

Abschiedsvorlesung

(Aus dem Leben eines Mathematikers)

Stand: 9. Juli 2016

Inhalt:

I. Jugendjahre (1959 – 1971)	1
II. Studienjahre (1971 – 1977)	39
III. Assistentenjahre (1977- 1986)	79
IV. Heisenberg-Stipendium (1986 -1987)	171
V. Berufung nach Oldenburg (1987 – 1995)	222
VI. Berufung nach Hamburg (1995 – 2000)	362
VII. Zweite Berufung nach Oldenburg (2000 – 2016)	392

I. Jugendjahre (1959-1971)



Textaufgabe (ca. 1962, 4. Klasse Volksschule):

Ein Zug fährt mit einer gleichbleibenden Geschwindigkeit von 72 km/h über eine 400 Meter lange Stahlbrücke. Man hört 25 Sekunden lang ein Fahrgeräusch. Wie lang ist der Zug?



Textaufgabe (ca. 1962, 4. Klasse Volksschule):

Ein Zug fährt mit einer gleichbleibenden Geschwindigkeit von 72 km/h über eine 400 Meter lange Stahlbrücke. Man hört 25 Sekunden lang ein Fahrgeräusch. Wie lang ist der Zug?

Lösung: Die Geschwindigkeit von 72 km/h entspricht 20 m/sec. Die Spitze des Zuges hat die Brücke also nach 20 sec durchfahren. In den restlichen 5 sec legt der Zug noch 100 m zurück, das entspricht genau der Länge des Zuges.

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

Gesucht sind die Lösungen der Gleichung

$$x^3 + ax^2 + bx + c = 0.$$

Vorgehen:



2. Eliminierung des quadratischen Glieds: Setze $x = y - \frac{a}{3}$, dann folgt

$$x^3 + ax^2 + bx + c = y^3 + \left(b - \frac{a^2}{3}\right)y + c - \frac{ab}{3} + \frac{2a^3}{27} = y^3 + py + q = 0.$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

2. Lösung der reduzierten kubischen Gleichung $y^3 + py + q = 0$:

Setze $y = u + v$, dann folgt

$$y^3 + py + q = u^3 + v^3 + (3uv + p)(u + v) + q = 0.$$

Mittels der Setzung $u^3 + v^3 = -q$ und $uv = -\frac{p}{3}$ bzw. $u^3v^3 = -\frac{p^3}{27}$

erhält man u^3 und v^3 als Lösung der quadratischen Gleichung

$$z^2 + qz - \frac{p^3}{27} = 0,$$

$$\text{also } u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

2. Lösung der reduzierten kubischen Gleichung $y^3 + py + q = 0$:

Insgesamt erhält man damit eine mögliche Lösung:

$$y = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

§ 43. Die Nullstellen der Funktionen dritten Grades.

175

Die Lösung des **Scipione del Ferro** (1465–1526).

45. Um die Lösung der reduzierten Form der kubischen Gleichung

$$x^3 + px + q = 0$$

zu gewinnen, benutzt man das oben angewandte Hilfsmittel. Man führt neue Unbekannte derart ein, daß man $x = u + v$ setzt; man hat dann zur Bestimmung von u und v noch eine zweite Bestimmungsgleichung frei, die man so wählt, daß die Untersuchung sich möglichst vereinfacht.

Wenn $x = u + v$ ist, so erhält man

$$u^3 + 3u^2v + 3uv^2 + v^3 + (u + v)p + q = 0 \text{ oder}$$

$$u^3 + v^3 + q + (u + v)(3uv + p) = 0.$$

Man setzt nun als zweite Gleichung für u und v fest:

$$3uv + p = 0. \text{ Daraus folgt dann } u^3 + v^3 = -q.$$

Aus dem System der beiden Gleichungen berechnet man u und v . Es ist

$$v = -\frac{p}{3u}, \text{ folglich } u^3 - \frac{p^3}{27u^3} = -q \text{ oder } u^6 + qu^3 - \left(\frac{p}{3}\right)^3 = 0.$$

$$\text{Hieraus erhält man } u = \sqrt[3]{-\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

$$\text{und } v = \sqrt[3]{-\frac{q}{2} \mp \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

§ 43. Die Nullstellen der Funktionen dritten Grades.

175

46. a) Um die drei Wurzelwerte für u und v zu erhalten, muß man die Hauptwurzeln mit den Wurzeln der Gleichung $x^3 - 1 = 0$, also mit

$$\varepsilon^3 = 1; \varepsilon = -\frac{1}{2} + \frac{i}{2} \sqrt{3} \text{ und } \varepsilon^2 = -\frac{1}{2} - \frac{i}{2} \sqrt{3}$$

multiplizieren. Wie heißen demnach die 3 Werte für u und die 3 Werte für v ?

b) Gib die möglichen Verknüpfungen der Wurzeln.

$$\begin{array}{cc} u & v \\ u \cdot \varepsilon & v \cdot \varepsilon \\ u \cdot \varepsilon^2 & v \cdot \varepsilon^2 \text{ an.} \end{array}$$

c) Stelle fest, welche von den möglichen Verknüpfungen hier zulässig sind. Beachte dabei, daß der zweiten Bestimmungsgleichung für u und v zufolge

$$uv = -\frac{p}{3}, \text{ also reell sein muß.}$$

Für die Gleichung $x^3 + px + q = 0$ erhält man demnach die 3 Wurzeln:

$$x_1 = u + v$$

$$x_2 = u\varepsilon + v\varepsilon^2 = -\frac{u+v}{2} + \frac{u-v}{2} i \sqrt{3}$$

$$x = u\varepsilon^2 + v\varepsilon = -\frac{u+v}{2} - \frac{u-v}{2} i \sqrt{3}, \text{ wo}$$

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \text{ und } v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \text{ ist.}$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

§ 43. Die Nullstellen der Funktionen dritten Grades.

177

Casus irreducibilis.

51. Es sei $x^3 - px + q = 0$ und $\left(\frac{q}{2}\right)^2 < \left(\frac{p}{3}\right)^3$.

In welcher Form bieten sich die Radikanden der drei Wurzeln dar?

Setze $-\frac{q}{2} + i \sqrt{\left(\frac{p}{3}\right)^3 - \left(\frac{q}{2}\right)^2} = r (\cos \varphi + i \sin \varphi)$

und $-\frac{q}{2} - i \sqrt{\left(\frac{p}{3}\right)^3 - \left(\frac{q}{2}\right)^2} = r (\cos (-\varphi) + i \sin (-\varphi)).$

Zeige, daß man Modul und Argument bestimmt nach den Formeln

$$r = \left(\frac{p}{3}\right)^{\frac{3}{2}} \text{ und } \cos \varphi = -\frac{\left(\frac{q}{2}\right)}{\left(\frac{p}{3}\right)^{\frac{3}{2}}}.$$

Ziehe die drei Wurzeln nach dem Moivre'schen Satze aus und begründe die Lösungen:

$$x_1 = r \cdot \cos \frac{\varphi}{3}, \quad x_2 = -r \cdot \cos \left(60^\circ - \frac{\varphi}{3}\right), \quad x_3 = -r \cdot \cos \left(60^\circ + \frac{\varphi}{3}\right).$$

Die Wurzeln, die im casus irreducibilis in imaginärer Form erscheinen, sind alle drei reell.

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

Alternative Behandlung des *Casus irreducibilis*

$$x^3 - px - q = 0 \text{ für den Fall } \frac{q^2}{4} < \frac{p^3}{27}:$$

Aus den bekannten trigonometrischen Formeln folgt:

$$\cos(3\alpha) = 4\cos^3(\alpha) - 3\cos(\alpha)$$

Setze $x = Ay$ und wähle dabei A so, dass $A^3 = 4B$ und $Ap = 3B$ gilt, also

$$A = \sqrt{\frac{4p}{3}}, \quad B = \sqrt{\frac{4p^3}{27}}. \text{ Es folgt}$$

$$x^3 - px - q = A^3 y^3 - Apy + q = B \cdot (4y^3 - 3y) - q = 0.$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

Alternative Behandlung des *Casus irreducibilis*

$$x^3 - px - q = 0 \text{ für den Fall } \frac{q^2}{4} < \frac{p^3}{27}:$$

Setze nun $y = \cos(\alpha)$, so folgt

$$x^3 - px - q = B \cdot (4y^3 - 3y) - q = \sqrt{\frac{4p^3}{27}} \cdot \cos(3\alpha) - q = 0.$$

Es folgt: $\cos(3\alpha) = \sqrt{\frac{27q^2}{4p^3}} < 1$ und somit als Lösung $x = \sqrt{\frac{4p}{3}} \cos(\alpha)$.

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

Beispiel: $x^3 - 3x - 1 = 0$

Hier ist $\cos(3\alpha) = \sqrt{\frac{27q^2}{4p^3}} = \frac{1}{2}$, also $3\alpha = 60^\circ$ und damit $\alpha = 20^\circ$.

Lösung:

$$x = \sqrt{\frac{4p}{3}} \cos(\alpha) = 2 \cos(20^\circ) = 2 \cdot 0,9397 = 1,8794$$

Lösung der allgemeinen Gleichung vom Grad 3 (ca. 1967, nach einem Schulbuch von 1927):

sin φ		45° ≤ φ ≤ 90°													
		0'	6'	12'	18'	24'	30'	36'	42'	48'	54'	60'			
Grad	→	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0			
45	→	0,7071	7083	7096	7108	7120	7133	7145	7157	7169	7181	7193	44		
46		7193	7206	7218	7230	7242	7254	7266	7278	7290	7302	7314	43		
47		7314	7325	7337	7349	7361	7373	7385	7396	7408	7420	7431	42		
48		7431	7443	7455	7466	7478	7490	7501	7513	7524	7536	7547	41		
49		7547	7559	7570	7581	7593	7604	7615	7627	7638	7649	7660	40		
50		7660	7672	7683	7694	7705	7716	7727	7738	7749	7760	7771	39		
51		7771	7782	7793	7804	7815	7826	7837	7848	7859	7869	7880	38		
52		7880	7891	7902	7912	7923	7934	7944	7955	7965	7976	7986	37		
53		7986	7997	8007	8018	8028	8039	8049	8059	8070	8080	8090	36		
54		8090	8100	8111	8121	8131	8141	8151	8161	8171	8181	8192	35		
55		8192	8202	8211	8221	8231	8241	8251	8261	8271	8281	8290	34		
56		8290	8300	8310	8320	8329	8339	8348	8358	8368	8377	8387	33		
57		8387	8396	8406	8415	8425	8434	8443	8453	8462	8471	8480	32		
58		8480	8490	8499	8508	8517	8526	8536	8545	8554	8563	8572	31		
59		8572	8581	8590	8599	8607	8616	8625	8634	8643	8652	8660	30		
60		8660	8669	8678	8686	8695	8704	8712	8721	8729	8738	8746	29		
61		8746	8755	8763	8771	8780	8788	8796	8805	8813	8821	8829	28		
62		8829	8838	8846	8854	8862	8870	8878	8886	8894	8902	8910	27		
63		8910	8918	8926	8934	8942	8949	8957	8965	8973	8980	8988	26		
64		8988	8996	9003	9011	9018	9026	9033	9041	9048	9056	9063	25		
65		9063	9070	9078	9085	9092	9100	9107	9114	9121	9128	9135	24		
66		9135	9143	9150	9157	9164	9171	9178	9184	9191	9198	9205	23		
67		9205	9212	9219	9225	9232	9239	9245	9252	9259	9265	9272	22		
68		9272	9278	9285	9291	9298	9304	9311	9317	9323	9330	9336	21		
69		9336	9342	9348	9354	9361	9367	9373	9379	9385	9391	9397	20		
70		9397	9403	9409	9415	9421	9426	9432	9438	9444	9449	9455	19		
71		9455	9461	9466	9472	9478	9483	9489	9494	9500	9505	9511	18		
72		9511	9516	9521	9527	9532	9537	9542	9548	9553	9558	9563	17		
73		9563	9568	9573	9578	9583	9588	9593	9598	9603	9608	9613	16		
74		9613	9617	9622	9627	9632	9636	9641	9646	9650	9655	9659	15		
75		9659	9664	9668	9673	9677	9681	9686	9690	9694	9699	9703	14		
76		9703	9707	9711	9715	9720	9724	9728	9732	9736	9740	9744	13		
77		9744	9748	9751	9755	9759	9763	9767	9770	9774	9778	9781	12		
78		9781	9785	9789	9792	9796	9799	9803	9806	9810	9813	9816	11		
79		9816	9820	9823	9826	9829	9833	9836	9839	9842	9845	9848	10		
80		9848	9851	9854	9857	9860	9863	9866	9869	9871	9874	9877	9		
81		9877	9880	9882	9885	9888	9890	9893	9895	9898	9900	9903	8		
82		9903	9905	9907	9910	9912	9914	9917	9919	9921	9923	9925	7		
83		9925	9928	9930	9932	9934	9936	9938	9940	9942	9943	9945	6		
84		9945	9947	9949	9951	9952	9954	9956	9957	9959	9960	9962	5		
85		9962	9963	9965	9966	9968	9969	9971	9972	9973	9974	9976	4		
86		9976	9977	9978	9979	9980	9981	9982	9983	9984	9985	9986	3		
87		9986	9987	9988	9989	9990	9990	9991	9992	9993	9993	9994	2		
88		9994	9995	9995	9996	9996	9997	9997	9997	9998	9998	9998	1		
89		9998	9999	9999	9999	9999	1,000	1,000	1,000	1,000	1,000	1,000	0		
		1.0	.9	.8	.7	.6	.5	.4	.3	.2	.1	.0	← Grad		
		60'	54'	48'	42'	36'	30'	24'	18'	12'	6'	0'			

sin φ

cos φ

tan φ

cot φ

sin x°

f (e°)

Log

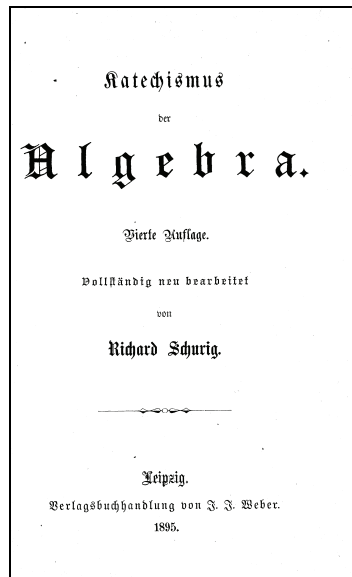
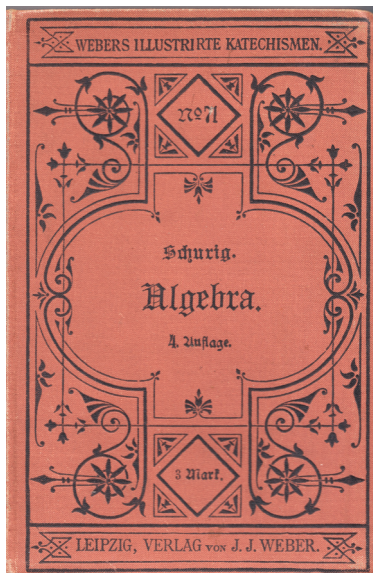
Beispiele:
sin φ = 0,9212 ergibt: φ = 67,1°
cos φ = 0,9981 ergibt: φ = 3,5°

cos φ

11

sin φ
cos φ
tan φ
cot φ
sin x
f (e°)
Log

Lösung der allgemeinen Gleichung vom Grad 3 (Exkurs):



Lösung der allgemeinen Gleichung vom Grad 3 (Exkurs):

234

§ 56. Die Gleichung dritten Grades.

Beispiel. $x^3 - 14x + 55 = 0$. Hier ist $b = -14$, $c = 55$, folglich:

$$\begin{aligned} x &= \sqrt[3]{-55/2 + \sqrt{(55/2)^2 + (-14/3)^3}} \\ &\quad + \sqrt[3]{-55/2 - \sqrt{(55/2)^2 + (-14/3)^3}}; \text{ d. i.} \\ x &= -\left[\sqrt[3]{55/2 - \sqrt{(55/2)^2 - (14/3)^3}}\right. \\ &\quad \left.+ \sqrt[3]{55/2 + \sqrt{(55/2)^2 - (14/3)^3}}\right] \\ &= -\left[\sqrt[3]{27,5 - \sqrt{756,25 - 101,62963}}\right. \\ &\quad \left.+ \sqrt[3]{27,5 + \sqrt{756,25 - 101,62963}}\right] \\ &= -\left[\sqrt[3]{27,5 - 25,58555} + \sqrt[3]{27,5 + 25,58555}\right] \\ &= -[1,2416943 + 3,7583057] \text{ oder } x = 5. \end{aligned}$$

VII. Ist $x^3 - bx \pm c = 0$ gegeben, wo $-b$ eine negative Zahl und $4b^3 > 27c^2$ sein soll, so wird die Cardanische Formel unbrauchbar, da alsdann die Quadratwurzel in derselben imaginär ist. Für diesen Fall, den man casus irreducibilis nennt, hat alsdann die Gleichung dritten Grades drei reelle Wurzeln, während sie nur eine reelle Wurzel und zwei imaginäre hat, wenn in der Cardanischen Formel die Quadratwurzel reell ist.

A. Der casus irreducibilis konnte bisher nicht algebraisch, sondern nur trigonometrisch in folgender Weise aufgelöst werden:

Bestimme den Winkel e aus

$$\sin e = \frac{3c}{b \sqrt{\frac{4b}{3}}}, \text{ wo } b \text{ und } c \text{ stets absolut (positiv) zu nehmen sind.}$$

§ 56. Die Gleichung dritten Grades.

235

$$\begin{aligned} \text{Alsdann ist } x_1 &= \sqrt[3]{\frac{4b}{3} \cdot \sin \frac{e}{3}} \\ x_2 &= \sqrt[3]{\frac{4b}{3} \cdot \cos \left(30^\circ + \frac{e}{3}\right)} \\ x_3 &= -(x_1 + x_2). \end{aligned} \left. \vphantom{\begin{aligned} x_1 &= \sqrt[3]{\frac{4b}{3} \cdot \sin \frac{e}{3}} \\ x_2 &= \sqrt[3]{\frac{4b}{3} \cdot \cos \left(30^\circ + \frac{e}{3}\right)} \right\} \text{ mit dem Vorzeichen von } c.$$

B. Dennoch kann die Auflösung der Gleichung dritten Grades für den casus irreducibilis auch algebraisch durch die nachstehende, vom Verfasser (H. Sch.) neuerdings gefundene Formel aufgelöst werden.

$$x = \frac{\pm mc}{b \left[n + p \sqrt{1 - \frac{27c^2}{4b^3} - \frac{c^2}{b^3}} \right]^r},$$

wo $\lg m = 0,7968036$; $n = 11,639016$; $\lg p = 1,0542793$; $r = 0,5853828$; $\lg r = 9,7674400 - 10$.

Die Berechnung von x ist am besten in folgender Weise auszuführen:

Man bestimme den Logarithmus von

$$\left[n + p \sqrt{1 - \frac{27c^2}{4b^3} - \frac{c^2}{b^3}} \right].$$

Suche von demselben (als Numerus gedacht) den Logarithmus und vermehre diesen um $9,7674400 - 10$. Der zur Summe dieser beiden Logarithmen gehörige Numerus ist der Log. von [...]. Alsdann ist

$$\lg x = \lg m + \lg c + d. E. \lg b + d. E. \lg [...];$$

x mit dem Vorzeichen von c .

Beispiel. $x^3 - 21x - 35 = 0$.

Da $4 \cdot 21^3 > 27 \cdot 35^2$, $27 \cdot 35^2 = 33075$, also $4b^3 > 27c^2$, so tritt hier der casus irreducibilis ein. Die Berechnung ist nun folgende:

$$\begin{aligned} \lg \left(n + p \sqrt{1 - \frac{27 \cdot 35^2}{4 \cdot 21^3} - \frac{35^2}{21^3}} \right) &= \lg \left(n + p \sqrt{\frac{3}{28} - \frac{25}{189}} \right) \\ &= \lg (11,639016 + 3,709036 - 0,132275) \\ &= \lg 15,215777 = 1,1822941. \end{aligned}$$

Lösung der allgemeinen Gleichung vom Grad 3 (Exkurs):

Lösung mit der Formel von Schurig:

$$x = -0,3472969254$$

Numerische Lösungen der Gleichung $x^3 - 3x - 1 = 0$:

$$-1,532088886 \quad -0,3472963553 \quad 1,879385242$$

Konstruktive Dreiteilung eines Winkels mit dem Tomahawk (Exkurs):

Der Tomahawk ist ein geometrisches „Werkzeug“, das aus einem Halbkreis und zwei Geradenabschnitten besteht. Die Grundform kann mit Zirkel und Lineal konstruiert werden. Mit einem Tomahawk kann man einen Winkel dreiteilen. Dies widerspricht nicht der Unlösbarkeit des klassischen Problems, da die erforderliche Verschiebung und Drehung des konstruierten Tomahawks in die nötige Lage über die dort erlaubten Konstruktionsmethoden hinausgeht.

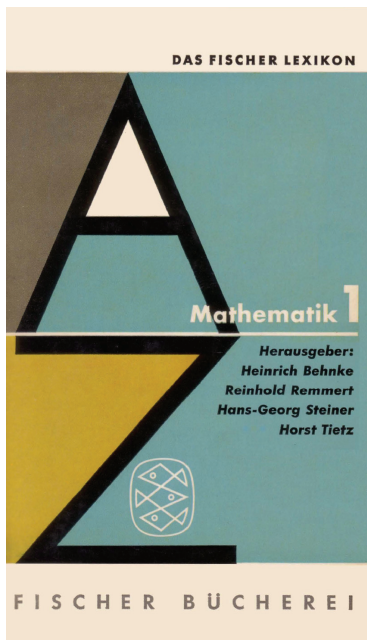
Man platziert den Tomahawk zu einem gegebenen Winkel so, dass sein Stiel durch die Spitze des Winkels läuft, seine Klinge den einen Schenkel des Winkels berührt und sein hinteres Ende den anderen Schenkel. Nun bildet der Stiel mit den beiden Schenkeln zwei Winkel, deren Maß ein Drittel beziehungsweise zwei Drittel des Ausgangswinkels beträgt.

Konstruktive Dreiteilung eines Winkels mit dem Tomahawk (Exkurs):

Es ist nicht gesichert, wer den Tomahawk erfunden hat. Die frühesten bekannten Erwähnungen stammen aus dem 19. Jahrhundert in Frankreich. Dort wird er in dem 1835 erschienenen Buch *Géométrie appliquée a l'industrie, a l'usage des artistes et des ouvriers* von Claude Lucien Bergery erwähnt. Henri Brocard beschrieb ihn 1877 in einer Publikation, in der er die Entdeckung dem Marineoffizier Pierre-Joseph Glotin zuschrieb, der ihn 1863 in einem Aufsatz für die *Mémoires de la Société des sciences physiques et naturelles de Bordeaux* behandelt hatte.

Quelle: Wikipedia

Lösung der allgemeinen Gleichung vom Grad 4 (Exkurs):



Enzyklopädie des Wissens

Die Mathematik, neben der Philosophie wohl die älteste Wissenschaft der Menschheit, hat sich in ihrer Forschung besonders in neuerer Zeit um die Klärung der eigenen Grundlagen bemüht.

Ihrer zentralen Stellung zwischen den Geistes- und Naturwissenschaften zufolge hat sie starke Impulse auf diese Nachbarwissenschaften übertragen. So basiert die Mathematisierung der Technik, der Physik, der Sozial- und Wirtschaftswissenschaften letztlich auf der Anwendung von Erkenntnissen, die aus der Erforschung der Grundlagen der gesamten Mathematik gewonnen wurden. Diese Grundlagen im Rahmen der einzelnen Sachwortartikel darzustellen, haben die Autoren des vorliegenden Mathematik-Lexikons (die zugleich auch Herausgeber sind) unternommen. Neben einer Einführung in die Grundlagen des dargestellten Sachgebiets wird ein Überblick über seine Probleme sowie seine jeweiligen modernen Anwendungsprobleme gegeben.

Die Anlage dieses Lexikons gibt schließlich dem Benutzer die Möglichkeit, einzelne mathematische Begriffe mit Hilfe des umfangreichen Registers im Text aufzusuchen und ihren Zusammenhang mit anderen mathematischen Sachgebieten zu erkennen. Wie in allen Bänden des Fischer Lexikons, so sind auch hier die enzyklopädischen Sachwortartikel alphabetisch angeordnet und Stichwörter, die im Register erscheinen, im Text kursiv gesperrt gedruckt.

Mit dem neuartigen Aufbau des Fischer Lexikons in einer Taschenbuch-Reihe verbindet sich eine alte Aufgabe, die schon d'Alembert 1751 in der Einleitung zur französischen Enzyklopädie formulierte: »Das Werk, das wir begonnen haben und zu Ende zu führen wünschen, hat einen doppelten Zweck: als Enzyklopädie soll es, soweit möglich, die Ordnung und Verkettung der menschlichen Kenntnisse erklären; und als methodisches Sachwörterbuch der Wissenschaften, Künste und Gewerbe soll es von jeder Wissenschaft und Kunst — gehöre sie zu den freien oder den technischen — die allgemeinen Grundsätze enthalten, auf denen sie beruhen, und die wesentlichen Besonderheiten, die ihren Umfang und Inhalt bedingen.«

1. Auflage erschienen 1964

Lösung der allgemeinen Gleichung vom Grad 4 (Exkurs):

f. Die *allgemeine Gleichung 4. Grades* wird durch $x = y - \frac{a_1}{4a_0}$ in die Gleichung $y^4 + py^2 + qy + r = 0$ transformiert; diese wird gelöst durch die Werte

$$\begin{aligned} & \sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}, \quad \sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}, \\ & -\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}, \quad -\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3}, \end{aligned}$$

wobei z_1, z_2, z_3 die Lösungen der *kubischen Resolvente* $z^3 + 2pz^2 + (4r - p^2)z - q^2 = 0$ sind.

Quelle : Fischer Lexikon Mathematik, Bd. 1 (1964), Stichwort: Gleichungen

Vgl. auch Bewersdorff (2013), Kapitel 3.

Konstruktionen mit Zirkel und Lineal (ca. 1968): Siebeneck

Approximative Konstruktion des regelmäßigen Siebenecks:

Kreisteilungsgleichung:

$$x^7 - 1 = 0$$

Nach Polynomdivision mit $x - 1$ erhält man die Gleichung

$$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 = 0$$

Nach Division durch x^3 und der Substitution $z = x + \frac{1}{x}$ erhält man weiter die Gleichung

$$z^3 + z^2 - 2z - 1 = 0.$$

Konstruktionen mit Zirkel und Lineal (ca. 1968): Siebeneck

Durch Reduktion auf die reduzierte Form mittels $z = y - \frac{1}{3}$ geht dies über in

$$f(y) := y^3 - \frac{7}{3}y - \frac{7}{27} = 0.$$

Dies entspricht dem *Casus irreducibilis*.

Eigener Ansatz: Approximative Lösung aus $-\frac{7}{3}y - \frac{7}{27} = 0$:

$$y = -\frac{1}{9} \text{ mit } f\left(-\frac{1}{9}\right) = -\frac{1}{729} = -0,00137\dots \text{ und somit } z = -\frac{4}{9}$$

Konstruktionen mit Zirkel und Lineal (ca. 1968): Siebeneck

Als eine approximative Lösung der Gleichung

$$x^7 - 1 = 0$$

erhält man somit

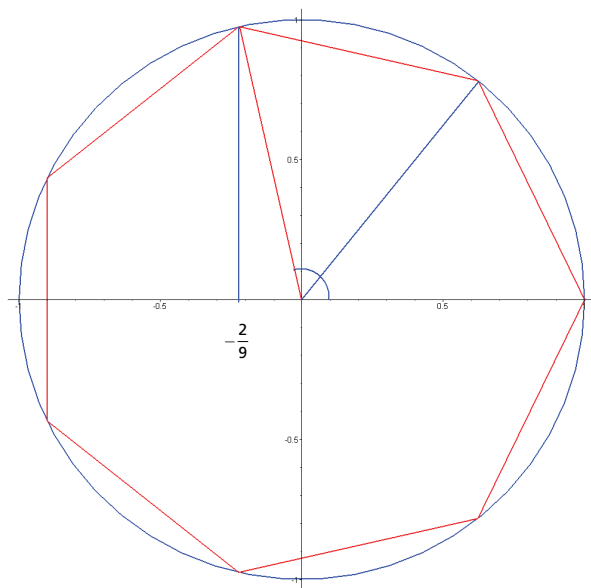
$$x = -\frac{2}{9} \pm \frac{i}{9} \sqrt{77}.$$

Fehlerabschätzung:

$$\left(-\frac{2}{9} \pm \frac{i}{9} \sqrt{77} \right)^7 = \frac{4.782.958}{4.782.969} \mp \frac{1.169}{4.782.969} \sqrt{77} i$$

$$= 0,9999976992 \mp 0,002144678967 i$$

Konstruktionen mit Zirkel und Lineal (ca. 1968): Siebeneck



Die Sehnenlänge beträgt exakt

$$L = 2 \sin\left(\frac{\pi}{7}\right) = \mathbf{0,8677674786}$$

und approximativ

$$\hat{L} = \frac{1}{3} \sqrt{18 - 3\sqrt{14}} = \mathbf{0,8676294550}$$

Der relative Fehler beträgt

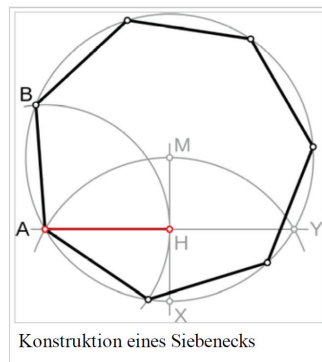
$$\mathbf{-0,0159\%}$$

Andere Näherungskonstruktionen (Quelle: Wikipedia)

Erste Näherung

Eine sehr einfache Näherungskonstruktion ist in folgender Zeichnung dargestellt:

1. Vom Mittelpunkt des Umkreises zeichnet man eine Gerade, die den Umkreis im Punkt X schneidet.
2. Dann zeichnet man einen Kreis um X , der durch M verläuft und den Umkreis in den Punkten A und Y schneidet.
3. Die Gerade AY schneidet die Strecke MX im Halbierungspunkt H .
4. Die rote Strecke AH ist eine gute Näherung für die Seitenlänge des Siebenecks.
5. Die Eckpunkte B bis G erhält man durch Abschlagen der Strecke AH .



Genau dieselbe Streckenlänge lässt sich folgendermaßen konstruieren:

1. Konstruiere das dem Umkreis einbeschriebene regelmäßige (gleichseitige) Dreieck.
2. Die Hälfte einer Dreiecksseite nimm als Näherung für die Seite des Siebenecks.

In dieser Form war sie bereits dem im 10. Jahrhundert in Bagdad wirkenden Gelehrten Abu l-Wafa bekannt.

Der relative Fehler beträgt hier $-0,2\%$.

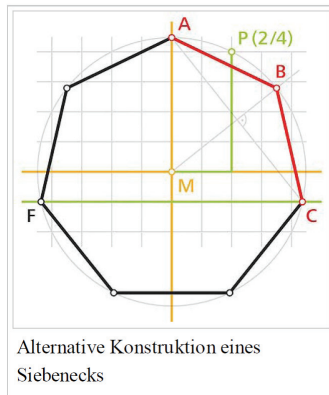
Andere Näherungskonstruktionen (Quelle: Wikipedia)

Zweite Näherung

Mit Koordinatensystem

Eine etwas aufwändigere, aber genauere Näherungskonstruktion ist in folgender Zeichnung dargestellt:

1. In einem rechtwinkligen Koordinatensystem zeichnet man einen Kreis, der seinen Mittelpunkt im Ursprung $(0/0)$ hat und genau durch den Punkt P mit den Koordinaten $(2/4)$ verläuft.
2. Der Schnittpunkt der positiven y -Achse mit der Kreislinie wird als Eckpunkt A des regelmäßigen Siebenecks festgelegt.
3. Die Gerade $y = -1$ (horizontale grüne Linie) schneidet die Kreislinie in unmittelbarer Nähe der Eckpunkte C und F .
4. Wenn man die Streckensymmetrale der Strecke AC mit dem Kreis schneidet, erhält man eine Näherung für den Eckpunkt B .
5. Die rote Strecke $\overline{AB}$ oder $\overline{BC}$ ist eine sehr gute Näherung für die Seitenlänge des regelmäßigen Siebenecks.
6. Die Eckpunkte D , E und G erhält man durch Spiegelung oder Abschlagen der Seitenlänge am Umkreis.

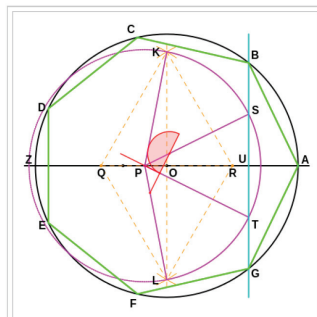


Der relative Fehler beträgt hier 0,0578%.

Andere Näherungskonstruktionen (Quelle: Wikipedia)

Exakte Konstruktion mittels Dreiteilung eines Winkels

- Konstruiere einen Kreis - den späteren Umkreis des Siebenecks - um einen Mittelpunkt (O) auf einer Grundlinie (AZ). Einer der Schnittpunkte mit dem Kreis ist der erste Eckpunkt (A) des späteren Siebenecks.
- Halbiere die beiden Radien des ersten Durchmessers (Punkte Q und R)
- Errichte auf der so erhaltenen Strecke zwei gleichseitige Dreiecke mit der Seitenlänge gleich dem Kreisradius. (Man erhält Punkte K und L).
- Trage auf der Grundlinie (AZ) vom Mittelpunkt aus $1/6$ des Radius in die dem auf der Grundlinie liegenden Eckpunkt entgegengesetzte Richtung ab (Punkt P).
- Zeichne um den so erhaltenen Punkt einen Hilfskreis durch die beiden nicht auf der Grundlinie liegenden Ecken der gleichseitigen Dreiecke.
- Zeichne in diesen Kreis die beiden Radien zu diesen beiden Punkten ein.
- Teile den von diesen Radien gebildeten Winkel unter Verwendung des Extrawerkzeugs in drei Teile (z.B. Tomahawk, in der Zeichnung rot dargestellt) und zeichne die so gewonnenen Geraden ein. Sie schneiden den Hilfskreis in zwei weiteren Punkten (Punkte S und T).
- Die Gerade durch diese Punkte - sie liegt senkrecht zur Grundlinie - schneidet den Umkreis des Siebenecks an den zum ersten Eckpunkt (A) benachbarten Ecken des Siebenecks (Punkte B und G).
- Ergänze die noch fehlenden Ecken durch Abtragen der Seiten.



Konstruktion mit Tomahawk (rot)

Weitere Quellen:

L. Bieberbach (1952): *Theorie der geometrische Konstruktionen*. Springer, Basel.

A.M. Gleason (1988): *Angle Trisection, the Heptagon, and the Triskaidecagon*. The American Mathematical Monthly 95 (3), 185-194.

R. Geretschläger (1997): *Folding the Regular Heptagon*. Crux Mathematicorum 23 (2), 81-88.

D. Dureisseix (1997): *Searching for Optimal Polygon - Remarks About a General Construction and Application to Heptagon and Nonagon*. Unpublished manuscript.

A. Johnson and A. Pimpinelli (2008): *Pegs and Ropes: Geometry at Stonehenge*. Nature Precedings:

<http://precedings.nature.com/documents/2153/version/1>

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Lösung nach C.F. Gauß:

Kreisteilungsgleichung:

$$x^{17} - 1 = 0 \quad (1)$$

Eine triviale Lösung ist $x = 1$; nach Polynomdivision ergibt sich aus (1) daher

$$\sum_{k=0}^{16} x^k = 0 \quad (2)$$

Setze

$$\begin{aligned} P_1 &:= x^{16} + x^{15} + x^{13} + x^9 + x^8 + x^4 + x^2 + x \\ P_2 &:= x^{14} + x^{12} + x^{11} + x^{10} + x^7 + x^6 + x^5 + x^3 \end{aligned} \quad (3)$$

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Es gilt für jede Lösung x wegen (1): $x^n = x^m$ mit $m \equiv n \pmod{17}$ und daher

$$\begin{aligned} P_1 + P_2 &= \sum_{k=1}^{16} x^k = -1 \\ P_1 \cdot P_2 &= 4 \sum_{k=1}^{16} x^k = -4 \end{aligned} \quad \text{mit der Lösung } P_{1,2} = -\frac{1}{2} \pm \frac{1}{2} \sqrt{17} \quad (4)$$

Setze weiter

$$\begin{aligned} P_{11} &:= x^{16} + x^{13} + x^4 + x \\ P_{12} &:= x^{15} + x^9 + x^8 + x^2 \\ P_{21} &:= x^{14} + x^{12} + x^5 + x^3 \\ P_{22} &:= x^{11} + x^{10} + x^7 + x^6 \end{aligned} \quad (5)$$

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Dann gilt

$$P_{11} + P_{12} = P_1$$

$$P_{11} \cdot P_{12} = \sum_{k=1}^{16} x^k = -1$$

$$P_{11,12} = \frac{P_1}{2} \pm \frac{1}{2} \sqrt{P_1^2 + 4}$$

mit den Lösungen

(6)

$$P_{21} + P_{22} = P_2$$

$$P_{21} \cdot P_{22} = \sum_{k=1}^{16} x^k = -1$$

$$P_{21,22} = \frac{P_2}{2} \pm \frac{1}{2} \sqrt{P_2^2 + 4}$$

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Setze weiter

$$\begin{aligned}P_{111} &:= x^{16} + x \\ P_{112} &:= x^{13} + x^4\end{aligned}\tag{7}$$

Dann gilt

$$\begin{aligned}P_{111} + P_{112} &= P_{11} \\ P_{111} \cdot P_{112} &= P_{21}\end{aligned}\quad \text{mit der Lösung } P_{111,112} = \frac{P_{11}}{2} \pm \frac{1}{2}\sqrt{P_{11}^2 - 4P_{21}}\tag{8}$$

Es bleibt zu lösen:

$$P_{111} = x^{16} + x \quad \text{bzw.} \quad P_{111}x = x^{17} + x^2 = 1 + x^2\tag{9}$$

mit dem Ergebnis

$$x_{1,2} = \frac{P_{111}}{2} \pm \frac{1}{2}\sqrt{P_{111}^2 - 4}\tag{10}$$

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Beide Lösungen sind komplex; nach Einsetzen aller Zwischenergebnisse aus (4), (6) und (8) erhält man

$$\Re(x_{1,2}) = \frac{1}{16} \cdot \left(-1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + 2\sqrt{17 + 3\sqrt{17} - \sqrt{34 - 2\sqrt{17}} - 2\sqrt{34 + 2\sqrt{17}}} \right) \quad (11)$$

bzw. numerisch

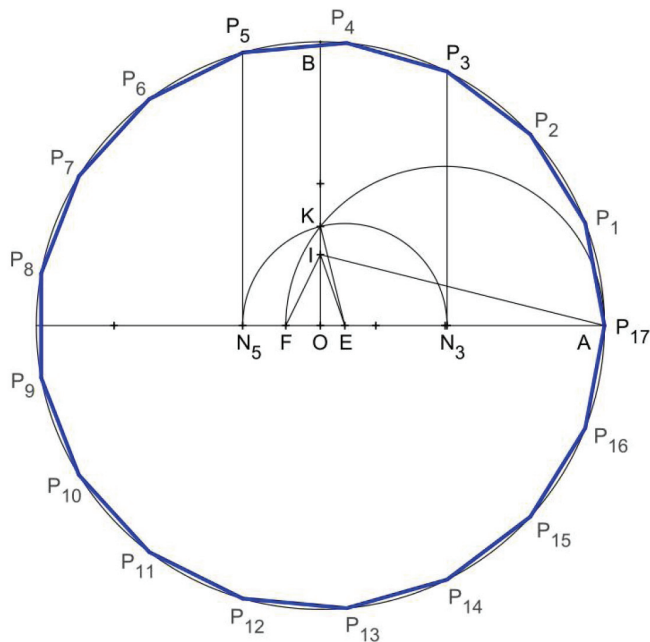
$$\Re(x_{1,2}) = 0,9324722294, \quad \Im(x_{1,2}) = \pm 0,3612416662. \quad (12)$$

Zum Vergleich: es ist

$$\cos\left(\frac{2\pi}{17}\right) = 0,9324722294, \quad \sin\left(\frac{2\pi}{17}\right) = 0,3612416662, \quad (13)$$

d.h. $x_1 = \frac{P_{111}}{2} + \frac{1}{2}\sqrt{P_{111}^2 - 4}$ entspricht der „ersten“ Ecke des regelmäßigen Siebzehnecks, gegen den Uhrzeigersinn gezählt.

