	0.9895	0.9898	0.9900	0.9903	0.9906	0.9908	0.9911	0.9913	0.9915
	759	559	296	969	581	133	625	060	437	758
2.4	0.9918	0.9920	0.9922	0.9924	0.9926	0.9928	0.9930	0.9932	0.9934	0.9936
	025	237	397	506	564	572	531	443	309	128
2.5	0.9937	0.9939	0.9941	0.9942	0.9944	0.9946	0.9947	0.9949	0.9950	0.9952
	903	634	323	969	574	139	664	151	600	012
2.6	0.9953	0.9954	0.9956	0.9957	0.9958	0.9959	0.9960	0.9962	0.9963	0.9964
	388	729	035	308	547	754	930	074	189	274
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9971	0.9972	0.9973
	330	358	359	333	280	202	099	972	821	646
2.8	0.9974	0.9975	0.9975	0.9976	0.9977	0.9978	0.9978	0.9979	0.9980	0.9980
	449	229	988	726	443	140	818	476	116	738
2.9	0.9981	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986
	342	929	498	052	589	111	618	110	588	051

Beispiel: $\Phi(2.01) = 0.9777844$.

Anhang D

Quantile der Standardnormalverteilung

$$u_p = \Phi^{-1}(p)$$

p	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.50	0.0000	0.0251	0.0502	0.0753	0.1004	0.1257	0.1510	0.1764	0.2019	0.2275
0.60	0.2533	0.2793	0.3055	0.3319	0.3585	0.3853	0.4125	0.4399	0.4677	0.4959
0.70	0.5244	0.5534	0.5828	0.6128	0.6433	0.6745	0.7063	0.7388	0.7722	0.8064
p	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.800	0.8416	0.8452	0.8488	0.8524	0.8560	0.8596	0.8633	0.8669	0.8705	0.8742
0.810	0.8779	0.8816	0.8853	0.8890	0.8927	0.8965	0.9002	0.9040	0.9078	0.9116
0.820	0.9154	0.9192	0.9230	0.9269	0.9307	0.9346	0.9385	0.9424	0.9463	0.9502
0.830	0.9542	0.9581	0.9621	0.9661	0.9701	0.9741	0.9782	0.9822	0.9863	0.9904
0.840	0.9945	0.9986	1.0027	1.0069	1.0110	1.0152	1.0194	1.0237	1.0279	1.0322
0.850	1.0364	1.0407	1.0450	1.0494	1.0537	1.0581	1.0625	1.0669	1.0714	1.0758
0.860	1.0803	1.0848	1.0893	1.0939	1.0985	1.1031	1.1077	1.1123	1.1170	1.1217
0.870	1.1264	1.1311	1.1359	1.1407	1.1455	1.1503	1.1552	1.1601	1.1650	1.1700
0.880	1.1750	1.1800	1.1850	1.1901	1.1952	1.2004	1.2055	1.2107	1.2160	1.2212
0.890	1.2265	1.2319	1.2372	1.2426	1.2481	1.2536	1.2591	1.2646	1.2702	1.2759
0.900	1.2816	1.2873	1.2930	1.2988	1.3047	1.3106	1.3165	1.3225	1.3285	1.3346
0.910	1.3408	1.3469	1.3532	1.3595	1.3658	1.3722	1.3787	1.3852	1.3917	1.3984
0.920	1.4051	1.4118	1.4187	1.4255	1.4325	1.4395	1.4466	1.4538	1.4611	1.4684
0.930	1.4758	1.4833	1.4909	1.4985	1.5063	1.5141	1.5220	1.5301	1.5382	1.5464
0.940	1.5548	1.5632	1.5718	1.5805	1.5893	1.5982	1.6072	1.6164	1.6258	1.6352
0.950	1.6449	1.6546	1.6646	1.6747	1.6849	1.6954	1.7060	1.7169	1.7279	1.7392
0.960	1.7507	1.7624	1.7744	1.7866	1.7991	1.8119	1.8250	1.8384	1.8522	1.8663
0.970	1.8808	1.8957	1.9110	1.9268	1.9431	1.9600	1.9774	1.9954	2.0141	2.0335
0.980	2.0537	2.0749	2.0969	2.1201	2.1444	2.1701	2.1973	2.2262	2.2571	2.2904
0.990	2.3263	2.3656	2.4089	2.4573	2.5121	2.5758	2.6521	2.7478	2.8782	3.0902
p	0.0000	0.0001	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009
0.9990	3.0902	3.1214	3.1559	3.1947	3.2389	3.2905	3.3528	3.4316	3.5401	3.7190

Für p -Werte mit $0 < p < 0.5$ gilt: $u_p = -u_{1-p}$

Anhang E

Zeichenerklärungen

\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{Q}	Menge der rationalen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{M}	Menge der maßdefinierenden Funktionen auf \mathbb{R} , die in $(-\infty, 0)$ verschwinden
\mathbb{B}	Menge der nichtnegativen reellen Funktionen, die auf jedem Intervall der Form $[0, t]$ beschränkt sind
\mathbb{I}^n	Menge der links offenen und rechts abgeschlossenen Intervalle im \mathbb{R}^n , $n \in \mathbb{N}$
\mathfrak{B}^n	$:= \sigma(\mathbb{I}^n)$ “ σ -Algebra der Borelschen Mengen des \mathbb{R}^n ”
$\overline{\mathbb{R}}$	$:= \mathbb{R} \cup \{-\infty, +\infty\}$
\mathfrak{B}	$:= \{B, B \cup \{-\infty\}, B \cup \{+\infty\}, B \cup \{-\infty, +\infty\} B \in \mathfrak{B}\}$
$\mathfrak{P}(M)$	Potenzmenge von M
$(a, b]$	$:= \{x a < x \leq b\}$ “links offenes, rechts abgeschlossenes Intervall”
$n!$	$:= n \cdot \dots \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$ “Fakultät von n ”
$(N)_n$	$:= \frac{N!}{n!} = N \cdot (N-1) \cdot \dots \cdot (N-n+1)$ “ n -te untere Faktorielle von N ”
$\binom{n}{k}$	$:= \frac{n!}{k! \cdot (n-k)!}$ “ n über k ”
$F(a-0)$	meint den linksseitigen Limes von $F(a)$
\uparrow	konvergiert von unten gegen
$X \stackrel{d}{=} \text{Exp}(\lambda)$	X ist exponential-verteilt
$X \stackrel{d}{=} Y$	X und Y sind identisch verteilt
$\text{Re}(x)$	Realteil der komplexen Zahl x
$\text{Im}(x)$	Imaginärteil der komplexen Zahl x
$O(n), o(n)$	seien die Landau-Symbole.

Anhang F

Literatur

Stochastik I

Folgende Bücher werden als begleitende Literatur zum Teil I des Skriptes empfohlen:

- H. BAUER:
Maß- und Integrationstheorie,
Walter de Gruyter, Berlin, 1990.
ISBN: 3110127725
Preis: 26.95 €
Kurzbeschreibung:
„Viele Gebiete der Mathematik und ihrer Anwendungen [...] erfordern solide Kenntnisse aus der Maß- und Integrationstheorie. Das Lehrbuch [...] führt den Leser [...] schnell, verlässlich und präzise zu den wichtigsten Ergebnissen der Maß- und Integrationstheorie hin. [...] Zahlreiche Beispiele erläutern die Bedeutung der erzielten Ergebnisse.[...] Übungsaufgaben laden den Leser zum vertieften Eindringen in den behandelten Stoff ein.“
- H. BAUER:
Wahrscheinlichkeitstheorie,
5. Auflage, Walter de Gruyter, Berlin, 2002.
ISBN: 3110172364
Preis: 36.95 €
Kurzbeschreibung:
„Das vorliegende Buch soll dem Studierenden als Wegführer in die Wahrscheinlichkeitstheorie dienen. Der Leser soll dabei mit den wichtigsten Ideen, Methoden und Resultaten dieser sich heute schnell entwickelnden und verzweigenden mathematischen Theorie bekanntgemacht werden. [...] Da heutzutage die Wahrscheinlichkeitstheorie unlöslich mit der Maß- und Integrationstheorie verbunden ist, verfolgt das Buch zugleich aber auch ein zweites Ziel, nämlich den Leser mit den Grundzügen der Maßtheorie vertraut zu machen. [...]“
- W. BEHNEN, G. NEUHAUS:
Grundkurs Stochastik,
3. Auflage, Teubner-Verlag, Stuttgart, 1995.

ISBN: 3930737698

Preis: 24.00 €

Kurzbeschreibung:

„Eine integrierte Einführung in die Wahrscheinlichkeitstheorie und Mathematische Statistik für Mathematiker, Wirtschaftsmathematiker, Informatiker und Physiker.

Es enthält: Wahrscheinlichkeitsmodelle, Anwendungspostulat und statistische Tests; mehrstufige Zufallsexperimente und grundlegende diskrete Modelle; Wahrscheinlichkeitsmodelle über euklidischen Räumen; Koppelung von allgemeinen Zufallsexperimenten (Satz von Fubini); Parameterschätzung (auch in approximativen Modellen); Konfidenzbereiche für Modellparameter; das Testen von Hypothesen.”

- M. A. BERGER:

An Introduction to Probability and Stochastic Processes,

Springer-Verlag, New York, 1992.

ISBN: 3540977848

Kurzbeschreibung:

„This is a textbook which will provide students with a straightforward introduction to the mathematical theory of probability. It is written with the aim of presenting the central results and techniques of the subject in a complete and self-contained account.

(...) Any student who has a familiarity with calculus and basic algebra will be able to use this text and throughout there are a wide variety of exercises to illustrate and to develop ideas. [...]”

- O. BEYER, H. HACKEL, V. PIEPER, J. TIEDGE:

Wahrscheinlichkeitsrechnung und mathematische Statistik,

7. Auflage, Teubner-Verlag, Stuttgart, 1995.

ISBN: 3-8154-2075-X

Kurzbeschreibung:

„Die Reihe „Mathematik für Ingenieure, Naturwissenschaftler, Ökonomen und Landwirte“ umfasst den [...] Lehrstoff für die Mathematikausbildung der genannten Disziplinen, bietet Möglichkeiten zur Vertiefung sowie Spezialisierung und unterstützt die Individualisierung des Studiums. [...] Das Lehrwerk ist nach modernen fachlichen und hochschulpädagogischen Prinzipien aufgebaut. [...]”

- P. BILLINGSLEY:

Probability and Measure,

2nd edition, John Wiley and Sons, New York, 1986.

ISBN: 0471007102

Preis: 102.90 €

Kurzbeschreibung:

„Intertwines measure theory and modern probability: probability problems generate an interest in measure theory and measure theory is then developed and applied to probability. Illustrates the connections probability theory has with applied mathematics on the one hand and with pure mathematics on the other.”

- M. FISZ:

Wahrscheinlichkeitsrechnung und mathematische Statistik,

VEB, Deutscher Verlag der Wissenschaften 1989.

ISBN: 3326000790

Kurzbeschreibung:

„Dieses Buch ist in der Hauptsache für Mathematiker bestimmt; es dürfte aber auch [...] solchen Lesern zugänglich sein, die [...] über gewisse Kenntnisse in der höheren Mathematik verfügen und sich für die Anwendungen der Wahrscheinlichkeitsrechnung interessieren. Der Leser findet in diesem Buch eine Einführung in die moderne Wahrscheinlichkeitsrechnung und die moderne mathematische Statistik. [...] Das Buch enthält zahlreiche Anwendungsbeispiele. [...]”

- P. GÄNSSLER UND W. STUTE:

Wahrscheinlichkeitstheorie,

Springer-Verlag, Berlin, 1977.

ISBN: 3540084185

Kurzbeschreibung:

„Für das Verständnis des vorliegenden Textes sind [...] Grundkenntnisse aus einer Vorlesung „Einführung in die Wahrscheinlichkeitstheorie und Mathematische Statistik” wünschenswert. [...] Auf eine Diskussion diskreter Modelle ist deshalb bewusst verzichtet worden. Die [...] getroffene Stoffauswahl umfasst eine zweisemestrige Vorlesung über Wahrscheinlichkeitstheorie. Neben der Vermittlung klassischer Grundlagen liegt der methodische Schwerpunkt auf der Konstruktion stochastischer Modelle unter besonderer Berücksichtigung einiger für die Anwendungen in der Mathematischen Statistik wichtigen Resultate. [...] ”

- M. GREINER/G. TINHOFFER:

Stochastik für Studienanfänger der Informatik,

Hanser, München, 1996.

ISBN: 3-446-18636-0

Kurzbeschreibung:

„Dieses Lehrbuch bietet einen Grundstock an Lehrstoff aus Wahrscheinlichkeitstheorie und Statistik, wie er in der Informatik benötigt wird und verbindet diesen Lehrstoff mit der Begriffswelt, die Informatiker in ihrem Berufsalltag vorfinden. Hierbei wird besonderes Gewicht auf die Aspekte Methodik und Modellierung gelegt. Der Leser soll [...] zukünftig in der Lage sein, Fragen aus seinem Berufsalltag in ein geeignetes stochastisches Modell umzusetzen und die ermittelten Resultate anschließend im Rahmen der ursprünglichen Fragestellung zu interpretieren.[...] Mehr als hundert Beispiele, Aufgaben und deren Lösungen sowie ein Kompromiss zwischen mathematischer Strenge und ausgewogener textlicher Darstellung des Stoffes motivieren den Leser zur aktiven Teilnahme an der Entwicklung und Lösung von Problemen aus der Stochastik.”