Exkurs: Konstruktion des regelmäßigen 17-Ecks



Konstruktion von Herbert William Richmond (1893)

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Ist der Umkreis um das entstehende Siebzehneck mit dem Mittelpunkt O gegeben, kann das Siebzehneck konstruiert werden durch:

1. Zeichnen des Durchmessers durch den Mittelpunkt O; Schnittpunkt mit Umkreis ist A.
2. Errichten des Radius senkrecht zu $\overline{AO}$ auf O bis zum Umkreis; Schnittpunkt mit Umkreis ist B.
3. Konstruktion des Punktes I durch Vierteln der Strecke $\overline{BO}$; I liegt näher an O.
4. Konstruktion des Punktes E durch Vierteln des Winkels OIA.
5. Konstruktion des Punktes F mithilfe einer Senkrechte zu $\overline{EI}$ auf I; Halbierung des 90° -Winkels; Schnittpunkt mit Durchmesser ist F und Winkel FIE ist 45° .
6. Konstruktion des Thaleskreises über $\overline{AF}$; Schnittpunkt mit $\overline{BO}$ ist K.

Exkurs: Konstruktion des regelmäßigen 17-Ecks

7. Zeichnen des Halbkreises um den Mittelpunkt E mit dem Radius $\overline{EK}$; Schnittpunkte mit dem Durchmesser sind N_3 und N_5 (dabei liegt N_3 sehr nahe beim Mittelpunkt des Thaleskreises über $\overline{AF}$).
8. Konstruktion der Tangente ab N_3 ; Schnittpunkt mit dem Umkreis ist der Eckpunkt P_3 des Siebzehneckes; der Kreisbogen $\widehat{OAP_3}$ ist somit $3/17$ des Umkreisumfanges.
9. Konstruktion der Tangente ab N_5 ; Schnittpunkt mit dem Umkreis ist der Eckpunkt P_5 des Siebzehneckes; der Kreisbogen $\widehat{OAP_5}$ ist somit $5/17$ des Umkreisumfanges.
10. Ein vierzehnmaliges Abtragen der Strecke $\overline{AP_3}$ ab dem Eckpunkt P_3 gegen dem Uhrzeigersinn ergibt der Reihe nach die Eckpunkte $P_6, P_9, P_{12}, P_{15}, P_1, P_4, P_7, P_{10}, P_{13}, P_{16}, P_2, P_8, P_{11}$ und P_{14} .
11. Verbinden der so gefundenen Punkte $P_1, P_2, \dots, P_{17}, P_1$ vervollständigt das 17-Eck.

Quelle: <https://de.wikipedia.org/wiki/Siebzehneck>

Exkurs: Konstruktion des regelmäßigen 17-Ecks

Literatur:

J. Bewersdorff (2013): *Algebra für Einsteiger*. Von der Gleichungslösung zur Galois-Theorie. 5. Aufl., Springer Spektrum, Wiesbaden, Kapitel 7.

H.W. Richmond (1893): *A Construction for a regular polygon of seventeen sides*. In: *The Quarterly Journal of Pure and Applied Mathematics*. 26, 206–207.

II. Studienjahre (1971-1977)



Studienjahre (1971–1977)

Studienort: RWTH Aachen

Akademische Lehrer:

- Walter Oberschelp (Mengenlehre und Logik)
- Herrmann Engels (Numerik)
- Friedhelm Erwe (Analysis)
- Jochen Reiner mann (Funktionentheorie)
- Joachim Neubüser (Lineare Algebra, Algebra)
- Adalbert Kerber (Lineare Algebra, Algebra)
- **Paul Leo Butzer (Analysis, Funktionalanalysis)**
- Rolf Joachim Nessel (Fourier-Analysis)
- Ernst Görlich (Topologie)
- **Burkhard Raухut (Stochastik, Extremwertstatistik, Spieltheorie)**
- Olaf Krafft (Statistik, Stochastische Prozesse)
- Rolf Kaerkes (Wahrscheinlichkeitstheorie)

Studienjahre (1971–1977)

Ab ca. 1975: Verstärktes Interesse an

- Extremwertstatistik (Rauhut; Diplomarbeit und Promotion 1979) und
- Operatorhalbgruppen (Butzer; Habilitation 1984)

Regelmäßig Nebentätigkeit als studentischer Tutor für C. Müller (Mathematik für Ingenieure I bis IV)

Studienjahre (1971–1977)

Überlegungen zur Herleitung von Extremwertverteilungen:

Sind $X_1, \dots, X_n$ stochastisch unabhängige Zufallsvariablen und bezeichnet $X_{n:n} = \max(X_1, \dots, X_n)$ für $n \in \mathbb{N}$, so gilt:

$$P(X_{n:n} \leq x) = P\left(\bigcap_{k=1}^n \{X_k \leq x\}\right) = F^n(x), \quad x \in \mathbb{R},$$

wobei F die Verteilungsfunktion der X_k bezeichne. Mit noch geeignet zu bestimmenden Konstantenfolgen $\{a_n\}$ (mit $a_n > 0$) und $\{b_n\}$ kann man nun versuchen, ähnlich zum Zentralen Grenzwertsatz einen Grenzwert für die normalisierten Maxima $Z_n = \frac{X_{n:n} - b_n}{a_n}$ zu finden. Wir zeigen hier exemplarisch am Beispiel der Exponentialverteilung $\mathcal{E}(\lambda)$ mit $\lambda > 0$, dass dies in gewissen Fällen funktioniert.

Studienjahre (1971–1977)

Wir wählen $a_n = \frac{1}{\lambda}$ und $b_n = \frac{\ln n}{\lambda}$. Dann folgt für $x \in \mathbb{R}$

$$\begin{aligned} P(Z_n \leq x) &= P\left(\frac{X_{n:n} - b_n}{a_n} \leq x\right) = P(X_{n:n} \leq a_n x + b_n) = F^n(a_n x + b_n) \\ &= F^n\left(\frac{x + \ln n}{\lambda}\right) = \left(1 - \exp\left(-\lambda \cdot \left(\frac{x + \ln n}{\lambda}\right)^+\right)\right)^n \\ &= \left(1 - \exp\left(-(x + \ln n)^+\right)\right)^n = \left\{\left(1 - \frac{e^{-x}}{n}\right)^+\right\}^n \rightarrow \exp(-\exp(-x)) \end{aligned}$$

für $n \rightarrow \infty$. Die zugehörige Grenzverteilung heißt *Gumbelverteilung* oder doppelte Exponentialverteilung.

Studienjahre (1971–1977)

Nach einem klassischen Satz von Fisher und Tippett (1928) gilt allgemein, dass im Falle der Konvergenz die (nicht-degenerierten) Grenzverteilungsfunktionen bis auf positiv-lineare Transformationen (Lage-Skalen-Familien) von genau einem der drei folgenden Typen sind:

$$\text{Fréchet-Klasse: } \Phi_{\alpha}(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad (\alpha > 0)$$

$$\text{Weibull-Klasse: } \Psi_{\alpha}(x) = \begin{cases} \exp\{-(-x)^{\alpha}\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad (\alpha > 0)$$

$$\text{Gumbel-Klasse: } \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}.$$

Studienjahre (1971–1977)

Man kann die drei Klassen von Verteilungsfunktionen auch durch eine einzige Verteilungsfunktionsklasse (mit nur einem Parameter) charakterisieren, und zwar folgendermaßen:

$$G_{\xi}(x) = \exp\left(-(1 + \xi x)^{-1/\xi}\right), \quad 1 + \xi x > 0 \quad (\xi \in \mathbb{R}),$$

wobei der Sonderfall $\xi = 0$ als Grenzfall aufzufassen ist.

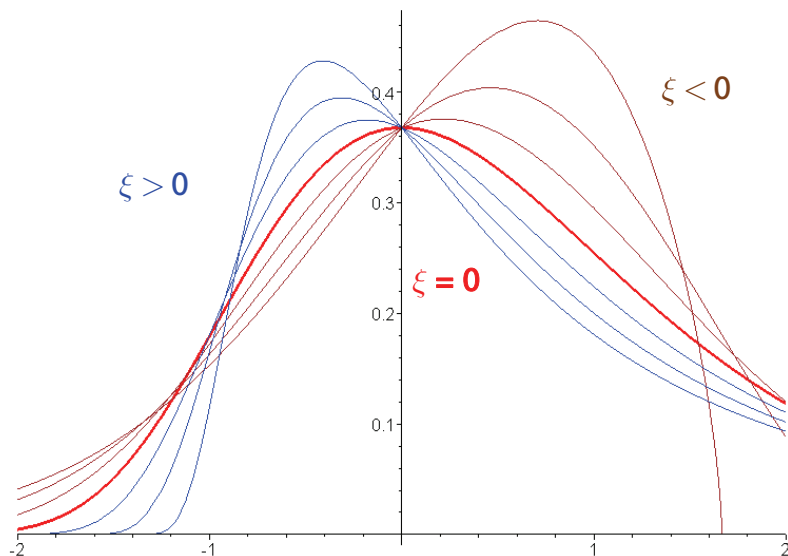
Für $\xi > 0$ ergibt sich die *Fréchet-Klasse*,

für $\xi < 0$ die *Weibull-Klasse* (in beiden Fällen mit $\alpha = 1/|\xi|$)

und für $\xi = 0$ die *Gumbel-Klasse*.

Der Parameter $\xi \in \mathbb{R}$ wird auch als *extremaler Index* bezeichnet.

Studienjahre (1971–1977)



Dichten der Standard-Grenzverteilungen in ξ -Parametrisierung

Studienjahre (1971–1977)

Bemerkung: Es ist nicht immer gewährleistet, dass man geeignete Folgen $\{a_n\}$ (mit $a_n > 0$) und $\{b_n\}$ finden kann, so dass $Z_n = \frac{X_{n:n} - b_n}{a_n}$ eine (nicht-degenerierte) Grenzverteilung besitzt.

Eine exakte Charakterisierung der (schwachen) Konvergenz kann wie folgt erhalten werden.

Studienjahre (1971–1977)

Bezeichnet

$x_{\sup} := \sup\{x \in \mathbb{R} \mid F(x) < 1\}$ und $\beta_n := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$ (Quantil), so gilt:

Die Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$ liegt in der *Gumbel-Klasse* genau dann, wenn eine Funktion g existiert mit der Eigenschaft

$$\lim_{t \uparrow x_{\sup}} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x} \text{ für alle } x \in \mathbb{R}.$$

Eine mögliche Wahl von g ist $g(t) = \int_t^{x_{\sup}} \frac{1 - F(u)}{1 - F(t)} du$ für $t < x_{\sup}$.

Als Normalisierungskonstanten können $a_n = g(\beta_n)$, $b_n = \beta_n$ für $n \in \mathbb{N}$ gewählt werden.

Studienjahre (1971–1977)

Bezeichnet

$x_{\sup} := \sup\{x \in \mathbb{R} | F(x) < 1\}$ und $\beta_n := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$ (Quantil), so gilt:

Die Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$ liegt in der *Fréchet-Klasse* genau dann, wenn $x_{\sup} = \infty$ ist und

$$\lim_{t \uparrow \infty} \frac{1 - F(t \cdot x)}{1 - F(t)} = x^{-\alpha} \text{ für alle } x > 0 \text{ gilt.}$$

Als Normalisierungskonstanten können $a_n = \beta_n$, $b_n = 0$ für $n \in \mathbb{N}$ gewählt werden.

Studienjahre (1971–1977)

Bezeichnet

$x_{\sup} := \sup\{x \in \mathbb{R} | F(x) < 1\}$ und $\beta_n := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$ (Quantil), so gilt:

Die Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$ liegt in der *Weibull-Klasse* genau dann, wenn $x_{\sup} < \infty$ ist und

$$\lim_{h \downarrow 0} \frac{1 - F(x_{\sup} - x \cdot h)}{1 - F(x_{\sup} - h)} = x^{\alpha} \text{ für alle } x > 0 \text{ gilt.}$$

Als Normalisierungskonstanten können $a_n = x_{\sup} - \beta_n$, $b_n = x_{\sup}$ für $n \in \mathbb{N}$ gewählt werden.

Studienjahre (1971–1977)

Beispiele:

Für die *Exponentialverteilung* $\mathcal{E}(\lambda)$ mit $\lambda > 0$ haben wir schon direkt gezeigt, dass die Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$ in der *Gumbel-Klasse*

liegt. Hier ist $x_{\sup} = \infty$ und $\beta_n = F^{-1}\left(1 - \frac{1}{n}\right) = \frac{\ln n}{\lambda}$, mit

$$g(t) = \int_t^{\infty} \frac{1 - F(u)}{1 - F(t)} du = \int_t^{\infty} \exp(-\lambda(u - t)) du = \int_0^{\infty} \exp(-\lambda v) dv = \frac{1}{\lambda} \text{ für } t < \infty.$$

Dies bestätigt den in der Herleitung gewählten Ansatz.

Studienjahre (1971–1977)

Beispiele:

Die *Pareto-Verteilung* $\mathcal{Pa}(\alpha)$ mit $\alpha > 0$ besitzt die Verteilungsfunktion

$$F(x) = 1 - \frac{1}{(1+x)^\alpha} \text{ für } x \geq 0.$$

Hier ist $x_{\sup} = \infty$ und $\beta_n = F^{-1}\left(1 - \frac{1}{n}\right) = n^{1/\alpha} - 1$. Es gilt

$$\lim_{t \uparrow \infty} \frac{1 - F(t \cdot x)}{1 - F(t)} = \lim_{t \uparrow \infty} \frac{\frac{1}{(1+tx)^\alpha}}{\frac{1}{(1+t)^\alpha}} = \lim_{t \uparrow \infty} \frac{(1+t)^\alpha}{(1+tx)^\alpha} = x^{-\alpha}, \text{ d.h. die Grenzverteilung}$$

von $Z_n = \frac{X_{n:n} - b_n}{a_n}$ liegt in der *Fréchet-Klasse*. Als Normalisierungs-konstanten können $a_n = \beta_n$, $b_n = 0$ für $n \in \mathbb{N}$ gewählt werden.

Studienjahre (1971–1977)

Beispiele:

Die *Pareto-Verteilung* $\mathcal{Pa}(\alpha)$ mit $\alpha > 0$ besitzt die Verteilungsfunktion

$$F(x) = 1 - \frac{1}{(1+x)^\alpha} \text{ für } x \geq 0.$$

Hier ist $x_{\sup} = \infty$ und $\beta_n = F^{-1}\left(1 - \frac{1}{n}\right) = n^{1/\alpha} - 1$.

Mit $a_n = \beta_n$, $b_n = 0$ für $n \in \mathbb{N}$ erhält man auch direkt

$$\begin{aligned} P(Z_n \leq x) &= P\left(\frac{X_{n:n} - b_n}{a_n} \leq x\right) = P(X_{n:n} \leq a_n x) = F^n(\beta_n x) \\ &= \left(1 - \frac{1}{(1 + \beta_n x)^\alpha}\right)^n \approx \left(1 - \frac{1}{(n^{1/\alpha} x)^\alpha}\right)^n = \left(1 - \frac{1}{n} \cdot \frac{1}{x^\alpha}\right)^n \rightarrow \exp(-x^{-\alpha}). \end{aligned}$$

Studienjahre (1971–1977)

Beispiele:

Die *stetige Gleicherteilung* $\mathcal{U}[a, b]$ mit $a < b < \infty$ besitzt die Verteilungsfunktion

$$F(x) = \frac{x-a}{b-a} \text{ für } a \leq x \leq b.$$

Hier ist $x_{\text{sup}} = b < \infty$ und $\beta_n = F^{-1}\left(1 - \frac{1}{n}\right) = a + (b-a) \cdot \frac{n-1}{n}$. Es gilt

$$\lim_{h \downarrow 0} \frac{1 - F(x_{\text{sup}} - x \cdot h)}{1 - F(x_{\text{sup}} - h)} = \lim_{h \downarrow 0} \frac{x \cdot h}{h} = x \text{ mit } \alpha = 1, \text{ d.h. die Grenzverteilung von}$$

$Z_n = \frac{X_{n:n} - b_n}{a_n}$ liegt in der *Weibull-Klasse*. Als Normalisierungskonstanten

können $a_n = b - \beta_n = \frac{b-a}{n}$, $b_n = b$ für $n \in \mathbb{N}$ gewählt werden.

Studienjahre (1971–1977)

- Ist die ursprüngliche Verteilung bereits aus der Fréchet-/ Weibull-/ Gumbel-Klasse, so auch die (dann stets existierende) Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$. Man nennt diese Verteilungsklassen daher auch *max-stabil*.
- Für diskrete Verteilungen über Teilmengen von $\mathbb{Z}$ existiert für *keine Wahl* normalisierender Konstanten eine nicht-degenerierte Grenzverteilung von $Z_n = \frac{X_{n:n} - b_n}{a_n}$.
- Das *theoretische* Quantil $\beta_n := F^{-1}\left(1 - \frac{1}{n}\right)$, $n \in \mathbb{N}$ entspricht in gewisser Weise dem *empirischen* Quantil $X_{n:n}$ ($\rightarrow$ Q-Q Plot).

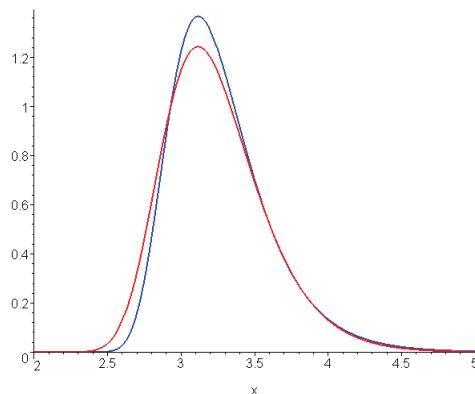
Studienjahre (1971–1977)

Weitere Beispiele:

Die Grenzverteilung der normalisierten Maxima von Zufallsvariablen mit einer *Standard-Normalverteilung* $\mathcal{N}(0,1)$ liegt in der *Gumbel-Klasse*, mit

$$a_n = \frac{1}{\sqrt{2\ln n}} \quad \text{und} \quad b_n = \sqrt{2\ln n} - \frac{1}{2} \cdot \frac{\ln(\ln n) + \ln(4\pi)}{\sqrt{2\ln n}} \quad \text{für } n \geq 2.$$

Rechts: Dichte von $X_{n:n}$ für $n = 1.000$
[rot] und passend skalierte Dichte der
Gumbelverteilung [blau]



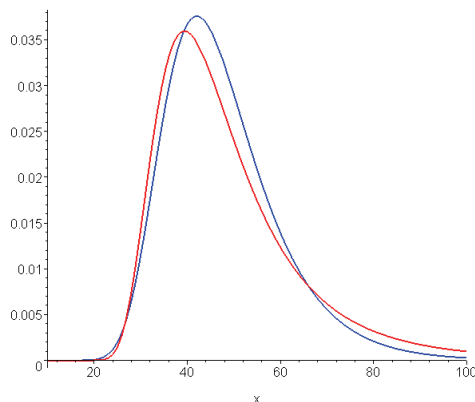
Studienjahre (1971–1977)

Weitere Beispiele:

Die Grenzverteilung der normalisierten Maxima von Zufallsvariablen mit einer *Standard-Lognormalverteilung* $\mathcal{LN}(0,1)$ liegt ebenfalls in der *Gumbel-Klasse*, mit

$\tilde{a}_n = a_n \cdot \exp(b_n)$ und $\tilde{b}_n = \exp(b_n)$ für $n \geq 2$ mit a_n, b_n wie zuvor.

Rechts: Dichte von $X_{n:n}$ für $n = 10.000$
[rot] und passend skalierte Dichte der
Gumbelverteilung [blau]



Studienjahre (1971–1977)

Empirische Verteilungsfunktion:

Interpretation:

Die empirische Verteilungsfunktion $\hat{F}_n$ ist die Verteilungsfunktion einer (zufälligen) diskreten Gleichverteilung auf den „Punkten“ $X_1, \dots, X_n$.

Satz von Glivenko-Cantelli (Zentralsatz der Statistik):

Die empirische Verteilungsfunktion $\hat{F}_n$ konvergiert gleichmäßig gegen die „wahre“ Verteilungsfunktion F für $n \rightarrow \infty$, d.h. es gilt

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| = 0.$$

Studienjahre (1971–1977)

Bemerkung: Die schwächere (punktweise) Aussage

$$\lim_{n \rightarrow \infty} |\hat{F}_n(x) - F(x)| = 0 \text{ für alle } x \in \mathbb{R}$$

ist eine direkte Konsequenz des Gesetzes der großen Zahlen (GGZ), angewendet auf die Folge $B(1, F(x))$ -binomialverteilter Zufallsvariablen $Y_n(x) := \mathbb{1}_{(-\infty, x]}(X_n)$ für $n \in \mathbb{N}$, $x \in \mathbb{R}$.

Studienjahre (1971–1977)

Ordnungsstatistiken:

Sind $X_1, \dots, X_n$ ($n \in \mathbb{N}$) unabhängige und identisch verteilte Zufallsvariablen mit der (theoretischen) Verteilungsfunktion F , so ist die k -te Ordnungsstatistik $X_{k:n}$ als ein Quantil der empirischen Verteilungsfunktion $\hat{F}_n$ definiert durch

$$X_{k:n} := \hat{F}_n^{-1}\left(\frac{k}{n}\right) \text{ für } k = 1, \dots, n.$$

Studienjahre (1971–1977)

Ordnungsstatistiken:

Interpretation: Die k -te Ordnungsstatistik entspricht dem k -größten Wert unter den $X_1, \dots, X_n$, d.h. es gilt

$$\min\{X_1, \dots, X_n\} = X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n} = \max\{X_1, \dots, X_n\}.$$

Falls die zu Grunde liegende Verteilung diskrete Anteile besitzt, können hierbei so genannte *Bindungen* (d.h. gleiche Werte) mit positiver Wahrscheinlichkeit auftreten. Im Fall stetiger Verteilungen geschieht dies nur mit Wahrscheinlichkeit Null.

Studienjahre (1971–1977)

Ordnungsstatistiken:

Hieraus ergibt sich im stetigen Fall eine alternative Darstellung der empirischen Verteilungsfunktion zu

$$\hat{F}_n(x) = \begin{cases} 0 & \text{wenn } x < X_{1:n} \\ \frac{k}{n} & \text{wenn } X_{k:n} \leq x < X_{k+1:n} \text{ für } k \in \{1, \dots, n-1\} \\ 1 & \text{wenn } x \geq X_{n:n} \end{cases}$$

Studienjahre (1971–1977)

Statistische Eigenschaften von Ordnungsstatistiken:

Für die *Verteilungsfunktion* einer Ordnungsstatistik gilt:

$$F_{X_{k:n}}(x) = P(X_{k:n} \leq x) = \sum_{i=k}^n \binom{n}{i} F^i(x) [1 - F(x)]^{n-i} \text{ für } x \in \mathbb{R}, k = 1, \dots, n.$$

Beweisskizze:

$$\begin{aligned} P(X_{k:n} \leq x) &= P\left(\hat{F}_n^{-1}\left(\frac{k}{n}\right) \leq x\right) = P\left(\frac{k}{n} \leq \hat{F}_n(x)\right) = P\left(\sum_{i=1}^n Y_i(x) \geq k\right) \\ &= \sum_{i=k}^n \binom{n}{i} F^i(x) [1 - F(x)]^{n-i}. \end{aligned}$$

Studienjahre (1971–1977)

Statistische Eigenschaften von Ordnungsstatistiken:

Für die *Verteilungsdichte* einer Ordnungsstatistik gilt, sofern F die Dichte f besitzt:

$$f_{X_{k:n}}(x) = \frac{f(x)}{B(k, n-k+1)} F^{k-1}(x) [1-F(x)]^{n-k} \quad \text{für } x \in \mathbb{R}, k = 1, \dots, n,$$

mit der Beta-Funktion $B(a, b) = \frac{\Gamma(a) \cdot \Gamma(b)}{\Gamma(a+b)}$ für $a, b > 0$.

Im Falle der Ganzzahligkeit gilt hier einfacher

$$B(a, b) = \frac{(a-1)!(b-1)!}{(a+b-1)!} = \frac{1}{b \cdot \binom{a+b-1}{a-1}} = \frac{1}{a \cdot \binom{a+b-1}{b-1}}.$$

$$\begin{aligned}f_{X_{k:n}}(x) &= \frac{d}{dx} F_{X_{k:n}}(x) = \sum_{i=k}^n \binom{n}{i} \frac{d}{dx} \left\{ F^i(x) [1 - F(x)]^{n-i} \right\} \\&= f(x) \sum_{i=k}^n \binom{n}{i} \left\{ i F^{i-1}(x) [1 - F(x)]^{n-i} - (n-i) F^i(x) [1 - F(x)]^{n-i-1} \right\} \\&= f(x) \left\{ n \sum_{i=k-1}^{n-1} \binom{n-1}{i} F^i(x) [1 - F(x)]^{n-1-i} - n \sum_{i=k}^{n-1} \binom{n-1}{i} F^i(x) [1 - F(x)]^{n-1-i} \right\} \\&= n \binom{n-1}{k-1} f(x) F^{k-1}(x) [1 - F(x)]^{n-k} = \frac{f(x)}{B(k, n-k+1)} F^{k-1}(x) [1 - F(x)]^{n-k} .\end{aligned}$$

Studienjahre (1971–1977)

Ist die Verteilungsfunktion F stetig, so gilt allgemein: Die Folge der Ordnungsstatistiken bildet eine (endliche) Markov-Kette mit Übergangswahrscheinlichkeiten

$$P(X_{k:n} \leq x | X_{k-1:n} = y) = F_{X_{k:n}}(x | X_{k-1:n} = y) = 1 - \left(\frac{1 - F(x)}{1 - F(y)} \right)^{n-k+1}, x > y, 2 \leq k \leq n.$$

Diese Eigenschaft gilt nicht bei diskreten Verteilungen!

Ordnungsstatistiken bei speziellen Verteilungen:

- Sind die Zufallsvariablen unabhängig und über $[0,1]$ stetig gleichverteilt, so ist $X_{k:n}$ $\mathcal{B}(k, n-k+1)$ Beta-verteilt mit der Dichte

$$f_{X_{k:n}}(x) = \frac{x^{k-1}(1-x)^{n-k}}{B(k, n-k+1)} \text{ für } 0 \leq x \leq 1.$$

- Sind die Zufallsvariablen unabhängig $\mathcal{E}(\lambda)$ -exponentialverteilt mit $\lambda > 0$, dann sind die Zuwächse (sukzessive Differenzen) $X_{1:n}, X_{2:n} - X_{1:n}, X_{3:n} - X_{2:n}, \dots, X_{n:n} - X_{n-1:n}$ ebenfalls unabhängig und exponentialverteilt, wobei $X_{k:n} - X_{k-1:n}$ einer $\mathcal{E}((n-k+1)\lambda)$ -Verteilung für $k = 1, \dots, n$ folgt. Dies impliziert u.a. für große Werte von k :

$$E(X_{n:n}) = \sum_{i=1}^n \frac{1}{i} \approx \frac{\ln n}{\lambda}, \quad \text{Var}(X_{n:n}) = \frac{1}{\lambda^2} \sum_{i=1}^n \frac{1}{i^2} \approx \frac{\pi^2}{6\lambda^2}.$$

Studienjahre (1971–1977)

Satz von Smirnov (I): Haben unter den Bedingungen des Satzes von Fisher-Tippett die normalisierten Maxima $Z_n = \frac{X_{n:n} - b_n}{a_n}$ eine Grenzverteilung in der Fréchet-/ Weibull-/ Gumbel-Klasse mit der Grenzverteilungsfunktion G , so haben die normalisierten Ordnungsstatistiken $Z_{n-k:n} = \frac{X_{n-k:n} - b_n}{a_n}$ für festes $k \in \mathbb{N}$ ebenfalls eine Grenzverteilung, und zwar mit der Grenzverteilungsfunktion ($\rightarrow$ Poisson-Verteilung)

$$\begin{aligned}\lim_{n \rightarrow \infty} P(Z_{n-k:n} \leq x) &= \lim_{n \rightarrow \infty} \sum_{i=n-k}^n \binom{n}{i} F^i(a_n x + b_n) [1 - F(a_n x + b_n)]^{n-i} \\ &= G(x) \sum_{j=0}^k \frac{[-\ln G(x)]^j}{j!} \quad \text{für } x \in \mathbb{R}.\end{aligned}$$

Studienjahre (1971–1977)

Satz von Smirnov (II): Seien $\{X_n\}_{n \in \mathbb{N}}$ unabhängige Zufallsvariablen und $\{k_n\}_{n \in \mathbb{N}}$ eine Folge natürlicher Zahlen mit $\lim_{n \rightarrow \infty} \sqrt{n} \cdot \left(\frac{k_n}{n} - \beta \right) = 0$ für ein $\beta > 0$. Falls normalisierende Konstanten $A_n(\beta) > 0, B_n(\beta) \in \mathbb{R}$ existieren mit

$$\lim_{n \rightarrow \infty} P\left(A_n(\beta) \cdot (X_{n-k_n:n} - B_n(\beta)) \leq x\right) = G^*(x) \quad \text{für alle } x \in \mathcal{C}(G^*),$$

dann korrespondiert G^* zu genau einer der folgenden vier (!) Typen:

$$\begin{aligned} G_{1,\alpha,c}(x) &= \begin{cases} 0, & \text{falls } x < 0 \\ \Phi(c x^\alpha) & \text{falls } x \geq 0 \end{cases} & G_{2,\alpha,c}(x) &= \begin{cases} \Phi(-c(-x)^\alpha), & \text{falls } x < 0 \\ 1, & \text{falls } x \geq 0 \end{cases} \\ G_{3,\alpha,c,d}(x) &= \begin{cases} \Phi(-c(-x)^\alpha), & \text{falls } x < 0 \\ \Phi(c x^\alpha) & \text{falls } x \geq 0 \end{cases} & G_4(x) &= \begin{cases} 0, & \text{falls } x < -1 \\ 1/2, & \text{falls } -1 \leq x \leq 1 \quad (\alpha, c, d > 0) \\ 1, & \text{falls } x > 1 \end{cases} \end{aligned}$$

Studienjahre (1971–1977)

Bemerkung: Zur Berechnung der normalisierenden Konstanten $A_n(\beta) > 0, B_n(\beta) \in \mathbb{R}$ in Smirnov's Satz II können wir von folgendem Hilfsresultat Gebrauch machen:

Falls für $p_n(x) := 1 - F\left(\frac{x}{A_n(\beta)} + B_n(\beta)\right)$, $n \in \mathbb{N}, x \in \mathbb{R}$ gilt:

$\lim_{n \rightarrow \infty} p_n(x) = \beta$ und $\lim_{n \rightarrow \infty} \sqrt{n} \frac{\beta - p_n(x)}{\sqrt{\beta(1-\beta)}} = \tau(x)$ für $x \in \mathbb{R}$, dann gilt auch

$$\lim_{n \rightarrow \infty} P\left(A_n(\beta) \cdot (X_{n-k_n:n} - B_n(\beta)) \leq x\right) = \Phi(\tau(x)) \text{ für alle } x \in \mathbb{R}.$$

Besitzt F eine Dichte f mit $f(F^{-1}(1-\beta)) > 0$, dann ist $\tau(x) = x$, und eine mögliche Wahl für die normalisierenden Konstanten ist

$$A_n(\beta) = \sqrt{n} \frac{f(F^{-1}(1-\beta))}{\sqrt{\beta(1-\beta)}} \text{ und } B_n(\beta) = -F^{-1}(1-\beta) \text{ für } 0 < \beta < 1 \text{ und } n \in \mathbb{N}.$$

Studienjahre (1971–1977)

Beispiel: Sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit einer $\mathcal{E}(1)$ -Exponentialverteilung. Mit

$$k_n := \lfloor \beta \cdot n \rfloor, A_n(\beta) := \sqrt{\frac{n\beta}{1-\beta}}, B_n(\beta) := -\ln \beta \text{ für } \beta > 0, n \in \mathbb{N} \text{ ergibt sich}$$

$$\lim_{n \rightarrow \infty} P\left(A_n(\beta) \cdot (X_{n-k_n:n} - B_n(\beta)) \leq x\right) = \Phi(x) \text{ für alle } x \in \mathbb{R}.$$

Beweis: Es gilt $\left|\frac{k_n}{n} - \beta\right| \leq \frac{1}{n}$ und daher $\lim_{n \rightarrow \infty} \sqrt{n} \cdot \left(\frac{k_n}{n} - \beta\right) = 0$. Ferner ist

$$p_n(x) = \exp\left(-\frac{x}{A_n(\beta)} - B_n(\beta)\right) = \beta \exp\left(-x \sqrt{\frac{1-\beta}{n\beta}}\right) \text{ mit } \lim_{n \rightarrow \infty} p_n(x) = \beta \text{ und}$$

$$\lim_{n \rightarrow \infty} \sqrt{n} \frac{\beta - p_n(x)}{\sqrt{\beta(1-\beta)}} = \sqrt{\frac{\beta}{1-\beta}} \lim_{n \rightarrow \infty} \sqrt{n} \left(1 - \exp\left(-x \sqrt{\frac{1-\beta}{n\beta}}\right)\right) = x.$$

Studienjahre (1971–1977)

Beispiel: Sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit einer $\mathcal{E}(1)$ -Exponentialverteilung. Mit

$$k_n := \lfloor \beta \cdot n \rfloor, A_n(\beta) := \sqrt{\frac{n\beta}{1-\beta}}, B_n(\beta) := -\ln \beta \text{ für } \beta > 0, n \in \mathbb{N} \text{ ergibt sich}$$

$$\lim_{n \rightarrow \infty} P\left(A_n(\beta) \cdot (X_{n-k_n:n} - B_n(\beta)) \leq x\right) = \Phi(x) \text{ für alle } x \in \mathbb{R}.$$

Alternativ:

$$A_n(\beta) = \sqrt{n} \frac{f(F^{-1}(1-\beta))}{\sqrt{\beta(1-\beta)}} = \sqrt{\frac{n}{\beta(1-\beta)}} \exp(-(-\ln(\beta))) = \beta \sqrt{\frac{n}{\beta(1-\beta)}} = \sqrt{\frac{n\beta}{1-\beta}}$$

und $B_n(\beta) = -F^{-1}(1-\beta) = -(-\ln(\beta)) = \ln \beta$ für $0 < \beta < 1$ und $n \in \mathbb{N}$,

wie zuvor.

Studienjahre (1971–1977)

Record Values und Record Times:

Sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit stetiger Verteilungsfunktion F . Die Folge $\{U_n\}_{n \in \mathbb{Z}^+}$ der (upper) Record Times ist rekursiv definiert durch

$$U_0 := 1, U_n := \inf \{k > U_{n-1} \mid X_k > X_{U_{n-1}}\} \text{ für } n \in \mathbb{N}.$$

Die $\{X_{U_n}\}_{n \in \mathbb{Z}^+}$ heißen (upper) Record Values.

Eigenschaften:

- U_n und X_{U_n} sind messbar und f.s. endlich für alle $n \in \mathbb{N}$
- Die $\{U_n\}_{n \in \mathbb{Z}^+}$ bilden eine homogene Markov-Kette mit

$$P(U_n = k \mid U_{n-1} = j) = \frac{j}{k(k-1)} \text{ für } 1 \leq j < k \in \mathbb{N}, n \in \mathbb{N}.$$

Studienjahre (1971–1977)

Record Values und Record Times:

Sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit stetiger Verteilungsfunktion F . Die Folge $\{U_n\}_{n \in \mathbb{Z}^+}$ der (upper) Record Times ist rekursiv definiert durch

$$U_0 := 1, U_n := \inf \{k > U_{n-1} \mid X_k > X_{U_{n-1}}\} \text{ für } n \in \mathbb{N}.$$

Die $\{X_{U_n}\}_{n \in \mathbb{Z}^+}$ heißen (upper) Record Values.

Eigenschaften:

- Die $\{(U_n, X_{U_n})\}_{n \in \mathbb{Z}^+}$ bilden eine homogene Markov-Kette mit

$$P(U_n = k, X_{U_n} > y \mid U_{n-1} = j, X_{U_{n-1}} = x) = F^{k-j-1}(x) \cdot (1 - F(y)) \text{ f.s.}$$

- Die $\{X_{U_n}\}_{n \in \mathbb{Z}^+}$ bilden eine homogene Markov-Kette mit

$$P(X_{U_n} > y \mid X_{U_{n-1}} = x) = \frac{1 - F(y)}{1 - F(x)} \text{ f.s.} \quad (1 \leq j < k \in \mathbb{N}, x \leq y \in \mathbb{R})$$

Studienjahre (1971–1977)

Record Values und Record Times:

Sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige Folge von Zufallsvariablen mit stetiger Verteilungsfunktion F . Die Folge $\{U_n\}_{n \in \mathbb{Z}^+}$ der (upper) Record Times ist rekursiv definiert durch

$$U_0 := 1, U_n := \inf \{k > U_{n-1} \mid X_k > X_{U_{n-1}}\} \text{ für } n \in \mathbb{N}.$$

Die $\{X_{U_n}\}_{n \in \mathbb{Z}^+}$ heißen (upper) Record Values.

Eigenschaften:

- Mit $\Delta_n := U_n - U_{n-1}$, $n \in \mathbb{N}$ und $\Delta_0 := 0$ gilt:

$$P(\Delta_n = k \mid X_{U_{n-1}} = x) = F^{k-1}(x) \cdot (1 - F(x)) \text{ f.s.}$$

Studienjahre (1971–1977)

Studienabschluss:

Diplomarbeit mit dem Thema

*Eine Charakterisierung der Poisson- und der Exponentialverteilung
durch Extremwerte unabhängiger Zufallsvariablen*

Erste Publikation daraus:

An Application to Stochastic Simulation, II. Symposium über Operations Research, RWTH Aachen 1977. In: *Methods of Operations Research* 29 (1978), 738 - 749.

An Application of Record Values to Stochastic Simulation

Dietmar Pfeifer, Aachen

Abstract:

Using record values rather than record times, a new method for testing the independence of random number generators with continuous cumulative distribution function (c.d.f.) is proposed as well as another algorithm for generating Poisson-distributed random variables (r.v.'s) from any continuous distribution.

Studienjahre (1971–1977)

Literatur:

M. R. Leadbetter, G. Lindgren and H. Rootzen (1983): *Extremes and Related Properties of Random Sequences and Processes*. Springer, N.Y.

Sidney I. Resnick (1987): *Extreme Values, Regular Variation and Point Processes*. Springer, N.Y.

B.C. Arnold, N. Balakrishnan and H.N. Nagaraja (1998): *Records*. Wiley, N.Y.

S. Coles (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.

V.B. Nevzorov (2001): *Records: Mathematical Theory*. Translations of Mathematical Monographs, Am. Math. Society, Vol. 194, Providence.

H.A. David, H.N. Nagaraja (2003): *Order Statistics*. 3. Aufl., Wiley, N.Y.

J. Beirlant, Y. Goegebeur, J. Segers and J. Teugels (2004): *Statistics of Extremes*. Theory and Applications. Wiley, N.Y.