- E. HENZE:

Einführung in die Maßtheorie,

Bibl. Institut, Mannheim, 1971.

ISBN: 341100505X

Kurzbeschreibung:

„Bei dieser Einführung in die Maß- und Integrationstheorie werden gleichzeitig die notwendigen Ergebnisse und Methoden für den Einstieg in die moderne Wahrscheinlich-

keitstheorie bereitgestellt. Das Buch wendet sich in erster Linie an Studenten der Mathematik, der Informatik und der Physik, kann aber auch anderen interessierten Lesern von Nutzen sein.”

- E. HENZE:

Stochastik für Einsteiger,
Bibl. Institut, Mannheim, 1997.
ISBN: 3528368942

Kurzbeschreibung:

„[...] Dieses Buch soll dem Leser einen Einstieg in die Stochastik, die Kunst des „geschickten Vermutens“, vermitteln und ihn in die Lage versetzen [...] kritisch und kompetent mitreden zu können. Es enthält 160 Übungsaufgaben mit Lösungen. [...] Als Lehrbuch zwischen gymnasialem Mathematikunterricht und Universität wendet es sich unter anderem an: [...] Studienanfänger an Universitäten, Fachhochschulen und Berufsakademien; Quereinsteiger aus Industrie und Wirtschaft.”

- H. HEUSER:

Lehrbuch der Analysis. Teil 2 Mathematische Leitfäden,
Teubner, Stuttgart, 2002.
ISBN: 3519522322

Kurzbeschreibung:

„Bei der Abfassung des zweiten Bandes [...] wollte ich die Theorie ausführlich und fasslich darstellen, ausgiebig motivieren und durch viele Beispiele und Übungen zum sicheren Besitz des Lesers machen. Außerdem wollte ich Brücken schlagen zu den Anwendungen analytischer Methoden in den allerverschiedensten Wissenschaften. [...] Dabei stehen diesmal im Vordergrund der Überlegungen Funktionen, deren Argumente und Werte Vektoren aus dem \mathbb{R}^p oder sogar Elemente aus noch viel allgemeineren Räumen sind. [...]”

- K. HINDERER:

Grundbegriffe der Wahrscheinlichkeitstheorie,
Springer-Verlag, 1980.
ISBN: 3540073094

Kurzbeschreibung:

„Das Buch [...] bietet eine solide, gut motivierte Darstellung mit einer Fülle konkreter Beispiele, ergänzt durch sorgfältig ausgesuchte Aufgaben nach jedem Paragraphen. Sowohl die historischen als auch die weiterführenden Bemerkungen geben eine gute Übersicht über Probleme und Fragestellungen aus der Wahrscheinlichkeitstheorie.”

- G. HÜBNER:

Stochastik. Eine Einführung für Mathematiker, Informatiker und Ingenieure.,
4. Auflage, Vieweg Verlag, 2003.
ISBN: 3528254432
Preis: 22.50 €

Kurzbeschreibung:

„Dieses Buch soll Informatiker, Ingenieure und Mathematiker in die Lage versetzen, konkrete Vorgänge mit Zufallseinfluss in den wesentlichen Aspekten zu verstehen, zu

modellieren und daraus Prognosen und Entscheidungshilfen abzuleiten. [...] Das Buch [...] richtet sich [...] an [...] Informatiker, Ingenieure, Mathematiker und Mathematik-Lehrer, die sich grundlegende Kenntnisse in stochastischer Modellierung und erste Einblicke in Anwendungsbereiche verschaffen wollen. [...] Besonders auf die Belange der Informatik zugeschnitten ist die Einbeziehung von Modellen und Bewertungen für Bedienungsprobleme und Kommunikationsnetze auf elementarem Niveau. [...]"

- U. KRENGEL:

Einführung in die Wahrscheinlichkeitstheorie und Statistik,

6. Auflage, Vieweg, 2002.

ISBN: 3528672595

Preis: 22.90 €

Kurzbeschreibung:

„Dieses Buch wendet sich an alle, die [...] in die Ideenwelt der Stochastik eindringen möchten. Stochastik ist die Mathematik des Zufalls. [...] Die beiden Hauptgebiete der Stochastik sind Wahrscheinlichkeitstheorie und Statistik. In der Wahrscheinlichkeitstheorie untersucht man zufällige Prozesse mit festen als bekannt angenommenen steuernden Wahrscheinlichkeiten. [...] Darüber hinaus liefert die Wahrscheinlichkeitstheorie Grundlagen für die Statistik, in der aus beobachteten Daten Schlüsse über unbekannte Wahrscheinlichkeiten und über zweckmäßiges Verhalten gezogen werden sollen. [...]"

- K. KRICKEBERG/H. ZIEZOLD:

Stochastische Methoden,

4. Auflage, Springer-Verlag, Berlin, 1995.

ISBN: 3-540-57792-0

Kurzbeschreibung:

„Im Vordergrund [...] stehen die eigentlichen „stochastischen“ Ideen und ihre praktischen Anwendungen, insbesondere in der Statistik, ohne dass mathematische strenge und Schönheit zu kurz kommen. Über die üblichen Grundlagen hinaus finden sich Kapitel über Simulation, nichtparametrische Statistik und Regression- und Varianzanalyse. [...] Besonderer Anziehungspunkt dieses Buches ist die „genetische“ Entwicklung der verschiedenen Typen von Wahrscheinlichkeitsverteilungen, ausgehend von der hypergeometrischen Verteilung. [...]"

- J. LEHN/H. WEGMANN:

Einführung in die Statistik,

4. Auflage, Teubner, 2004.

ISBN: 3519320711

Preis: 22.90 €

Kurzbeschreibung:

„Eine elementare Darstellung statistischer Schätz- und Testverfahren einschließlich der zugrundeliegenden Modellbildung für Mathematiker, Informatiker, Wirtschaftswissenschaftler, Naturwissenschaftler und Ingenieure.

Es enthält: Methoden der Beschreibenden Statistik; Zufallsvariablen und ihre Verteilungen; Gesetze der Großen Zahlen und ihre Eigenschaften; Tests bei Normalverteilungsannahmen; χ^2 -Tests und Kontingenztafeln; verteilungsunabhängige Tests; einfache Varianzanalyse und Regression.“

- R. MATHAR/D. PFEIFFER:
Stochastik für Informatiker,
Teubner, 1990.
ISBN: 3519022400

Kurzbeschreibung:

„Das vorliegende Buch [...] wendet sich vor allem an Informatikstudenten und Mathematikstudenten mit Nebenfach Informatik mit dem Ziel, stochastische Grundbegriffe unter besonderer Berücksichtigung Informatik-spezifischer Aspekte zu vermitteln. [...] Ziel des Buches ist es daher, eine einheitliche und möglichst geschlossene Übersicht über die zum Verständnis benötigten Grundlagen zu geben. [...] Trotz des überwiegenden Lehrbuchcharakters dieses Textes haben wir uns allerdings auch bemüht, neuere Entwicklungen, die z.T. bisher nur in Originalarbeiten vorliegen, mit einzubeziehen, um dort, wo es im Rahmen unseres Zugangs möglich ist, Anschluss an Fragestellungen der aktuellen Forschung zu erlangen. [...]“

- J. PFANZAGL:
Elementare Wahrscheinlichkeitsrechnung,
Gruyter, Berlin, 1988.
ISBN: 3110114194

Kurzbeschreibung:

„Die vorliegende Einführung der Wahrscheinlichkeitsrechnung ist „elementar“ in dem Sinne, dass weder Kenntnisse aus der Maßtheorie noch aus der Funktionentheorie vorausgesetzt werden. [...] Das Anliegen des Buches ist die Entwicklung anwendungsbezogenen stochastischen Denkens. Diesem Ziel dient eine verhältnismäßig große Anzahl von Beispielen, die [...] zeigen sollen, dass es sich bei der Wahrscheinlichkeitsrechnung um ein Teilgebiet der Mathematik handelt, das durch Anwendungen immer wieder neue Facetten erhält. [...]“

- P. P. SPIES:
Grundlagen stochastischer Modelle,
Hanser, München, 1982.
ISBN: 3446137114

Literatur speziell zu Kapitel 4

- J. BANKS:
Principles of Quality Control,
John Wiley and Sons, New York, 1989.
ISBN: 0471635510
- D.C. MONTGOMERY:
Introduction to Statistical Quality Control,
2nd edition, John Wiley and Sons, New York, 1991.
ISBN: 0471656313

Kurzbeschreibung:

„This book is about the use of modern statistical methods for quality control and improvement. It provides comprehensive coverage of the subject from basic principles to

state-of-art concepts and applications. The objective is to give the reader a sound understanding of the principles and the basis for applying them in a variety of both product and nonproduct situations. While statistical techniques are emphasized throughout, the book has a strong engineering and management orientation. [...] By presenting theory, and supporting the theory with clear and relevant examples, Montgomery helps the reader to understand the big picture of important concepts. [...]"

- H. RINNE UND H.-J. MITTAG:
Statistische Methoden der Qualitätssicherung,
3. Auflage, Carl Hanser Verlag, München, 1995.
ISBN: 3446180060

Kurzbeschreibung:

„Dieses Buch ist bewusst anwendungsorientiert geschrieben und zeichnet sich durch eine besonders sorgfältige didaktische Gestaltung aus. Es enthält: Zahlreiche Abbildungen und Fotos; mehr als 100 Übungsaufgaben mit ausführlichen Lösungen; viele durchgerechnete Anwendungsbeispiele; verständnisfördernde, zusammenfassende Übersichten; kommentierte Literaturangaben.“

- W. UHLMANN:
Statistische Qualitätskontrolle,
2. Auflage, Teubner-Verlag, Stuttgart, 1982.
ISBN: 3519123061

Kurzbeschreibung:

„Ein Lehrbuch für Statistiker, Mathematiker, Ingenieure und Wirtschaftswissenschaftler. Es enthält: Wahrscheinlichkeitstheoretische Grundlagen; statistische Grundlagen; Eingangs- und Endkontrolle; kostenoptimale Prüfpläne; sequentielle Tests; Kontrollkarten; Kosten und Kontrollabstand; kontinuierliche Stichprobenpläne.“

Literatur speziell zu Kapitel 8

- F. BEICHELT:
Zuverlässigkeits- und Instandhaltungstheorie,
Teubner, Stuttgart, 1993.
ISBN: 3519029855

Kurzbeschreibung:

„Das Buch ist eine moderne Einführung in die Zuverlässigkeits- und Instandhaltungstheorie auf der Grundlage stochastischer Modellbildung.[...] Zahlreiche numerische Beispiele erleichtern das inhaltliche Verständnis. Das Buch wendet sich an Praktiker und Studierende mathematisch-naturwissenschaftlich-technischer Fachrichtungen. [...]"

- K.-W. GAEDE:
Zuverlässigkeit — Mathematische Modelle,
Hanser, München, 1977.
ISBN: 3446123709

Kurzbeschreibung:

„Dieses Buch behandelt in leicht verständlicher und mathematisch sauberer Form die

Problemstellungen der Zuverlässigkeitstheorie und die zu ihrer Lösung nötigen mathematischen Modelle. Die erforderlichen Begriffe und die ausführlich beschriebenen Lösungsverfahren werden an praxisorientierten Beispielen erläutert. [...] Damit ist das Buch [...] auch zum Selbststudium für alle Ingenieure, Informatiker, Physiker und Mathematiker geeignet, die in der Praxis Zuverlässigkeitsprobleme zu bearbeiten haben.”

- P. GESSNER/H. WACKER:

Dynamische Optimierung,

Hansa, München, 1972.

ISBN: 3446115390

Kurzbeschreibung:

„Dieses Buch wendet sich an den Praktiker. Es versucht im Bereich der dynamischen Optimierung ähnlich wie bei der linearen Optimierung das Modelldenken einzuführen. [...] Alle jene Optimierungsprobleme der Unternehmensforschung werden behandelt, bei denen es darum geht, zeitabhängige Prozesse optimal zu steuern oder zu beeinflussen. [...] Hierzu werden viele realistische Beispiele diskutiert. [...] Für den an der Theorie weniger interessierten Praktiker sind für jedes Modell die Rechenschritte für sich verständlich aufgelistet. [...]”

- K. NEUMANN:

Dynamische Optimierung,

Bibliographisches Institut AG, Mannheim, 1969.