III. Assistentenjahre (1977-1986)



Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

1. Bekannte Grenzwertsätze:

$$\lim_{n \rightarrow \infty} \frac{\ln U_n}{n} = \lim_{n \rightarrow \infty} \frac{\ln \Delta_n}{n} = 1 \text{ f.s.} \quad (\text{Rényi 1962})$$

$$\text{w-lim}_{n \rightarrow \infty} \frac{\ln U_n - n}{\sqrt{n}} = \text{w-lim}_{n \rightarrow \infty} \frac{\ln \Delta_n - n}{\sqrt{n}} = \mathcal{N}(0,1) \quad (\text{Neuts 1967})$$

$$\lim_{n \rightarrow \infty} \sup_{\inf} \frac{\ln U_n - n}{\sqrt{2n \ln \ln n}} = \lim_{n \rightarrow \infty} \sup_{\inf} \frac{\ln \Delta_n - n}{\sqrt{2n \ln \ln n}} = \pm 1 \text{ f.s.} \quad (\text{Holmes/Strawderman 1969})$$

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

1. Bekannte Grenzwertsätze:

Neue eigene Resultate:

$$E(\ln \Delta_n) = n - \gamma + \mathcal{O}\left(\frac{n}{2^n}\right), \quad \text{Var}(\ln \Delta_n) = n - \frac{\pi^2}{6} + \mathcal{O}\left(\frac{n^2}{2^n}\right) \quad (1980)$$

$$E(\ln U_n) = n + 1 - \gamma + \mathcal{O}\left(\frac{1}{2^n}\right), \quad \text{Var}(\ln U_n) = n + 1 - \frac{\pi^2}{6} + \mathcal{O}\left(\frac{n^3}{2^n}\right) \quad (1984)$$

Hier bezeichnet $\gamma = 0,577215\dots$ die Euler'sche Konstante (in anderen Publikationen auch mit C bezeichnet).

Z. Wahrscheinlichkeitstheorie verw. Gebiete
66, 293–296 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete
© Springer-Verlag 1984

A Note on Moments of Certain Record Statistics

D. Pfeifer

Institut für Statistik und Wirtschaftsmathematik, Technische Hochschule,
Wüllnerstr. 3, 5100 Aachen

Summary. Asymptotic expansions for the mean and variance of the logarithms of record and inter-record times from i.i.d. random variables are given, refining rough estimations from Rényi's and Neuts' Central Limit Theorem for these record statistics.

Assistentenjahre (1977–1986)

n	k	$P(\Delta_n \leq k)$	$\Phi\left(\frac{\ln k - n}{\sqrt{n}}\right)$	$\Phi\left(\frac{\ln k - (n - C)}{\sqrt{n + \frac{1}{6}}}\right)$
2	1	0,2500	0,0786	0,2281
	2	0,3889	0,1777	0,3512
	3	0,4792	0,2619	0,4326
	4	0,5433	0,3322	0,4924
	5	0,5917	0,3912	0,5389
	10	0,7255	0,5847	0,6775
	20	0,8264	0,7593	0,7950
3	50	0,9114	0,9118	0,9039
	1	0,1250	0,0416	0,1305
	2	0,2126	0,0915	0,2111
	5	0,3755	0,2110	0,3529
	10	0,5147	0,3436	0,4778
	20	0,6455	0,4990	0,6048
	50	0,7839	0,7007	0,7552
4	100	0,8582	0,8230	0,8444
	1	0,0625	0,0228	0,0748
	5	0,2209	0,1160	0,2227
	10	0,3325	0,1980	0,3186
	20	0,4577	0,3078	0,4287
	50	0,6186	0,4825	0,5816
	100	0,7223	0,6189	0,6906
5	500	0,8837	0,8659	0,8800
	1	0,0313	0,0127	0,0431
	5	0,1234	0,0647	0,1376
	10	0,2002	0,1138	0,2054
	20	0,2992	0,1850	0,2899
	50	0,4494	0,3133	0,4215
	100	0,5631	0,4299	0,5282
5	500	0,7782	0,7065	0,7565
	1000	0,8426	0,8032	0,8325

(Φ denotes the distribution function of the standard normal distribution)

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

25.1. Pfeifer's model

Pfeifer (1982, 1984c) considered records in a scheme of series of i.i.d. random variables $X_{n1}, X_{n2}, \dots$, $n = 0, 1, 2, \dots$. Let F_n be the distribution function of random variables $X_{n1}, X_{n2}, \dots$. The sequence $\Delta(n)$, $n = 0, 1, 2, \dots$, of inter-record times is recursively defined as follows:

$$\Delta(0) = 0 \text{ and } \Delta(n) = \min\{j : X_{nj} > X_{n-1:\Delta(n-1)}\}, \quad n = 1, 2, \dots$$

Then, following Pfeifer, $L(n) = 1 + \Delta(1) + \Delta(2) + \dots + \Delta(n)$ and $X(n) = X_{u,\Delta(n)}$, $n = 0, 1, 2, \dots$, form sequences of record times and record values. It has been shown that under some additional restrictions, $(\Delta(n), X(n))$, $n = 0, 1, \dots$ is a Markov chain. Pfeifer has also showed that if $F_n(x) = 1 - \exp(-\lambda_n x)$, $x > 0$, where $\lambda_n > 0$, $n = 1, 2, \dots$ then the random variables $X(n-1)$ and $\delta(n) = X(n) - X(n-1)$ are independent, and the distribution function of $\delta(n)$ coincides with F_n . Pfeifer (1984c) proved in addition some limit theorems for $\log \Delta(n)$.

The record model described above reflects situations where conditions of an experiment change after the occurrence of a new record value. For example, after the destruction of an old component, a modified one has to be used. Pfeifer's model was further discussed by Deheuvels (1984b) and Gupta (1984). Reliability properties of record values in Pfeifer's scheme have been studied by Kamps (1994). Kamps (1995, Section 1.7) has also considered k_n -records in Pfeifer's record model.

Handbook of Statistics 16: *Order Statistics: Theory & Methods*, North Holland, 1998

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

6.5 THE PFEIFER MODEL

The F^α record model assumed that the distribution changes with each observation. Pfeifer (1982) introduced an interesting variation of the classical model wherein the successive (upper) record values constitute a Markov chain with nonstationary transition distribution given by

$$P(R_n > y \mid R_{n-1} = x) = \frac{1 - F_n(y)}{1 - F_n(x)}, \quad y \geq x. \quad (6.5.1)$$

Such a dependence structure for the record value sequence can be produced as follows: Suppose we have a double array of independent random variables $\{X_{0j}, X_{nj}; n, j \geq 1\}$ such that X_{nj} has cdf F_n , $n \geq 0$. Now take $R_0 = X_{01}$ and define $\Delta_n = \min\{j : X_{nj} > R_{n-1}\}$ and $R_n = X_{n, \Delta_n}$ for $n \geq 1$. We call this setting the *Pfeifer record model*. When the F_n 's are continuous, in order that all these record statistics are well defined, it is necessary to assume that the $F_n^{-1}(1)$'s form a nondecreasing sequence. If an F_n is discrete such that $P(X_{nj} = F_n^{-1}(1))$ is positive, then one needs to assume that $F_{n+1}^{-1}(1) > F_n^{-1}(1)$. To keep our discussion simple, we will assume hereinafter that the F_n 's are all continuous cdf's with nondecreasing upper limits. Note that while the X_{nj} 's are independent, the record generating observations are not.

B.C. Arnold et al. (1998): *Records*. Wiley, N.Y.

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

Pfeifer's scheme. The next model of records can also be connected with sports. The German athlete Uwe Hohn on July 20, 1984, was the first to exceed the 100 meter mark in javelin throwing. His javelin landed at 104.8 m, which made this sport dangerous for spectators. Then the design of the javelin was changed by shifting its center of gravity forward, which reduced the length of throws. The records for the new javelin have not yet surpassed Hohn's. But the best throws nowadays come close to 100 m, and undoubtedly the rules will change again as soon as this mark is exceeded. The idea of varying distributions of the random variables in the sequence $X_1, X_2, \dots$ is basic for the scheme by Pfeifer (1982, 1984).

Let $\{X_{nk}, n \geq 1, k \geq 1\}$ be a double array of independent random variables having in each row $X_{n1}, X_{n2}, \dots$ a common d.f. F_n . It will be convenient to define first the inter-record times $\Delta(n) = L(n) - L(n-1)$ rather than the record times $L(n)$. They are defined as

$$\Delta(1) = 1, \quad \Delta(n+1) = \min\{k: X_{n+1,k} > X_{n,\Delta(n)}\}, \quad n = 1, 2, \dots$$

Then the record times $L(n)$ and the record values $X(n)$ in Pfeifer's scheme are given by

$$L(n) = \Delta(1) + \dots + \Delta(n)$$

and

$$X(n) = X_{n,\Delta(n)}, \quad n = 1, 2, \dots$$

If $F_1 = F_2 = \dots$, then these definitions coincide with the definitions of the classical records statistics.

V.B. Nevzorov (2001): *Records. Mathematical Theory*. Transl. of Math. Monographs 194, AMS, Providence.

D. Pfeifer: Abschiedsvorlesung (Sommersemester 2016)

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

**"Record Values"
in einem stochastischen Modell
mit nicht-identischen Verteilungen**

Dietmar Pfeifer

"Record Values" in einem
stochastischen Modell mit nicht-identischen
Verteilungen

Von der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinisch-Westfälischen Technischen Hochschule Aachen

zur Erlangung
des akademischen Grades
eines Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von

Diplom-Mathematiker
Dietmar Pfeifer
aus Wuppertal-Elberfeld

Referent: Professor Dr. B. Rauhut

Korreferent: Professor Dr. O. Krafft

Tag der mündlichen Prüfung: 27. November 1979

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

J. Appl. Prob. **19**, 127–135 (1982)

Printed in Israel

0021–9002/82/010127–09 \$01.15

© *Applied Probability Trust* 1982

**CHARACTERIZATIONS OF EXPONENTIAL DISTRIBUTIONS BY
INDEPENDENT NON-STATIONARY RECORD INCREMENTS**

DIETMAR PFEIFER,* *Technical University Aachen*

Abstract

A non-homogeneous version of the classical record process is presented which allows two different characterizations of exponential distributions by independent non-stationary record increments. A connection with the interarrival times of the corresponding record counting process (which is pure birth) is also pointed out.

RECORD VALUES; MARKOV ADDITIVE CHAIN; EXPONENTIAL DISTRIBUTION; INDEPENDENT NON-STATIONARY INCREMENTS; COUNTING PROCESS; PURE BIRTH PROCESS

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

2. Verallgemeinerungen des „klassischen“ Modells:

J. Appl. Prob. **19**, 664–667 (1982)

Printed in Israel

0021–9002/82/030664–04 \$00.65

© *Applied Probability Trust* 1982

THE STRUCTURE OF ELEMENTARY PURE BIRTH PROCESSES

DIETMAR PFEIFER,* *Technical University, Aachen*

Abstract

A complete characterization of elementary pure birth processes is given by means of record counting processes from independent (non-identically) distributed random variables.

PURE BIRTH PROCESS; RECORD VALUES; COUNTING PROCESS

Assistentenjahre (1977–1986)

Motivationen für eigene Forschungen:

3. Halbgruppen von Operatoren:

MATH. SCAND. 10 (1962), 153–162

ON THE EXPONENTIAL FORMULAS OF SEMI-GROUP THEORY

KAI LAI CHUNG¹

1.

The purpose of this paper is to present a simple unified approach to a group of theorems in semi-group theory called the “exponential formulas”, due to Hille, Phillips, Widder and D. G. Kendall [1, p. 354]. A more general and apparently new formula is arrived at, which includes some known cases. It turns out that these formulas are in essence summability methods which are best comprehended from the point of view of elementary probability theory. They are all in the spirit of S. Bernstein’s proof of Weierstrass’s approximation theorem, the same idea being present in M. Riesz’s proof of Hille’s first exponential formula (see [1, p. 314]). Whereas the details here are just a little simpler than in [1], it seems of some interest to exhibit the general pattern and to reduce the proofs to routine verifications. The reader who is not acquainted with the language of probability should have no difficulty in translating everything into the language of classical analysis. But the probability way of thinking is really germane to the subject.

1982年 第2卷 第4期
Vol. 2, No. 4, 1982

数学研究与评论
JOURNAL OF MATHEMATICAL
RESEARCH & EXPOSITION

On a General Probabilistic Representation Formula for Semigroups of Operators*

D. Pfeifer

Postal address:
Dietmar Pfeifer
Institut für Statistik und Wirtschaftsmathematik
Rheinisch-westfälische Technische Hochschule Aachen
Wüllnerstr. 3
D 5100 Aachen
West Germany

Abstract

In this paper we present a general probabilistic representation theorem for semigroups of operators which essentially comprises almost all known representation theorems of this kind. Especially, so-called first and second main theorems turn out to be special cases of the same general formula. The general theorem also permits the calculation of rates of convergence involving the second modulus of continuity.

Exkurs: Halbgruppen von Operatoren

Definition. Eine Familie $\mathcal{T} = \{T(t) \mid t \geq 0\} \subseteq \mathcal{L}[\mathcal{X}, \mathcal{X}]$ beschränkter linearer Operatoren auf einem Banach-Raum $\mathcal{X}$ heißt *(Operator-)Halbgruppe*, wenn gilt:

- a) $T(s) \circ T(t) = T(s+t)$ für alle $s, t \geq 0$
- b) $T(0) = I$.

Die Operator-Halbgruppe $\mathcal{T}$ heißt *von der Klasse (C_0)* , wenn zusätzlich gilt:

- c) $T(t)f \xrightarrow{s} f$ bei $t \downarrow 0$, für alle $f \in \mathcal{X}$

(d.h. die Abbildung $T(t)f$ ist für jedes $f \in \mathcal{X}$ stark stetig in $t = 0$).

Exkurs: Halbgruppen von Operatoren

Lemma. Für eine Operator-Halbgruppe der Klasse (C_0) bestehen folgende Eigenschaften:

- a) Die Operator-Normen $\|T(t)\|$ sind in jedem endlichen Teilintervall von $\mathbb{R}^+$ beschränkt.
- b) Die Abbildung $T(t)f$ ist für jedes $f \in \mathcal{X}$ stark stetig in jedem Punkt $t = t_0 \in (0, \infty)$.
- c) Es gilt

$$\omega_0 := \inf_{t>0} \left\{ \frac{1}{t} \ln \|T(t)\| \right\} = \lim_{t \rightarrow \infty} \left\{ \frac{1}{t} \ln \|T(t)\| \right\} \in [-\infty, \infty)$$

- d) Zu jedem $\omega > \omega_0$ existiert ein $M(\omega) \geq 1$ mit

$$\|T(t)\| \leq M(\omega) e^{\omega t} \text{ für alle } t \geq 0.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG1 (Translations-Halbgruppe). Wir bezeichnen mit $UCB(\mathbb{R})$ den Banach-Raum der auf $\mathbb{R}$ gleichmäßig stetigen beschränkten Funktionen mit der üblichen Supremums-Norm

$$\|f\|_{UCB} := \sup \{ |f(x)| \mid x \in \mathbb{R} \} \text{ für } f \in UCB(\mathbb{R}).$$

Die durch

$$T(t)f(x) := f(x+t), \quad x \in \mathbb{R}, \quad t \geq 0 \quad \text{für } f \in UCB(\mathbb{R})$$

definierte Halbgruppe der Klasse (C_0) heißt *Halbgruppe der Linkstranslationen*. Hier gilt für alle $t \geq 0$ die Gleichheit $\|T(t)f\|_{UCB} = \|f\|_{UCB}$ für alle $f \in UCB(\mathbb{R})$ und somit $\|T(t)\| = 1$ für alle $t \geq 0$. Dies entspricht der Situation $\omega_0 = 0$ in dem obigen Lemma.

Exkurs: Halbgruppen von Operatoren

Beispiel HG2. Es sei $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein beliebiger Banach-Raum und $\lambda \in \mathbb{R}$ fest gewählt. Für die durch

$$T(t)f := e^{\lambda t} \cdot f, \quad t \geq 0 \quad \text{für } f \in \mathcal{X}$$

definierte Halbgruppe der Klasse (C_0) gilt offensichtlich:

$$\|T(t)f\|_{\mathcal{X}} = e^{\lambda t} \|f\|_{\mathcal{X}} \quad \text{und somit} \quad \|T(t)\| = e^{\lambda t} \quad \text{für alle } t \geq 0. \quad \text{Hier ist also } \omega_0 = \lambda.$$

Zum Nachweis der (C_0) -Eigenschaft beachte:

$$\lim_{t \downarrow 0} \|T(t)f - f\|_{\mathcal{X}} = \left\{ \lim_{t \downarrow 0} |e^{\lambda t} - 1| \right\} \|f\|_{\mathcal{X}} = 0 \quad \text{für } f \in \mathcal{X}.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG3. Wir betrachten den Banach-Raum $\mathcal{X} = \mathbb{R}^2$ mit der $\ell^1(2)$ -Norm. Dann wird für feste $\mu, \lambda \in \mathbb{R}$ durch

$$T(t)f := \frac{1}{2} \begin{bmatrix} e^{\mu t} + e^{\lambda t} & e^{\mu t} - e^{\lambda t} \\ e^{\mu t} - e^{\lambda t} & e^{\mu t} + e^{\lambda t} \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad t \geq 0 \quad \text{für alle } f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \in \mathcal{X}$$

eine Halbgruppe der Klasse (C_0) definiert. Dies folgt aus der leicht nachprüfbaren Beziehung

$$\begin{bmatrix} e^{\mu s} + e^{\lambda s} & e^{\mu s} - e^{\lambda s} \\ e^{\mu s} - e^{\lambda s} & e^{\mu s} + e^{\lambda s} \end{bmatrix} \cdot \begin{bmatrix} e^{\mu t} + e^{\lambda t} & e^{\mu t} - e^{\lambda t} \\ e^{\mu t} - e^{\lambda t} & e^{\mu t} + e^{\lambda t} \end{bmatrix} = 2 \begin{bmatrix} e^{\mu(s+t)} + e^{\lambda(s+t)} & e^{\mu(s+t)} - e^{\lambda(s+t)} \\ e^{\mu(s+t)} - e^{\lambda(s+t)} & e^{\mu(s+t)} + e^{\lambda(s+t)} \end{bmatrix}$$

für $s, t \geq 0$ sowie

$$\lim_{t \downarrow 0} \frac{1}{2} \begin{bmatrix} e^{\mu t} + e^{\lambda t} & e^{\mu t} - e^{\lambda t} \\ e^{\mu t} - e^{\lambda t} & e^{\mu t} + e^{\lambda t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Exkurs: Halbgruppen von Operatoren

Hier gilt

$$\|T(t)f\|_1 \leq \frac{1}{2} \left\{ |e^{\mu t} + e^{\lambda t}| + |e^{\mu t} - e^{\lambda t}| \right\} \|f\|_1 \text{ für } t \geq 0 \text{ und } f \in \mathcal{X},$$

wobei Gleichheit bei geeigneter Wahl der Vorzeichen von f_1 und f_2 erreicht werden kann. Die (C_0) -Eigenschaft folgt aus

$$\lim_{t \downarrow 0} \|T(t)f - f\|_1 \leq \frac{1}{2} \left\{ \lim_{t \downarrow 0} |e^{\lambda t} + e^{\mu t} - 2| + \lim_{t \downarrow 0} |e^{\lambda t} - e^{\mu t}| \right\} \|f\|_1 = 0, f \in \mathcal{X}.$$

Hieraus ergibt sich

$$\|T(t)\| = \begin{cases} e^{\mu t}, & \lambda \leq \mu \\ e^{\lambda t}, & \lambda \geq \mu \end{cases} \text{ für } t \geq 0 \quad \text{und somit } \omega_0 = \begin{cases} \mu, & \lambda \leq \mu \\ \lambda, & \lambda \geq \mu \end{cases} = \max\{\mu, \lambda\}.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG4. Es sei $UCB(\mathbb{R}^+)$ der Banach-Raum der auf $\mathbb{R}^+$ gleichmäßig stetigen beschränkten Funktionen (mit der gleichen Norm wie $UCB(\mathbb{R})$). Wir betrachten den Teilraum

$$\mathcal{X} := \left\{ f \in UCB(\mathbb{R}^+) \mid f(x) = 0 \text{ für alle } x \geq 1 \right\}$$

und hier die Translations-Halbgruppe wie in Beispiel HG1:

$$T(t)f(x) := f(x+t), \quad x \geq 0, \quad t \geq 0 \quad \text{für } f \in \mathcal{X}.$$

Im Gegensatz zu vorher ist jetzt aber $T(t)f = \mathbf{0}$ für $f \in \mathcal{X}$ und $t \geq 1$, also gilt $\omega_0 = -\infty$.

Exkurs: Halbgruppen von Operatoren

Der infinitesimale Erzeuger

In diesem Abschnitt soll näher untersucht werden, ob und in welchem Sinn für gewisse $f \in \mathcal{X}$ "der" Grenzwert $Af := \lim_{t \downarrow 0} \frac{1}{t}(T(t) - I)f$ existiert.

Man gelangt damit zum Begriff des *infinitesimalen Erzeugers* A einer Halbgruppe, der i. Allg. nur auf einem Teilraum $D \subseteq \mathcal{X}$ definiert und dort auch i. Allg. nur abgeschlossen und nicht beschränkt ist. Allerdings spielt er für die Charakterisierung von Halbgruppen eine fundamentale Rolle, gerade auch im Zusammenhang mit Markoff-Prozessen (wie dem Wiener-Prozess), und hat damit z.B. Auswirkungen auf das Berechnen von Optionspreisen in zeitstetigen Finanzmarktmodellen.

Exkurs: Halbgruppen von Operatoren

Abstrakte Differentiation und Integration

Definition (starke / schwache Stetigkeit). Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f: E \rightarrow \mathcal{X}$ eine Abbildung. f heißt

a) *stark stetig* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} \|f(t) - f(t_0)\|_{\mathcal{X}} = 0;$$

b) *schwach stetig* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} |f^*(f(t)) - f^*(f(t_0))| = 0 \text{ für alle } f^* \in \mathcal{X}^*.$$

Exkurs: Halbgruppen von Operatoren

Ist $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ ein weiterer Banach-Raum und $T : E \rightarrow \mathcal{L}[\mathcal{X}, \mathcal{Y}]$, so heit T

c) *gleichmig stetig* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} \|T(t) - T(t_0)\| = 0;$$

d) *stark stetig* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} \|T(t)f - T(t_0)f\|_{\mathcal{Y}} = 0 \text{ fr alle } f \in \mathcal{X};$$

e) *schwach stetig* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} |f^*(T(t)f) - f^*(T(t_0)f)| = 0 \text{ fr alle } f \in \mathcal{X} \text{ und } f^* \in \mathcal{Y}^*.$$

Exkurs: Halbgruppen von Operatoren

Definition (starke / schwache Differenzierbarkeit). Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f: E \rightarrow \mathcal{X}$ eine Abbildung. f heißt

a) *stark differenzierbar* in $t_0 \in E$, wenn ein $g_0 \in \mathcal{X}$ existiert mit

$$\lim_{h \rightarrow 0} \left\| \frac{1}{h} \{f(t_0 + h) - f(t_0)\} - g_0 \right\|_{\mathcal{X}} = 0;$$

$g_0 = \frac{sd}{dt} f(t_0)$ heißt *starke Ableitung von f im Punkt t_0* ;

b) *schwach differenzierbar* in $t_0 \in E$, wenn ein $g_0 \in \mathcal{X}$ existiert mit

$$\lim_{h \rightarrow 0} \left| \frac{1}{h} \{f^*(f(t_0 + h)) - f^*(f(t_0))\} - f^*(g_0) \right| = 0 \text{ für alle } f^* \in \mathcal{X}^*;$$

$g_0 = \frac{wd}{dt} f(t_0)$ heißt *schwache Ableitung von f im Punkt t_0* .

Exkurs: Halbgruppen von Operatoren

Ist $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ ein weiterer Banach-Raum und $T : E \rightarrow \mathcal{L}[\mathcal{X}, \mathcal{Y}]$, so heit T

c) *gleichmig differenzierbar* in $t_0 \in E$, wenn ein $G_0 \in \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ existiert mit

$$\lim_{h \rightarrow 0} \left\| \frac{1}{h} \{T(t_0 + h) - T(t_0)\} - G_0 \right\| = 0;$$

$G_0 = \frac{ud}{dt} T(t_0)$ heit *gleichmige Ableitung von T im Punkt t_0* ;

d) *stark differenzierbar* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} \left\| \frac{1}{h} \{T(t_0 + h) - T(t_0)\} f - G_0 f \right\|_{\mathcal{Y}} = 0 \text{ fr alle } f \in \mathcal{X};$$

$G_0 = \frac{sd}{dt} T(t_0)$ heit *starke Ableitung von T im Punkt t_0* ;

Exkurs: Halbgruppen von Operatoren

e) *schwach differenzierbar* in $t_0 \in E$, wenn gilt:

$$\lim_{t \rightarrow t_0} \left| \frac{1}{h} f^* \left[\{T(t_0 + h) - T(t_0)\} f \right] - f^*(G_0 f) \right| = 0 \text{ für alle } f \in \mathcal{X} \text{ und } f^* \in \mathcal{Y}^*;$$

$G_0 = \frac{wd}{dt} T(t_0)$ heißt *schwache Ableitung* von T im Punkt t_0 .

Bemerkung: Sofern keine Missverständnisse entstehen (können), wird die starke (schwache) Ableitung auch nur mit $\frac{d}{dt} f(t_0)$ bezeichnet.

Exkurs: Halbgruppen von Operatoren

Bemerkung: Beim Begriff der Stetigkeit bzw. der Differenzierbarkeit gilt jeweils, dass die gleichmäßige die starke und diese die schwache Eigenschaft impliziert; jedoch gilt im Allgemeinen nicht die Umkehrung. Beispielsweise impliziert die starke Stetigkeit die schwache wegen

$$\lim_{t \rightarrow t_0} \|f^*(f(t)) - f^*(f(t_0))\| \leq \|f^*\| \cdot \lim_{t \rightarrow t_0} \|f(t) - f(t_0)\|_{\mathcal{X}} = 0 \text{ für alle } f^* \in \mathcal{X}^*.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG5. Wir betrachten für $1 \leq p \leq \infty$ den Raum $\mathcal{X} := L^p[-1,1]$ mit

$$g(x;t) := \mathbb{1}_{(t,1]}(x), \quad x \in [-1,1], \quad t \in (0,1).$$

Dann definiert mit $E := (0,1)$ die Abbildung

$$f : t \mapsto g(\cdot; t) = \mathbb{1}_{(t,1]}$$

eine Abbildung von E nach $\mathcal{X}$. Diese Abbildung ist für $1 \leq p < \infty$ stark und damit auch schwach stetig in $t_0 := 0$, denn es gilt:

$$\lim_{t \rightarrow 0} \|f(t) - f(0)\|_{\mathcal{X}} = \lim_{t \rightarrow 0} \|\mathbb{1}_{(0,t]}\|_p = \lim_{t \rightarrow 0} t^{1/p} = 0.$$

Für $p = \infty$ ist diese Abbildung dagegen nicht stark stetig wegen

$$\|f(t) - f(0)\|_{\infty} = \|\mathbb{1}_{(0,t]}\|_{\infty} = 1 \quad \text{für alle } t \in (0,1).$$

(Bemerkung: Sie ist auch nicht schwach stetig!)

Exkurs: Halbgruppen von Operatoren

Lemma. Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f: E \rightarrow \mathcal{X}$ eine Abbildung. Ist die Abbildung $f(t)$ in $t_0 \in E$ stark (schwach) differenzierbar, so ist $f(t)$ dort auch stark (schwach) stetig.

Beweis: Die starke Stetigkeit folgt aus der Ungleichung

$$\begin{aligned}\|f(t) - f(t_0)\|_{\mathcal{X}} &= \left\| h \cdot \left\{ \frac{1}{h} [f(t_0 + h) - f(t_0)] - g_0 \right\} + h \cdot g_0 \right\|_{\mathcal{X}} \\ &\leq h \left\| \frac{1}{h} [f(t_0 + h) - f(t_0)] - g_0 \right\|_{\mathcal{X}} + h \|g_0\|_{\mathcal{X}} \quad \text{für alle } h > 0;\end{aligned}$$

analog argumentiert man bei schwacher Stetigkeit. ■

Exkurs: Halbgruppen von Operatoren

Lemma. Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f: E \rightarrow \mathcal{X}$ eine Abbildung. Ist die Abbildung $f(t)$ in jedem Punkt $t \in E$ stark bzw. schwach differenzierbar mit $\frac{wd}{dt}f = \mathbf{0}$, so existiert ein $g \in \mathcal{X}$ mit $f(t) = g$ für alle $t \in E$.

Beweis: Nach Definition der schwachen Ableitung ist für jedes beschränkte lineare Funktional $f^* \in \mathcal{X}^*$ die Funktion $f^*(f(\cdot))$ im üblichen Sinne differenzierbar mit

$$\frac{d}{dt}f^*(f(t)) = f^*\left(\frac{wd}{dt}f(t)\right) = f^*(\mathbf{0}) = 0 \text{ für alle } t \in E.$$

Also ist die Abbildung $f^*(f(\cdot))$ auf E konstant. Für festes $t_0 \in E$ ist somit $f^*(f(t)) = f^*(f(t_0))$ für alle $t \in E$ bzw.

$$f^*(f(t)) - f^*(f(t_0)) = f^*(f(t) - f(t_0)) = 0 \text{ für alle } t \in E$$

und somit $f(t) - f(t_0) = \mathbf{0}$ bzw. $f(t) = f(t_0) =: g$ für alle $t \in E$. ■

Exkurs: Halbgruppen von Operatoren

Definition (abstraktes Riemann-Integral) Es sei $E = [a, b]$, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f: E \rightarrow \mathcal{X}$ eine Abbildung, die stark (schwach) stetig auf E sei. Für jede Zerlegung $\Delta_n \in \mathcal{D}[a, b]$ und Zwischenpunkte $\xi_i \in [t_{i-1}, t_i]$, $i = 1, \dots, n$ sei ferner

$$S(f, \{\xi_i\}_{i=1}^n, \Delta_n) := \sum_{i=1}^n f(\xi_i) \cdot (t_i - t_{i-1}).$$

Die Größe $\delta_n := \max\{t_i - t_{i-1} \mid 1 \leq i \leq n\}$ bezeichne die *Feinheit* der Zerlegung $\Delta_n \in \mathcal{D}[a, b]$. Existiert dann

$$\lim_{\delta_n \rightarrow 0} S(f, \{\xi_i\}_{i=1}^n, \Delta_n) = \lim_{\delta_n \rightarrow 0} \sum_{i=1}^n f(\xi_i) \cdot (t_i - t_{i-1})$$

unabhängig von der Wahl der Zwischenpunkte im starken (schwachen) Sinn in $\mathcal{X}$, so heißt f *stark (schwach) Riemann-Stieltjes-integrierbar*, und der obige Grenzwert wird mit

$$\mathbf{s} \int_a^b f(t) dt \quad \text{bzw.} \quad \mathbf{w} \int_a^b f(t) dt \quad \text{bezeichnet.}$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

- a) Ist f stark (schwach) stetig, so ist f auch stark (schwach) Riemann-integrierbar.
- b) Das starke (schwache) Riemann-Integral ist linear, d.h. sind f_1, f_2 stark (schwach) Riemann-integrierbar, so auch $\alpha_1 f_1 + \alpha_2 f_2$ für alle $\alpha_1, \alpha_2 \in \mathbb{R}$ mit

$$\int_a^b \alpha_1 f_1(t) + \alpha_2 f_2(t) dt = \int_a^b \alpha_1 f_1(t) dt + \int_a^b \alpha_2 f_2(t) dt.$$

- c) Ist f stark (schwach) Riemann-integrierbar über $[a, b]$, so auch über jedem kompakten Teilintervall von $[a, b]$. Insbesondere gilt:

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt \quad \text{für alle } c \in (a, b).$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

d) Ist f stark Riemann-integrierbar über $[a, b]$, so auch $\|f(\cdot)\|_{\mathcal{X}}$ (im üblichen Sinne), und es gilt

$$\left\| \int_a^b f(t) dt \right\|_{\mathcal{X}} \leq \int_a^b \|f(t)\|_{\mathcal{X}} dt \leq (b-a) \max_{a \leq x \leq b} \|f(t)\|_{\mathcal{X}}.$$

e) Ist f stark (schwach) Riemann-integrierbar über $[a, b]$, so auch $f^*(f(\cdot))$ für alle $f^* \in \mathcal{X}^*$ (im üblichen Sinne), und es gilt

$$f^* \left(\int_a^b f(t) dt \right) = \int_a^b f^*(f(t)) dt \text{ für alle } f^* \in \mathcal{X}^*.$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

f) Ist f stark (schwach) stetig, so ist die durch

$$F(t) := \int_a^t f(s) ds, \quad a \leq t \leq b$$

definierte Abbildung $F : E \rightarrow \mathcal{X}$ stark (schwach) differenzierbar mit

$$\frac{d}{dt} F(t) = f(t) \text{ für alle } a \leq t \leq b.$$

g) Besitzt f eine stark stetige Ableitung f' auf $[a, b]$, so gilt

$$\int_a^t f'(s) ds = f(t) - f(a), \quad a \leq t \leq b.$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

h) Ist $\{f_n(\cdot)\}_{n \in \mathbb{N}}$ eine Folge Riemann-integrierbarer Funktionen auf $[a, b]$ mit

$$\lim_{n \rightarrow \infty} \sup_{a \leq t \leq b} \|f_n(t) - f(t)\| = 0,$$

so gilt

$$\int_a^b f(t) dt = \text{s-}\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt.$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

- i) Ist $D \subseteq \mathcal{X}$ ein linearer Teilraum, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ ein weiterer Banach-Raum, $T : D \rightarrow \mathcal{Y}$ ein *beschränkter* linearer Operator und f stark stetig mit Werten in D , so ist die Abbildung $(Tf)(t) := T(f(t))$ für $t \in E$ stark stetig mit

$$\int_a^b Tf(t) dt = T \left(\int_a^b f(t) dt \right).$$

Exkurs: Halbgruppen von Operatoren

Lemma (Eigenschaften des abstrakten Integrals). Es sei $E = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \rightarrow \mathcal{X}$ seien Abbildungen. Dann gilt:

- j) Ist $D \subseteq \mathcal{X}$ ein linearer Teilraum, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ ein weiterer Banach-Raum, $T : D \rightarrow \mathcal{Y}$ ein *abgeschlossener* linearer Operator und f stark stetig mit Werten in D , so ist die Abbildung $(Tf)(t) := T(f(t))$ für $t \in E$ Riemann-integrierbar mit

$$\int_a^b Tf(t) dt = T \left(\int_a^b f(t) dt \right).$$

Exkurs: Halbgruppen von Operatoren

Lemma. Sind $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ Banach-Räume und ist der Operator $T: E = [a, b] \rightarrow \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ stark stetig, dann existiert das Riemann-Integral $\int_a^b T(t) dt$ im Sinne der gleichmäßigen Konvergenz, und es gilt

$$\left(\int_a^b T(t) dt \right) f = \int_a^b (T(t)) f dt.$$

Ferner existiert die Abbildung

$$F(t) := \int_a^t T(s) ds, \quad a \leq t \leq b$$

punktweise im Sinne der gleichmäßigen Konvergenz, und ist gleichmäßig differenzierbar mit

$$\frac{d}{dt} F(t) = \frac{d}{dt} \int_a^t T(s) ds = T(t), \quad a \leq t \leq b.$$

Exkurs: Halbgruppen von Operatoren

Potenzen des infinitesimalen Erzeugers

Für das Folgende sei wieder grundsätzlich angenommen, dass $\mathfrak{T} = \{T(t) \mid t \geq 0\} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ eine Operator-Halbgruppe der Klasse (C_0) ist.

Definition (infinitesimaler Erzeuger). Der *inifinitesimale Erzeuger* A einer Operator-Halbgruppe der Klasse (C_0) ist definiert durch

$$Af := \text{s-lim}_{h \downarrow 0} \frac{1}{h} [T(h) - I]f$$

für alle $f \in \mathcal{X}$, für die dieser Grenzwert existiert. Die Menge dieser $f \in \mathcal{X}$ wird mit $D(A)$ bezeichnet (*Definitionsbereich von A*).

Aus Vereinfachungsgründen wird im Folgenden die Abkürzung

$$A_h := \frac{1}{h} [T(h) - I] \text{ für alle } h > 0$$

verwendet.

Exkurs: Halbgruppen von Operatoren

Potenzen des infinitesimalen Erzeugers

Lemma (Eigenschaften des infinitesimalen Erzeugers). Der infinitesimale Erzeuger A und sein Definitionsbereich $D(A)$ besitzen folgende Eigenschaften:

a) $D(A)$ ist ein linearer Teilraum von $\mathcal{X}$, und A ist ein linearer Operator auf $D(A)$.

b) Es ist $T(t)f \in D(A)$ für alle $f \in D(A)$ und $t \geq 0$, mit

$$\frac{d}{dt}T(t)f = A(T(t)f) = T(t)(Af) \text{ für alle } f \in D(A) \text{ und } t \geq 0.$$

c) Es gilt

$$T(t)f - f = \int_0^t T(s)Af \, ds \text{ für alle } t \geq 0.$$

d) $D(A)$ ist dicht in $\mathcal{X}$ und A ist abgeschlossen.

Exkurs: Halbgruppen von Operatoren

Potenzen des infinitesimalen Erzeugers

Die Nacheinander-Ausführung des infinitesimalen Erzeugers A führt damit in natürlicher Weise auf höhere Potenzen von A , wie die nachfolgende Definition zeigt.

Definition (höhere Potenzen von A). Für $r \in \mathbb{Z}^+ = \{0, 1, 2, \dots\}$ werden die (höheren) Potenzen des infinitesimalen Erzeugers A einer Operator-Halbgruppe der Klasse (C_0) folgendermaßen rekursiv definiert:

$$A^0 := I \quad \text{und} \quad A^r f := A(A^{r-1}f) \quad \text{für } r \in \mathbb{N}, f \in D(A^r) \quad \text{mit}$$

$$D(A^r) := \{f \in \mathcal{X} \mid f \in D(A^{r-1}) \wedge A^{r-1}f \in D(A)\}.$$

Exkurs: Halbgruppen von Operatoren

Potenzen des infinitesimalen Erzeugers

Satz (Eigenschaften der Potenzen des Erzeugers). Die Potenzen A^r des infinitesimalen Erzeugers A und die zugehörigen Definitionsbereiche $D(A^r)$ besitzen folgende Eigenschaften:

- a) $D(A^r)$ ist ein linearer Teilraum von $\mathcal{X}$, und A^r ist ein linearer Operator auf $D(A^r)$ für alle $r \in \mathbb{N}$.
- b) Für jedes $r \in \mathbb{N}$ ist $T(t)f \in D(A^r)$ für alle $f \in D(A^r)$ und $t \geq 0$, mit

$$T(t)f - \sum_{k=0}^{r-1} \frac{t^k}{k!} A^k f = \frac{1}{(r-1)!} \int_0^t (t-s)^{r-1} T(s) A^r f ds \quad \text{für alle } f \in D(A^r), t \geq 0.$$

Ferner gilt für jedes $r \in \mathbb{N}$ und alle $h > 0$:

$$A_h^r f = \frac{1}{h^r} \int_0^h \cdots \int_0^h T(s_1 + \cdots + s_r) A^r f ds_1 \cdots ds_r \quad \text{für alle } f \in D(A^r).$$

Exkurs: Halbgruppen von Operatoren

Potenzen des infinitesimalen Erzeugers

Satz (Eigenschaften der Potenzen des Erzeugers). Die Potenzen A^r des infinitesimalen Erzeugers A und die zugehörigen Definitionsbereiche $D(A^r)$ besitzen folgende Eigenschaften:

- c) $D(A^r)$ ist dicht in $\mathcal{X}$ für jedes $r \in \mathbb{N}$, ebenso $\bigcap_{r \in \mathbb{N}} D(A^r)$, und A^r ist abgeschlossen.
- d) $D(A^r)$ ist ein Banach-Raum unter der Norm

$$\|f\|_{D(A^r)} := \sum_{k=0}^r \|A^k f\|_{\mathcal{X}} \quad \text{für alle } f \in D(A^r).$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG6 (Fortsetzung von Beispiel HG1). Hier gilt offensichtlich:

$$\lim_{h \downarrow 0} A_h f(x) = \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{h} = f'(x) = Af(x) \text{ für } x \in \mathbb{R},$$

mit

$$D(A) = \{f \in UCB(\mathbb{R}) \mid f' \in UCB(\mathbb{R})\}$$

(man beachte, dass die Konvergenz hier nicht nur punktweise, sondern gleichmäßig sein muss). Für die höheren Potenzen ergibt sich entsprechend

$$A^r f = f^{(r)} \text{ mit } D(A^r) = \{f \in UCB(\mathbb{R}) \mid f^{(r)} \in UCB(\mathbb{R})\}, \quad r \in \mathbb{N}.$$

Beispiel HG7 (Fortsetzung von Beispiel HG2). Hier gilt offensichtlich:

$$\text{s-lim}_{h \downarrow 0} A_h f = \lim_{h \downarrow 0} \frac{e^{\lambda h} - 1}{h} f = \lambda f = Af,$$

mit $D(A) = \mathcal{X}$. Für die höheren Potenzen ergibt sich analog

$$A^r f = \lambda^r f \text{ mit } D(A^r) = \mathcal{X}, \quad r \in \mathbb{N}.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG8 (Fortsetzung von Beispiel HG3).