ISBN: 3411007141

Kurzbeschreibung:

„Das vorliegende Taschenbuch ist aus einer Vorlesung [...] entstanden. Der Stoff der Vorlesung ist durch Anwendungsbeispiele aus den verschiedensten Gebieten (Lagerhaltung, chemische Reaktortechnik, Regelungstechnik, Zuverlässigkeitstheorie, Produktions- und Wirtschaftssteuerung u.a.) [...] ergänzt worden. [...]”

Stochastik II

Die nachfolgenden Bücher eignen sich zum Nacharbeiten des Teil II:

- A.O. ALLEN:

Probability, Statistics and Queueing Theory,

Academic Press, London, 2nd. Edition, 1990.

ISBN: 0120510510

- G. ALSMEYER:

Erneuerungstheorie,

Teubner-Verlag, Stuttgart, 1991.

ISBN: 3519027305

- G. AUMANN/O. HAUPT:

Einführung in die reelle Analysis, Band I,

de Gruyter, Bln., 1981.

ISBN: 3110019701

- F. BEICHELT/P. FRANKEN:
Zuverlässigkeit und Instandhaltung: Mathematische Methoden,
Hanser-Verlag, München, 1984.
ISBN: 3446139060
- U.N. BHAT:
Sixty years of queueing theory,
Management Science 15,280-294, 1969.
- E. BROCKMEYER, H.L. HALSTROM, A. JENSEN:
The Life and Works of A.K. Erlang,
Acta Polytechnica Scandinavia, Mathematics and Computing Machinery Series, 6,
1960.
- J.A. BUZACOTT/J.G. SHANTIKUMAR:
Stochastic Models of Manufacturing Systems,
Prentice Hall, 1993.
ISBN: 0138475679
Preis: 90.90 €
- Y.S. CHOW/H. TEICHER:
Probability Theory - Independence, Interchangeability, Martingales,
Springer-Verlag, New York, 1988.
ISBN: 0387406077
- E. CINLAR:
Introduction to stochastic processes,
Prentice-Hall, 1975.
- J. W. COHEN:
The single server queue,
North Holland Publishing Company, Amsterdam, 1969.
- G.H. FICHTENHOLZ:
Differential- und Integralrechnung II,
VEB Deutscher Verlag der Wissenschaften, Berlin, 1974.
ISBN: 3817112793
Preis: 34.80 €
- D. GROSS/C.M. HARRIS:
Fundamentals of Queueing Theory,
John Wiley and Sons, New York, 2nd. Edition, 1986.
ISBN: 0471170836
Preis: 108.50 €
- D. L. IGLEHART/G. S. SHEDLER:
An Introduction to the Regenerative Method for Simulation Analysis,
Springer-Verlag, Berlin, 1977.
ISBN: 3540084088

- D. L. IGLEHART/G. S. SHEDLER:
Regenerative Simulation of Response Times in Networks of Queues,
Springer-Verlag, Berlin, 1980.
ISBN: 3540099425
- S. KARLIN/H.M. TAYLOR:
A first course in stochastic processes,
Academic Press, 1975.
Preis: 83.50 €
- L. KLEINROCK:
Queueing Systems, Volume 1,
John Wiley and Sons, 1975.
- L. KLEINROCK:
Queueing Systems, Volume 2,
John Wiley and Sons, 1976.
- J. KOHLAS:
Stochastische Methoden des Operations Research,
Teubner, 1977.
ISBN: 3519023423
- M. KOLONKO:
Stochastik I, Skriptum zur Vorlesung,
TU Clausthal, 2005.
- M. KOLONKO:
Stochastik II, Skriptum zur Vorlesung,
TU Clausthal, 2006.
- S.S. LAVENBERG:
Computer Performance Modeling Handbook,
Academic Press, San Diego, 1983.
ISBN: 0124387209
- S. I. RESNICK:
Adventures in Stochastic Processes,
Birkhäuser, Boston, 1992.
ISBN: 0817635912
Preis: 65.50 €
- T.L. SAATY:
Elements of Queueing Theory with Applications,
Mc Graw Hill, New York, 1961.
ISBN: 0486645533
- P. TRAN-GIA:
Analytische Leistungsbewertung verteilter Systeme,

Springer-Verlag, Berlin, 1996.

ISBN: 3540606661

Kurzbeschreibung:

„Dieses Buch vermittelt die gängigen Methoden der Modellbildung und Analyse verteilter Systeme, die in der Leistungsbewertung von Rechner- und Kommunikationssystemen sowie von Fertigungssystemen angewendet werden. Zunächst werden die Grundlagen der Wahrscheinlichkeitsrechnung, der stochastischen Prozesse, der Markov- und der Erneuerungsprozesse behandelt, die zum Verständnis der analytischen Leistungsbewertungsverfahren erforderlich sind. Grundmodelle der klassischen Nachrichtenverkehrstheorie und des Operations Research werden eingehend beschrieben, wobei Modellierungs- und Anwendungsaspekte im Vordergrund stehen. Abschließend werden moderne Analysemethoden vorgestellt, z.B. zeitdiskrete Analyseverfahren und Algorithmen sowie die Klasse der matrixanalytischen Methoden.“

Anhang G

Historie

In der folgenden Auflistung werden einige für die Entwicklung der Stochastik bedeutende Mathematiker mit ihren Lebensdaten und kurzen Beschreibungen ihrer Wirkungsfelder aufgeführt. Die Liste erhebt keinen Anspruch auf Vollständigkeit. Insbesondere fehlen wichtige Mathematiker, die nicht direkt im Bereich der Stochastik gewirkt haben, durch ihre Arbeiten zur Maß- oder Integrationstheorie die moderne Stochastik aber erst ermöglichten.

- **Thomas Bayes**

(* 1702 in London, England; † 17. April 1761 in Tunbridge Wells, Kent, England)

Thomas Bayes studierte ab 1719 Theologie an der Universität in Edinburgh und beschäftigte sich nebenbei mit Mathematik. 1733 wurde er Pfarrer der presbyterianischen Kapelle in Tunbridge Wells, 35 Meilen südöstlich von London. 1742 ernannte man Bayes zum Mitglied der Royal Society, obwohl der bis zu diesem Zeitpunkt noch keinerlei mathematische Arbeiten veröffentlicht hatte. Insgesamt publizierte Bayes selbst nur 2 Arbeiten. Seine wichtigsten Forschungsergebnisse, die unter anderem auch den später als „Formel von Bayes“ benannten Satz enthielten, wurden erst aus seinem Nachlass bekannt.

- **Richard Ernest Bellman**

(* 26. August 1920 in New York; † 19. März 1984 in Santa Monica, Californien)

Bellman studierte bis 1943 am Brooklyn College (B.A.) sowie an der University of Wisconsin (M.A.) Mathematik. Danach arbeitete er 2 Jahre in Los Alamos in der theoretischen Physik. Seit 1965 lehrte er an der Universität von Southern California, Los Angeles, als Professor für Mathematik, Elektroingenieurwesen und Medizin.

1953 stellte Bellman die Methode der dynamischen Programmierung auf, die für die Entscheidungstheorie sowie für die Variationsrechnung und optimalen Steuerung wesentlich ist. Bellman beschäftigte sich auch mit der Modellierung biologischer Prozesse und der Theorie der unscharfen Mengen.

- **Familie Bernoulli**

Die schweizer Gelehrtenfamilie Bernoulli hat über mehrere Generationen hinweg sehr große Beiträge zur Mathematik geleistet.

Im Folgenden werden nur die beiden Mitglieder der Bernoulli-Familie aufgeführt, die sich wesentlich mit stochastischen Fragestellungen beschäftigt haben. Insbesondere werden Jakob Bernoulli II, Johann Bernoulli I, Johann Bernoulli II, Johann Bernoulli III,

Niklaus Bernoulli I und Niklaus Bernoulli II, die allesamt bedeutende Mathematiker waren, hier nicht näher erwähnt.

– **Daniel Bernoulli**

(* 8. Februar 1700 in Groningen; † 17. März 1782 in Basel)

Daniel Bernoulli interessierte sich hauptsächlich für Anwendungen der Mathematik. Er entwickelte das Prinzip zur Lösung algebraischer Gleichungen mit Hilfe von rekurrenten Reihen („Methode von Bernoulli“) und untersuchte Kettenbrüche. Außerdem lieferte er wichtige Beiträge zur Wahrscheinlichkeitstheorie, die später teilweise von Laplace in seine Theorie aufgenommen wurden.

– **Jakob Bernoulli I**

(* 27. Dezember 1654 in Basel; † 16. August 1705 in Basel)

Jakob Bernoulli I ist der erste Gelehrte in der Familie der Bernoullis und überhaupt der erste bekannte Schweizer Mathematiker. Er befasste sich überwiegend mit analytischen Fragestellungen (er stand u.a. mit Leibniz, der gerade eine Infinitesimalmethoden aufgestellt hatte, in Kontakt), sowie mit stochastischen Problemen. Seine Arbeit baute auf den Ergebnissen von Huygens über das Glücksspiel auf. In einer, erst nach seinem Tode durch seinen Neffen Niklaus Bernoulli I veröffentlichten Arbeit, stellte Jakob Bernoulli I bereits das Gesetz der großen Zahlen auf und verallgemeinerte viele kombinatorische Ansätze von Huygens.

• **Emile Borel**

(* 7. Januar 1871 Saint-Affrique; † 3. Februar 1956 in Paris)

Borel beschäftigte sich zunächst mit Funktionentheorie. Nach seiner Tätigkeit als Forschungsbeirat im Kriegsministerium von 1914–1918 übernahm er den Lehrstuhl für Wahrscheinlichkeitsrechnung und mathematische Physik an der Sorbonne. Während seiner Arbeit in der Funktionentheorie prägte Borel den Begriff des Maßes und der überabzählbaren Überdeckung. Ab 1905 befasste sich Borel mit den Nutzungsmöglichkeiten seiner Maßtheorie in der Wahrscheinlichkeitstheorie. Außerdem ist Borel Mitbegründer der Spieltheorie und bewies das Minimax-Theorem für 3 Spieler.

• **Guido Fubini**

(* 19. Januar 1879 in Venedig; † 6. Juni 1943 in New York)

Zu den wichtigsten Arbeiten Fubinis gehört der 1907 von ihm bewiesene und später nach ihm benannte Satz. Darüber hinaus befasste sich Fubini mit projektiver Differentialgeometrie sowie der Theorie diskontinuierlicher Gruppen und automorpher Funktionen.

• **Andrej Nikolajewitsch Kolmogorov**

(* 25. April 1903 in Tambow; 20. Oktober 1987 in Moskau)

Kolmogorov gilt als einer der bedeutendsten Mathematiker der Gegenwart. Er befasste sich vorwiegend mit Wahrscheinlichkeitstheorie, mathematischer Statistik und Logik, Maß- und Integrationstheorie, Funktionalanalysis sowie Informations- und Algorithmentheorie. Nebenbei entwarf er Lehrpläne und Schulbücher für den Mathematikunterricht und prägte so zu großen Teilen den Mathematikunterricht in der Sowjetunion.

Mit seiner Arbeit „Grundbegriffe der Wahrscheinlichkeitsrechnung“ von 1933 löste er das 6. Problem der berühmten 23 von Hilbert gestellten mathematischen Probleme.

- **Pierre Simon Marquis de Laplace**

(* 28. März 1749 in Beaumont-en-Auge; † 5. März 1827 in Paris)

Laplace befasste sich sehr viel mit partiellen Differential- und Differenzengleichungen. Seine Entwicklung der Laplace-Transformation diente ihm dazu, Naturerscheinungen analytisch zu erfassen. Neben vielen Arbeiten zu physikalischen Themen befasste er sich mit Themen der Wahrscheinlichkeitsrechnung. Seine 1812 erschienene Theorie stellte eine umfassende Darstellung der damals bekannten Wahrscheinlichkeitstheorie dar. In ihr wurde der Begriff der Wahrscheinlichkeit definiert, sowie die mathematische Erwartung erörtert. Zudem greift Laplace in seiner Arbeit das, von J. Bernoulli gefundene Gesetz der großen Zahlen auf.

Auf Laplace geht auch die Idee zurück, dass das Geschehen in einem physikalischen System exakt vorherbestimmbar sei, wenn nur alle Anfangszustände bekannt sind („Laplacescher Determinismus“).

Laplace war neben seiner Tätigkeit als Forscher ab 1794 Vorsitzender der Kommission für Maße und Gewichte und unter Napoleon Bonaparte Minister des Inneren.

- **Henri Lebesgue**

(* 28. Juni 1875 in Beauvais (Frankreich); † 26. Juli 1941 in Paris)

Lebesgue erkannte, dass viele zu seiner Zeit gültigen Theorien für eine Reihe von Fragestellungen unzureichend waren. 1902 verallgemeinerte er den Riemannschen Integralbegriff zu dem wesentlich leistungsfähigeren Lebesgueschen Integral. Lebesgues Resultate wurden zunächst nur zögernd aufgenommen, stellen heute aber die Grundlage für die moderne Analysis dar.