Hier gilt, mit $B_t := \frac{1}{2} \begin{bmatrix} e^{\mu t} + e^{\lambda t} & e^{\mu t} - e^{\lambda t} \\ e^{\mu t} - e^{\lambda t} & e^{\mu t} + e^{\lambda t} \end{bmatrix}$, $t \geq 0$:

$$\lim_{h \downarrow 0} \frac{1}{t} (B_t - I) = \frac{1}{2} \lim_{h \downarrow 0} \begin{bmatrix} \frac{e^{\mu t} - 1}{t} + \frac{e^{\lambda t} - 1}{t} & \frac{e^{\mu t} - 1}{t} - \frac{e^{\lambda t} - 1}{t} \\ \frac{e^{\mu t} - 1}{t} - \frac{e^{\lambda t} - 1}{t} & \frac{e^{\mu t} - 1}{t} + \frac{e^{\lambda t} - 1}{t} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \mu + \lambda & \mu - \lambda \\ \mu - \lambda & \mu + \lambda \end{bmatrix} =: B.$$

Damit erhält man für den infinitesimalen Erzeuger und seine Potenzen:

$$A^r f = B^r \cdot f, \text{ mit } D(A^r) = \mathcal{X}, r \in \mathbb{N}.$$

Exkurs: Halbgruppen von Operatoren

Beispiel HG8 (Fortsetzung von Beispiel HG3).

Die Potenzen der Matrix B können in diesem Fall explizit angegeben werden; es gilt:

$$B^r = \frac{1}{2} \begin{bmatrix} \mu^r + \lambda^r & \mu^r - \lambda^r \\ \mu^r - \lambda^r & \mu^r + \lambda^r \end{bmatrix} \text{ für } r \in \mathbb{N}.$$

Dies kann z.B. durch vollständige Induktion leicht verifiziert werden.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Satz (Darstellungssatz für beschränkte Erzeuger). Es sei $A \in \mathcal{L}[\mathcal{X}, \mathcal{X}]$ ein beschränkter linearer Operator. Dann wird durch

$$T(t) := e^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k, \quad t \geq 0$$

eine Halbgruppe $\mathfrak{T}$ der Klasse (C_0) definiert, deren infinitesimaler Erzeuger gerade A ist, mit $D(A) = \mathcal{X}$. Ferner gilt dann

$$\|T(t)\| \leq e^{t\|A\|}, \quad t \geq 0.$$

Ist umgekehrt $\mathfrak{T} = \{T(t) \mid t \geq 0\} \subseteq \mathcal{L}[\mathcal{X}, \mathcal{X}]$ eine Halbgruppe der Klasse (C_0) mit beschränktem infinitesimalem Erzeuger A , so ist notwendig $D(A) = \mathcal{X}$, und die Halbgruppe besitzt die oben gegebene eindeutige Darstellung.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Bemerkung: Die die Halbgruppe $\mathfrak{T} = \{e^{tA} \mid t \geq 0\}$ darstellende Potenzreihe ist sogar gleichmäßig konvergent und die Halbgruppe ist auch gleichmäßig differenzierbar, d.h. es gilt

$$\frac{1}{t}(T(t) - I) = \frac{1}{t}(e^{tA} - I) \rightarrow A \text{ gleichmäßig für } t \downarrow 0 \text{ (in der Operator-Norm),}$$

Ferner gilt: Sind $\mathfrak{T}_i = \{T_i(t) \mid t \geq 0\}$ kommutierende Operator-Halbgruppen der Klasse (C_0) auf demselben Banach-Raum $\mathcal{X}$ mit den infinitesimalen

Erzeugern A_i , $i = 1, \dots, n$, so ist auch $\mathfrak{T} = \left\{ \prod_{i=1}^n T_i(t) \mid t \geq 0 \right\}$ eine Operator-

Halbgruppe der Klasse (C_0) , wobei das „Produkt“ die Hintereinanderausführung der Operatoren bezeichne. Diese Halbgruppe besitzt den

infinitesimalen Erzeuger $A = \sum_{i=1}^n A_i$ mit $D(A) = \bigcap_{i=1}^n D(A_i)$.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Beispiel HG9 (Fortsetzung von Beispiel HG3). Der hier gegebene infinitesimale Erzeuger ist offensichtlich darstellbar als Summe zweier beschränkter linearer Operatoren, in Matrix-Form gegeben durch

$$B_1 := \frac{\mu}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B_2 := \frac{\lambda}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \text{ mit } B_1^k := \frac{\mu^k}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B_2^k := \frac{\lambda^k}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, k \in \mathbb{N}.$$

Die hierdurch einzeln erzeugten Halbgruppen sind (in Matrix-Form):

$$\begin{aligned} e^{tB_1} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} B_1^k = I + \frac{1}{2} \left(\sum_{k=1}^{\infty} \frac{(\mu t)^k}{k!} \right) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = I + \frac{e^{\mu t} - 1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{\mu t} + 1 & e^{\mu t} - 1 \\ e^{\mu t} - 1 & e^{\mu t} + 1 \end{bmatrix} \\ e^{tB_2} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} B_2^k = I + \frac{1}{2} \left(\sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} \right) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = I + \frac{e^{\lambda t} - 1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{\lambda t} + 1 & -e^{\lambda t} + 1 \\ -e^{\lambda t} + 1 & e^{\lambda t} + 1 \end{bmatrix} \end{aligned}$$

für $t \geq 0$, mit der erwarteten Produkt-Darstellung

$$e^{t(B_1+B_2)} = e^{tB_1} e^{tB_2} = \frac{1}{4} \begin{bmatrix} e^{\mu t} + 1 & e^{\mu t} - 1 \\ e^{\mu t} - 1 & e^{\mu t} + 1 \end{bmatrix} \cdot \begin{bmatrix} e^{\lambda t} + 1 & -e^{\lambda t} + 1 \\ -e^{\lambda t} + 1 & e^{\lambda t} + 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{\mu t} + e^{\lambda t} & e^{\mu t} - e^{\lambda t} \\ e^{\mu t} - e^{\lambda t} & e^{\mu t} + e^{\lambda t} \end{bmatrix}.$$

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Beispiel HG11. Ist $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) mit infinitesimalem Erzeuger A , so ist offensichtlich auch die Halbgruppe $\mathfrak{T}_\lambda = \{e^{\lambda t}T(t) \mid t \geq 0\}$ von der Klasse (C_0) für beliebiges $\lambda \in \mathbb{R}$, mit dem infinitesimalem Erzeuger $A_\lambda = \lambda I + A$ wegen

$$\frac{1}{t}(e^{\lambda t}T(t) - I)f = \left\{ (T(t) - I)\frac{1}{t}(e^{\lambda t}I - I) \right\}f + \frac{1}{t}(T(t) - I)f + \frac{1}{t}(e^{\lambda t}I - I)f \rightarrow Af + \lambda f$$

für $t \downarrow 0$ und $f \in D(A)$, mit $D(A_\lambda) = D(A)$. Wählt man speziell $\lambda \leq -\omega$ mit der Konstanten $M(\omega)$ (siehe das Lemma auf Folie 92), so erhält man noch die Abschätzung

$$\|e^{\lambda t}T(t)\| \leq M(\omega) \quad \text{für alle } t \geq 0.$$

Halbgruppen mit dieser Eigenschaft heißen *gleichmäßig beschränkt* (engl.: equi-bounded); im Fall $M(\omega)=1$ heißen sie auch *Kontraktions-Halbgruppe*.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Definition (Stetigkeitsmodul). Ist $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) und $L > 0$, so heißt für $\delta > 0$ und $f \in \mathcal{X}$ die Größe

$$\omega_L(f; \delta) := \sup \left\{ \|(T(t) - T(s))f\| \mid |t - s| < \delta, 0 \leq s, t \leq L \right\}$$

der rektifizierte Stetigkeitsmodul (der Halbgruppe) für $f \in \mathcal{X}$ im Intervall $[0, L]$.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Satz (E. Hille's erste Exponentialformel (1942)). Es sei $\mathcal{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Dann gilt:

$$T(t)f = \mathbf{s}\text{-}\lim_{h \downarrow 0} \exp(tA_h)f \quad \text{für alle } f \in \mathcal{X},$$

wobei der Grenzwert gleichmäßig bzgl. t in jedem kompakten Teilintervall von $\mathbb{R}^+$ existiert. Ist ferner $L > 1$ vorgegeben, so gibt es eine – nur von L und dem charakteristischen Exponenten ω_0 der Halbgruppe abhängige – universelle Konstante K derart, dass genauer gilt:

$$\|T(t)f - \exp(tA_h)f\|_{\mathcal{X}} \leq \omega_L(f; h^{1/3}) + h^{1/3}K\|f\|_{\mathcal{X}} \quad \text{für alle } t \in [0, L-1] \text{ und } f \in \mathcal{X},$$

für genügend kleine $h > 0$.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Bemerkung: Der Bereich für h , in dem die im Satz angegebene Ungleichung gültig ist, kann explizit angegeben werden, indem man die Gleichung $\omega h = \ln(1 + \sqrt{h})$ nach h mit der (unter den gegebenen Voraussetzungen) eindeutigen Lösung, etwa $h = h_\omega$, auflöst. Die Ungleichung gilt dann im Intervall $h \in [0, h_\omega]$. Für (etwa durch $M \geq 1$) gleichmäßig beschränkte bzw. Kontraktions-Halbgruppen kann die Abschätzung im Satz von Hille noch vereinfacht werden zu

$$\|T(t)f - \exp(tA_h)f\|_{\mathcal{X}} \leq \omega_L(f; h^{1/3}) + 2h^{1/3}M\|f\|_{\mathcal{X}}$$

für alle $t \in [0, L-1]$, $h > 0$ und $f \in \mathcal{X}$.

Exkurs: Halbgruppen von Operatoren

Analytische Darstellungssätze für Operatorhalbgruppen

Beispiel. Es sei $\mathcal{X} = UCB(\mathbb{R})$ der Banach-Raum der auf $\mathbb{R}$ definierten, gleichmäßig stetigen und beschränkten Funktionen. Dann gilt die verallgemeinerte Taylor-Formel:

$$f(x+t) = \lim_{h \downarrow 0} \sum_{k=0}^{\infty} \frac{t^k}{k!} \delta_h^k f(x) \quad \text{für jedes } f \in \mathcal{X} \text{ und } x \in \mathbb{R},$$

gleichmäßig bzgl. t in jedem kompakten Teilintervall von $\mathbb{R}^+$. Dabei ist

$$\delta_h f(x) := \frac{f(x+h) - f(x)}{h} \quad \text{für } x \in \mathbb{R} \text{ und } h > 0.$$

Bemerkung: Dies ist deswegen bemerkenswert, weil hier eine approximative Reihenentwicklung ohne jegliche Differenzierbarkeitsvoraussetzungen möglich ist. Für die iterierten Differenzen gilt dabei noch folgende, aus dem binomischen Lehrsatz folgende Darstellung:

$$\delta_h^k f(x) = \frac{1}{h^k} \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} f(x+ih) \quad \text{für alle } x \in \mathbb{R}, k \in \mathbb{N} \text{ und } h > 0.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Für das Folgende betrachten wir verschiedene Borel'sche σ -Algebren: zum einen die von den offenen Mengen in $\mathcal{X}$ erzeugte σ -Algebra $\mathfrak{B}(\mathcal{X})$, zum anderen die von den offenen Mengen der Banach-Algebra $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$ erzeugte σ -Algebra $\mathfrak{B}[\mathcal{X}, \mathcal{X}]$.

Lemma. Es sei $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ist die Halbgruppe in einem genügend kleinen Intervall $[0, \delta]$ mit $\delta > 0$ injektiv und existiert ein $t_0 > 0$ mit der Eigenschaft $\liminf_{t \downarrow t_0} \|T(t) - T(t_0)\| > 0$, so ist die Operator-wertige Abbildung $t \mapsto T(t)$ nicht $\mathfrak{B}[\mathcal{X}, \mathcal{X}]$ -messbar.

Stochastische Darstellungssätze für Operatorhalbgruppen

Beispiel. Eine Halbgruppe, auf die die Voraussetzungen des Lemmas zutrifft, ist die Halbgruppe der Linkstranslationen. Hier gilt nämlich:

$$\|T(t) - T(s)\| = 2 \text{ für alle } s, t \geq 0, s \neq t.$$

Für die Bildung von "Erwartungswerten" für Abbildungen der Art $T(X)$ mit Zufallsvariablen X , die Werte in $\mathbb{R}^+$ annehmen, benötigen wir allerdings noch weitere Integralbegriffe, die über das oben definierte Riemann-Integral hinausgehen. Wir beginnen mit dem so genannten *Bochner-Integral* (eingeführt 1932 von Salomon Bochner, polnischer Mathematiker, 1899 – 1982).

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Definition. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum. Eine Abbildung $\varphi : \Omega \rightarrow \mathcal{X}$ heißt *Elementarfunktion*, wenn sie in der Form

$$\varphi(\omega) = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(\omega), \quad \omega \in \Omega$$

mit $\alpha_1, \dots, \alpha_n \in \mathcal{X}$ und $n \in \mathbb{N}$ darstellbar ist. Sie heißt *stark messbar*, wenn es eine Folge $\{\varphi_n\}_{n \in \mathbb{N}}$ von Elementarfunktionen gibt mit $\varphi = \text{s-lim}_{n \rightarrow \infty} \varphi_n$. Das *Bochner-Integral* einer Elementarfunktion φ ist definiert als

$$B \int \varphi d\mu := \sum_{i=1}^n \mu(A_i) \alpha_i.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Eine stark messbare Abbildung φ heißt *Bochner-integrierbar*, wenn es eine Folge $\{\varphi_n\}_{n \in \mathbb{N}}$ von Elementarfunktionen gibt mit

$$\varphi = \mathbf{s}\text{-}\lim_{n \rightarrow \infty} \varphi_n \quad \mu\text{-fast überall} \quad \text{und} \quad \lim_{n \rightarrow \infty} \int \|\varphi - \varphi_n\|_{\mathcal{X}} d\mu = 0.$$

In diesem Fall existiert der Grenzwert $\mathbf{s}\text{-}\lim_{n \rightarrow \infty} B \int \varphi_n d\mu$ in $\mathcal{X}$ und wird mit $B \int \varphi d\mu$ bezeichnet (*Bochner-Integral* von φ bzgl. μ).

Bemerkung: Das Bochner-Integral ähnelt in seinem Aufbau dem Lebesgue-Integral; es verfügt auch über analoge Eigenschaften.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Es gilt:

- a) Ist φ stark messbar, so ist $\|\varphi\|_{\mathcal{X}}$ messbar im üblichen Sinne.
- b) Für Elementarfunktionen φ gilt: $\left\| B \int \varphi d\mu \right\|_{\mathcal{X}} \leq \int \|\varphi\|_{\mathcal{X}} d\mu.$
- c) Sind φ_1, φ_2 Elementarfunktionen und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $\varphi = \beta_1 \varphi_1 + \beta_2 \varphi_2$ ebenfalls eine Elementarfunktion, und es gilt

$$B \int \varphi d\mu = \beta_1 \cdot B \int \varphi_1 d\mu + \beta_2 \cdot B \int \varphi_2 d\mu.$$

- d) Das Bochner-Integral ist eindeutig bestimmt.

Bemerkung: Ähnlich wie beim Lebesgue-Integral wird als Konvention definiert:

$$B \int_A \varphi d\mu := B \int \mathbb{1}_A \cdot \varphi d\mu.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Satz (Bochner 1932). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum. Eine stark messbare Abbildung φ ist genau dann Bochner-integrierbar, wenn $\|\varphi(\cdot)\|_{\mathcal{X}}$ im gewöhnlichen Sinne integrierbar ist. Ferner gilt: Ist $\{\varphi_n\}_{n \in \mathbb{N}}$ eine Folge von Elementarfunktionen mit

$$\varphi = \mathbf{s}\text{-}\lim_{n \rightarrow \infty} \varphi_n \quad \mu\text{-fast überall} \quad \text{und} \quad \lim_{n \rightarrow \infty} \int \|\varphi - \varphi_n\|_{\mathcal{X}} d\mu = 0,$$

so gilt auch

$$\int \|\varphi\|_{\mathcal{X}} d\mu = \lim_{n \rightarrow \infty} \int \|\varphi_n\|_{\mathcal{X}} d\mu.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Für eine Bochner-integrierbare Abbildung φ gilt:

$$\left\| B \int \varphi d\mu \right\|_{\mathcal{X}} \leq \int \|\varphi\|_{\mathcal{X}} d\mu.$$

Ist g eine nicht-negative Lebesgue-integrierbare Abbildung und φ stark messbar mit $\|\varphi(\cdot)\|_{\mathcal{X}} \leq g$ μ -f.ü., so ist φ Bochner-integrierbar. Sind ferner $\{\psi_n\}_{n \in \mathbb{N}}$ und ψ stark messbar mit $\|\psi_n(\cdot)\|_{\mathcal{X}} \leq g$ und $\varphi = \mathbf{s}\text{-}\lim_{n \rightarrow \infty} \psi_n$ μ -f.ü., so sind alle ψ_n und ψ Bochner-integrierbar. Ferner gilt:

$$\lim_{n \rightarrow \infty} \int \|\psi - \psi_n\|_{\mathcal{X}} d\mu = 0 \quad \text{impliziert} \quad B \int \psi d\mu = \mathbf{s}\text{-}\lim_{n \rightarrow \infty} B \int \psi_n d\mu.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Ist $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ ein weiterer Banach-Raum und $T \in \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ ein beschränkter linearer Operator, dann ist für eine $\mathcal{X}$ -wertige Bochner-integrierbare Abbildung φ die $\mathcal{Y}$ -wertige Abbildung $T\varphi$ ebenfalls Bochner-integrierbar mit

$$B \int T\varphi d\mu = T \left(B \int \varphi d\mu \right).$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Ist der Maßraum $(\Omega, \mathcal{A}, \mu)$ vollständig und σ -endlich, dann sind folgende Aussagen für eine Abbildung $\varphi: \Omega \rightarrow \mathcal{X}$ äquivalent:

- a) Es gibt eine μ -Nullmenge $\mathcal{N} \subseteq \mathcal{X}$ derart, dass $\varphi: \Omega \setminus \mathcal{N} \rightarrow \mathcal{X}$ $\mathfrak{B}(\mathcal{X})$ -messbar ist und $\varphi(\Omega \setminus \mathcal{N})$ eine abzählbare und in $\mathcal{X}$ dichte Teilmenge enthält (d.h. in diesem Sinn ist φ μ -f.ü. separabel-wertig).
- b) Es existiert eine μ -Nullmenge $\mathcal{N} \subseteq \mathcal{X}$ und eine Folge $\{\varphi_n\}_{n \in \mathbb{N}}$ von Elementarfunktionen mit $\varphi(\omega) = \mathbf{s}\text{-}\lim_{n \rightarrow \infty} \varphi_n(\omega)$ für alle $\omega \in \Omega \setminus \mathcal{N}$.

Insbesondere ist die Abbildung φ genau dann stark messbar, wenn sie $\mathfrak{B}(\mathcal{X})$ -messbar ist und μ -f.ü. separabel-wertig.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Es sei X eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E(e^{tX})$ für $t = \eta$ mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe. Dann existiert für jedes $f \in \mathcal{X}$ das Bochner-Integral (Erwartungswert)

$$E[T(X)f] := B \int T(X)f \, dP \text{ mit}$$

$$\|E[T(X)f]\|_{\mathcal{X}} \leq \int \|T(X)f\|_{\mathcal{X}} \, dP \leq M(\eta)\psi_X(\eta)\|f\|_{\mathcal{X}}.$$

Bemerkung: Ohne die Anwendung auf ein Element $f \in \mathcal{X}$ ist die Abbildung $\omega \mapsto T(X(\omega))$ i. Allg. nicht stark messbar, da die Abbildung $t \mapsto T(t)$ i. Allg. nicht separabel-wertig ist. Daher kann auch auf diesem Weg i. Allg. kein "Erwartungswert" $E[T(X)]$ definiert werden. Hierzu ist ein weiterer Integral-Begriff nötig, nämlich der des *Pettis-Integrals*.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Definition (Pettis-Integral, B.J. Pettis 1938). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ ein Banach-Raum. Eine Abbildung $\varphi: \Omega \rightarrow \mathcal{X}$ heißt *schwach messbar*, wenn $f^*(\varphi)$ messbar ist für alle beschränkten linearen Funktionale $f^* \in \mathcal{X}^*$. φ heißt *Pettis-integrierbar*, wenn φ schwach messbar ist und ein Element $J \in \mathcal{X}$ existiert mit

$$f^*(J) = \int f^*(\varphi) d\mu \text{ für alle } f^* \in \mathcal{X}^*.$$

In diesem Fall heißt J das *Pettis-Integral* von φ bezüglich μ , in Zeichen:
 $J = P \int \varphi d\mu.$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Das Pettis-Integral besitzt u.a. folgende Eigenschaften:

- a) Sind φ_1, φ_2 Pettis-integrierbar und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $\varphi = \beta_1 \varphi_1 + \beta_2 \varphi_2$ ebenfalls Pettis-integrierbar, und es gilt

$$P \int \varphi d\mu = \beta_1 \cdot P \int \varphi_1 d\mu + \beta_2 \cdot P \int \varphi_2 d\mu.$$

- b) Das Pettis-Integral ist im Falle der Existenz eindeutig bestimmt.

- c) Ist φ Bochner-integrierbar, so auch Pettis-integrierbar, und es gilt

$$B \int \varphi d\mu = P \int \varphi d\mu.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Das Pettis-Integral ist im Unterschied zum Bochner-Integral nicht "konstruktiv"; seine Bestimmung hängt wesentlich von der Struktur des Dualraums $\mathcal{X}^*$ ab. Dies kann z.B. bei Banach-Algebren wie $\mathcal{L}[\mathcal{X}, \mathcal{Y}]$ und deren Dualräumen zu Problemen führen. Man kann diesen Integralbegriff aber abschwächen, wenn man sich auf einen geeigneten Teilraum des Dualraums zurückzieht, der noch die Eindeutigkeit im Fall der Existenz garantiert. Dies führt zum Begriff des modifizierten Pettis-Integrals.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Definition (modifiziertes Pettis-Integral). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ seien Banach-Räume. Eine Abbildung $S: \Omega \rightarrow \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ heißt *modifiziert schwach messbar*, wenn $f^*(S(\cdot)(f))$ messbar ist für alle beschränkten linearen Funktionale $f^* \in \mathcal{Y}^*$ und alle $f \in \mathcal{X}$. S heißt *modifiziert Pettis-integrierbar*, wenn S modifiziert schwach messbar ist und ein Element $J \in \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ existiert mit

$$f^*(J(f)) = \int f^*(S(\cdot)(f)) d\mu = \int f^*(S(\omega)(f)) \mu(d\omega) \text{ für alle } f^* \in \mathcal{Y}^* \text{ und } f \in \mathcal{X}.$$

In diesem Fall heißt J das *modifizierte Pettis-Integral* von S bezüglich μ , in Zeichen: $J = P^* \int S d\mu$.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Das modifizierte Pettis-Integral besitzt u.a. folgende Eigenschaften:

- a) Ist S modifiziert Pettis-integrierbar, so ist $S(\cdot)(f): \Omega \rightarrow \mathcal{Y}$ Pettis-integrierbar für alle $f \in \mathcal{X}$. In diesem Falle ist

$$\left(P^* \int S d\mu\right)(f) = P \int S(\cdot)(f) d\mu.$$

- b) Ist S Pettis-integrierbar im üblichen Sinn oder Bochner-integrierbar, dann auch modifiziert Pettis-integrierbar, und beide Integrale sind gleich.

- c) Sind S_1, S_2 Pettis-integrierbar und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $S = \beta_1 S_1 + \beta_2 S_2$ ebenfalls Pettis-integrierbar, und es gilt

$$P^* \int S d\mu = \beta_1 \cdot P^* \int S_1 d\mu + \beta_2 \cdot P^* \int S_2 d\mu.$$

- d) Das modifizierte Pettis-Integral ist im Falle der Existenz eindeutig bestimmt.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ seien Banach-Räume. Ferner sei eine Abbildung $S: \Omega \rightarrow \mathcal{L}[\mathcal{X}, \mathcal{Y}]$ derart gegeben, dass $S(\cdot)(f)$ stark messbar sei für jedes $f \in \mathcal{X}$, und dass $\|S(\cdot)\|$ durch eine μ -integrierbare Funktion $g \geq 0$ dominiert sei. Dann ist S modifiziert Pettis-integrierbar, und es gilt

$$\left\| P^* \int S d\mu \right\| \leq \int g d\mu.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Satz. Es sei X eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) | t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E(e^{tX})$ für $t = \eta$ mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe. Dann existiert für $T(X)$ das modifizierte Pettis-Integral (Erwartungswert)

$$E[T(X)] := P^* \int T(X) dP,$$

mit

$$\|E[T(X)]\| \leq M(\eta)\psi_X(\eta) \quad \text{und} \quad E[T(X)]f = E[T(X)f] \quad \text{für alle } f \in \mathcal{X}.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Satz. Es seien X und Y nicht-negative, stochastisch unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E(e^{tx})$ für $t = \eta$ mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe. Dann sind $T(X)$, $T(Y)$ und $T(X) \circ T(Y)$ modifizierte Pettis-integrierbar, und es gilt die "Produkt-Formel"

$$E[T(X) \circ T(Y)] = E[T(X)] \circ E[T(Y)].$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Bemerkung: Diese Beziehung wurde von K.L. Chung in seiner Arbeit 1962 verwendet, allerdings ohne die hierbei auftretenden Schwierigkeiten („zufälliger Operator“) zu thematisieren.

LEMMA 2. For $\lambda > w$, we have

$$E\{T(\sigma_n(\lambda))\} = [\lambda R(\lambda)]^n .$$

PROOF. Clearly for each k ,

$$E\{T(\tau_k(\lambda))\} = \int_0^\infty \lambda e^{-\lambda t} T(t) dt = \lambda R(\lambda) .$$

Since the $\tau_k(\lambda)$'s are independent, so are the $T(\tau_k(\lambda))$'s as operator-valued r.v.'s, and we have

$$E\{T(\sigma_n(\lambda))\} = E\left\{\prod_{k=1}^n T(\tau_k(\lambda))\right\} = \prod_{k=1}^n E\{T(\tau_k(\lambda))\} = [\lambda R(\lambda)]^n .$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Satz. Es sei N eine nicht-negative ganzzahlige und Y eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ferner existiere ein $\eta > \omega_0$, so dass die momenterzeugende Funktion $\psi_Y(t) = E(e^{tY})$ sowie die Komposition $\varphi_N(\psi_Y(t))$ für $t = \eta$ existiere, wobei $\varphi_N(s) = E(s^Y)$ die wahrscheinlichkeitserzeugende Funktion von N bezeichne. Ist dann $\{Y_k\}_{k \in \mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie Y verteilt und auch von N unabhängig sind, und bezeichnet $X := \sum_{k=1}^N Y_k$, so ist $T(X)$ modifiziert Pettis-integrierbar, und es gilt

$$E[T(X)] = \varphi_N(E[T(Y)]) = \sum_{k=0}^{\infty} P(N = k) \{E[T(Y)]\}^k,$$

mit der üblichen Konvention $\{E[T(Y)]\}^0 = I$.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Es sei $\{N(\tau) \mid \tau \geq 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ derart, dass

$\frac{1}{\tau}N(\tau)$ für $\tau \rightarrow \infty$ stochastisch gegen eine reelle Zahl $\zeta \geq 0$ konvergiere.

Ist dann $\{Y_k\}_{k \in \mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie eine nicht-negative Zufallsvariable Y mit $E(Y) = \gamma \geq 0$ verteilt und auch von der Familie $\{N(\tau) \mid \tau \geq 0\}$ unabhängig sind, und

bezeichnet $X(\tau) := \sum_{k=1}^{N(\tau)} Y_k$ für jedes $\tau \geq 0$, so gilt:

$$\frac{1}{\tau}X(\tau) = \frac{1}{\tau} \sum_{k=1}^{N(\tau)} Y_k \rightarrow \zeta \gamma \text{ stochastisch für } \tau \rightarrow \infty.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Lemma. Es sei $\{Z(\tau) \mid \tau \geq 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, die für $\tau \rightarrow \infty$ stochastisch gegen eine reelle Zahl $\xi \geq 0$ konvergiere, und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) . Ferner existiere ein $\eta > \omega_0$, so dass für die momenterzeugenden Funktionen der Ausdruck $s(\eta, \tau_0) := \sup_{\tau \geq \tau_0} \{\psi_{Z(\tau)}(\eta)\} < \infty$ für ein $\tau_0 \geq 0$ ist. Dann gilt:

$$E[T(Z(\tau))]f \xrightarrow{s} T(\xi)f \quad \text{für } \tau \rightarrow \infty, \text{ für alle } f \in \mathcal{X}.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Satz (Stochastischer Darstellungssatz für Operatorhalbgruppen, Pfeifer (1984)). Es sei $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) und $\{N(\tau) \mid \tau \geq 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ derart, dass $\frac{1}{\tau}N(\tau)$ für $\tau \rightarrow \infty$ stochastisch gegen eine reelle Zahl $\zeta \geq 0$ konvergiere. Ferner sei $\{Y_k\}_{k \in \mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie eine nicht-negative Zufallsvariable Y mit $E(Y) = \gamma \geq 0$ verteilt und auch von der Familie $\{N(\tau) \mid \tau \geq 0\}$ unabhängig sind. Existiert dann ein $\eta > \omega_0$, so dass für die momenterzeugenden Funktionen der Ausdruck $s(\eta, \tau_0) := \sup_{\tau \geq \tau_0} \{\psi_{Z(\tau)}(\eta)\} < \infty$ für ein $\tau_0 \geq 0$ ist, so gilt:

$$\varphi_{N(\tau)} \left(E \left[T \left(\frac{Y}{\tau} \right) \right] \right) f \xrightarrow{s} T(\xi) f \quad \text{für } \tau \rightarrow \infty, \text{ für alle } f \in \mathcal{X}, \text{ mit } \xi = \zeta \gamma.$$

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Dieses Resultat erlaubt eine Vielzahl von Darstellungssätzen, die praktisch alle in der Halbgruppentheorie bekannten Ergebnisse umfassen (darunter Kendall (1954), Hille and Phillips (1957), Chung (1962), Ditzian (1971), Shaw (1980)). Wir zeigen hier nur exemplarisch, wie *Hille's erste Exponentialformel* leicht hieraus gewonnen werden kann.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

Dazu betrachten wir einen homogenen Poisson-Prozess $\{N(\tau) \mid \tau \geq 0\}$ mit Parameter $\xi > 0$, für den die Voraussetzung des Darstellungssatzes mit $\zeta = \xi$ erfüllt ist, sowie die konstante Zufallsvariable $Y = 1$. Da für jedes $\tau > 0$ die Zufallsvariable $N(\tau)$ eine Poisson-Verteilung mit Parameter $\tau\xi$ besitzt, folgt

$$\begin{aligned}s(\eta, \tau_0) &:= \sup_{\tau \geq \tau_0} \left\{ \varphi_{N(\tau)}(\psi_{Y|_{\tau}}(\eta)) \right\} = \sup_{\tau \geq \tau_0} \left\{ \exp(\tau\xi(\psi_{Y|_{\tau}}(\eta) - 1)) \right\} \\ &= \sup_{\tau \geq \tau_0} \left\{ \exp(\tau\xi(e^{\eta/\tau} - 1)) \right\} \leq \sup_{\tau \geq \tau_0} \left\{ \exp(\xi\eta e^{1/\tau}) \right\} = \exp(\xi\eta e^{1/\tau_0}) < \infty\end{aligned}$$

mit beliebigem $\tau_0 > 0$. Es folgt

$$T(\xi)f = s\text{-}\lim_{\tau \rightarrow \infty} \varphi_{N(\tau)} \left(E \left[T \left(\frac{Y}{\tau} \right) \right] \right) f = s\text{-}\lim_{\tau \rightarrow \infty} \exp \left(\xi\tau \left[T \left(\frac{1}{\tau} \right) - I \right] \right) f = s\text{-}\lim_{\tau \rightarrow \infty} \exp(\xi A_{1/\tau}) f,$$

für $\tau \rightarrow \infty$ und für alle $f \in \mathcal{X}$, was gerade der Aussage des Satzes von Hille entspricht.

Exkurs: Halbgruppen von Operatoren

Stochastische Darstellungssätze für Operatorhalbgruppen

**STOCHASTISCHE METHODEN
IN DER THEORIE DER HALBGRUPPEN
LINEARER OPERATOREN**

Dr. rer. nat. Dietmar Pfeifer

Stochastische Methoden in der Theorie
der Halbgruppen linearer Operatoren

Von der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinisch-Westfälischen Technischen Hochschule Aachen
genehmigte Habilitationsschrift zur Erlangung der *venia legendi*

vorgelegt von
Dr. rer. nat. Dietmar Pfeifer
aus Wuppertal

Referenten: Prof. Dr. P. L. Butzer
Prof. Dr. B. Rauhut

Tag der Habilitation: 9. Mai 1984

Exkurs: Halbgruppen von Operatoren

Weitere Inhalte:

- Abschätzungen der Konvergenzgeschwindigkeit
- Anwendungen auf die Poisson-Approximation

Exkurs: Halbgruppen von Operatoren

Reprinted from JOURNAL OF APPROXIMATION THEORY
All Rights Reserved by Academic Press, New York and London

Vol. 43, No. 3, March 1985
Printed in Belgium

Approximation-Theoretic Aspects of Probabilistic Representations for Operator Semigroups

DIETMAR PFEIFER

*Institut für Statistik und Wirtschaftsmathematik,
Technical University, Aachen, West Germany*

Communicated by P. L. Butzer

Received May 8, 1984

Umfang: 25 Seiten

Exkurs: Halbgruppen von Operatoren

THEOREM 3.1. *Let X be a non-negative real random variable with expectation $E(X) = \xi$ and with $\psi_X(\delta) < \infty$ for some $\delta > 0$. Let further $\{X_n; n \in \mathbb{N}\}$ be an i.i.d. sequence of random variables, identically distributed with X . Then, if $\bar{X}_n = (1/n) \sum_{k=1}^n X_k$ denotes the arithmetic mean of $X_1, \dots, X_n$, we have*

$$e^{t\xi} \leq \psi_{\bar{X}_n}^*(t) \leq e^{t\xi} \exp \left\{ \frac{2t^2 \psi_X^*(\eta)}{e^2 n (\eta - |t|/n)^2} \right\} \quad (3.8)$$

for $|t| < n\eta$, and $0 < \eta \leq \delta$. Further,

$$E(|\bar{X}_n - \xi|^\alpha) \leq 2n^{-\alpha/2} (\alpha \psi_X^*(\eta) / e \eta^2)^{\alpha/2} \quad (3.9)$$

for $n \geq 15\alpha / \psi_X^*(\eta)$ ($\alpha > 0$).

Exkurs: Halbgruppen von Operatoren

THEOREM 4.1. *Let X be a non-negative real-valued random variable with $\psi_X^*(\delta) < \infty$ for some $\delta > 0$. Let further $\xi = E(X)$ denote the expectation of X and $\sigma^2 = \sigma^2(X)$ its variance. Then the following relations hold:*

$$\begin{aligned}
 & \|E[T(X)]f - T(\xi)f\| \\
 & \leq M \|Af\| \{e^{\omega\xi}E(|X - \xi|) + \omega\{E(X - \xi)^4\}^{1/2}\{\psi_X^*(2\omega)\}^{1/2}\} \\
 & \leq M \|Af\| \{\sigma e^{\omega\xi} + \omega\{E(X - \xi)^4\}^{1/2}\{\psi_X^*(2\omega)\}^{1/2}\} \\
 & \text{for } f \in D(A), \text{ if } \delta \geq 2\omega;
 \end{aligned} \tag{4.3}$$

$$\begin{aligned}
 & \|E[T(X)]f - T(\xi)f\| \\
 & \leq \frac{M}{2} \|A^2f\| \{\sigma^2 e^{\omega\xi} + \omega\{E(X - \xi)^4\}^{3/4}\{\psi_X^*(4\omega)\}^{1/4}\} \\
 & \text{for } f \in D(A^2), \text{ if } \delta \geq 4\omega;
 \end{aligned} \tag{4.4}$$

$$\begin{aligned}
 & \left\| E[T(X)]f - T(\xi)f - \frac{\sigma^2}{2} T(\xi) A^2f \right\| \\
 & \leq \frac{M}{6} \|A^3f\| \{e^{\omega\xi}E(|X - \xi|^3) + \omega\{E(X - \xi)^6\}^{2/3}\{\psi_X^*(3\omega)\}^{1/3}\} \\
 & \text{for } f \in D(A^3), \text{ if } \delta \geq 3\omega.
 \end{aligned} \tag{4.5}$$

Here M and ω denote the characteristic constants of the semigroup given by (2.1).

Exkurs: Halbgruppen von Operatoren

THEOREM 4.3. *Let $\{T(t) \mid t \geq 0\}$ be an arbitrary strongly continuous operator semigroup, and let $A_h = (1/h)(T(h) - I)$ for $h > 0$. Then*

$$\begin{aligned} & \|T(\xi)f - e^{\xi A_{1/\tau}}f\| \\ & \leq M e^{\omega\xi} \|Af\| \left\{ \sqrt{\frac{\xi}{\tau}} + 2\omega \frac{\xi}{\tau} \exp\left\{\frac{\omega^2\xi}{\tau} e^{2\omega/\tau}\right\} \right\} \\ & \text{for } f \in D(A) \text{ and } \tau \geq \frac{1}{\xi}; \end{aligned} \quad (4.11)$$

$$\begin{aligned} & \|T(\xi)f - e^{\xi A_{1/\tau}}f\| \\ & \leq \frac{M}{2} e^{\omega\xi} \|A^2f\| \left\{ \frac{\xi}{\tau} + 2\omega \sqrt{\frac{2\xi^3}{\tau^3}} \exp\left\{\frac{2\omega^2\xi}{\tau} e^{4\omega/\tau}\right\} \right\} \\ & \text{for } f \in D(A^2) \text{ and } \tau \geq \frac{1}{\xi}; \end{aligned} \quad (4.12)$$

$$\begin{aligned} T(\xi)f - e^{\xi A_{1/\tau}}f &= -\frac{\xi}{2\tau} T(\xi) A^2f + \mathcal{O}\left(\frac{1}{\tau\sqrt{\tau}}\right) \\ & \text{for } f \in D(A^3) \text{ and } \tau \rightarrow \infty, \end{aligned} \quad (4.13)$$

Exkurs: Halbgruppen von Operatoren und Poisson-Approximation

Semigroup Forum Vol. 31 (1985) 201-205
© 1985 Springer-Verlag New York Inc.

RESEARCH ARTICLE

A SEMIGROUP SETTING FOR DISTANCE
MEASURES IN CONNEXION WITH POISSON
APPROXIMATION

D. Pfeifer

Communicated by Jerome A. Goldstein

Exkurs: Halbgruppen von Operatoren und Poisson-Approximation

Semigroup Forum Vol. 34 (1986) 203–224
© 1986 Springer-Verlag New York Inc.

RESEARCH ARTICLE

OPERATOR SEMIGROUPS AND POISSON CONVERGENCE
IN SELECTED METRICS

Paul Deheuvels and Dietmar Pfeifer

Communicated by J.A. Goldstein

The Annals of Probability
1986, Vol. 14, No. 2, 663–676

A SEMIGROUP APPROACH TO POISSON APPROXIMATION

BY P. DEHEUVELS AND D. PFEIFER

Université Paris VI and Technical University Aachen

The aim of this paper is twofold: first, to show that Poisson approximation problems for independent summands can in a natural way be treated in a suitable operator semigroup framework, allowing at the same time for an asymptotically precise evaluation of the leading term with respect to the total variation distance; second, to determine asymptotically those Poisson distributions which minimize this distance for given Bernoulli summands. Besides semigroup methods, coupling techniques as well as direct computations are used.

OXFORD STUDIES IN PROBABILITY · 2

Poisson Approximation

A. D. BARBOUR
LARS HOLST
and
SVANTE JANSON



OXFORD SCIENCE PUBLICATIONS

The methods used by Kerstan and Le Cam, although quite different, both rely on the multiplicative properties of convolution. Kerstan compares the probability generating function of the sum of indicators with the Poisson probability generating function in an ingenious way, and then uses Cauchy's formula as a basic tool in obtaining the estimate of total variation distance from the difference of the probability generating functions. Le Cam also uses complex variables, in this case Fourier methods, for the detailed estimates, but his basic approach is to represent distributions over the non-negative integers $\mathbb{Z}$ as operators on the space of bounded measurable functions over $\mathbb{Z}$, with multiplication of operators corresponding to convolution. This method has since been developed by Deheuvels and Pfeifer, who have used it to obtain the sharpest results yet for the Poisson approximation of sums of independent indicators: in particular, in Deheuvels and Pfeifer (1988), the complex variable techniques of Uspensky (1931) and Shorgin (1977) are combined with the operator method to yield asymptotic expansions of arbitrary order for set probabilities, together with explicit bounds on the error involved. Indeed, by choosing the space of operators appropriately, the method gives good estimates of the distance between the distribution of W and $Po(\lambda)$ in at least three different metrics. For a more detailed discussion, see Witte (1990).

Sonstiges:

- Gastprofessur an der University of North Carolina, Chapel Hill (September 1985 bis Dezember 1985)

Reisetätigkeiten und Gastbesuche:

- The University of Iowa (Kooperation mit Ursula Gather)
- The Pennsylvania State University, University Park
- Johns Hopkins University, Baltimore (Kooperation mit Robert Serfling und Allan Karr)
- Colorado State University, Fort Collins (Berufungsangebot)

Pub. Inst. Stat. Univ.
XXXII, fasc. 3, 1987, 71 à 79

A NOTE ON STABILITY OF MAXIMA
AND RECORDS OF AN IID SEQUENCE

Ursula GATHER and Dietmar PFEIFER

ACKNOWLEDGEMENTS

A first draft of the paper was written when the authors were visiting the University of Iowa, Iowa, and the University of North Carolina at Chapel Hill, resp.. The second author is also indebted to the Deutsche Forschungsgemeinschaft for financial support by a Heisenberg research grant. Finally, we would like to thank P. Deheuvels for some interesting discussions in the area of extremal statistics and records.

A MARTINGALE CHARACTERIZATION OF MIXED POISSON PROCESSES

DIETMAR PFEIFER* AND
URSULA HELLER,** *Technical University Aachen*

Abstract

It is shown that an elementary pure birth process is a mixed Poisson process iff the sequence of post-jump intensities forms a martingale with respect to the σ -fields generated by the jump times of the process. In this case, the post-jump intensities converge almost surely to the mixing random variable of the process.

INTENSITIES; COMPLETELY MONOTONIC FUNCTIONS; BERNSTEIN'S THEOREM

* Postal address: Institut für Statistik und Wirtschaftsmathematik, RWTH Aachen, Wüllnerstrasse 3, D-5100 Aachen, W. Germany.

Research supported in part by US AFOSR Contract No. F 49620 85 C 0144 at the University of North Carolina at Chapel Hill.

** Present address: Gothaer Lebensversicherung a.G., Gothaer Platz, 3400 Göttingen, W. Germany.

**POISSON APPROXIMATIONS IN SELECTED METRICS BY
COUPLING AND SEMIGROUP METHODS WITH
APPLICATIONS**

Paul DEHEUELS

Université Paris VI, Paris, France

Alan KARR

Johns Hopkins University, Baltimore, MD 21218, U.S.A.

Dietmar PFEIFER

Universität Oldenburg, Oldenburg, West Germany

Robert SERFLING

Johns Hopkins University, Baltimore, MD 21218, U.S.A.

Received 7 July 1987; revised manuscript received 6 October 1987

Recommended by P. Révész

IV. Heisenberg-Stipendium (1986-87)



Reisetätigkeit und Kooperationen:

- Université Paris VI (Pierre et Marie Curie), 1986
(Poisson-Approximation)
- University of California, Santa Barbara (Sommersemester 1987)
(Extremwertstatistik, Problem der besten Wahl)
- Indiana University, Bloomington, Sommer 1987
(Poisson-Approximation)
- Stanford University, Palo Alto (August 1987)
(Extremwertstatistik)

[Publikation der Arbeiten zeitlich verzögert durch Begutachtungsprozesse]

Poisson Approximations of Multinomial Distributions and Point Processes

PAUL DEHEUVELS

Université Paris VI, Paris, France

AND

DIETMAR PFEIFER

University of Oldenburg, West Germany

Communicated by P. Hall

We present an extension to the multinomial case of former estimations for univariate Poisson binomial approximation problems and generalize a result obtained by N. K. Arenbaev (*Theory Probab. Appl.* **21** (1976), 805–810). As an application, we evaluate the total variation distance between superpositions of independent Bernoulli point processes and a suitable Poisson process. The main tool will be a multiparameter semigroup approach. © 1988 Academic Press, Inc.

RESEARCH ARTICLE

**A NEW SEMIGROUP TECHNIQUE IN
POISSON APPROXIMATION**

P. Deheuvels, D. Pfeifer and Madan L. Puri

ABSTRACT. We present a unified and self-contained approach to Poisson approximation problems for independent Bernoulli summands with respect to several metrics by a general semigroup technique, expanding and completing earlier work on this subject by the first two authors [4], [5], [6].

D. PFEIFER, *Universität Oldenburg*

On a relationship between record times and record values of an i.i.d. sequence

Let $\{X_n\}$ be an i.i.d. sequence of random variables with a continuous c.d.f. F . Define upper and lower record times by

$$\begin{aligned}U_0 &= 1, & U_{n+1} &= \inf \{k; X_k > X_{U_n}\}; & n &\geq 0; \\L_0 &= 1, & L_{n+1} &= \inf \{k; X_k < X_{L_n}\}; & n &\geq 0.\end{aligned}$$

If F is the c.d.f. of an exponential distribution with unit mean, then we have the following result.

Theorem. (a) L_n and $L_n X_{L_n}$ are independent for all $n \geq 0$.
(b) $L_n X_{L_n}$ is exponentially distributed with unit mean.

This theorem allows the following conclusion.

Corollary. (a) $X_{U_n} - \log U_n$ is asymptotically Λ -distributed for $n \rightarrow \infty$ where Λ denotes the c.d.f. of a doubly-exponentially distributed random variable.

(b) $\log U_n = X_{U_n} + O(\log n)$ a.s. ($n \rightarrow \infty$).

Similar results can be derived also for the case of a general c.d.f. F , as well as strong approximations jointly for record times, inter-record times and record values by Poisson and Wiener processes.

Workshop on Extremes of Random Processes in Applied Probability, Santa Barbara 1987. In: *Advances in Applied Probability* 20 (1988), 12.

EXTREMAL PROCESSES, SECRETARY PROBLEMS AND THE $1/e$ LAW

DIETMAR PFEIFER,* *University of Oldenburg*

Abstract

We consider a class of secretary problems in which the order of arrival of candidates is no longer uniformly distributed. By a suitable embedding in a time-transformed extremal process it is shown that the asymptotic winning probability is again $1/e$ as in the classical situation. Extensions of the problem to more than one choice are also considered.

NON-HOMOGENEOUS POISSON PROCESS; POISSON BINOMIAL DISTRIBUTION; COUPLING;
RECORD VALUES

Semigroups and Probability:

From representation theorems to Poisson approximation

Dietmar Pfeifer

Institut für Mathematik

Carl von Ossietzky University, Oldenburg

Conference in honor of Paul Deheuvels
Paris, June 20-21, 2013

Agenda

- Introduction
- Probabilistic representations of operator semigroups
- The semigroup approach to Poisson approximation
- Bibliography

Introduction

Important relationships between

semigroups	↔	probability
Feller semigroups	↔	Markov processes
convolution semigroups	↔	Poisson approximation
representation theorems	↔	Bochner and Pettis integral
approximation theorems	↔	law of large numbers

Introduction

Semigroups $\{T(t)|t \geq 0\} \subseteq \mathcal{L}[\mathcal{X}, \mathcal{X}]$ of class (C_0) on a Banach space $\mathcal{X}$:

- $T(0) = I$ (identity)
- $T(s+t) = T(s) \circ T(t)$ for $s, t \geq 0$
- $\lim_{t \downarrow 0} \|T(t)f - f\| = 0$ for all $f \in \mathcal{X}$

There exist constants $M \geq 1$, $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq M e^{\omega t}$ for $t \geq 0$

Infinitesimal generator:

$$Af(x) = \lim_{t \downarrow 0} \frac{1}{t} [T(t) - I]f(x) \text{ for } f \in \mathcal{D}(A) \quad \text{and} \quad AT(t)f = T(t)Af = \frac{d}{dt} T(t)f$$

Introduction

Feller semigroups $\{T(t)|t \geq 0\}$ for standard Lévy processes $\{X_t|t \geq 0\}$:

$$T(t)f(x) = E[f(x + X_t)] \text{ for } f \in \mathcal{X} = UCB(\mathbb{R})$$

Infinitesimal generator:

$$Af(x) = \lim_{t \downarrow 0} \frac{1}{t} [T(t) - I]f(x) = \lim_{t \downarrow 0} \frac{E[f(x + X_t) - f(x)]}{t} \text{ for } f \in \mathcal{D}(A)$$

Introduction

Feller semigroups $\{T(t)|t \geq 0\}$ for standard Lévy processes $\{X_t|t \geq 0\}$:

$$T(t)f(x) = E[f(x + X_t)] \text{ for } f \in \mathcal{X} = UCB(\mathbb{R})$$

Example: Brownian motion: $P^{X_t} = \mathcal{N}(\mu t, \sigma^2 t)$ with $\mu \in \mathbb{R}$, $\sigma > 0$:

$$Af(x) = \lim_{t \downarrow 0} \frac{E[f(x + X_t) - f(x)]}{t} = \mu f'(x) + \frac{\sigma^2}{2} f''(x)$$

for $f \in \mathcal{D}(A) = \{f \in \mathcal{X} | f', f'' \in \mathcal{X}\}$

Introduction

Feller semigroups $\{T(t)|t \geq 0\}$ for standard Lévy processes $\{X_t|t \geq 0\}$:

$$T(t)f(x) = E[f(x + X_t)] \text{ for } f \in \mathcal{X} = UCB(\mathbb{R})$$

Example: Brownian motion: $P^{X_t} = \mathcal{N}(\mu t, \sigma^2 t)$ with $\mu \in \mathbb{R}$, $\sigma > 0$:

$$\begin{aligned} \frac{E[f(x + X_t) - f(x)]}{t} &= \frac{1}{t} E[X_t] f'(x) + \frac{1}{2t} E[X_t^2] f''(x) + \mathcal{O}(t^2) \\ &= \mu f'(x) + \frac{\sigma^2}{2} f''(x) + \mathcal{O}(t^2) \end{aligned}$$

Introduction

Feller semigroups $\{T(t)|t \geq 0\}$ for standard Lévy processes $\{X_t|t \geq 0\}$:

$$T(t)f(x) = E[f(x + X_t)] \text{ for } f \in \mathcal{X} = UCB(\mathbb{R})$$

Example: Gamma process: $P^{X_t} = \Gamma(t, \lambda)$ with $\lambda > 0$:

$$Af(x) = \lim_{t \downarrow 0} \frac{E[f(x + X_t) - f(x)]}{t} = \int_0^\infty e^{-\lambda u} \frac{f(x + u) - f(x)}{u} du$$

for $f \in D(A) = \{f \in \mathcal{X} | f' \in \mathcal{X}\}$ [pure jump process]

Introduction

Feller semigroups $\{T(t)|t \geq 0\}$ for Lévy processes $\{X_t|t \geq 0\}$:

$$T(t)f(x) = E[f(x + X_t)] \text{ for } f \in \mathcal{X} = UCB(\mathbb{R})$$

Example: Poisson process: $P^{X_t} = \mathcal{P}(\lambda t)$ with $\lambda > 0$:

$$Af(x) = \lim_{t \downarrow 0} \frac{E[f(x + X_t) - f(x)]}{t} = \lambda[f(x+1) - f(x)]$$

for $f \in \mathcal{D}(A) = \mathcal{X}$

Introduction

Convolution semigroups $\{T(t)|t \geq 0\}$ on $\mathcal{X} = \ell^p$ with $p \in \{1, \infty\}$:

$$T(t)f = p(t) * f \text{ for } f \in \mathcal{X}$$

where $p(t) = (p(t)(0), p(t)(1), p(t)(2), \dots) \in \ell^1$ is an infinitely divisible discrete distribution, i.e.

$$p(s+t) = p(s) * p(t) \text{ for } s, t \geq 0.$$

Introduction

Convolution semigroups $\{T(t)|t \geq 0\}$ on $\mathcal{X} = \ell^p$ with $p \in \{1, \infty\}$:

$$T(t)f = p(t) * f \text{ for } f \in \mathcal{X}$$

Example: Negative binomial convolution semigroup: $p(t) = NB(t, p)$ with $0 < p < 1$:

$$p(t)(n) = \binom{t+n-1}{n} p^t (1-p)^n = \frac{\Gamma(t+n)}{\Gamma(t) \cdot n!} p^t (1-p)^n$$

$$Af(n) = \lim_{t \downarrow 0} \frac{1}{t} [T(t) - I]f(n) = \begin{cases} -\ln(p) \cdot f(0) & \text{if } n = 0 \\ \sum_{k=1}^n \frac{(1-p)^k}{k} [f(n-k) - f(n)] & \text{if } n > 0 \end{cases} \text{ for } f \in \mathcal{X}$$

Introduction

Convolution semigroups $\{T(t)|t \geq 0\}$ on $\mathcal{X} = \ell^p$ with $p \in \{1, \infty\}$:

$$T(t)f = p(t) * f \text{ for } f \in \mathcal{X}$$

Example: Poisson convolution semigroup: $p(t) = \mathcal{P}(\lambda t)$ with $\lambda > 0$:

$$Af(n) = \lim_{t \downarrow 0} \frac{1}{t} [T(t) - I]f(n) = \begin{cases} -\lambda f(0) & \text{if } n = 0 \\ \lambda [f(n-1) - f(n)] & \text{if } n > 0 \end{cases} \text{ for } f \in \mathcal{X}$$

Introduction

Representation theorems for $\{T(t)|t \geq 0\}$:

History (excerpt): Hille (1942), Widder (1946), Yosida (1948), Kendall (1954), Trotter (1958), Kato (1959), Chung (1962), Ditzian (1969), Butzer & Hahn (1980), Shaw (1980), Pfeifer (1984-1986)

If A is bounded: $T(t)f = e^{tA}f = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k f$ for $f \in \mathcal{D}(A) = \mathcal{X}$

$$T(t)f = e^{tA}f = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^n f \text{ for } f \in \mathcal{D}(A) = \mathcal{X}$$

$$T(t)f = e^{tA}f = \lim_{n \rightarrow \infty} \left(I - \frac{t}{n} A \right)^{-n} f \text{ for } f \in \mathcal{D}(A) = \mathcal{X}$$

Introduction

General case (A unbounded):

Hille's first exponential formula: with $A_h = \frac{1}{h}[T(h) - I]$:

$$T(t)f = \lim_{h \downarrow 0} \exp(tA_h)f \text{ for } f \in \mathcal{X}$$

Hille-Yosida: with resolvent $R(\lambda; A)f = (\lambda I - A)^{-1}f = \int_0^\infty e^{-\lambda t} T(t)f dt, \lambda > \omega$:

$$T(t)f = \lim_{\lambda \rightarrow \infty} \exp(tB_\lambda)f \text{ for } f \in \mathcal{X}$$

with

$$B_\lambda = \lambda[\lambda R(\lambda; A) - I] \text{ for } \lambda > \omega$$

Introduction

General case (A unbounded): with $A_h = \frac{1}{h}[T(h) - I]$:

Kendall:

$$T(t)f = \lim_{h \downarrow 0} \left(I + \frac{t}{n} A_{1/n} \right)^n f \text{ for } f \in \mathcal{X}$$

Shaw:

$$T(t)f = \lim_{h \downarrow 0} \left(I - \frac{t}{n} A_{1/n} \right)^{-n} f \text{ for } f \in \mathcal{X}$$

[particular cases of Chernov's product formula]

Probabilistic representations of operator semigroups

Some notation:

$$\varphi_N(t) := E[t^N] = \sum_{n=0}^{\infty} t^n P(N = n)$$

[probability generating function for a non-negative integer valued random variable N]

$$\psi_X(t) := E[e^{tx}] = \sum_{n=0}^{\infty} \frac{t^n}{n!} E[X^n]$$

[moment generating function for a non-negative real valued random variable X]

Probabilistic representations of operator semigroups

Main Representation Theorem (Pfeifer 1984): Let N be a non-negative integer-valued random variable with $E(N) = \zeta$ and $Y > 0$ be a real-valued random variable with $E(Y) = \gamma$ such that $\varphi_N(\delta_1) < \infty$ and $\psi_Y(\delta_2) < \infty$ for some $\delta_1 > 1$ and $\delta_2 > 0$. Then for sufficiently large n ,

$$\varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \in \mathcal{L}[\mathcal{X}, \mathcal{X}] \text{ with}$$

$$\left\| \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\| \leq M \varphi_N \left(\psi_Y \left(\frac{\omega}{n} \right) \right),$$

and

$$T(\xi)f = \lim_{n \rightarrow \infty} \left\{ \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\}^n f \text{ for } f \in \mathcal{X} \text{ with } \xi = \zeta\gamma.$$

Probabilistic representations of operator semigroups

Corollary: comprises (all) known representation theorems: for example:

Choose $P^N = B(1, \xi)$ [binomial distribution] and $Y \equiv 1$, then

$$\left\{ \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\}^n = \left(I + \frac{\xi}{n} A_{1/n} \right)^n \quad (\text{Kendall})$$

Choose $P^N = NB(1, \xi)$ [negative binomial distribution] and $Y \equiv 1$, then

$$\left\{ \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\}^n = \left(I - \frac{\xi}{n} A_{1/n} \right)^{-n} \quad (\text{Shaw})$$

Probabilistic representations of operator semigroups

Corollary: comprises (all) known representation theorems: for example:

Hille's first exponential formula: Choose $P^N = \mathcal{P}(\xi)$ and $Y \equiv 1$, then

$$\left\{ \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\}^n = \exp \left(\xi n \left[T \left(\frac{1}{n} \right) - I \right] \right) = \exp(\xi A_{1/n})$$

Probabilistic representations of operator semigroups

Corollary: comprises (all) known representation theorems: for example:

Hille-Yosida: Choose $P^N = \mathcal{P}(\xi)$ and $P^Y = \mathcal{E}(1)$, then

$$E\left[T\left(\frac{Y}{n}\right)\right]f = \int_0^\infty e^{-u}T\left(\frac{u}{n}\right)f du = n \int_0^\infty e^{-nv}T(v)f dv = nR(n; A)f \text{ for } f \in \mathcal{X}$$

and hence

$$\left\{ \varphi_N \left(E \left[T \left(\frac{Y}{n} \right) \right] \right) \right\}^n = \exp(\xi n [nR(n; A) - I]) = \exp(\xi B_n)$$

Probabilistic representations of operator semigroups

Idea of proof: Let $\{Y_n\}_{n \in \mathbb{N}}$ be i.i.d. as Y , independent of N . Consider the random sum $X = \sum_{k=1}^N Y_k$ with $\psi_X(t) = \varphi_N(\psi_Y(t))$. Then in some sense

$$\begin{aligned}
 E[T(X)] &= \sum_{n=0}^{\infty} E\left[T\left(\sum_{k=1}^n Y_k\right)\right] \cdot P(N = n) \\
 &= \sum_{n=0}^{\infty} E[T(Y_1) \circ T(Y_2) \circ \cdots \circ T(Y_n)] \cdot P(N = n) \\
 &= \sum_{n=0}^{\infty} E[T(Y)]^n \cdot P(N = n) = \varphi_N(E[T(Y)])
 \end{aligned}$$

Probabilistic representations of operator semigroups

Now take i.i.d. copies $\{X_n\}_{n \in \mathbb{N}}$ of X and consider $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ which converges a.s. (and thus also in probability) to $\xi = \zeta\gamma$. By the law of large numbers, it follows that

$$T(\xi)f = \lim_{n \rightarrow \infty} E[T(\bar{X}_n)]f = \lim_{n \rightarrow \infty} E\left[T\left(\frac{X}{n}\right)\right]^n f = \lim_{n \rightarrow \infty} \left\{ \varphi_N\left(E\left[T\left(\frac{Y}{n}\right)\right]\right) \right\}^n f \text{ for } f \in \mathcal{X}.$$

Idea of proof: modulus of continuity

$$\omega(f, \xi, \delta) := \sup \left\{ \|(T(t) - T(\xi))f\| \mid |t - \xi| < \delta, t > 0 \right\}$$

[Note that for $f \in \mathcal{X}$ and $\varepsilon > 0$, there exists a $\delta > 0$ with $\omega(f, \xi, \delta) < \varepsilon$]

Probabilistic representations of operator semigroups

For all $\delta > 0$, we have

$$\begin{aligned}
 \|E[T(\bar{X}_n)]f - T(\xi)f\| &= \|E[T(\bar{X}_n) - T(\xi)]f\| \leq E\|T(\bar{X}_n) - T(\xi)\|f\| \\
 &= \int \| [T(\bar{X}_n) - T(\xi)]f \| dP = \int_{\{|\bar{X}_n - \xi| < \delta\} \cup \{|\bar{X}_n - \xi| \geq \delta\}} \| [T(\bar{X}_n) - T(\xi)]f \| dP \\
 &\leq \omega(f, \xi, \delta) + \|f\| \int_{\{|\bar{X}_n - \xi| \geq \delta\}} \{ \|T(\bar{X}_n)\| + \|T(\xi)\| \} dP \\
 &\leq \omega(f, \xi, \delta) + \|f\| P(|\bar{X}_n - \xi| \geq \delta) M \{ E[e^{\omega \bar{X}_n}] + e^{\omega \xi} \} \leq 2\omega(f, \xi, \delta)
 \end{aligned}$$

for sufficiently large n .

Probabilistic representations of operator semigroups

Crucial point: In which sense does $E[T(X)]$ exist, and does there hold

$$E[T(X + Y)] = E[T(X) \circ T(Y)] = E[T(X)] \circ E[T(Y)] \quad (*)$$

for independent random variables X, Y ?

Theorem (Pfeifer 1984): If $\liminf_{t \downarrow z} \|T(t) - T(z)\| > 0$ for some $z > 0$ and the semigroup is injective in a neighbourhood of zero, then $t \mapsto T(t)$ is neither Borel-measurable nor separably valued, hence $E[T(X)]$ does in general not exist as a Bochner expectation.

Example: semigroup of translations: $T(t)f = f(\bullet + t)$ for $f \in \mathcal{X} = UCB(\mathbb{R})$ with $\|T(t) - T(z)\| = 2$ whenever $t \neq z$.

Solution: (modified) Pettis integral (uses a suitable subset of the dual space of $\mathcal{L}[\mathcal{X}, \mathcal{X}]$ and the Hahn-Banach-Theorem) $\rightarrow (*)$ can be justified!

Probabilistic representations of operator semigroups

Consequence: Main Representation Theorem (and extensions) can be used to find estimates for the rate of convergence, central idea:

$$E[T(X)]f - T(\xi)f = E(X - \xi)T(\xi)Af + \frac{1}{2}E(X - \xi)^2 T(\xi)A^2f + R \approx \frac{\sigma^2}{2}T(\xi)A^2f$$

with $\sigma^2 = \text{Var}(X)$; more precisely (among other results):

$$\left\| E[T(X)]f - T(\xi)f - \frac{\sigma^2}{2}T(\xi)A^2f \right\| \leq \frac{M}{6} \|A^3f\| \left\{ e^{\omega\xi} E|X - \xi|^3 + \omega \left\{ E(X - \xi)^6 \right\}^{2/3} \left\{ \psi_X(3\omega) \right\}^{1/3} \right\}$$

for $f \in \mathcal{D}(A^3) \rightarrow$ starting point for joint work on Poisson approximation

The semigroup approach to Poisson approximation

History (excerpt): Le Cam (1960), Franken (1964), Chen (1974), Serfling (1975, 1978), Arenbaev (1976), Shorgin (1977), Presman (1984), Barbour & Hall (1984), Serfozo (1985), Deheuvels and Pfeifer (1986-1989)

Startup framework: Let $X_1, \dots, X_n$ be independent binomially distributed over $\{0,1\}$ with $P(X_i = 1) = p_i \in (0,1)$ and T be Poisson distributed with parameter $\lambda > 0$. Define $S := \sum_{i=1}^n X_i$. Then for a large class of probability metrics ρ , we have

$$\rho(P^S, P^T) = \left\| \left\{ \prod_{i=1}^n (I + p_i A) - e^{\lambda A} \right\} f \right\|_\rho$$

for a suitable Banach space $\mathcal{X}$ with norm $\|\cdot\|_\rho$ and a suitable $f \in \mathcal{X}$. Here A is the generator of the Poisson convolution semigroup.

The semigroup approach to Poisson approximation

Examples:

total variation:

$$\rho(P^S, P^T) = \sup_{A \subseteq \mathbb{Z}^+} |P(S \in A) - P(T \in A)|: \quad \mathcal{X} = \ell^1, \|\cdot\|_\rho = \frac{1}{2} \|\cdot\|_{\ell^1}, f = (1, 0, 0, \dots) = g$$

Kolmogorov metric:

$$\rho(P^S, P^T) = \sup_{m \in \mathbb{Z}^+} |P(S \leq m) - P(T \leq m)|: \quad \mathcal{X} = \ell^\infty, \|\cdot\|_\rho = \|\cdot\|_{\ell^\infty}, f = (1, 1, 1, \dots) = h$$

Fortet-Mourier (Wasserstein) metric:

$$\rho(P^S, P^T) = \sum_{k=0}^{\infty} |P(S \leq k) - P(T \leq k)|: \quad \mathcal{X} = \ell^1, \|\cdot\|_\rho = \|\cdot\|_{\ell^1}, f = (1, 1, 1, \dots) = h$$

The semigroup approach to Poisson approximation

Theorem (Deheuvels & Pfeifer 1986): Under the assumptions of the startup framework, it holds

$$\rho(P^S, P^T) = \left\| \left\{ \prod_{i=1}^n (I + p_i A) - e^{\lambda A} \right\} f \right\|_\rho = \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho + \mathcal{O}(\max\{v, s^2, (\lambda - t)^2\})$$

if $s = \mathcal{O}(1)$ with $t = \sum_{k=1}^n p_k$, $s = \sum_{k=1}^n p_k^2$, $v = \sum_{k=1}^n p_k^3$ and $\delta = \frac{\lambda - t}{s}$.

The semigroup approach to Poisson approximation

Particular cases: $\lambda = \sum_{k=1}^n p_k = t$ [i.e. $E(S) = E(T)$]:

$$\rho(P^S, P^T) = \left\| \left\{ \prod_{i=1}^n (I + p_i A) - e^{\lambda A} \right\} f \right\|_{\rho} \sim \frac{s}{2} \|e^{tA} A^2 f\|_{\rho}$$

total variation:

$$\|e^{tA} A^2 f\|_{\rho} = \frac{1}{2} \sum_{k=0}^{\infty} e^{-t} \frac{t^{k-2}}{k!} |t^2 - 2kt + k(k-1)| = e^{-t} \left\{ \frac{t^{a-1}(a-t)}{a!} + \frac{t^{b-1}(b-t)}{b!} \right\} \sim \frac{2}{t\sqrt{2\pi e}}$$

with $a = \left\lfloor t + \frac{1}{2} + \sqrt{t + \frac{1}{4}} \right\rfloor$, $b = \left\lfloor t + \frac{1}{2} - \sqrt{t + \frac{1}{4}} \right\rfloor$, so

$$\rho(P^S, P^T) \sim \frac{1}{\sqrt{2\pi e}} \frac{\sum_{k=1}^n p_k^2}{\sum_{k=1}^n p_k}$$

The semigroup approach to Poisson approximation

Particular cases: $\lambda = \sum_{k=1}^n p_k = t$ [i.e. $E(S) = E(T)$]:

$$\rho(P^S, P^T) = \left\| \left\{ \prod_{i=1}^n (I + p_i A) - e^{\lambda A} \right\} f \right\|_{\rho} \sim \frac{s}{2} \|e^{tA} A^2 f\|_{\rho}$$

Kolmogorov metric:

$$\|e^{tA} A^2 f\|_{\rho} = e^{-t} \sup_{k \in \mathbb{Z}^+} \frac{t^{k-1}}{k!} |t - k| = e^{-t} \max \left\{ \frac{t^{a-1}(a-t)}{a!}, \frac{t^{b-1}(b-t)}{b!} \right\} \sim \frac{1}{t\sqrt{2\pi e}}$$

with $a = \left\lfloor t + \frac{1}{2} + \sqrt{t + \frac{1}{4}} \right\rfloor$, $b = \left\lfloor t + \frac{1}{2} - \sqrt{t + \frac{1}{4}} \right\rfloor$, so

$$\rho(P^S, P^T) \sim \frac{1}{2\sqrt{2\pi e}} \frac{\sum_{k=1}^n p_k^2}{\sum_{k=1}^n p_k}$$

The semigroup approach to Poisson approximation

Particular cases: $\lambda = \sum_{k=1}^n p_k = t$ [i.e. $E(S) = E(T)$]:

$$\rho(p^S, p^T) = \left\| \left\{ \prod_{i=1}^n (I + p_i A) - e^{\lambda A} \right\} f \right\|_{\rho} \sim \frac{s}{2} \|e^{tA} A^2 f\|_{\rho}$$

Fortet-Mourier (Wasserstein) metric:

$$\|e^{tA} A^2 f\|_{\rho} = \sum_{k=0}^{\infty} e^{-t} \frac{t^{k-1}}{k!} |t - k| = 2e^{-t} \frac{t^{\lfloor t \rfloor}}{\lfloor t \rfloor!} \sim \frac{2}{\sqrt{2\pi t}}$$

$$\text{so } \rho(p^S, p^T) \sim \frac{1}{\sqrt{2\pi}} \frac{\sum_{k=1}^n p_k^2}{\sqrt{\sum_{k=1}^n p_k}}$$

The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

The solution is of the form $\delta = \delta(t) \in \left[0, \frac{1}{2}\right]$ with $\lambda = t + \delta(t)s$

total variation:

$$\delta(t) = \begin{cases} \frac{1}{2} & \text{if } 0 < t \leq \sqrt{2} \\ \frac{1}{2} - \frac{3}{2t} \frac{2-t}{3-t} & \text{if } \sqrt{2} < t \leq \sqrt[3]{6} \\ 0 & \text{if } \sqrt[3]{6} < t \leq 2 \end{cases}$$

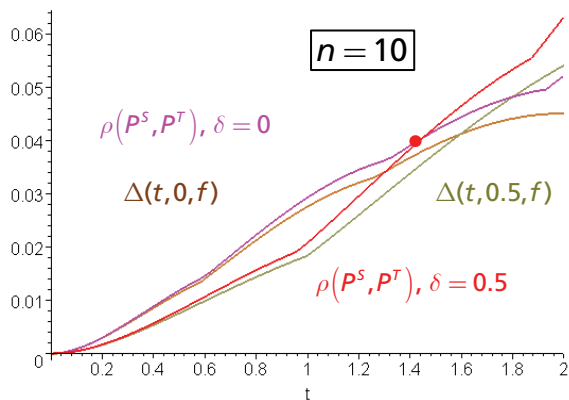
The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, \quad s = \frac{t^2}{n}$$



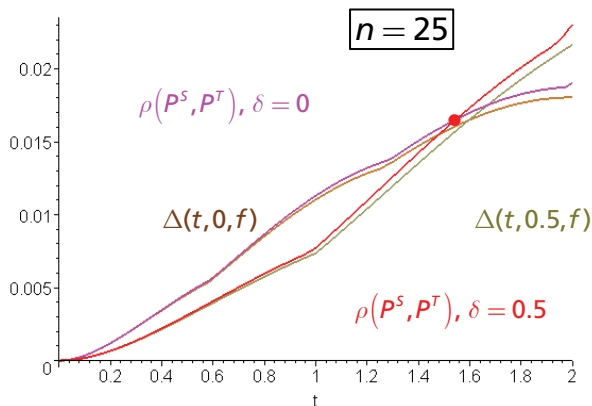
The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, \quad s = \frac{t^2}{n}$$



The semigroup approach to Poisson approximation

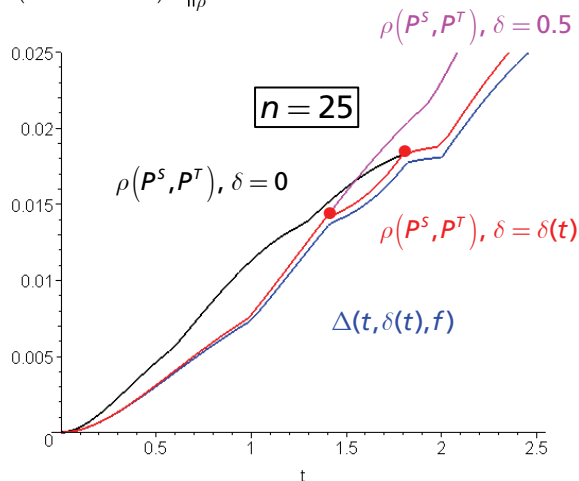
Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, s = \frac{t^2}{n}$$

$$\delta(t) = \begin{cases} \frac{1}{2} & \text{if } 0 < t \leq \sqrt{2} \\ \frac{1}{2} - \frac{3}{2t} \frac{2-t}{3-t} & \text{if } \sqrt{2} < t \leq \sqrt[3]{6} \\ 0 & \text{if } \sqrt[3]{6} < t \leq 2 \end{cases}$$



The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

The solution is of the form $\delta = \delta(t) \in \left[0, \frac{1}{2}\right]$ with $\lambda = t + \delta(t)s$

Kolmogorov metric:

$$\delta(t) = \begin{cases} \frac{1}{2} - \frac{1}{2(1+t)} & \text{if } 0 < t \leq \sqrt{3} - 1 \\ \frac{1}{2} - \frac{t}{2+t^2} & \text{if } \sqrt{3} - 1 < t \leq 1 \end{cases}$$

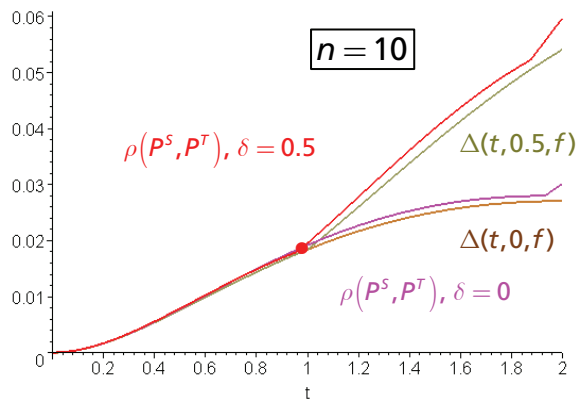
The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, \quad s = \frac{t^2}{n}$$



The semigroup approach to Poisson approximation

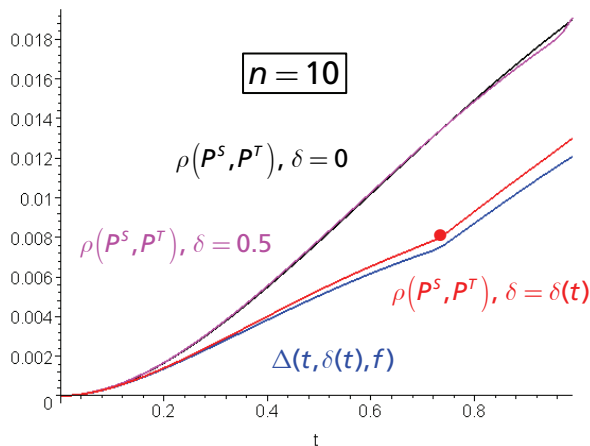
Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, s = \frac{t^2}{n}$$

$$\delta(t) = \begin{cases} \frac{1}{2} - \frac{1}{2(1+t)} & \text{if } 0 < t \leq \sqrt{3} - 1 \\ \frac{1}{2} - \frac{t}{2+t^2} & \text{if } \sqrt{3} - 1 < t \leq 1 \end{cases}$$



The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

The solution is of the form $\delta = \delta(t) \in \left[0, \frac{1}{2}\right]$ with $\lambda = t + \delta(t)s$

Fortet-Mourier metric:

$$\delta(t) = \begin{cases} \frac{1}{2} & \text{if } 0 < t \leq \ln 2 \\ 0 & \text{if } \ln 2 < t \leq 1 \\ \frac{1}{2} - \frac{1}{2t} & \text{if } 1 < t \leq \alpha \\ 0 & \text{if } \alpha < t \leq 2 \end{cases}$$

where $\alpha = 1.6784\dots$ is the positive root of $2(1 + \alpha) = e^\alpha$.

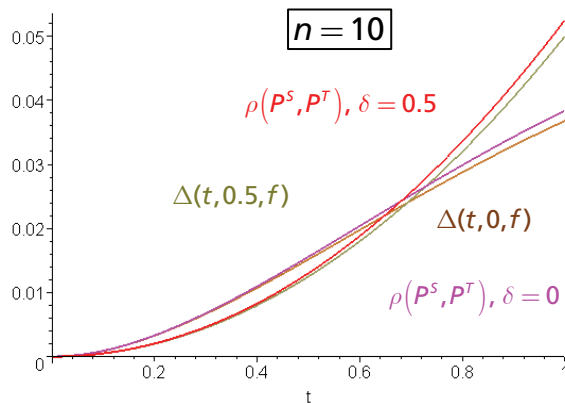
The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, \quad s = \frac{t^2}{n}$$



The semigroup approach to Poisson approximation

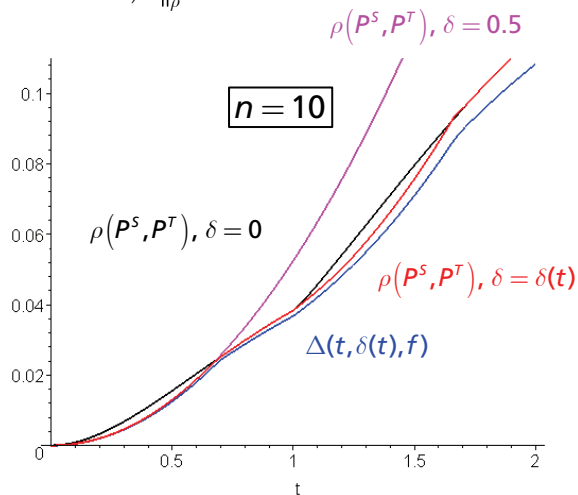
Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

Examples:

$$p_k = \frac{t}{n}, k = 1, \dots, n, s = \frac{t^2}{n}$$

$$\delta(t) = \begin{cases} \frac{1}{2} & \text{if } 0 < t \leq \ln 2 \\ 0 & \text{if } \ln 2 < t \leq 1 \\ \frac{1}{2} - \frac{1}{2t} & \text{if } 1 < t \leq \alpha \\ 0 & \text{if } \alpha < t \leq 2 \end{cases}$$



The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{s}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

The solution is of the form $\delta = \delta(t)$ with $\lambda = t + \delta(t)s$
(dependent on the underlying metric!)