- **Andrej Andrejewitch Markov**

(* 14. Juni 1856 in Gouvernement Rjasan; † 20. Juli 1922 in Petrograd)

Markov studierte von 1874–1878 unter anderem bei Tschebyscheff und beschäftigte sich zunächst hauptsächlich mit Fragestellungen der Zahlen- und Funktionentheorie. Später befasste er sich überwiegend mit Wahrscheinlichkeitsrechnung. Dabei legte er wichtige Grundlagen zur Entwicklung der Theorie der stochastischen Prozesse. Außerdem entwickelte Markov die Theorie der später nach ihm benannten Markovschen Prozesse bzw. Ketten.

- **Pafnuti Lwowitch Tschebyscheff**

(* 16. Mai 1821 in Okatowo; † 8. Dezember 1894 in Petersburg)

Tschebyscheff befasste sich zunächst überwiegend mit Zahlentheorie. Unter anderem wirkte er an der Herausgabe der zahlentheoretischen Manuskripte Eulers mit. Später beschäftigte er sich dann überwiegend mit wahrscheinlichkeitstheoretischen Fragestellungen. Insbesondere erarbeitete er die Gesetzmäßigkeiten von Summen unabhängiger Summanden. Er verdeutlichte die Wichtigkeit solcher Begriffe wie Zufallsgröße oder Erwartungswert, verallgemeinerte das Gesetz der großen Zahlen und vereinfachte dessen Beweis erheblich.

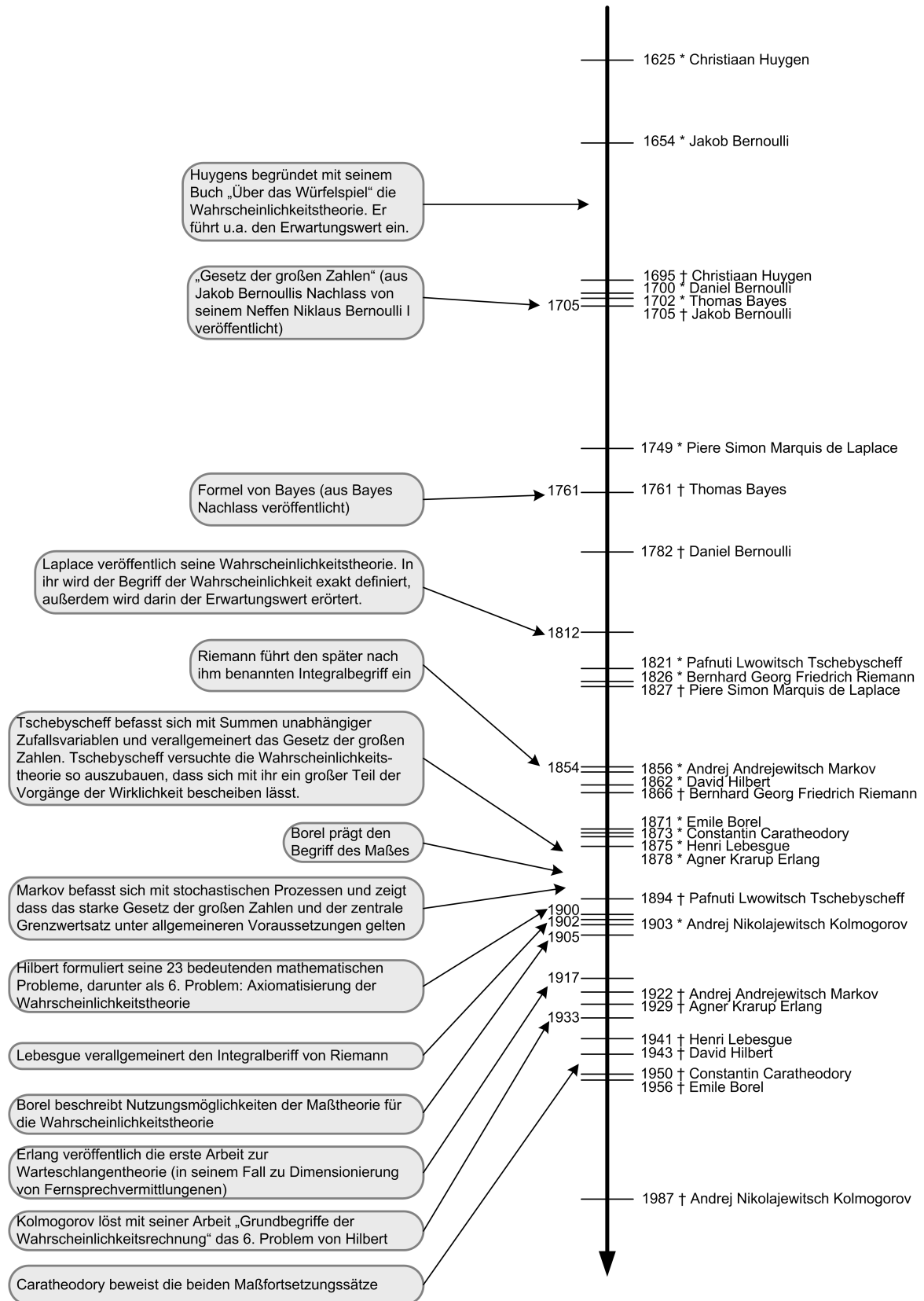
- **Bernhard Georg Friedrich Riemann**

(* 17. September 1826 in Breselenz bei Dannenberg; † 20. Juli 1866 in Salses in Italien)

Riemann studierte ab 1846 an der Universität in Göttingen zunächst Theologie und Philosophie, wechselte dann aber bald zur Mathematik. In seiner 1854 vorgelegten Habilitationsschrift führte Riemann das später nach ihm benannte Riemann-Integral ein. Neben der Integrationstheorie befasste er sich mit vielen weiteren mathematischen Gebieten. So forschte Riemann u.a. auf dem Gebiet der partiellen Differentialgleichungen, sowie in der Zahlentheorie und der nichteuklidischen Geometrie. Die Ideen Riemanns sind bis heute von großer Bedeutung: Die Riemannsche Hypothese über die Nullstellen der ζ -Funktion wird in sehr vielen Sätzen der Zahlentheorie verwendet. Beweisen lies sich die Riemannsche Hypothese allerdings bis heute nicht.