For the total variation and the Fortet-Mourier metric, $\delta(t) = \frac{1}{2}$ is optimal for small values of t with $\lambda = t + \frac{s}{2} \sim \sum_{k=1}^n -\ln(1 - p_k)$ [Coupling approach, Serfling 1978] but e.g. not for the Kolmogorov metric!

The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{S}{2} \|e^{tA}(2\delta A + A^2)f\|_\rho$ w.r.t. δ !

$$\begin{aligned}
 & \|e^{tA}(2\delta A + A^2)g\|_{\ell^1} = \\
 & e^{-t} \sum_{k=0}^{\infty} \frac{t^{k-2}}{k!} \left| k^2 - 2k\left(t + \frac{1}{2} - \delta t\right) + t(t - 2\delta t) \right| = \\
 & 2 e^{-t} \left\{ \frac{t^{c-1}(c - (1-2\delta)t)}{c!} + \frac{t^{d-1}((1-2\delta)t - d)}{d!} \right\} \\
 & \sim \frac{2}{t\sqrt{2\pi}} \left\{ \zeta \exp\left(-\frac{1}{2\zeta^2}\right) + \frac{1}{\zeta} \exp\left(-\frac{1}{2\zeta^2}\right) \right\} \\
 & \geq \frac{4}{t\sqrt{2\pi}e} \quad (t \rightarrow \infty)
 \end{aligned}$$

$$\begin{aligned}
 & \text{where } c = \lceil t - r + \sqrt{t+r^2} \rceil, \quad d = \lceil t - r - \sqrt{t+r^2} \rceil \\
 & \text{and } r = \delta t - \frac{1}{2}, \quad \zeta = \delta\sqrt{t} + \sqrt{1+\delta^2 t}
 \end{aligned}$$

$$\begin{aligned}
 & \|e^{tA}(2\delta A + A^2)h\|_{\ell^\infty} = e^{-t} \sup_{n \geq 0} \frac{t^{n-1}}{n!} |(1-2\delta)t - n| = \\
 & e^{-t} \max \left\{ \frac{t^{c-1}(c - (1-2\delta)t)}{c!}, \frac{t^{d-1}((1-2\delta)t - d)}{d!} \right\} \\
 & \sim \frac{1}{t\sqrt{2\pi}} \max \left\{ \zeta \exp\left(-\frac{1}{2\zeta^2}\right), \frac{1}{\zeta} \exp\left(-\frac{1}{2\zeta^2}\right) \right\} \\
 & \sim \frac{\zeta}{t\sqrt{2\pi}} \exp\left(-\frac{1}{2\zeta^2}\right) \quad (t \rightarrow \infty)
 \end{aligned}$$

The semigroup approach to Poisson approximation

Question: What is an (asymptotically) “optimal” choice of λ ?

Answer: Minimize $\Delta(t, \delta, f) := \frac{S}{2} \|e^{tA} (2\delta A + A^2) f\|_\rho$ w.r.t. δ !

$$\|e^{tA} (2\delta A + A^2) f\|_1 = e^{-t} \sum_{k=0}^{\infty} \frac{t^{k-1}}{k!} |k - (1-2\delta)t| =$$

$$2\delta - 4\delta \sum_{k=0}^N e^{-t} \frac{t^k}{k!} + 2 e^{-t} \frac{t^N}{N!}$$

$$\sim \frac{2}{\sqrt{2\pi(1-2\delta)t}} \exp\left(-\frac{2\delta^2 t}{1-2\delta}\right) + 2\delta - 4\delta \phi(-2\delta\sqrt{t}) \quad (t \rightarrow \infty)$$

$$\text{where } N = \lfloor (1-2\delta)t \rfloor \quad \text{and } \phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du,$$

and for $\delta > 0$,

$$4\delta \phi(-2\delta\sqrt{t}) \sim \frac{2}{\sqrt{2\pi t}} \exp(-2\delta^2 t) \quad (t \rightarrow \infty).$$

The semigroup approach to Poisson approximation

Extensions:

- multinomial distributions
- point processes
- Markov chains
- mixed distributions
- Poisson-stopped sums
- weighted metrics
- signed measures
- ...

Bibliography

- [1] PFEIFER, D. (1984): *Probabilistic representations of operator semigroups - a unifying approach*. Semigroup Forum 30, 17 - 34.
- [2] PFEIFER, D. (1985): *Approximation-theoretic aspects of probabilistic representations for operator semigroups*. J. Approx. Theory 43, 271 - 296.
- [3] DEHEUVELS, P. AND PFEIFER, D. (1984): *A semigroup approach to Poisson approximation*. Ann. Prob. 14, 663 - 676.
- [4] DEHEUVELS, P. AND PFEIFER, D. (1986): *Operator semigroups and Poisson convergence in selected metrics*. Semigroup Forum 34, 203 - 224.
- [5] DEHEUVELS, P. AND PFEIFER, D. (1987): *Semigroups and Poisson approximation*. In: New Perspectives in Theoretical and Applied Statistics, M.L. Puri, J.P. Vilaplana and W.Wertz (Eds.), 439 - 448.
- [6] DEHEUVELS, P. AND PFEIFER, D. (1988): *On a relationship between Uspenky's theorem and Poisson approximations*. Ann. Inst. Stat. Math. 40, 671 - 681.
- [7] DEHEUVELS, P. AND PFEIFER, D. (1988): *Poisson approximations of multinomial distributions and point processes*. J. Multivar. Analysis 25, 65 - 89.
- [8] DEHEUVELS, P., KARR, A., SERFLING, R. AND PFEIFER, D. (1988): *Poisson approximations in selected metrics by coupling and semigroup methods with applications*. J. Stat. Plann. Inference 20, 1 - 22.
- [9] DEHEUVELS, P., PFEIFER, D. AND PURI, M.L. (1989): *A new semigroup technique in Poisson approximation*. Semigroup Forum 38, 189 - 201.

V. Berufung nach Oldenburg (1987-95)



Nachfolge Prof. Dr. Reiner Horst
(Lineare Optimierung, Operations Research)

Lehraufgaben:

Mathematik für Ökonomen (Analysis und Lineare Algebra)

Unterstützung der Arbeitsgruppe Stochastik

Mitwirkung an der Ausbildung von Lehramtstudenten

Mathematik für Ökonomen

(Lineare Algebra)

Prof. Dr. D. Pfeifer

Universität Oldenburg

© 1990

III. Elemente der Spieltheorie

1. Das Grundmodell

Wir betrachten hier nur sog. **Zweipersonen-Nullsummenspiele**, d.h. bei zwei beteiligten **Spielern** gewinnt der eine, was der andere verliert.

Mathematisch werden solche Spiele beschrieben durch die Angabe der **Strategiemengen** $\mathcal{X}_1$ und $\mathcal{X}_2$ für die beiden Spieler (welche hier als **endliche** Mengen mit m bzw. n Elementen vorausgesetzt seien) sowie der **Auszahlungsmatrix** $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, welche den Gewinn angibt, den Spieler 1 von Spieler 2 erhält, wenn der erste seine i -te und der zweite Spieler seine j -te Strategie wählt. In der Regel werden dabei von Spiel zu Spiel verschiedene Strategien zum Einsatz kommen; daher bezeichne in dem $1 \times m$ -Vektor $q = (q_1, \dots, q_m)$ mit $0 \leq q_i \leq 1, \sum_{i=1}^m q_i = 1$ die Zahl q_i die **Wahrscheinlichkeit**, mit

der Spieler 1 seine i -te Strategie wählt; analog bezeichne in dem $n \times 1$ -Vektor $p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$

mit $0 \leq p_j \leq 1, \sum_{j=1}^n p_j = 1$ die Zahl p_j die Wahrscheinlichkeit dafür, daß Spieler 2 seine j -te Strategie wählt. Die Wahrscheinlichkeitsvektoren q bzw. p heißen dabei **gemischte Strategien** der Spieler. Die Spieler werden dann versuchen, durch geeignete Bestimmung gemischter Strategien ihren **mittleren Gewinn** zu maximieren (Spieler 1) bzw. den **mittleren Verlust** zu minimieren (Spieler 2). Der mittlere Gewinn für Spieler 1 beträgt dabei bei Wahl des Strategienpaares q, p

$$W = \sum_{i=1}^m \sum_{j=1}^n q_i a_{ij} p_j = qAp.$$

Sind hier q^* bzw. p^* wieder die entsprechenden **optimalen** gemischten Strategien, so heißt der zugehörige mittlere Gewinn W für Spieler 1 auch **Wert des Spiels**. Ein Spiel heißt **fair**, wenn der Wert des Spiels Null beträgt, da dann keiner der Spieler gegenüber dem anderen (im Mittel) einen Vorteil besitzt.

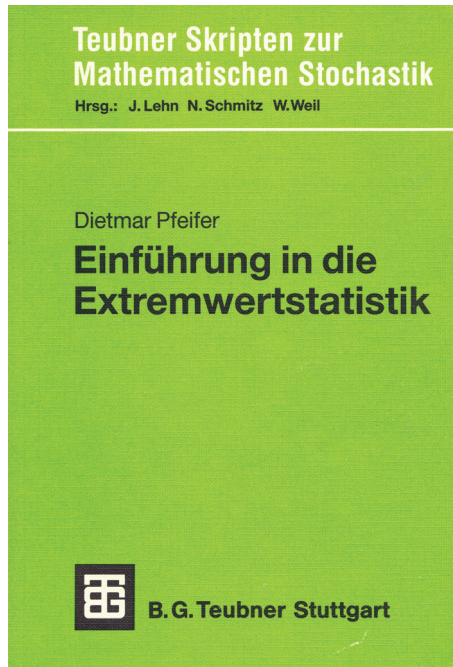
Mathematisch ist ein optimales gemischtes Strategienpaar also durch die Beziehung

$$q^*Ap^* = \max_q \min_p qAp$$

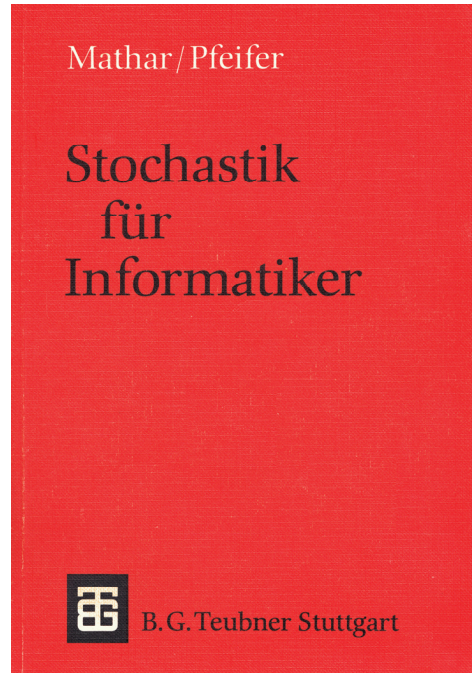
charakterisiert bzw. durch die Beziehung

$$qAp^* \leq q^*Ap^* \leq q^*Ap$$

für alle gemischten Strategien q, p .



1989



1990

Einblicke

Wissenschaft und Forschung an der
Carl von Ossietzky Universität Oldenburg



• Das Kölner Dom-Mosaik als Kunstwerk und Dokument
• Anekdote und Bauernwirtschaft
• Biblische und afrikanische Weisheit im Sprichwort • Vegetationsstrefen früh erkennen

Nr. 14

• Sanierung industrieller Altlasten am Beispiel Pöchl/Nordhorn • Forschungsprogramm Selektive Theorien • Tschernobyl aus der Sicht der Mathematik DM 3,-

Mathematik

Tschernobyl und die Folgen aus der Sicht der Mathematik

von Dietmar Pfeifer

A ls ich Ende April 1986 mit meiner Familie im Garten unserer heutigen Freunde in der Nähe unserer damaligen Weinreize Aachen bei strahlendem Weiter einen unbescherten Nachmittag verbrachte, war mir noch nicht klar, welche verheerenden Auswirkungen die Reaktor Katastrophe von Tschernobyl nicht nur in der Sowjetunion selbst, sondern flächendeckend in großen Teilen Europas haben sollte. Damals war es für mich unvorstellbar, daß ein Unfall in einem Kernkraftwerk in über 2000 km Entfernung derartig gravierende Konsequenzen für das alltägliche Leben haben konnte: Versäulter oder starke Einschränkung von bestimmten Lebensmittel wie Milch und Gemüse, Spielverbot für Kinder in Sandkästen oder Außenanlagen von Schulen, eingeschränkter Aufenthalt im Freien bei regnerischem Wetter... Tägliche Meldungen von Röntgen des Bodens und der Nahrungsmittel in Becquerel pro m² oder kg oder Liter wurden bald nur Gewohnheit. Eigene Messungen der Kontamination des Bodens mit Geräten von Kollegen des physikalischen Instituts der PWWI Aachen bestätigten die Meldungen in vollem Umfang: Mehrere Tausend Bq/m² auf dem Spielplatz hinter dem Haus, die Nahrungsmittel in den Gärten der Nachbarn, dem Gelände des städtischen Kindergartens, des unsere Tochter damals besuchte. Die anfängliche Beunruhigung über langfristige mögliche Folgen schien wie eine Zunahme von Krebskrankungen, oder schlimmer noch - da nicht kalkulierbar - gesundheitliche Veränderungen bei künftigen Generationen, sich mittlerweile allerdings gelegt. Erst die anfängliche des fünften Jahres der Katastrophe auflebende Beunruhigung in Zeitungen und Fernsehen läßt wieder aufleben; da geht es Dirfer in der Nähe des Unfalls, die sogar von höchsten politischen Stellen für die nächsten 700 Jahre als unbewohnbar angesehen werden; über 88000 umgesiedelte Menschen haben keine Heimat mehr. Wissenschaftler sprechen allein für die Sowjetunion von über 5000 Toten als unmittelbare Konsequenz der Katastrophe. Der für ursprünglich auf mindestens 30 Jahre Schutz ausgelegt Spätkörper der Röhre des Reaktors droht bereits jetzt zusammenzubrechen - eine zweite Verfallswelle blüht radioaktiven Staubs könnte Westeuropas erneut bedröhen.

Ein mathematisches Modell des radioaktiven Zerfalls

V on den zahlreichen bei dem Reaktorunfall von Tschernobyl freigesetzten Spaltprodukten sind in erster Linie die flüchtigen radioaktiven Isotope ¹³¹I (Jod), ¹³⁷Cs (Cäsium) und ⁹⁰Sr (Strontium) mit Halbwertszeiten von ca. 8 Tagen bzw. 30,2 bzw. 28,5 Jahren Hauptverursacher der hohen flächendeckenden Kontamination. Nach dem bereits vor 25 Jahren existierenden vierjährigen Chemie-Lexikon von Rönne ist ¹³⁷Cs eines der potentell gefährlichsten Radionuklide, da es von Magen-Darm-Trakt vollständig resorbiert wird und beim Menschen eine Halbwertszeit bezogen auf die Verweildauer im Körper von 100 bis 150 Tagen aufweist. ¹³⁷Cs, das sich bevorzugt in der Schilddrüse

ansammelt, liegt hier etwa im selben zytischen Bereich, während ⁹⁰Sr beim Menschen eine Halbwertszeit von mehreren Jahren besitzt, wodurch es wegen der Entfaltung im Knochenmark besonders gefährlich wird. Im folgenden wollen wir ein einfaches mathematisches Modell vorstellen, welches das Langzeitverhalten radioaktiver Stoffe hinsichtlich genau beschreibt, wobei spezifische physikalische Besonderheiten (z.B. Wechselwirkungen mit anderen Elementarteilchen, Kaskadenprozesse) vernachlässigt werden können, ohne die tatsächlichen Gegebenheiten qualitativ oder quantitativ wesentlich zu verändern. An den Beispielen (Jod) und Cäsium (langfristig) können wir dabei zwei statisch grundlegende verschiedene Gesichtspunkte des Zerfalls modellmäßig erfassen. In der Physik geht man im allgemeinen davon aus, daß der Zerfall eines radioaktiven Teilchens spontan, d.h. ohne erkennbare äußere Wirkung geschieht. Dem zufälligen Charakter dieses Vorgangs kann man dann mit Hilfe der Wahrscheinlichkeitstheorie dadurch modellieren, daß man für die Lebensdauer X_i eines individuellen Teilchens eine sogenannte Exponentialverteilung annimmt, d.h. es gilt:

$$P(X_i > x) = e^{-\lambda x}, \quad x \geq 0. \quad (1)$$

H erbei bezeichnet P das Symbol für Wahrscheinlichkeit (das probabilistisch ist) und λ eine positive Materialkonstante, die von der Halbwertszeit des Isotops abhängt. Der in (1) angegebene Ausdruck ist die formalistische Beschreibung der Wahrscheinlichkeit dafür, daß das Teilchen i länger als x Zeiteinheiten überlebt. Die Annahme einer solchen Exponentialverteilung ist mathematisch zwingend, wenn die Spontanität des Zerfalls gleichgesetzt wird mit der Bedingung

$$P(X_i > x + y | X_i > y) = P(X_i > x), \quad x, y \geq 0. \quad (2)$$

d.h. die bedingte Wahrscheinlichkeit dafür, daß das Teilchen i weitere x Zeiteinheiten überlebt, wenn es bereits y Zeiteinheiten

Abkürzung

We present a statistical model of radioactive decay which allows for an exhaustive study of the radiation time for a source of nuclear particles. The results are in particular applied to the nuclear power plant accident of Chernobyl in 1986.

überlebt hat, unabhängig von y ist und gerade genau so groß wie die -unbedingte- Wahrscheinlichkeit dafür, daß es überhaupt x Zeiteinheiten überlebt. Man sagt auch, die Exponentialverteilung

Einblicke Nr. 14 / November 1991
Carl von Ossietzky Universität Oldenburg

KETTENBRIEFE — WAS SIE VERSPRECHEN, WAS SIE HALTEN

von Dietmar Pfeifer, Oldenburg

(erschieden in: Stochastik in der Schule 12 (1992), Nr. 3, 37 – 47)

Zusammenfassung

Es wird ein einfaches stochastisches Modell eines Kettenbriefsystems vorgestellt, welches hinreichend genau erklärt, warum solche Systeme meist nicht das halten, was sie versprechen, wobei auch auf statistische Fragestellungen eingegangen wird. Eine Veranschaulichungsmöglichkeit des Modells durch den Einsatz eines PC in der Schule wird ebenfalls diskutiert.

A probabilistic variant of Chernoff's product formula

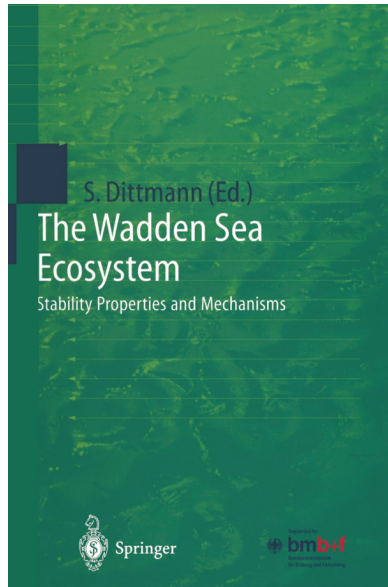
Dietmar Pfeifer

Abstract

In this paper we present a version of Chernoff's product formula based on probabilistic representations for equi-bounded operator semigroups which covers e.g. Hille's first exponential formula as well as the Post–Widder real inversion formula. General rates of convergence are given in terms of the rectified modulus of continuity.

Applicable Analysis, 1993

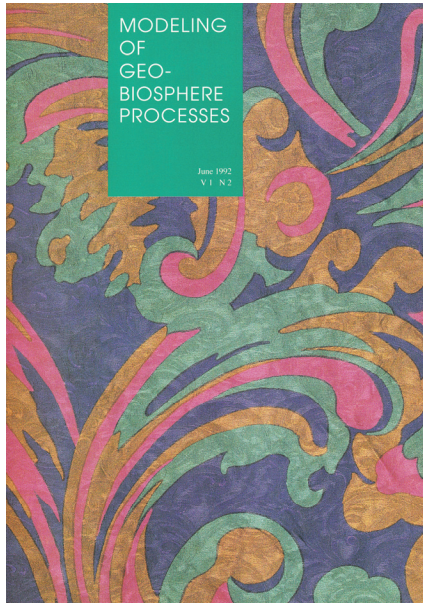
Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



The “Ecosystem Research Wadden Sea” consisted of four large joint projects carried out since 1989 on the southern North Sea coast of Germany. The projects had two major goals which were closely related. One was to develop a protection and management strategy for the Wadden Sea to preserve the diversity of its biota and landscape. This aspect also addressed the sustainable use of the Wadden Sea, i.e. a use of resources in such a way that no long-term decline of biological diversity results and the needs and aspirations of present and future generations are maintained. This goal was the task of two applied projects within the “Ecosystem Research Wadden Sea”, one in the state of Schleswig-Holstein and the other in Lower Saxony (Stock et al. 1996; Dittmann et al. 1997; Kellermann et al. 1997). The second goal was to achieve a fundamental understanding of this ecosystem. This was the task of two research projects, namely SWAP (“Sylter Wattenmeer Austausch Prozesse” = “Exchange processes of the Sylt Rømø-bight”), which was carried out in Schleswig-Holstein (Gätje & Reise 1998), and ELAWAT (“Elastizität des Ökosystems Wattenmeer” = “Resilience of the Wadden Sea ecosystem”; Dittmann et al. 1997), carried out in Lower Saxony.

Although both basic research projects were directed towards the level of the ecosystem, they reflected two different conceptual approaches to ecosystem research. Ecosystem research can mean an assessment of the flow of energy and matter. SWAP followed in the tradition of this approach, which is based on the insight that all ecological processes ultimately constitute a transformation of matter and energy. However, critics of this approach, which goes back to Odum (1971), doubt that energy and matter are the only valid currencies in ecology (Manson & McGlade 1993). They believe it fails to pay attention to the “elements” of ecological systems, the organisms (Jax 1996). SWAP was aware of this critique and did not neglect the study of species.

Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



Modeling Geo-Biosphere Processes vol. 1, p. 145-161 Cremlingen 1992

Spatial Point Processes and their Applications to Biology and Ecology

D. Pfeifer, H.-P. Bäumler & M. Albrecht

Summary

Random spatial point patterns occur in several areas in biology and ecology, e.g. in connection with the spreading of insect larvae (Richter & Söndgerath 1990, Example 3.1), the distribution of trees in woodlands (Cressie 1991, Chapter 8.2: Spatial Data Analysis of Longleaf Pines; Mecke, Schneider, Stoyan & Weil 1990, p. 136 ff.; Ripley 1981, Chapter 8.6: Examples), the distribution of bird's nests (Stoyan, Kendall & Mecke 1989, Example 5.1; Ripley 1981, p. 180 ff.), or experiments with revitalization of eroded areas in the wadden sea (Pfeifer, Bäumler & Albrecht 1992), to mention some typical examples.

Although the mathematical theory of random point patterns and pattern processes derived from such models is quite well developed and also a sufficiently application-oriented literature is available, the general importance of the field for mathematical modeling in the applied sciences has seemingly not yet been fully recognized.

Therefore, some of the most important features of stochastic point process theory and some further generalizations thereof like Boolean models and mosaics will be presented in the sequel, with particular emphasis on possible applications in biology and ecology.

Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)

Senckenbergiana maritima | 27 | (3/6) | 129–136 | Frankfurt am Main, 30. 12. 1996

Modeling Dynamics and Spatial Aggregation of Biological Populations by Stochastic Networks

KONSTANTIN BOROVKOV & DIETMAR PFEIFER & HANS-PETER BAUMER

With 2 Text-Figures

Abstract

[BOROVKOV, K. & PFEIFER, D. & BAUMER, H.-P. (1996): Modeling dynamics and spatial aggregation of biological populations by stochastic networks. – *Senckenbergiana marit.*, 27: (3/6): 129–136, 2 figs.; Frankfurt a. M.]

A mathematical model of an open stochastic network is presented which can be used to describe dynamics of a biological population, consisting of individuals of several species. The variables in the model are the numbers of individuals within disjoint observation windows, so that in a sense the model describes a 'macroscopic' behaviour of the population instead of dealing with individual behaviour of its members (i.e. their exact locations, what is the case when one exploits e.g. a point process model). Our approach allows to model quantitatively both immigration-emigration and (for some species and to some extent) reproduction processes, and also the simplest forms of interaction between individuals.

The model can easily be implemented as a computer program for simulation studies. Moreover, its remarkable feature is that (under natural conditions) there exists an equilibrium distribution of the process in the model, which is given by a simple explicit formula. Varying parameters of the model, one gets the whole spectrum of different equilibrium distributions, which provide rather good possibilities for fitting experimental data. In particular, the model (under a natural choice of parameters) exhibits the same peculiar variance-to-expectation relationships, as is usually discovered in experimental data sets.

Senckenbergiana maritima | 29 | (1/6) | 63–76 | Frankfurt am Main, 15. 12. 1998

Statistical Tools for Monitoring Benthic Communities

DIETMAR PFEIFER & HANS-PETER BAUMER & ROB DEKKER & ULRICH SCHLEIER

With 12 Text-Figures and 6 Tables

Key words: Wadden Sea, monitoring, statistics, non-metric multidimensional scaling, minimal areas, diversity indices, conservation policy, economic development, multi criteria model for policy decisions.

Abstract

[PFEIFER, D. & BAUMER, H.-P. & DEKKER, R. & SCHLEIER, U. (1998): Statistical tools for monitoring benthic communities. – *Senckenbergiana marit.*, 29 (1/6): 63–76, 12 figs., 6 tabs.; Frankfurt a. M.]

On the basis of long-term data from the monitoring program of the *Rijksinstituut voor Kust en Zee* (RIKZ), The Netherlands, we discuss how various statistical procedures can be successfully applied to the monitoring of benthic communities, especially in the Wadden Sea area. In particular, the following topics will be treated:

- (i) *Non-Metric Multidimensional Scaling* (NMDS) for visualization of trends and/or changes in species composition;
- (ii) *Minimal Areas* for checking the representativity of samples w.r.t. the number of species present;
- (iii) *Diversity* of species communities and its statistical estimation from samples including estimates for the estimation error in terms of their variance and standard deviation.

Although most of the underlying statistical techniques are in principle well-established [see, e.g. DUDAY & KNEPPEL (1987) or STRELLERBERG (1991)] and applications in marine ecology are numerous, some new statistical results are added from an application-oriented point of view.

Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



The "Minimal Area" Problem in Ecology: A Spatial Poisson Process Approach

D. Pfeifer¹, H.-P. Bäumer², U. Schleier³

¹ FB Mathematik, Institut für Mathematische Stochastik,
Universität Hamburg, D-20146 Hamburg, FRG

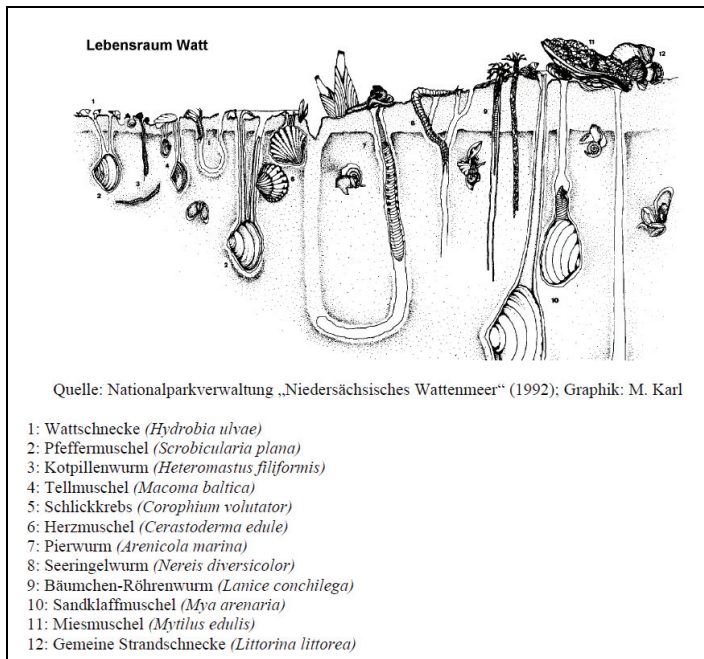
² Hochschulrechenzentrum, ³ Fachbereich Mathematik,
Carl von Ossietzky Universität, D-26111 Oldenburg, FRG

Summary

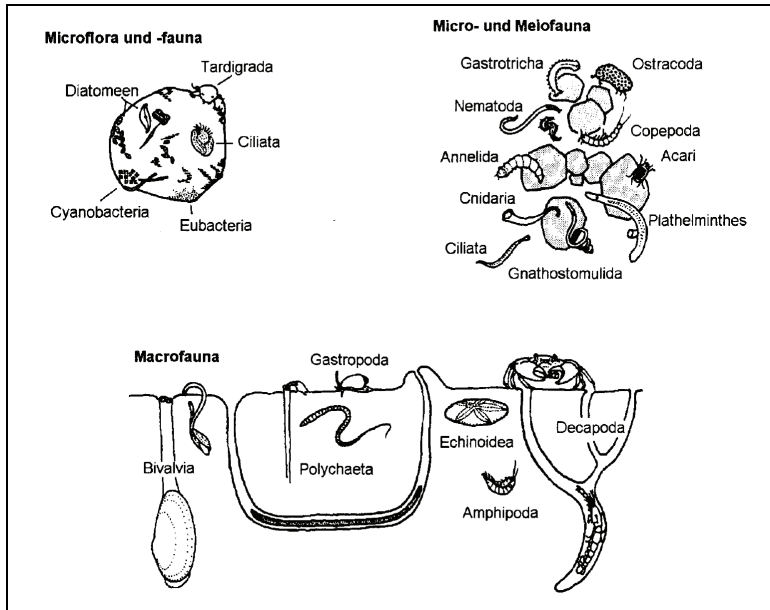
Ecologists frequently deal with the problem of finding an appropriate size of the observation area such that all species of a certain community are represented in that area. A *minimal area* is – in a certain sense – the smallest observation area with this property. In this paper we show that modelling the spatial distribution of biological communities by superpositions of homogeneous Poisson point processes not only allows for a simple explanation of "typical" forms of observed species–area curves, but gives also rise to explicit calculations of the distribution and moments (expectation, variance) of the size of the corresponding minimal area.

Keywords: Species–Area Curve, Sampling, Statistics of Extremes

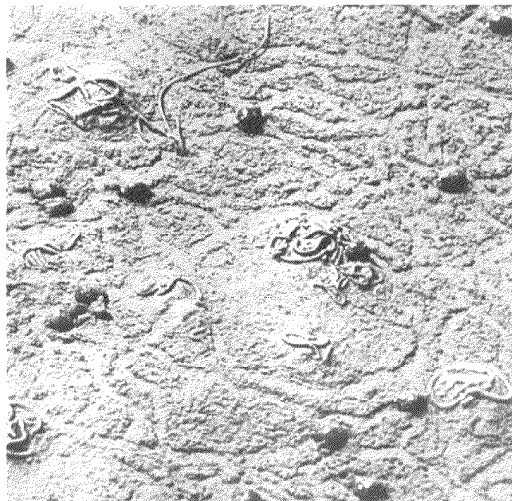
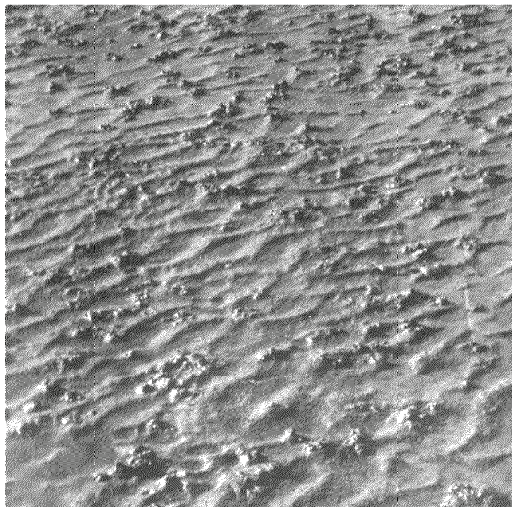
Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



Projekt: Ökosystemforschung Niedersächsisches Wattenmeer (ELAWAT: Elastizität des Ökosystems Wattenmeer), Teilprojekt B (1993 - 2000)



Räumliche Verteilung von *Arenicola marina* (links)
und *Littorina littorea* (rechts)

Grundzüge der Statistischen Ökologie

D. Pfeifer
H.-P. Bäumler
U. Schleier
V. de Valk

Räumliche Statistik

I. Punktprozess-Modelle (nach Skript *Räumliche Statistik*)

Punktprozesse bilden den Einstieg in die einfachsten zufälligen räumlichen Strukturen mit diskretem Charakter. Ein wichtiger Punktprozess, der die Grundlage zahlreicher Verallgemeinerungen bildet, ist der (homogene und inhomogene) *räumliche Poisson-Prozess*, der in seiner eindimensionalen Version als *Markov-Prozess* in der Warteschlangentheorie, der Informatik und der Versicherungs- und Finanzmathematik von fundamentaler Bedeutung ist.

I.1. Maßtheoretische Grundlagen

Lemma 1: Es sei $\mathfrak{X}$ eine nicht-leere Menge, $(\mathfrak{Y}_i, \mathcal{B}_i)$, $i \in I$ seien Messräume mit einer nicht-leeren Indexmenge I . Ferner seien $T_i: \mathfrak{X} \rightarrow \mathfrak{Y}_i$ für $i \in I$ beliebige Abbildungen. $\mathcal{B} := \sigma(T_i; i \in I)$ bezeichne wie üblich die von den Abbildungen T_i erzeugte σ -Algebra über $\mathfrak{X}$. Ist $(\Omega, \mathcal{A})$ ein weiterer Messraum und $T: \Omega \rightarrow \mathfrak{X}$ eine Abbildung, so gilt: T ist genau dann $\mathcal{A}$ - $\mathcal{B}$ -messbar, wenn alle Kompositionen $T_i \circ T$, $i \in I$ $\mathcal{A}$ - $\mathcal{B}_i$ -messbar sind.

Lemma 2: Es sei $(\mathfrak{X}, \mathcal{B})$ ein Messraum und $\mathcal{M}$ eine nicht-leere Menge von Maßen auf $\mathcal{B}$. Für jedes $B \in \mathcal{B}$ sei die Abbildung $\tau_B: \mathcal{M} \rightarrow \mathbb{R}$ definiert durch $\tau_B(\mu) := \mu(B)$ für jedes $\mu \in \mathcal{M}$ (die so definierten Abbildungen heißen *Evaluationsabbildungen*). $\mathfrak{M}$ bezeichne die von der Gesamtheit der Evaluationsabbildungen erzeugte σ -Algebra über $\mathcal{M}$. Dann gilt:

- a) Für jedes $A \in \mathcal{B}$ und jede Borel-Menge $D \in \mathcal{B}^1$ gehört die Menge $\{\mu \in \mathcal{M} \mid \mu(A) \in D\}$ zu $\mathfrak{M}$.
- b) Besitzt die σ -Algebra $\mathcal{B}$ einen abzählbaren, erzeugenden Semiring $\mathcal{S}$, der eine Folge von Mengen enthält, deren Vereinigung $\mathfrak{X}$ überdeckt und auf dem alle $\mu \in \mathcal{M}$ endlich sind (d.h. alle Maße $\mu \in \mathcal{M}$ sind σ -endlich), so gehören alle einelementigen Mengen $\{\mu\}$ für $\mu \in \mathcal{M}$ zu $\mathfrak{M}$, d.h. $\mathcal{T} := \{\{\mu\} \mid \mu \in \mathcal{M}\}$ ist das zu $\mathfrak{M}$ gehörige Atomsystem. Ferner gilt dann: $\{\tau_E \mid E \in \mathcal{S}\}$ ist ein Erzeuger von $\mathfrak{M}$. Diese Aussage bleibt auch gültig, wenn $\mathcal{S}$ lediglich ein erzeugender Semiring von $\mathcal{B}$ ist.

Lemma 2: Es sei $(\mathfrak{X}, \mathcal{B})$ ein Messraum und $\mathcal{M}$ eine nicht-leere Menge von Maßen auf $\mathcal{B}$. Für jedes $B \in \mathcal{B}$ sei die Abbildung $\tau_B : \mathcal{M} \rightarrow \mathbb{R}$ definiert durch $\tau_B(\mu) := \mu(B)$ für jedes $\mu \in \mathcal{M}$ (die so definierten Abbildungen heißen *Evaluationsabbildungen*). $\mathfrak{M}$ bezeichne die von der Gesamtheit der Evaluationsabbildungen erzeugte σ -Algebra über $\mathcal{M}$. Dann gilt:

c) Für alle $\alpha, \beta \geq 0$ ist die Abbildung $S_{\alpha\beta} : \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{M} : (\mu, \nu) \mapsto \alpha\mu + \beta\nu$ $\mathfrak{M} \otimes \mathfrak{M}$ -messbar. Allgemeiner gilt:

Ist $\alpha = \{\alpha_n\}_{n \in \mathbb{N}}$ eine Folge nichtnegativer reeller Zahlen, so ist die

Abbildung $S_\alpha : \bigtimes_{n=1}^{\infty} \mathcal{M} \rightarrow \mathcal{M} : \{\mu_n\}_{n \in \mathbb{N}} \mapsto \sum_{n=1}^{\infty} \alpha_n \mu_n$ $\bigotimes_{n=1}^{\infty} \mathfrak{M}$ -messbar.

d) Es sei einschränkend $\mathcal{M}$ eine Menge σ -endlicher Maße auf $\mathcal{B}$. Ist $(\mathfrak{Y}, \mathcal{C})$ ein weiterer Messraum und $\mathcal{L}$ eine nicht-leere Menge σ -endlicher Maße auf $\mathcal{C}$, so ist die Abbildung $S_{\otimes} : \mathcal{M} \times \mathcal{L} \rightarrow \mathcal{N} : (\mu, \nu) \mapsto \mu \otimes \nu$ $\mathfrak{M} \otimes \mathfrak{L}$ -messbar, wobei $\mathfrak{L}$ die von den Evaluationsabbildungen auf $\mathcal{C}$ und $\mathfrak{N}$ die von den Evaluationsabbildungen auf $\mathcal{N}$, der Menge aller σ -endlichen Maße auf $\mathcal{B} \otimes \mathcal{C}$, erzeugten σ -Algebren bezeichnen.

Lemma 2: Es sei $(\mathfrak{X}, \mathcal{B})$ ein Messraum und $\mathcal{M}$ eine nicht-leere Menge von Maßen auf $\mathcal{B}$. Für jedes $B \in \mathcal{B}$ sei die Abbildung $\tau_B : \mathcal{M} \rightarrow \mathbb{R}$ definiert durch $\tau_B(\mu) := \mu(B)$ für jedes $\mu \in \mathcal{M}$ (die so definierten Abbildungen heißen *Evaluationsabbildungen*). $\mathfrak{M}$ bezeichne die von der Gesamtheit der Evaluationsabbildungen erzeugte σ -Algebra über $\mathcal{M}$. Dann gilt:

e) Ist $D \in \mathcal{B}$ fest, so ist die Spurmaß-Abbildung $S_D : \mathcal{M} \rightarrow \mathcal{M} : \mu \mapsto \mu(\cdot \cap D)$ $\mathfrak{M}$ - $\mathfrak{M}$ -messbar.