Weitere Informationen und diverse Biographien finden sich unter:

- S. GOTTWALD, H.-J. ILGAUDS, K.-H. SCHLOTE:
Lexikon bedeutender Mathematiker,
Verlag Harri Deutsch, Thun, 1990.
ISBN: 3-8171-1164-9
- Turnbell Server, Biographies
<http://www-groups.dcs.st-and.ac.uk/>
- Mathematik.ch: Bedeutende Mathematiker
<http://www.mathematik.ch/mathematiker/>
- Wikipedia (Kategorie: Mathematiker)
<http://de.wikipedia.org/wiki/Kategorie:Mathematiker>



Stichwortverzeichnis

- P -integrierbar, 8
- α -Quantil, 65
- Überlagerungsprozess, 140
- Überlebensfunktion, 156

- Ablehnungsbereich, 70
- absolut-stetig, 79
- Alter, 110
- Alternative, 69
- Ankunftsrate, 195
- Annahmebereich, 70
- Anzahl der Kunden im System, 196
- aperiodisch, 142
- arithmetisch verteilt, 118
- asymptotisch erwartungstreu, 60
- aufgespanntes Gitter, 182
- Ausfallrate, 156

- Bayes, Thomas, 231
- Bedienrate, 195
- Bedienungsregel, 195
- Bedienungsschalter, 193
- Bedienungszeiten, 195
- bedingte Überlebenswahrscheinlichkeit, 157
- bedingte Dichte, 88
- bedingte Erwartung, 76
- bedingte Restlebensdauer, 157
- bedingte Verteilung, 88
- bedingter Erwartungswert, 76, 96
- Bellman, Richard Ernest, 231
- Bernoulli, Daniel, 232
- Bernoulli, Jakob I, 232
- Bernoulliverteilung, 11
- Beyer-Quotient, 184
- Bias, 54
- Borel, Emile, 232
- Borelraum, 85

- charakteristische Funktion, 10
- Cramér-Rao-Schranke, 58

- Decreasing Failure Rate, 157
- definit
 - positiv-semi-, 13
- Diskrepanz
 - n -dimensionale, 176
- Dreiecksverteilung, 12
- Durchlaufzeiten, 196

- Einpunktverteilung, 11
- einseitiges Testproblem, 69
- Erneuerungsargument, 108
- Erneuerungsfolge, 100
- Erneuerungsfunktion, 102
- Erneuerungsgleichung, 108
- Erneuerungsprozess, 100, 101
 - modifiziert, 137
 - stationär, 138
- Erneuerungszeitpunkt, 100
- Erwartung
 - bedingt, 79
- erwartungstreu, 54
- erwartungstreuer Schätzer mit minimaler Varianz, 56
- Erwartungswert, 8

- Faltung, 105
- Filtration, 146
- Fisher-Information, 58
- Fourier-Transformierte, 10
- Fubini, Guido, 232

- Gütefunktion, 70
- gemischter Kongruenzgenerator, 174
- Gitterbasis, 182
- Gitterkonstante, 118
- gleichmäßig besser, 66

- Havarieerneuerung, 158
- Havarieinstandsetzung, 156
- Hit or Miss Methode, 191

- homogen, 140
- Hypothese, 69
- Hypothesentest, 70
- Increasing Failure Rate, 157
- inhomogen, 140
- Kennfunktion, 66
- Kolmogorov, Andrej Nikolajewitsch, 232
- Konfidenzintervall, 65, 201
- Konfidenzoeffizient, 65
- Konfidenzniveau, 65
- Konvergenz in Verteilung, 26, 33
- kritischen Bereich, 70
- kumulative Ausfallrate, 161
- kumulativen Kosten des i -ten Zyklus, 158
- Kunde, 194
- Länge des i -ten Zyklus, 158
- Laplace, Pierre Simon Marquis de, 233
- Laplace-Transformierte, 100
- Lebensdauer, 100, 156
- Lebensdauerfunktion, 156
- Lebesgue, Henri, 233
- Likelihood-Funktion, 62
- lineare Kongruenzmethode, 174
- Markov, Andrej Andrejewitsch, 233
- Maximum-Likelihood-Schätzer, 62
- minimale Reparatur, 158
- Minkowski-reduzierte Basis, 184
- Mittelwertmethode, 191
- multiplikativer Kongruenzgenerator, 174
- natürliches Gitter, 182
- Operationscharakteristik, 70
- Periodenlänge, 176
- periodisch, 142
- Poissonprozess, 140
- Poissonverteilung, 11
- positiv-semidefinit, 13
- prophylaktisch, 158
- prophylaktischer Instandsetzung, 156
- Pseudo-Inverse, 185
- Punktschätzer, 53
- Quantil, 65
- Rückwärtsrekurrenzzeit, 110
- Raummittel, 151
- Regenerationspunkte, 146
- Regenerationszyklus, 146, 202
- regenerativer Prozess, 146
- rekurrent, 142
- Restlebenszeit, 110
- Riemann, Bernhard Georg Friedrich, 233
- Riemann-integrierbar
 - direkt, 126
- Sample Mean Methode, 191
- Satz
 - von Bochner, 14
- Schätzer, 53, 201
 - gleichmäßig bester, 56
- Schätzwerte, 53
- Schalter, 194
- schwach konsistent, 53
- schwache Konvergenz, 26, 27, 33
- semidefinit
 - positiv-, 13
- stark konsistent, 53
- statistisches Modell, 52
- Stichprobe vom Umfang n , 52
- Stichprobenkovarianz, 55
- Stichprobenmittelwert, 53
- Stichprobenmoment, 53
- Stichprobenvarianz, 53
- Stoppzeit, 146
- straff, 27
- Test zum (Signifikanz-) Niveau α , 71
- Testproblem, 69
 - einseitiges, 69
 - zweiseitiges, 69
- totale Erneuerung, 158
- transient, 142
- trennscharf, 71
- Tschebyscheff, Pafnuti Lwowitch, 233
- Umkehrformel, 14
- Umkehrmethode, 185
- unverfälscht, 66
- unverzerrt, 54
- vage Konvergenz, 27
- Verteilung

bedingte, [85](#)
regulär bedingte, [85](#)
Verweilzeiten, [196](#)
verzerrt, [54](#)
Verzerrung, [54](#)
vollständig gleichverteilt, [176](#)
Vorwärtsrekurrenzzeit, [110](#)

Wahrscheinlichkeit
bedingte, [84](#)
regulär bedingte, [85](#)
Wahrscheinlichkeitsverteilung
Bernoulli, [11](#)
Einpunkt, [11](#)
Poisson, [11](#)
Waldsche Gleichung, [83](#)
Warteraum, [194](#)
Wartezeiten, [196](#)
Wirkungsgrad, [56](#)

Zählprozess, [101](#)
Zeitmittel, [151](#)
Zeitpunkt der n -ten Erneuerung, [100](#)
Zufallsvariable
komplexwertig, [8](#)
Zufallszahlengeneratoren, [173](#)
Zuverlässigkeitsfunktion, [156](#)
zweiseitiges Testproblem, [69](#)
Zwischenankunftszeit, [195](#)