Bemerkung:

Das Atomsystem $\mathcal{T} := \{\{\mu\} \mid \mu \in \mathcal{M}\}$ ist sicher dann *erschöpfend*, wenn $\mathcal{M}$ abzählbar ist. Für überabzählbare $\mathcal{M}$ ist das nicht unbedingt so; Gegenbeispiel: $\mathcal{M} = \{\varepsilon_x \mid x \in \mathfrak{X} := \mathbb{R}\}$ (Menge aller Einpunktmaße über $\mathbb{R}$). Hier gilt $\tau_A^{-1}(\{1\}) = \{\varepsilon_x \mid x \in A\}$, $\tau_A^{-1}(\{0\}) = \{\varepsilon_x \mid x \in A^c\} = \tau_{A^c}^{-1}(\{1\})$ für alle $A \in \mathcal{B}^1$ und somit $\mathfrak{M} = \{\{\varepsilon_x \mid x \in A\} \mid A \in \mathcal{B}^1\}$; $\mathfrak{M}$ ist damit in gewisser Weise „isomorph“ zur Borel’schen σ -Algebra $\mathcal{B}^1$ (über die Abbildung $x \mapsto \varepsilon_x$ für $x \in \mathbb{R}$), für die das entsprechende Atomsystem der elementigen Mengen gerade *nicht* erschöpfend ist.

Definition 1: In der Situation von Lemma 2 sei zusätzlich $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Ein Zufallselement $\xi : (\Omega, \mathcal{A}) \rightarrow (\mathcal{M}, \mathfrak{M})$ heißt *zufälliges Maß* über $\mathcal{B}$ (mit Werten in $\mathcal{M}$).

Lemma 3: Es sei I eine nicht-leere Indexmenge und $\{\xi_i\}_{i \in I}$ eine stochastisch unabhängige Familie von zufälligen Maßen über $\mathcal{B}$ (mit Werten in $\mathcal{M}$). Dann ist auch jede Familie $\{\xi_i(B_i)\}_{i \in I}$ mit Mengen $\{B_i\}_{i \in I} \subseteq \mathcal{B}$ stochastisch unabhängig.

Bemerkung: Lemma 3 kann direkt auf Familien von Zufallsvektoren $\left\{ \left(\xi_i(B_{i1}), \dots, \xi_i(B_{ik_i}) \right) \right\}_{i \in I}$ mit Mengen $B_{i1}, \dots, B_{ik_i} \in \mathcal{B}$ im Sinne der Bemerkung zu Definition 1 erweitert werden. Sind umgekehrt alle solchen Familien stochastisch unabhängig, so ist unter den Voraussetzungen in der zweiten Bemerkung zu Definition 1 auch die Familie $\{\xi_i\}_{i \in I}$ stochastisch unabhängig.

Lemma 4: In der Situation von Definition 1 sei ξ ein zufälliges Maß über $\mathcal{B}$ (mit Werten in $\mathcal{M}$). Dann wird durch die „punktweise“ Definition

$$(E\xi)(B) := E[\xi(B)] = \int \xi(B) dP, \quad B \in \mathcal{B}$$

ein Maß auf $\mathcal{B}$ definiert. $E\xi$ heißt (das) *Intensitätsmaß* zu ξ .

Bemerkung: Im Sinne des Satzes 1 ist eine reguläre bedingte Verteilung $P(\cdot | \mathcal{C})$ für eine beliebige Teil- σ -Algebra $\mathcal{C} \subseteq \mathcal{A}$ ein zufälliges Maß $\xi: (\Omega, \mathcal{A}, P) \rightarrow (\mathcal{M}, \mathfrak{M})$, wobei $\mathcal{M}$ die Menge aller Wahrscheinlichkeitsmaße auf $\mathcal{A}$ ist. Denn eine reguläre bedingte Verteilung $P(\cdot | \mathcal{C})$ besitzt gerade die genannten Eigenschaften (Wahrscheinlichkeitsmaß in der einen, messbare Abbildung in der anderen Variablen), ist also ein zufälliges (Wahrscheinlichkeits-)Maß der angegebenen Art.

Für das Intensitätsmaß ergibt sich hier noch

$$(E\xi)(B) := E[\xi(B)] = \int \xi(B) dP = \int P(B | \mathcal{C}) dP = P(B), \quad B \in \mathcal{B}, \text{ also } E\xi = P.$$

Definition 2: Ein zufälliges Maß ξ im Sinne der Definition 1 heißt *Punktprozess*, wenn gilt: $\xi(B) \in \mathbb{Z}^+$ für alle $B \in \mathcal{B}$ (d.h. $\xi(B)$ „zählt“ die Anzahl der „Punkte“ in der Menge $B \in \mathcal{B}$).

Definition 3: Es sei X ein Zufallselement auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Werten in einem Messraum $(\mathfrak{X}, \mathcal{B})$. Dann heißt das zufällige Maß $\xi = \varepsilon_X$ *Einpunkt- oder Dirac-Prozess*.

Bemerkung: Für jedes $B \in \mathcal{B}$ gilt

$$\xi(B) = \varepsilon_X(B) = \mathbb{1}_B(X)$$

(mit der üblichen Indikatorfunktion $\mathbb{1}_B$), und dies ist eine $\mathcal{A}$ -messbare Abbildung. Daher ist gemäß der ersten Bemerkung zu Definition 1 $\xi = \varepsilon_X$ ein zufälliges Maß. Hier gilt insbesondere

$$(E\xi)(B) = E[\xi(B)] = E[\mathbb{1}_B(X)] = P(X \in B) = P^X(B), \text{ also } E\xi = P^X.$$

I.2. Der Poisson-Punktprozess

Die in diesem Abschnitt betrachteten Punktprozess-Modelle sind fundamental für alle Anwendungen, in denen gewisse räumliche Unabhängigkeiten eine wesentliche Rolle spielen. Wir behalten hier der Einfachheit halber die Bezeichnungen aus Abschnitt I.1 bei, d.h. $(\mathcal{X}, \mathcal{B})$ ist ein Messraum und $\mathcal{M}$ eine geeignete nicht-leere Menge von Maßen auf $\mathcal{B}$, $\mathfrak{M}$ die von den Evaluationsabbildungen über $\mathcal{M}$ erzeugte σ -Algebra und ein zufälliges Maß $\xi: (\Omega, \mathcal{A}) \rightarrow (\mathcal{M}, \mathfrak{M})$ eine messbare Abbildung mit Werten in $\mathcal{M}$.

Definition 4: Ein zufälliges Maß ξ heißt *Poisson-Punktprozess* mit Intensitätsmaß μ (auf $\mathcal{B}$), wenn die folgenden Eigenschaften erfüllt sind:

- a) Für jedes $B \in \mathcal{B}$ ist $\xi(B)$ Poisson-verteilt mit Parameter $\mu(B)$.
- b) Ist I eine beliebige Indexmenge und $\{B_i\}_{i \in I}$ eine Familie paarweise disjunkter Mengen aus $\mathcal{B}$, so ist die Familie der Zufallsvariablen $\{\xi(B_i)\}_{i \in I}$ stochastisch unabhängig.

Bemerkungen:

- Falls ein solches zufälliges Maß existiert, ist seine Verteilung gemäß der zweiten Bemerkung zu Definition 1 eindeutig bestimmt.
- Die Frage der Existenz lässt sich zumindest im Fall, dass das Intensitätsmaß μ σ -endlich ist, positiv (und konstruktiv) beantworten. Dabei werden einige wesentliche Eigenschaften der Familie der Poisson-Verteilungen ausgenützt, nämlich die *Faltungsstabilität* und die *Zerlegungsstabilität*.

Satz 2: Es sei $I = \{1, 2, \dots, n\}$ mit $n \in \mathbb{N}$ eine endliche Indexmenge und $\{\lambda_i\}_{i \in I}$ eine Familie positiver reeller Zahlen. Ferner sei $\lambda := \sum_{i \in I} \lambda_i$.

- a) Ist $\{X_i\}_{i \in I}$ eine unabhängige Familie Poisson-verteilter Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit $P^{X_i} = \mathcal{P}(\lambda_i)$, $i \in I$, so gilt für $S := \sum_{i \in I} X_i$: $P^S = \mathcal{P}(\lambda)$ bzw. $\mathcal{P}\left(\sum_{i \in I} \lambda_i\right) = \ast_{i \in I} \mathcal{P}(\lambda_i)$ (Faltungsstabilität der Poisson-Verteilung).
- b) Ist S eine $\mathcal{P}(\lambda)$ -verteilte Zufallsvariable auf $(\Omega, \mathcal{A}, P)$, unabhängig von einer weiteren Familie $\mathbf{Z}_k = (Z_{k1}, \dots, Z_{kn})$, $k \in \mathbb{N}$ stochastisch unabhängiger, n -dimensionaler multinomialverteilter Zufallsvektoren mit $\sum_{i \in I} Z_{ki} = 1$ und $P(Z_{ki} = 1) = p_i := \frac{\lambda_i}{\lambda} = 1 - P(Z_{ki} = 0)$ für alle $k \in \mathbb{N}$ und $i \in I$, so sind die Zufallsvariablen $X_i := \sum_{k=1}^S Z_{ki}$ für $i \in I$ stochastisch unabhängig und jeweils $\mathcal{P}(\lambda_i)$ -verteilt (Zerlegungsstabilität der Poisson-Verteilung).

Bemerkung: Die Aussagen von Satz 2 lassen sich zum einen auf den Fall verallgemeinern, dass die λ_i lediglich nicht-negativ sind mit der Setzung $\mathcal{P}(0) = \varepsilon_0$ (Einpunktverteilung in Null); zum anderen bleibt die Aussage a) sinngemäß auch für unendliche (aber abzählbare) Indexmengen I richtig.

Satz 3: Es sei $\{X_n\}_{n \in \mathbb{N}}$ eine unabhängige, identisch (wie X) verteilte Familie von Zufallselementen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Werten in dem Messraum $(\mathcal{X}, \mathcal{B})$. $Q = P^X$ bezeichne die Verteilung von X . Ferner sei N eine von $\{X_n\}_{n \in \mathbb{N}}$ stochastisch unabhängige Poisson-verteilte Zufallsvariable auf $(\Omega, \mathcal{A}, P)$ mit Parameter $\lambda \geq 0$. Dann ist der durch

$$\xi := \sum_{n=1}^N \varepsilon_{X_n}$$

definierte Punktprozess ein *Poisson-Punktprozess* mit dem Intensitätsmaß $\mu = \lambda \cdot Q$. (Man sagt auch, ξ sei ein Poisson-Punktprozess mit den *Parametern* λ und Q .)

Ist umgekehrt μ ein beliebiges, endliches Maß auf $\mathcal{B}$, welches ein vom Nullmaß verschiedenes Intensitätsmaß eines Punktprozesses ξ ist, und ist der Grundraum $(\Omega, \mathcal{A}, P)$ genügend „groß“ (d.h. er erlaubt die Existenz beliebig vieler stochastisch unabhängiger Ereignisse), so wird durch

$$\xi^* := \sum_{n=1}^N \varepsilon_{X_n}$$

mit N und $\{X_n\}_{n \in \mathbb{N}}$ wie oben für $Q = \frac{1}{\lambda} \mu$ ein Poisson-Punktprozess definiert, der dieselbe Verteilung wie ξ besitzt (und daher stochastisch nicht von ξ „unterscheidbar“ ist). ξ^* heißt deshalb auch *Repräsentant* von ξ .

Zwischenbemerkungen:

- Für den Fall $N \equiv 0$ (d.h. $\lambda = E(N) = 0$) wird ξ auf Grund der dann leeren Summe zum *Nullmaß*. Ist umgekehrt das Intensitätsmaß μ vom Nullmaß verschieden, so ist zwingend $\lambda = E(N) > 0$. Daher ist Q in diesem Fall wohldefiniert.
- ξ ist definitionsgemäß ein zufälliges Maß, denn für beliebiges $B \in \mathcal{B}$ sind alle Abbildungen

$$\xi(B) = \sum_{n=1}^N \varepsilon_{X_n}(B) = \sum_{n=1}^N \mathbb{1}_B(X_n)$$

$\mathcal{A}$ -messbar.

Bemerkungen:

- Ein Poisson-Punktprozess mit endlichem Intensitätsmaß μ kann offenbar als zufällige Mischung von Einpunkt-Prozessen nach einer Poisson-Verteilung aufgefasst werden.
- Ist die Verteilung Q der Ausgangs-„Punkte“ $\{X_n\}_{n \in \mathbb{N}}$ atomlos in dem Sinne, dass $\{x\} \in \mathcal{B}$ gilt für alle $x \in \mathcal{X}$ mit $Q(\{x\}) = 0$, so ist auch die Verteilung P^ξ des Poisson-Punktprozesses atomlos. Insbesondere gilt dann $P(\xi = \mu) = 0$ für alle nicht-trivialen Maße $\mu \in \mathcal{M}$. Eine „Visualisierung“ des Poisson-Punktprozesses (z.B. im Fall $\mathcal{X} = \mathbb{R}^d$ mit $d = 1, 2, 3$) erfolgt dann üblicherweise durch Plotten der zufälligen „Punkt“-Menge Γ . Punktprozesse dieser Art heißen auch *einfach*.
- Ein Poisson-Punktprozess mit Parametern λ und atomloser Verteilung Q kann auch mit der zufälligen (fast sicher endlichen) „Punkt“-Menge $\Gamma = \{X_1, \dots, X_N\}$ identifiziert werden (mit $\Gamma = \emptyset$ für $N = 0$). Im Sinne der Maßtheorie ist das eine so genannte „zufällige abgeschlossene Menge“ (ZAM).

Bemerkungen:

- Ist die Verteilung Q der Ausgangs-„Punkte“ $\{X_n\}_{n \in \mathbb{N}}$ rein diskret, d.h. gibt es eine abzählbare Teilmenge $\mathcal{D} \in \mathcal{B}$ mit $Q(\mathcal{D}) = 1$, so ist auch die Verteilung P^ξ des Poisson-Punktprozesses rein diskret. In diesem Fall ist $P(\xi = \mu) > 0$ für gewisse Maße $\mu \in \mathcal{M}$.

Beispiel: $(\mathcal{X}, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}^d)$ mit $d \in \mathbb{N}$ und $\mathcal{D} = \{\mathbf{0}\}$. Hier gilt $\xi = N \cdot \varepsilon_0$ und somit $P(\xi = \varepsilon_0) = P(N = 1) = \lambda \cdot e^{-\lambda} > 0$, falls $\lambda > 0$. Die Visualisierung eines solchen Poisson-Punktprozesses geschieht dann üblicherweise durch Plotten der zufälligen „Punkt“-Menge Γ unter Anbringung einer „Marke“, die die Vielfachheit der jeweiligen Realisation bezeichnet.

- Ein Poisson-Punktprozess ξ mit Parametern λ und Q ist faktisch ein *Zählprozess* in dem Sinne, dass die Größe $\xi(B) = \sum_{n=1}^N \mathbb{1}_B(X_n)$ für $B \in \mathcal{B}$ „zählt“, wie viele „Punkte“ in die Menge B fallen.

Satz 4: Es seien ξ_1 und ξ_2 stochastisch unabhängige Poisson-Punktprozesse mit Parametern λ_i und Q_i , $i=1,2$. Dann ist $\xi = \xi_1 + \xi_2$ ein Poisson-Punktprozess mit den Parametern $\lambda := \lambda_1 + \lambda_2$ und dem Parametermaß $Q := \frac{\lambda_1}{\lambda_1 + \lambda_2} Q_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} Q_2$ (Überlagerungseigenschaft).

Ist umgekehrt ξ ein Poisson-Punktprozess mit Parametern λ und Q und sind λ_1 und λ_2 zwei nicht-negative reelle Zahlen mit $\lambda_1 + \lambda_2 = \lambda$, so existieren zwei stochastisch unabhängige Poisson-Punktprozesse ξ_1 und ξ_2 mit Parametern λ_i und $Q_i = Q$, $i=1,2$, so dass $\xi = \xi_1 + \xi_2$ gilt (Teilungseigenschaft).

Bemerkung:

- Satz 4 kann offensichtlich sofort auf *endlich viele* Überlagerungen bzw. Teilungen verallgemeinert werden. Interessant ist hierbei, dass auch die unabhängige Überlagerung *völlig unterschiedlicher* Poisson-Punktprozesse wieder zu einem Poisson-Punktprozess führt.

Satz 5: Der im vorigen Absatz konstruierte Punktprozess $\xi := \sum_{i \in I} \xi_i$ ist in Bezug auf seine Verteilung wohldefiniert und ist die Realisierung eines Poisson-Punktprozesses mit Intensitätsmaß μ .

Bemerkung: Ist μ ein beliebiges σ -endliches Intensitätsmaß, so existiert immer ein Poisson-Punktprozess ξ , der μ als Intensitätsmaß besitzt, d.h. mit der Eigenschaft $\mu = E\xi$. Eine kanonische Konstruktion für ξ ist folgendermaßen möglich:

Es sei $\mathfrak{X} = \bigoplus_{i \in I} \mathfrak{X}_i$ eine disjunkte (messbare) Zerlegung von $\mathfrak{X}$ mit $0 < \mu(\mathfrak{X}_i) < \infty$ und $\mu_i := \mu(\cdot \cap \mathfrak{X}_i)$ für alle $i \in I$. Sei ferner $\lambda_i = \mu(\mathfrak{X}_i)$ und $Q_i = \frac{1}{\lambda_i} \mu_i$ für alle $i \in I$. Die Zufallsvariablen $\{N_i\}_{i \in I}$ und die Zufallselemente $\{X_{in}\}_{(i,n) \in I \times \mathbb{N}}$ (mit Werten in $\mathfrak{X}$) seien gemeinsam stochastisch unabhängig, wobei die N_i jeweils $\mathcal{P}(\lambda_i)$ -verteilt seien und die X_{in} für alle $n \in \mathbb{N}$ die Verteilung Q_i besitzen mögen, für alle $i \in I$.

Bemerkung: Ist μ ein beliebiges σ -endliches Intensitätsmaß, so existiert immer ein Poisson-Punktprozess ξ , der μ als Intensitätsmaß besitzt, d.h. mit der Eigenschaft $\mu = E\xi$. Eine kanonische Konstruktion für ξ ist folgendermaßen möglich:

Dann ist

$$\xi = \sum_{i \in I} \sum_{n=1}^{N_i} \varepsilon_{X_{in}}$$

ein Repräsentant des Poisson-Punktprozesses ξ mit dem Intensitätsmaß μ .

Für den allgemeinen Poisson-Punktprozess bleiben die in Satz 4 formulierten Eigenschaften sinngemäß bestehen.

Satz 6: Es seien ξ_1 und ξ_2 stochastisch unabhängige Poisson-Punktprozesse mit σ -endlichen Intensitätsmaßen $\mu_i, i=1,2$. Dann ist $\xi = \xi_1 + \xi_2$ ein Poisson-Punktprozess mit σ -endlichem Intensitätsmaß $\mu = \mu_1 + \mu_2$ (Überlagerungseigenschaft).

Ist umgekehrt ξ ein Poisson-Punktprozess mit σ -endlichem Intensitätsmaß μ und sind μ_1 und μ_2 zwei σ -endliche Maße mit $\mu_1 + \mu_2 = \mu$, so existieren zwei stochastisch unabhängige Poisson-Punktprozesse ξ_1 und ξ_2 mit den Intensitätsmaßen μ_1 und μ_2 , so dass $\xi = \xi_1 + \xi_2$ gilt (Teilungseigenschaft).

Bemerkungen:

- Jeder allgemeine Poisson-Punktprozess ξ mit σ -endlichem Intensitätsmaß μ kann lokal (d.h. bei Einschränkung auf eine Teilmenge $B \in \mathcal{B}$ endlichen Maßes $\mu(B) < \infty$) vermöge $\xi(\cdot \cap B)$ kanonisch als einfacher Poisson-Punktprozess im Sinne von Satz 3 dargestellt werden.
- Im Gegensatz zu den endlichen Poisson-Punktprozessen im Sinne des Satzes 3 können bei allgemeinen Poisson-Punktprozessen mit σ -endlichem Intensitätsmaß μ Punkthäufungen auftreten; in diesem Fall kann der Poisson-Punktprozess u.U. nicht mehr mit einer zufälligen abgeschlossenen Menge identifiziert werden.

Bemerkungen:

Beispiel:

$\mathfrak{X} = \mathbb{R}$, $\mathfrak{X}_0 = (-\infty, 0] \cup (1, \infty)$, $\mathfrak{X}_i = \left(\frac{1}{i+1}, \frac{1}{i} \right)$ für $i \in I = \mathbb{Z}^+$ mit dem Intensitätsmaß μ mit der Lebesgue-Dichte

$$f(x) = \begin{cases} 1/x, & x > 0 \\ 0, & \text{sonst.} \end{cases}$$

Hier gilt $\mu(\mathfrak{X}_i) = \int_{1/(i+1)}^{1/i} \frac{1}{x} dx = \ln\left(1 + \frac{1}{i}\right) < \infty$ für alle $i \in \mathbb{N}$, aber

$E\xi((0, b]) = \int_0^b \frac{1}{x} dx = \infty$ für alle $b > 0$ (der Nullpunkt ist ein

Häufungspunkt des Punktprozesses). Jede „Realisierung“ des Punktprozesses ist damit fast sicher *keine* abgeschlossene Menge, denn 0 ist fast sicher kein möglicher „Punkt“ des Prozesses.

Satz 7: Es sei ξ ein Poisson-Punktprozess mit Intensitätsmaß μ . $C \in \mathcal{B}$ sei fest. Dann ist das zufällige Maß $\xi_C = \xi(\cdot \cap C)$ ebenfalls ein Poisson-Punktprozess mit dem Intensitätsmaß $\mu_C = \mu(\cdot \cap C)$.

Bemerkung: Satz 7 besagt anschaulich, dass die „Einschränkung“ eines Poisson-Punktprozesses auf eine feste Menge $C \in \mathcal{B}$ wieder zu einem Poisson-Prozess führt mit dem Spurmaß als neuem Intensitätsmaß. Eine verwandte Aussage beinhaltet der folgende Satz.

Satz 8: Es sei ξ ein Poisson-Punktprozess mit σ -endlichem Intensitätsmaß μ auf dem Produktraum $(\mathcal{X}, \mathcal{B}) = (\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{B}_1 \otimes \mathcal{B}_2)$. $C \in \mathcal{B}_2$ sei fest. Dann ist das zufällige Maß $\xi^C := \xi(\cdot \times C)$ wieder ein Poisson-Punktprozess [über $(\mathcal{X}_1, \mathcal{B}_1)$] mit dem σ -endlichen Intensitätsmaß $\mu^C = \mu(\cdot \times C)$.

Bemerkungen:

- Im speziellen Fall, dass $\mu = \mu_1 \otimes \mu_2$ ein Produktmaß auf $\mathcal{B}_1 \otimes \mathcal{B}_2$ ist, gilt noch $\mu^C = \mu_1 \otimes \mu_2(\cdot \times C) = \mu_2(C) \cdot \mu_1$.
- Anschaulich bedeutet Satz 8, dass sich auch nach einer Projektion der „Punkte“ des Poisson-Punktprozesses auf einen niedriger dimensional Teilraum wieder ein Poisson-Punktprozess ergibt.

I.3. Der homogene Poisson-Punktprozess

Definition 5: Im Fall $\mathcal{X} = \mathbb{R}^d$ und $\mathcal{B} = \mathcal{B}^d$ (Borel'sche σ -Algebra) mit $\mu = \lambda \cdot m^d$, $\lambda > 0$ (m^d : Lebesgue-Maß), $d \in \mathbb{N}$ heißt der Poisson-Punktprozess ξ mit dem Intensitätsmaß μ *homogener Poisson-Punktprozess* (oder kürzer: homogener Poisson-Prozess) mit Parameter λ (über $\mathbb{R}^d$).

Bemerkungen:

- Der homogene Poisson-Prozess kann als Repräsentant einer stetigen „Gleichverteilung“ über $\mathcal{X} = \mathbb{R}^d$ aufgefasst werden, weil nach der obigen kanonischen Konstruktion die in eine beliebige Menge $B \in \mathcal{B}^d$ positiven Lebesgue-Maßes fallenden (zufällig vielen) „Punkte“ einer stetigen Gleichverteilung über B genügen. Der Parameter λ steuert dabei, wie „dicht“ die Punkte liegen.
- Überlagerungen und Aufteilungen von homogenen Poisson-Prozessen (im Sinne von Satz 6) führen wieder auf homogene Poisson-Prozesse.

Satz 9: Es sei ξ ein d -dimensionaler, homogener Poisson-Punktprozess mit Parameter $\lambda > 0$ ($d \in \mathbb{N}$). Die der Größe nach sortierten Entfernungen $R_{(i)}$, $i \in \mathbb{N}$ der „Punkte“ vom Nullpunkt sind dann gegeben durch

$$R_{(i)} := \inf \{r \geq 0 \mid \xi(K(r)) \geq i\}, \quad i \in \mathbb{N}.$$

Dabei bezeichne $K(r)$ die abgeschlossene Kugel mit Radius $r > 0$ um den Nullpunkt in $\mathbb{R}^d$.

Es gilt:

a) Für jedes $n \in \mathbb{N}$ besitzt der Zufallsvektor $\mathbf{R}_n = (R_{(1)}, \dots, R_{(n)})$ die Dichte

$$f_n(x_1, \dots, x_n) = \begin{cases} (\lambda d v_d)^n \exp(-\lambda v_d x_n^d) \prod_{i=1}^n x_i^{d-1} & \text{für } 0 < x_1 < \dots < x_n \\ 0 & \text{sonst.} \end{cases}$$

Dabei bezeichnet $v_d := \frac{2\sqrt{\pi^d}}{d\Gamma\left(\frac{d}{2}\right)}$ das d -dimensionale Volumen $m^d(K(1))$

der Einheitskugel.

Satz 9: Es sei ξ ein d -dimensionaler, homogener Poisson-Punktprozess mit Parameter $\lambda > 0$ ($d \in \mathbb{N}$). Die der Größe nach sortierten Entfernungen $R_{(i)}$, $i \in \mathbb{N}$ der „Punkte“ vom Nullpunkt sind dann gegeben durch

$$R_{(i)} := \inf \{r \geq 0 \mid \xi(K(r)) \geq i\}, \quad i \in \mathbb{N}.$$

Dabei bezeichne $K(r)$ die abgeschlossene Kugel mit Radius $r > 0$ um den Nullpunkt in $\mathbb{R}^d$.

Es gilt:

- b) Die Zufallsvariablen $R_{(1)}^d, R_{(2)}^d - R_{(1)}^d, \dots, R_{(n)}^d - R_{(n-1)}^d$ sind stochastisch unabhängig und jeweils $\mathcal{E}(\lambda v_d)$ -exponentialverteilt.

Bemerkungen:

- Die Verteilung(sfunktion) von $R_{(1)}$ wird *sphärische Kontaktverteilung (sfunktion)* genannt, auch bei allgemeineren Punktprozessen. Für einen Poisson-Punktprozess ξ mit beliebigem σ -endlichem Intensitätsmaß μ erhält man beispielsweise

$$P(R_{(1)} > r) = P(\xi(K(r)) = 0) = \exp(-\mu(K(r))) \text{ für } r \geq 0.$$

Bemerkungen:

- Die oben erwähnte alternative Konstruktion ($\rightarrow$ Simulation) eines homogenen Poisson-Prozesses beschreiben wir zunächst für den speziellen Fall $\mathfrak{X} = \mathbb{R}^2$:

1. Schritt: Erzeuge eine Folge $\{Z_n\}_{n \in \mathbb{N}}$ stochastisch unabhängiger, je $\mathcal{U}[0,1]$ -verteilter Zufallsvariablen (Standard-Zufallszahlen); setze

$$R_n := \sqrt{-\frac{1}{\lambda\pi} \ln \left(\prod_{i=1}^n Z_i \right)}.$$

Dies sind die sukzessiven Radien der Kreise, auf denen die „Punkte“ liegen.

Bemerkungen:

- Die oben erwähnte alternative Konstruktion ($\rightarrow$ Simulation) eines homogenen Poisson-Prozesses beschreiben wir zunächst für den speziellen Fall $\mathfrak{X} = \mathbb{R}^2$:

2. Schritt: Erzeuge eine weitere, von der ersten unabhängige Folge $\{V_n\}_{n \in \mathbb{N}}$ stochastisch unabhängiger, je $\mathcal{U}[0,1]$ -verteilter Zufallsvariablen; setze $\varphi_n := 2\pi V_n$. Dies sind die sukzessiven Winkel der „Punkte“ auf den Kreisen. Setze nun

$$X_n := R_n \begin{pmatrix} \cos \varphi_n \\ \sin \varphi_n \end{pmatrix}.$$

Dies ist die Folge der „Punkte“ des Poisson-Prozesses.

Bemerkungen:

Diese Konstruktion kann auf beliebige Dimensionen $d \in \mathbb{N}$ erweitert werden; dazu ist im ersten Schritt

$$R_n := \sqrt[d]{-\frac{1}{\lambda \nu_d} \ln \left(\prod_{i=1}^n Z_i \right)}$$

zu setzen (mit dem d -dimensionalen Volumen $\nu_d = m^d(K(\mathbf{1})) = \frac{2\sqrt{\pi^d}}{d\Gamma\left(\frac{d}{2}\right)}$

der Einheitskugel), im zweiten Schritt müssen auf der d -dimensionalen Einheitssphäre stetig gleichverteilte Zufallsvektoren erzeugt werden. Diese erhält man am einfachsten durch eine Normierung unabhängiger Zufallsvektoren $\mathbf{V}_n$ mit einer multivariaten Normalverteilung $\mathcal{N}(\mathbf{0}, \mathbf{I})$ mit der Einheitsmatrix $\mathbf{I}$ als Varianz-Kovarianzmatrix (d.h. alle Komponenten von $\mathbf{V}_n$ sind unabhängig. Die „Punkte“ des Poisson-Prozesses erhält man dann vermöge

$$X_n := \frac{R_n}{\|\mathbf{V}_n\|} \mathbf{V}_n.$$

I.4. Besonderheiten des homogenen Poisson-Punktprozesses

Einige der speziellen Eigenschaften des (mehrdimensionalen) Lebesgue-Maßes „vererben“ sich in natürlicher Weise auf einen homogenen Poisson-Prozess ξ mit Parameter $\lambda > 0$. So ist ξ z.B. *stationär* (translationsinvariant) und *isotrop* (drehungsinvariant), d.h. die Prozesse ξ_{+a} und ξ_{dr} mit

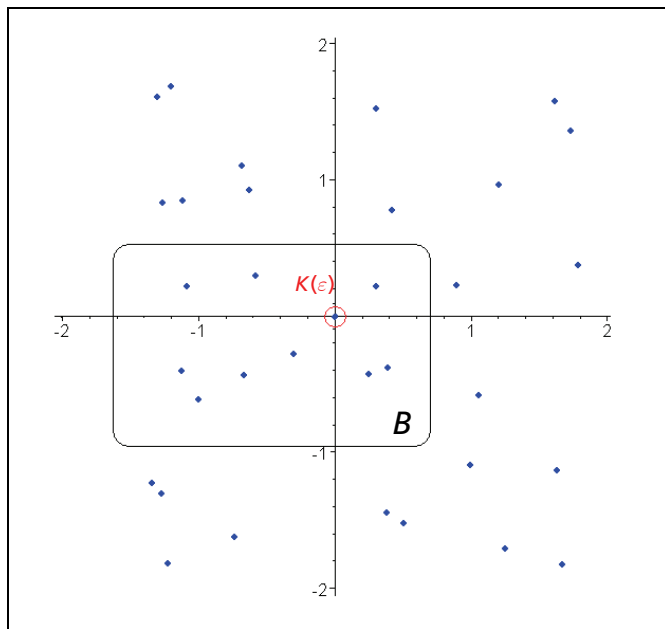
$$\xi_{+a}(B) = \xi(a + B) \text{ und } \xi_{dr}(B) = \xi(D(B)) \text{ für } B \in \mathcal{B}^d$$

besitzen dieselbe Verteilung wie ξ , sind also ebenfalls homogene Poisson-Prozesse mit Parameter $\lambda > 0$. Hierbei ist $a \in \mathbb{R}^d$ und D eine Drehungsabbildung. Anders ausgedrückt: verschiebt man die „Punkte“ eines homogenen Poisson-Prozesses um den festen Vektor $a \in \mathbb{R}^d$ oder dreht man sie in eine beliebige Richtung, so lässt sich die neue Punktkonfiguration statistisch nicht von einem homogenen Poisson-Prozess mit demselben Parameter unterscheiden. Noch anders ausgedrückt: die „Punkte“ in einem Beobachtungsfenster $B \in \mathcal{B}^d$ verhalten sich statistisch gleich für alle zueinander kongruenten Mengen B .

Eine weitere Besonderheit des homogenen Poisson-Prozesses ist die folgende: Fügt man den „Punkten“ des Prozesses den Nullpunkt $\mathbf{0}$ als Punkt hinzu, ergibt sich ein Punktmuster, das statistisch identisch ist mit jedem Punktmuster, das sich ergibt, wenn man die „Punkte“ des Prozesses gleichmäßig so verschiebt, dass einer von ihnen den Nullpunkt trifft (Satz von СЛИВНЯК (Slivnyak), vgl. STOYAN, KENDALL AND MECKE (1987), Section 4.4). In der Sprache der Wahrscheinlichkeitstheorie bedeutet dies exakter:

$$P^{\xi}(\bullet \mid \xi(\{\mathbf{0}\}) = 1) = P^{\xi + \varepsilon_{\mathbf{0}}}.$$

Dies lässt sich folgendermaßen anschaulich motivieren:



Ist $\varepsilon > 0$ klein und $\mathbf{0} \in K(\varepsilon) \subseteq B \in \mathcal{B}$, so gilt für $k \in \mathbb{N}$

$$\begin{aligned}
 P(\xi(B) = k | \xi(K(\varepsilon)) = 1) &= \frac{P(\{\xi(B) = k\} \cap \{\xi(K(\varepsilon)) = 1\})}{P(\xi(K(\varepsilon)) = 1)} \\
 &= \frac{P(\{\xi(B \setminus K(\varepsilon)) = k - 1\} \cap \{\xi(K(\varepsilon)) = 1\})}{P(\xi(K(\varepsilon)) = 1)} \\
 &= P(\{\xi(B \setminus K(\varepsilon)) = k - 1\}) \approx P(\{\xi(B) = k - 1\}) \\
 &= P(\{\xi(B) + \varepsilon_0(B) = k\}),
 \end{aligned}$$

und für $K(\varepsilon) \cap B = \emptyset$ entsprechend

$$\begin{aligned}
 P(\xi(B) = k | \xi(K(\varepsilon)) = 1) &= \frac{P(\{\xi(B) = k\} \cap \{\xi(K(\varepsilon)) = 1\})}{P(\xi(K(\varepsilon)) = 1)} \\
 &= P(\{\xi(B) = k\}) = P(\{\xi(B) + \varepsilon_0(B) = k\}),
 \end{aligned}$$

so dass nach Grenzübergang $\varepsilon \downarrow 0$ folgt

$$P(\xi(B) = k | \xi(\{\mathbf{0}\}) = 1) = P(\{\xi(B) + \varepsilon_0(B) = k\}) \text{ für alle } B \in \mathcal{B}^d.$$

I.5. Statistik für den homogenen Poisson-Punktprozess

Da der homogene Poisson-Prozess nur durch einen einzigen Parameter $\lambda > 0$ charakterisiert ist, ist es vergleichsweise leicht, auf der Basis bereits einer einzigen (genügend großen) Realisation des Punktprozesses geeignete Schätzer für λ zu finden. Wir stellen im Folgenden einige solcher Schätzer, die auch in der Praxis verwendet werden, vor und diskutieren ihre Güteeigenschaften.

Wir setzen für den Rest dieses Abschnitts voraus, dass ein homogener d -dimensionaler Poisson-Prozess mit Parameter $\lambda > 0$ und $d \in \mathbb{N}$ gegeben ist, der in einer hinreichend großen Menge $B \in \mathcal{B}^d$ beobachtet werden kann.

Methode 1: Proportional-Statistik: $\hat{\lambda} = \frac{\xi(B)}{m^d(B)}$

Dieser Schätzer ist erwartungstreu wegen

$$E(\hat{\lambda}) = E\left(\frac{\xi(B)}{m^d(B)}\right) = \frac{E(\xi(B))}{m^d(B)} = \lambda \frac{m^d(B)}{m^d(B)} = \lambda.$$

Er besitzt die Varianz

$$\text{Var}(\hat{\lambda}) = \text{Var}\left(\frac{\xi(B)}{m^d(B)}\right) = \frac{\text{Var}(\xi(B))}{\{m^d(B)\}^2} = \lambda \frac{m^d(B)}{\{m^d(B)\}^2} = \frac{\lambda}{m^d(B)}.$$

Der Schätzer ist konsistent, wenn die Menge B vergrößert wird, d.h. für $m^d(B) \rightarrow \infty$.

Methode 2: Leerzellen-Statistik: $\hat{\lambda} = \frac{n \ln\left(\frac{n}{N}\right)}{m^d(B)}$

mit $N = \#\{i | \xi(B_i) = 0\}$ für eine ausreichend feine disjunkte Zerlegung $B = \bigoplus_{i=1}^n B_i$ in Borel-Mengen B_i mit $m^d(B_i) = \frac{1}{n} m^d(B)$, $1 \leq i \leq n \in \mathbb{N}$. Dieser Schätzer ist motiviert durch die Beziehung

$$\exp\left(-\frac{\lambda}{n} m^d(B)\right) = \exp(-\lambda m^d(B_i)) = P(\xi(B_i) = 0) \approx \frac{N}{n};$$

Auflösen nach λ ergibt den angegebenen Schätzer. Dieser ist nicht exakt, aber asymptotisch (für $n \rightarrow \infty$) erwartungstreu mit der Abschätzung (Jensen'sche Ungleichung bei Konkavität)

$$E(\hat{\lambda}) = -\frac{nE\left(\ln\left(\frac{N}{n}\right)\right)}{m^d(B)} \geq -\frac{n \ln\left(\frac{E(N)}{n}\right)}{m^d(B)} = -\frac{n \ln\left(\exp\left(-\frac{\lambda}{n} m^d(B)\right)\right)}{m^d(B)} = \lambda.$$

Methode 3: Abstands-Statistik: $\hat{\lambda} = \frac{n-1}{\nu_d R_{(n)}^d},$

wobei $R_{(n)}$ den n -größten Abstand der „Punkte“ vom Nullpunkt bezeichnet. (Hierfür ist Voraussetzung, dass B den Nullpunkt enthält und genügend „groß“ ist.) $\lambda \nu_d R_{(n)}^d$ besitzt als Summe unabhängiger, identisch $\mathcal{E}(1)$ -exponentialverteilter Zufallsvariablen eine Erlang-Verteilung. Damit ergibt sich:

$$E(\hat{\lambda}) = \lambda E\left(\frac{n-1}{\lambda \nu_d R_{(n)}^d}\right) = \lambda(n-1) \int_0^{\infty} \frac{1}{x} \frac{x^{n-1}}{(n-1)!} e^{-x} dx = \lambda \int_0^{\infty} \frac{x^{n-2}}{(n-2)!} e^{-x} dx = \lambda,$$

d.h. der Schätzer ist erwartungstreu für $n \geq 2$ (der Schätzer hat den Erwartungswert Null für $n = 1!$), mit der Varianz

$$\text{Var}(\hat{\lambda}) = \frac{\lambda^2}{n-2} \text{ für } n \geq 3$$

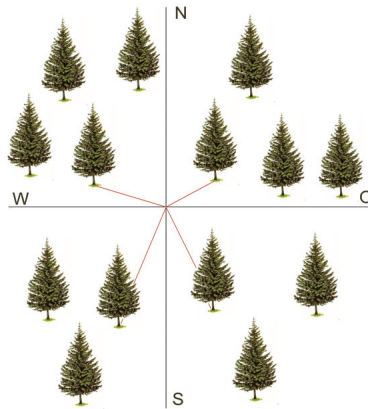
wegen

$$\begin{aligned}
 E(\hat{\lambda}^2) &= \lambda^2 E \left\{ \left(\frac{n-1}{\lambda \nu_d R_{(n)}^2} \right)^2 \right\} = \lambda^2 (n-1)^2 \int_0^\infty \frac{1}{x^2} \frac{x^{n-1}}{(n-1)!} e^{-x} dx \\
 &= \lambda^2 \frac{n-1}{n-2} \int_0^\infty \frac{x^{n-3}}{(n-3)!} e^{-x} dx = \frac{n-1}{n-2} \lambda^2.
 \end{aligned}$$

Der Schätzer ist also konsistent für $n \rightarrow \infty$.

Methode 4: Die so genannte PCQ-Methode:

Diese Methode ist eine historische Variante der Methode 3 (*PCQ = Point Center Quarter*). Sie wurde beim Bau der amerikanischen transkontinentalen Eisenbahnlinien angewendet, um den erwarteten Holzeinschlag in bewaldeten Gebieten zu schätzen.



Bei der PCQ-Methode werden in den 4 (als paarweise disjunkt angenommenen) Quadranten Q_i , die durch die Himmelsrichtungen vorgegeben sind ($\rightarrow$ Kompass), die kürzesten Abstände R_i der Bäume vom Nullpunkt (Standort) bestimmt. Diese Abstände sind auf Grund der zentralen Eigenschaft der stochastischen Unabhängigkeit der Poisson-Prozesse $\xi(\bullet \cap Q_i), i = 1, \dots, 4$ ebenfalls stochastisch unabhängig; ihre Verteilung ist identisch und gegeben durch

$$\begin{aligned}
 P(R_i > x) &= P(\xi(Q_i \cap K(x)) = 0) = \exp(-\lambda \cdot m^2(Q_i \cap K(x))) \\
 &= \exp\left(-\frac{\lambda\pi}{4}x^2\right) \text{ für } x > 0;
 \end{aligned}$$

d.h. jedes R_i^2 ist exponentialverteilt mit Parameter $\frac{\lambda\pi}{4}$. Die originäre PCQ-Methode verwendet nun den Schätzer

$$\hat{\lambda}_{PCQ} = \frac{16}{\left(\sum_{i=1}^4 R_i\right)^2}$$

als Schätzer für den unbekannten Parameter λ . Dieser Schätzer ist allerdings nicht erwartungstreu, wie man anhand folgender Überlegungen leicht einsieht:

Es gilt für $i = 1, 2, 3, 4$:

$$\begin{aligned} E(R_i) &= E(\sqrt{R_i^2}) = \frac{\lambda\pi}{4} \int_0^\infty \sqrt{x} e^{-\frac{\lambda\pi}{4}x} dx = \frac{2}{\sqrt{\lambda\pi}} \int_0^\infty \sqrt{y} e^{-y} dy = \frac{2}{\sqrt{\lambda\pi}} \Gamma\left(\frac{3}{2}\right) \\ &= \frac{2}{\sqrt{\lambda\pi}} \cdot \frac{\sqrt{\pi}}{2} = \frac{1}{\sqrt{\lambda}} \end{aligned}$$

und damit auf Grund der Jensen'schen Ungleichung für konvexe Funktionen:

$$E(\hat{\lambda}_{PCQ}) = E\left(\frac{16}{\left(\sum_{i=1}^4 R_i\right)^2}\right) > \frac{16}{\left(\sum_{i=1}^4 E(R_i)\right)^2} = \lambda.$$

Der Schätzer $\hat{\lambda}_{PCQ}$ überschätzt den wahren Parameter λ also systematisch. Mit Hilfe von Monte-Carlo-Simulationen erhält man genauer:

$$E(\hat{\lambda}_{PCQ}) \approx 1,268 \lambda, \text{ Var}(\hat{\lambda}_{PCQ}) = 0,847 \lambda^2.$$

Es wäre demnach besser, den modifizierten erwartungstreuen Schätzer

$$\hat{\lambda}_{mod} = \frac{12,618}{\left(\sum_{i=1}^4 R_i\right)^2} \text{ mit } \text{Var}(\hat{\lambda}_{mod}) = 0,528 \lambda^2$$

zu verwenden. Alternativ kann auch der (erwartungstreue) Schätzer

$$\hat{\lambda}_{4,alt} = \frac{12}{\pi \sum_{i=1}^4 R_i^2} \text{ verwendet werden, denn } \sum_{i=1}^4 R_i^2 \text{ ist } \mathcal{E}\left(4, \frac{\lambda\pi}{4}\right)\text{-Erlang-verteilt,}$$

und damit gilt

$$\begin{aligned}
 E\left(\hat{\lambda}_{4,alt}\right) &= E\left(\frac{12}{\pi \sum_{i=1}^4 R_i^2}\right) = \frac{12}{\pi} \cdot \left(\frac{\lambda\pi}{4}\right)^4 \int_0^{\infty} \frac{1}{x} \cdot \frac{x^3}{3!} e^{-\frac{\lambda\pi}{4}x} dx \\
 &= \frac{12}{\pi} \cdot \left(\frac{\lambda\pi}{4}\right)^4 \int_0^{\infty} \frac{x^2}{2!} e^{-\frac{\lambda\pi}{4}x} dx = \frac{4}{\pi} \cdot \frac{\lambda\pi}{4} = \lambda
 \end{aligned}$$

mit

$$\begin{aligned}
 E\left(\hat{\lambda}_{4,alt}^2\right) &= E\left(\frac{144}{\pi^2 \left(\sum_{i=1}^4 R_i^2\right)^2}\right) = \frac{144}{\pi^2} \cdot \left(\frac{\lambda\pi}{4}\right)^4 \int_0^{\infty} \frac{1}{x^2} \cdot \frac{x^3}{3!} e^{-\frac{\lambda\pi}{4}x} dx \\
 &= \frac{24}{\pi^2} \cdot \left(\frac{\lambda\pi}{4}\right)^4 \int_0^{\infty} x \cdot e^{-\frac{\lambda\pi}{4}x} dx = \frac{24}{\pi^2} \cdot \frac{\lambda^2 \pi^2}{16} = \frac{3}{2} \lambda^2
 \end{aligned}$$

oder

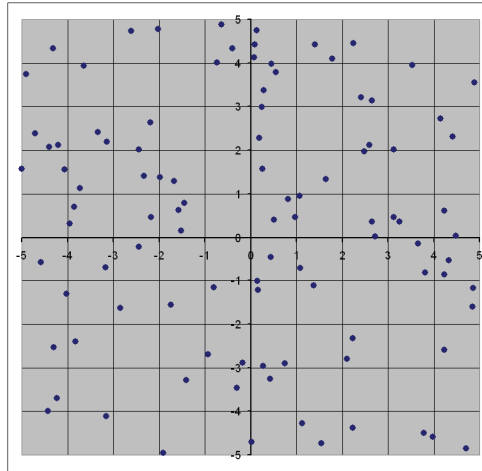
$$\text{Var}(\hat{\lambda}_{4,alt}) = \left(\frac{3}{2} - 1\right) \lambda^2 = \frac{1}{2} \lambda^2.$$

Der alternative Schätzer hat damit eine leicht geringere Varianz als der modifizierte PCQ-Schätzer.

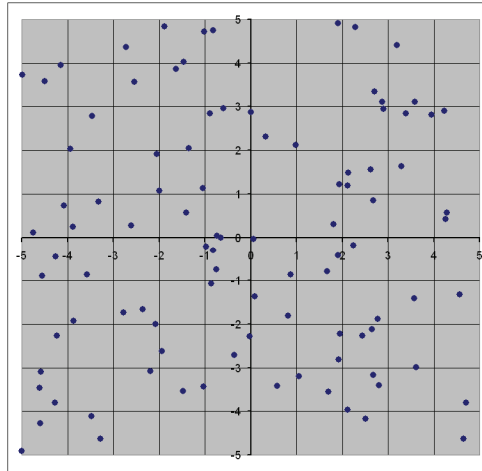
Das Abstands-Verfahren kann leicht auf eine höhere Anzahl von Sektoren erweitert werden. Vergrößert man die Zahl der Beobachtungssektoren (mit gleichen Sektor-Winkeln) von 4 auf $m > 4$ (z.B. $m = 8$; $\rightarrow$ Oktanten), so ist der passende erwartungstreue Schätzer gegeben durch

$$\hat{\lambda}_{m,alt} = \frac{m(m-1)}{\pi \sum_{i=1}^m R_i^2} \quad \text{mit} \quad \text{Var}(\hat{\lambda}_{m,alt}) = \frac{\lambda^2}{m-2}.$$

Die folgenden Graphiken zeigen zwei Simulationen eines homogenen Poisson-Prozesses mit Parameter 1 und die zugehörigen Schätzer.



Methode	Proportional	Leerzellen	PCQ modifiziert	PCQ 4 alternativ	PCQ 8 alternativ
Schätzer	1,000	1,109	0,705	0,742	0,700



Methode	Proportional	Leerzellen	PCQ modifiziert	PCQ 4 alternativ	PCQ 8 alternativ
Schätzer	1,020	0,968	1,156	0,881	1,120

Neben Schätzverfahren für den Parameter eines homogenen Poisson-Prozesses sind auch Testverfahren zur Überprüfung der Homogenitätsannahme wichtig. Ein häufig verwendeter und einfach durchzuführender Test ist der

Dispersionsindex-Test: Dazu wird - ähnlich wie bei der Leerzellenstatistik - ein Beobachtungsfenster B für den Punktprozess ξ in n gleich große disjunkte Bereiche B_i (bzgl. des Lebesgue-Maßes) aufgeteilt: $B = \bigoplus_{i=1}^n B_i$ mit Borel-Mengen B_i und $m^d(B_i) = \frac{1}{n} m^d(B)$, $1 \leq i \leq n \in \mathbb{N}$. Der Dispersionsindex I ist dann definiert als

$$I = \frac{\sum_{i=1}^n (\xi(B_i) - \bar{\xi})^2}{\bar{\xi}} \quad \text{mit} \quad \bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi(B_i).$$

Unter der (Null-)Hypothese, dass ξ ein homogener Poisson-Prozess ist, ist der Dispersionsindex I asymptotisch (für $n \rightarrow \infty$) χ^2_{n-1} -verteilt (Chi-Quadrat-verteilt mit $n-1$ Freiheitsgraden; $\rightarrow$ Chi-Quadrat-Anpassungstest). Ist I zu groß, deutet das auf Klumpenbildung hin, ist I zu klein, deutet das auf eine zu starke Regelmäßigkeit hin. (Als Fausregel zur Anwendbarkeit des Dispersionsindex-Tests sollte $n \geq 6$ und $\bar{\xi} \geq 1$ sein.) Mit Hilfe der Chi-Quadrat-Verteilung kann man also entweder zwei verschiedene einseitige oder einen zweiseitigen Signifikanz-Test formulieren.

Für die oberste Graphik wählen wir $n = 25$ (statt 100) mit der Auszählung

3	2	9	5	2
8	5	3	3	3
4	5	4	4	8
3	2	7	3	3
3	2	3	3	3

mit $\bar{\xi} = 4$ und dem Dispersionsindex $I = \frac{1}{4} \sum_{i=1}^n (\xi(B_i) - 4)^2 = 24$. Für den zweiseitigen Niveau- α -Test ergeben sich folgende Annahmebereiche $A(\alpha)$:

α	0,10	0,05	0,01
$A(\alpha)$	[13,85 36,42]	[12,40 39,36]	[9,89 45,56]

Der Dispersionsindex-Test würde also hier (wie zu erwarten) zu keinem der drei angegebenen α -Werte ablehnen.

Poisson-Cluster-Prozesse

Für viele naturwissenschaftlich motivierte Fragestellungen bietet es sich an, einen Poisson-Punktprozess („Elternprozess“) als Ausgangspunkt für eine Clusterbildung von Punkten um die Elternpunkte herum zu verwenden („Tochterprozesse“). Dabei können die Tochterprozesse z.B. selbst wieder (vom Elternprozess stochastisch unabhängige) Poisson-Punktprozesse sein. Als (lokales) Modell in $\mathbb{R}^d$ bietet sich beispielhaft etwa folgende Konstruktion an:

$$\xi = \sum_{k=1}^N \sum_{i=1}^{N_k} \varepsilon_{X_k + Z_{ki}},$$

wobei $\{N, N_k\}_{k \in \mathbb{N}}$ eine Familie stochastisch unabhängiger, Poisson-verteilter Zufallsvariablen ist mit nicht-negativen Parametern λ_E für N (Elternprozess) und λ_T für die $\{N_k\}_{k \in \mathbb{N}}$ (Tochterprozesse), unabhängig von den (ebenfalls untereinander unabhängigen) d -dimensionalen Zufallsvektoren $\{X_k, Z_{ki}\}_{k, i \in \mathbb{N}}$. Ferner werde hier angenommen, dass die $\{X_k\}_{k \in \mathbb{N}}$ die Verteilung Q^X und die $\{Z_{ki}\}_{k, i \in \mathbb{N}}$ die Verteilung Q^Z besitzen.

Für das resultierende Intensitätsmaß ergibt sich dann:

Satz 11: Unter den obigen Voraussetzungen gilt

$$E\xi(B) = \lambda_E \cdot \lambda_T \cdot \int_{\mathbb{R}^d} Q^x(B-z) Q^z(dz) \text{ für alle } B \in \mathcal{B}^d.$$

Bemerkungen:

- Für die mittlere Anzahl aller Tochterpunkte ergibt sich damit $E\xi(\mathbb{R}^d) = \lambda_E \cdot \lambda_T$.
- In vielen Anwendungsfällen ist die Verteilung Q^z symmetrisch um den Nullpunkt herum, so dass die Tochterpunkte dann symmetrisch um die Elternpunkte herum verteilt sind.

Anwendungsbeispiel Thomas-Prozesse:

Das im Folgenden ausführlicher besprochene Anwendungsbeispiel hat seinen Ursprung in einer biologisch motivierten Fragestellung im Rahmen des Projekts *Ökosystemforschung Niedersächsisches Wattenmeer*, grundlagenorientierter Teil ELAWAT (1989 - 1999).

Auszug aus dem Synthese-Bericht des ICBM vom März 1999 (nach WIRTZ UND NIESEL, modifiziert; vgl auch DITTMANN (1999), Kapitel 4.3.2):

Im Wattsediment findet sich ein großer Artenreichtum an kleineren Organismen (Meiobenthos). Diese Organismen sind nicht gleichmäßig im Sediment verteilt. In Abhängigkeit von äußeren Bedingungen wie z.B. der Korngröße des Sediments, der bodennahen Advektion oder von den ökologischen Ansprüchen einer Art kann es auf unterschiedlichen räumlichen Skalen zu verschiedenen Verteilungsmustern kommen. Dies gilt beispielsweise für den Copepoden *Harpacticus obscurus*.

Die Charakterisierung der räumlichen Heterogenität des Meiobenthos ist von Bedeutung. Sie läßt Rückschlüsse auf ökologische Eigenschaften der jeweils betrachteten Art zu. Insbesondere räumliche Wechselwirkungen mit anderen Arten oder innerhalb der Population (z.B. Vermehrung) werden über die Verteilungsmuster indirekt erkennbar.

Allerdings führen Dichteschwankungen in der ökologischen Freilandarbeit zu dem Problem, eine geeignete Meßstrategie zu finden. Die Probenahmedichte darf aus praktischen Gründen nicht zu groß sein. Sie sollte aber auch nicht zu gering gewählt werden, da sich sonst die Verteilungsmuster nicht charakterisieren lassen. Folglich müssen die Anzahl und Größe von erfaßten Flächen auch über einen minimalen Grenzwert liegen.

Aus der Problematik ergeben sich insgesamt zwei Fragestellungen:

- Liegt den Dichteschwankungen meiobenthischer Organismen ein Muster zugrunde und falls ja, kann das Muster mit wenigen Größen beschrieben werden?
- Welche Probenahmestrategie ist notwendig, um das Muster zu erfassen?

Prinzipiell können die beiden Fragestellungen mit **statistischen Verfahren** bearbeitet werden. Doch vor allem die anvisierte, ökologische Interpretation der Verteilungsmuster macht den Einsatz eines Modells sinnvoll. Denn in dem Modell können bereits *Hypothesen zur Musterbildung* getestet werden.

Einen solchen hypothesenbasierten Weg beschreiten die in der Ökosystemforschung Wattenmeer eingesetzten stochastischen Verteilungsmodelle des Meiobenthos. Sie leiten sich aus Annahmen zu der Verteilungsart der betrachteten Arten ab. Das Modell zur Musterbildung des Copepoden *Harpacticus obscurus* (Ruderfußkrebs) beruht beispielsweise darauf, daß die Copepoden stark von der makrobenthische Art *Lanice conchilega* (Bäumchenröhrenwurm) gefördert werden. Im Detail basiert das Modell auf folgenden Hypothesen:

- Die Copepoden gruppieren sich um einzelne Individuen des Bäumchenröhrenwurms gemäß einer radialsymmetrischen Normalverteilung (Tochterprozesse).
- Die Individuen des Bäumchenröhrenwurms sind zufällig gleichmäßig verteilt (Elternprozess).

Die folgenden Tabellen zeigen die Auszählungen der Anzahlen von *Harpacticus obscurus* an drei verschiedenen Standorten; die gesamte ausgezählte Fläche beträgt jeweils 100 cm² [= 1 Flächeneinheit]. μ bezeichnet jeweils die durchschnittliche Anzahl Individuen pro Zelle und I den jeweiligen Dispersionsindex.

95	1	3	0	42
1	1	1	4	2
0	5	8	81	24
11	1	6	71	116
1	5	116	2	10

Probe 6

$$\mu = 24,28$$

$$I = 1464,21$$

165	22	1	94	68
11	82	111	97	153
0	0	24	13	15
31	1	46	22	11
2	0	5	8	6

Probe 8

$$\mu = 39,52$$

$$I = 1490,13$$

245	142	326	52	293
222	368	84	18	67
239	25	477	213	204
18	570	183	494	47
238	119	126	591	20

Probe 18

$$\mu = 215,24$$

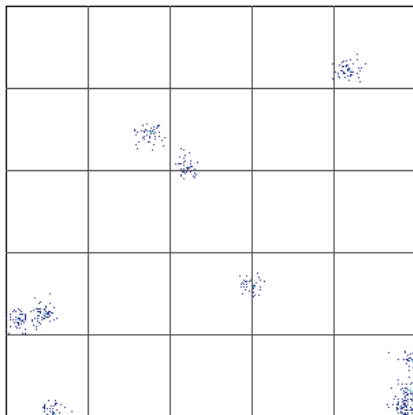
$$I = 3377,47$$

Die sehr großen Werte der Dispersionsindices deuten auf ein stark geklumpertes Punktmuster hin. Insofern erscheint es naheliegend, für die Modellierung einen räumlichen Cluster-Prozess zu verwenden. In einem ersten Ansatz kann man einen Thomas-Prozess betrachten, bei dem die Tochter-Prozesse Poisson-Punktprozesse sind, deren Intensitätsmaße durch ein Vielfaches $\lambda_T > 0$ bivariater radialsymmetrischer Normalverteilungen gegeben sind (d.h. mit stochastisch unabhängigen Komponenten). Die folgenden Graphiken zeigen Simulationen solcher Prozesse mit den Parametern $\lambda_E = 10$ für den Elternprozess und $\lambda_T = 50$ für die Tochterprozesse. Die Standardabweichung für die Normalverteilung beträgt jeweils 5 [in Pixel-Einheiten], die Auszählung erfolgt wie in den obigen drei Beispielen.

intensity of parent process = 10
 intensity of daughter process = 50
 standard deviation = 5 [Pixels]

mean value = 17.55 index-of-dispersion = 1519.59

estimate for intensity of parent process = 9.64 realized: 11
 estimate for intensity of daughter process = 45.53 realized: 39.9

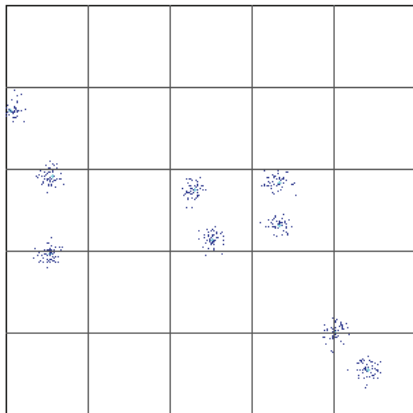


0	0	0	2	46
0	46	27	0	0
0	0	13	0	0
104	0	20	21	0
31	0	0	0	129

intensity of parent process = 10
 intensity of daughter process = 50
 standard deviation = 5 [Pixels]

mean value = 17.71 index-of-dispersion = 1238.32

estimate for intensity of parent process = 8.21 realized: 9
 estimate for intensity of daughter process = 53.94 realized: 49.22

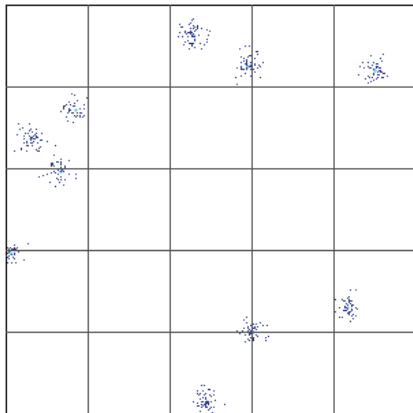


0	0	0	0	0
47	0	0	0	0
62	0	98	89	0
44	0	3	9	18
0	0	0	12	61

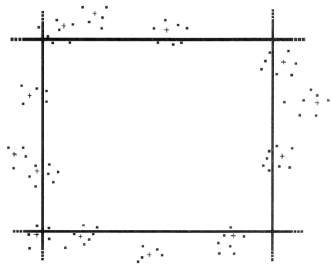
intensity of parent process = 10
 intensity of daughter process = 50
 standard deviation = 5 [Pixels]

mean value = 19.71 index-of-dispersion = 1279.26

estimate for intensity of parent process = 9.64 realized: 10
 estimate for intensity of daughter process = 51.13 realized: 49.29



0	0	95	15	51
117	2	0	0	0
41	0	0	0	0
20	0	12	11	49
0	0	67	13	0



Für die Schätzung des Parameters λ_E verwenden wir eine geeignete Modifikation der *Leerschellenstatistik*, da diese nicht direkt angewendet werden kann, weil Zellen, die zwar von Elternpunkten frei sind, mit Tochterpunkten besetzt sein können. Wenn die Parameter λ_E und λ_T in einem „passenden“ Verhältnis zur Streuung σ

der umgebenden Normalverteilung stehen, bietet sich ein *Schwellenwert-Verfahren* der folgenden Art an:

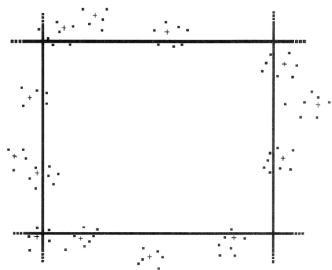
Es sei $S > 0$ der Schwellenwert für eine „schwach besetzte Zelle“ (vgl. die Graphik links oben). Mit

$$L_S := \#\{i \mid \xi(B_i) \leq S\}$$

sei die Anzahl derjenigen Zellen B_i bezeichnet, die höchstens S Tochterpunkte enthalten. Dann ist

$$\hat{\lambda}_E := \frac{n}{m^d(B)} \ln \left(\frac{n}{\hat{L}_S} \right)$$

ein geeigneter Schätzer für λ_E .



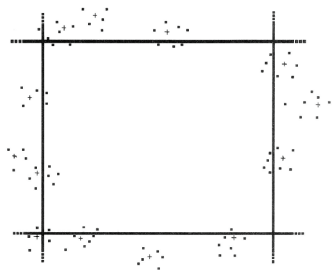
Für die Schätzung des Parameters λ_E verwenden wir eine geeignete Modifikation der *Leierzellenstatistik*, da diese nicht direkt angewendet werden kann, weil Zellen, die zwar von Elternpunkten frei sind, mit Tochterpunkten besetzt sein können. Wenn die Parameter λ_E und λ_T in einem „passenden“ Verhältnis zur Streuung σ

der umgebenden Normalverteilung stehen, bietet sich ein *Schwellenwert-Verfahren* der folgenden Art an:

Hierbei bezeichnet $\hat{L}_S$ die empirische Anzahl „schwach besetzter Zellen“. Diese werden ideell mit Zellen identifiziert, die ursprünglich frei von „Elternpunkten“ sind. Als heuristisch motivierte Wahl für den Schwellenwert S verwenden wir im Fall der Dimension 2 die Größe

$$S := \frac{N}{3(n - \ell)},$$

wobei $N = \xi(B)$ die Gesamtzahl aller Tochterpunkte im Gebiet $B = \bigcup_{i=1}^n B_i$ bezeichnet und ℓ die Anzahl der Zellen, die frei von Tochterpunkten sind.

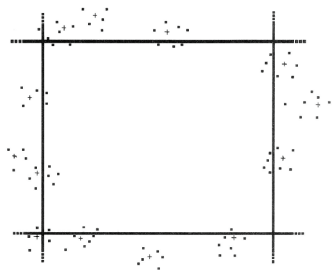


Für die Schätzung des Parameters λ_E verwenden wir eine geeignete Modifikation der *Leierzellenstatistik*, da diese nicht direkt angewendet werden kann, weil Zellen, die zwar von Elternpunkten frei sind, mit Tochterpunkten besetzt sein können. Wenn die Parameter λ_E und λ_T in einem „passenden“ Verhältnis zur Streuung σ

der umgebenden Normalverteilung stehen, bietet sich ein *Schwellenwert-Verfahren* der folgenden Art an:

Motivation: Tochterpunkte von Elternpunkten, die nahe an einer Kante einer Zelle liegen, streuen in etwa zur Hälfte in die Zelle, während Tochterpunkte für Elternpunkte, die nahe an einer Ecke einer Zelle liegen, in etwa zu einem Viertel in die Zelle streuen. Ferner ist $n - \ell$ die Anzahl derjenigen Zellen, die von Tochterpunkten besetzt sind, so dass im Mittel $\frac{N}{n - \ell}$ Tochterpunkte in einer besetzten Zelle liegen. Für den

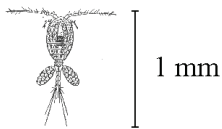
Parameter λ_T bietet sich daran anschließend der Schätzer $\hat{\lambda}_T = \frac{N}{\hat{\lambda}_E}$ an.



Für die Schätzung des Parameters λ_E verwenden wir eine geeignete Modifikation der *Leierzellenstatistik*, da diese nicht direkt angewendet werden kann, weil Zellen, die zwar von Elternpunkten frei sind, mit Tochterpunkten besetzt sein können. Wenn die Parameter λ_E und λ_T in einem „passenden“ Verhältnis zur Streuung σ

der umgebenden Normalverteilung stehen, bietet sich ein *Schwellenwert-Verfahren* der folgenden Art an:

Die Parameterschätzungen in den drei obigen Simulationsstudien wurden auf genau diese Weise ermittelt. Wie man sieht, werden die ursprünglichen Parameter $\lambda_E = 10$ und $\lambda_T = 50$ relativ gut identifiziert.



Ruderfußkrebs

Vergleicht man die simulierten Auszählungsergebnisse mit der Probe 6, so sieht man zwar, dass die Anzahlen in der gleichen Größenordnung liegen und auch die Dispersionsindices vergleichbar hoch sind, allerdings enthält Probe 6 wesentlich weniger Leerzellen. Es ist daher davon auszugehen, dass neben einer geklumpten „Ansiedlung“ von *Copepoden* um *Lanice*-Röhren herum auch noch ein zufälliges Auftreten von *Copepoden* „im freien Gelände“ anzunehmen ist. Für die Modellierung bedeutet das die Annahme einer stochastisch unabhängigen Überlagerung eines Thomas-Prozesses für die Röhren-orientierten *Copepoden* mit einem homogenen Poisson-Prozess mit Intensität $\lambda_z > 0$ für die „frei wandernden“ *Copepoden*.



Lanice-Feld

Für die mittlere Anzahl aller Tochterpunkte des modifizierten Gesamtprozesses ergibt sich damit der Wert

$$\lambda_{\text{tot}} = E\xi(\mathbb{R}^d) = \lambda_E \cdot \lambda_T + \lambda_Z.$$

Nimmt man an, dass ein Anteil $p \in [0,1]$ aller Tochterpunkte den frei wandernden Teil der Gesamtpopulation bezeichnet, so erhält man die Gleichung

$$\lambda_Z = p \cdot \lambda_{\text{tot}} \quad \text{bzw.} \quad \lambda_E \cdot \lambda_T = (1-p) \cdot \lambda_{\text{tot}},$$

So dass der Schätzer für den Parameter λ_T modifiziert werden muss zu

$$\hat{\lambda}_T^{\text{mod}} = (1-p) \cdot \frac{\hat{\lambda}_{\text{tot}}}{\hat{\lambda}_E}.$$

Das Konzept der „schwach besetzten Zelle“ kann dabei weitgehend unverändert bleiben, solange der Anteil frei wandernder Copepoden nicht zu groß wird.

Ggf. bietet es sich an, den Schwellenwert geeignet anzupassen:

$$S_{tot} = \frac{S}{1-p}.$$

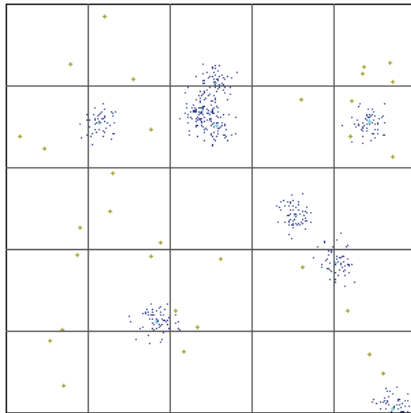
Ein geeignetes adaptives Verfahren zur Schätzung aller drei Parameter wird in PFEIFER, BÄUMER, ORTLEB, SCHLEIER UND SACH (1996) beschrieben.

Die folgenden drei Graphiken zeigen Simulationen nach dem modifizierten Thomas-Prozess für die drei Proben 6, 8 und 18. Die frei wandernden Copepoden sind mit einem grünen Kreuz dargestellt.

intensity of parent process = 10
 intensity of daughter process = 60
 intensity of noise process = 31.57
 standard deviation = 7 [Pixels]

mean value = 22.8 index-of-dispersion = 1383.94

estimate for intensity of parent process = 11.15	realized: 9
estimate for intensity of daughter process = 48.53	realized: 59.88
estimate for intensity of noise process = 28.5	per cell: 1.26



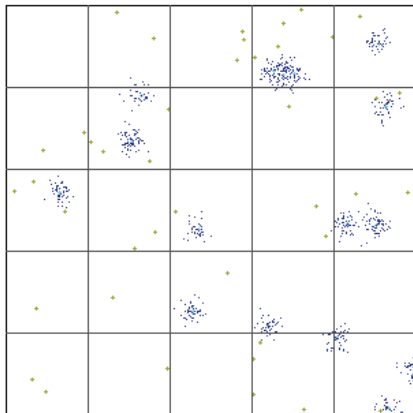
1	2	42	0	4
8	43	162	1	59
1	3	0	68	5
2	51	8	16	37
2	7	2	0	46

Simulation für Probe 6

intensity of parent process = 18
 intensity of daughter process = 50
 intensity of noise process = 47.36
 standard deviation = 5 [Pixels]

mean value = 31.71 index-of-dispersion = 1351.92

estimate for intensity of parent process = 14.49	realized: 16
estimate for intensity of daughter process = 51.97	realized: 47
estimate for intensity of noise process = 39.65	per cell: 1.89



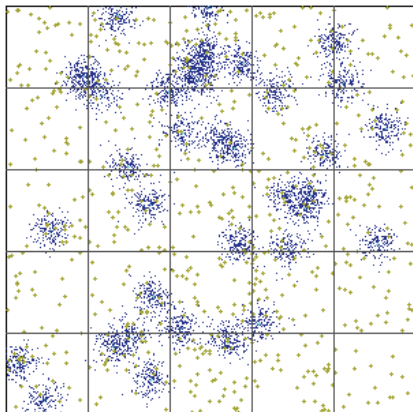
0	8	3	153	44
2	98	0	6	48
54	2	37	3	118
1	1	49	40	11
2	1	0	24	88

Simulation für Probe 8

intensity of parent process (max. 100) ? 35
 intensity of daughter process (max. 250) ? 170
 proportion p of noise intensity ? 0.1
 standard deviation [Pixels] ? 7

mean value = 272.39 index-of-dispersion = 3593.75

estimate for intensity of parent process = 40.23	realized: 37
estimate for intensity of daughter process = 152.32	realized: 165.27
estimate for intensity of noise process = 681	per cell: 26.44



192	346	897	222	210
68	322	607	263	280
198	265	170	613	151
23	285	271	238	57
337	471	254	54	16

Simulation für Probe 18

ECOLOGICAL MODELLING

INTERNATIONAL JOURNAL ON ECOLOGICAL MODELLING AND SYSTEMS ECOLOGY

Ecological Modelling 87 (1996) 285–294

Modeling spatial distributional patterns of benthic meiofauna species by Thomas and related processes

D. Pfeifer ^{a,*}, H.-P. Bäumer ^b, H. Ortleb ^c, G. Sach ^d, U. Schleier ^c

^a Institut für Mathematische Stochastik, Universität Hamburg, Bundesstr. 55, D-20146 Hamburg, Germany

^b Carl von Ossietzky Universität Oldenburg, HRZ – Angewandte Statistik, Postfach 25 03, D-26111 Oldenburg, Germany

^c Carl von Ossietzky Universität Oldenburg, Fachbereich Mathematik, Postfach 25 03, D-26111 Oldenburg, Germany

^d Christian-Albrechts-Universität Kiel, Zoologisches Institut, Olshausenstr. 40, D-24098 Kiel, Germany

Received 8 February 1995; accepted 2 October 1995



Ecological Modelling 87 (1996) 285–294

ECOLOGICAL
MODELLING

Modeling spatial distributional patterns of benthic meiofauna species by Thomas and related processes

D. Pfeifer ^{a,*}, H.-P. Bäumer ^b, H. Ortleb ^c, G. Sach ^d, U. Schleier ^c

^a Institut für Mathematische Stochastik, Universität Hamburg, Bundesstr. 55, D-20146 Hamburg, Germany

^b Carl von Ossietzky Universität Oldenburg, HRZ – Angewandte Statistik, Postfach 25 03, D-26111 Oldenburg, Germany

^c Carl von Ossietzky Universität Oldenburg, Fachbereich Mathematik, Postfach 25 03, D-26111 Oldenburg, Germany

^d Christian-Albrechts-Universität Kiel, Zoologisches Institut, Olshausenstr. 40, D-24098 Kiel, Germany

Received 8 February 1995; accepted 2 October 1995

Abstract

Spatial distributional patterns of benthic meiofauna in the presence of certain macrofauna species are of particular interest in the ecological research of tidal flats. The promotion of benthic meiofauna by macrofauna species is frequently observed in field experiments and is in part responsible for the spatial patchiness of such distributions. It is the aim of this paper to show how Thomas and related cluster processes can be used for a stochastic modeling of such patterns when only aggregate abundances of individuals from quadrat counts are available. In particular, three data samples of *Harpacticus obscurus* (Crustacea: Copepoda) from sites where the polychaete *Lanice conchilega* is predominant are analyzed.

Keywords: Cluster analysis; Marine ecosystems; Spatial patterns; Zooplankton

II. Zufällige abgeschlossene Mengen

Punktprozesse sind häufig der Ausgangspunkt für „kompliziertere“ räumliche Strukturen - zufällige abgeschlossene Mengen - , die z.B. durch geeignete zufällige Überlagerung abgeschlossener Mengen entstehen und die insbesondere ein positives Lebesgue-Maß (Fläche, Volumen) besitzen. Ihre theoretischen Grundlagen sind noch tiefer liegend als die der Punktprozesse. Wir beschränken uns hier auf den Fall des d -dimensionalen Euklidischen Raumes $(\mathcal{X}, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}^d)$ mit $d \in \mathbb{N}$.

II.1. Maßtheoretische Grundlagen

Wir verwenden hier durchgängig die folgende Notation:

$$\mathcal{I}^d := \left\{ I_{\mathbf{a}, \mathbf{b}} := \prod_{i=1}^d (a_i, b_i] \subseteq \mathbb{R}^d \mid \mathbf{a}_i, \mathbf{b}_i \in \mathbb{R}, i = 1, \dots, d \right\}$$

$$\mathcal{O}^d := \{ O \subseteq \mathbb{R}^d \mid O \text{ offen} \}$$

$$\mathcal{F}^d := \{ A \subseteq \mathbb{R}^d \mid A \text{ abgeschlossen} \}$$

$$\mathcal{K}^d := \{ K \subseteq \mathbb{R}^d \mid K \text{ kompakt} \}$$

sowie

$$\mathcal{F}_K^d := \{ F \in \mathcal{F}^d \mid F \cap K \neq \emptyset \} \quad \text{für } K \subseteq \mathfrak{X}$$

und

$$\mathcal{E}_0^d := \{ \mathcal{F}_K^d \mid K \in \mathcal{K}^d \}, \quad \mathfrak{F}^d := \sigma(\mathcal{E}_0^d).$$

Sofern keine Konfusion besteht, lassen wir den nach oben gestellten Index d im Folgenden meist weg, schreiben also $\mathcal{F}_K := \{ F \in \mathcal{F} \mid F \cap K \neq \emptyset \}$ statt $\mathcal{F}_K^d := \{ F \in \mathcal{F}^d \mid F \cap K \neq \emptyset \}$ usw.

Lemma 6: Es sei $\{A_n\}_{n \in \mathbb{N}}$ eine Mengenfolge mit $A_n \subseteq \mathfrak{X}$ für alle $n \in \mathbb{N}$.

a) Es gilt ohne weitere Einschränkungen: $\mathcal{F}_{\bigcup_{n \in \mathbb{N}} A_n} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_{A_n}$.

b) Unter der Voraussetzung $A_n \downarrow A = \bigcap_{k \in \mathbb{N}} A_k$ für ein $A \subseteq \mathfrak{X}$ und $A_n \in \mathcal{K}$ für alle $n \in \mathbb{N}$ gilt auch: $\mathcal{F}_{\bigcap_{n \in \mathbb{N}} A_n} = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{A_n}$.

Bemerkung: Die Annahme der Kompaktheit der Mengen A_n ist wesentlich, wie das folgende Gegenbeispiel zeigt: Sei $A \in \mathcal{K}$ beliebig. Dann ist $\mathbb{Q}^d \setminus A = \{q_1, q_2, \dots\}$ eine unendliche, abzählbare Menge. Setze $A_n := A \cup \{q_{n+1}, q_{n+2}, \dots\}$ für alle $n \in \mathbb{N}$. Diese Mengen sind sämtlich *nicht* kompakt, aber es gilt offensichtlich $A_n \downarrow A$. Sei nun $F \subseteq A^c$ ein beliebiges abgeschlossenes d -dimensionales Intervall (ein solches existiert immer). Dann enthält die Menge $F \cap \mathbb{Q}^d = F \cap (\mathbb{Q}^d \setminus A)$ unendlich viele Elemente, so dass $F \cap A_n \neq \emptyset$ für alle $n \in \mathbb{N}$ gilt. Es folgt also $F \in \bigcap_{n \in \mathbb{N}} \mathcal{F}_{A_n}$, nicht aber $F \in \mathcal{F}_{\bigcap_{n \in \mathbb{N}} A_n} = \mathcal{F}_A$, womit hier der Fall $\mathcal{F}_{\bigcap_{n \in \mathbb{N}} A_n} \neq \bigcap_{n \in \mathbb{N}} \mathcal{F}_{A_n}$ vorliegt.

Lemma 7: Es bezeichne $\mathcal{G}_K := \{F \in \mathcal{F} \mid F \subseteq K\}$ für $K \subseteq \mathfrak{X}$. Dann gilt $\mathcal{G}_K \in \mathfrak{F}$ für alle $K \in \mathcal{F} \cup \mathcal{O}$, und folgende Mengensysteme sind neben $\mathcal{E}_0$ ebenfalls Erzeuger von $\mathfrak{F}$:

$$\mathcal{E}_1 := \{\mathcal{F}_O \mid O \in \mathcal{O}\}$$

$$\mathcal{E}_2 := \{\mathcal{F}_A \mid A \in \mathcal{F}\}$$

$$\mathcal{E}_3 := \{\mathcal{F}_I \mid I \in \mathcal{I}\}$$

$$\mathcal{E}_4 := \{\mathcal{G}_O \mid O \in \mathcal{O}\}$$

$$\mathcal{E}_5 := \{\mathcal{G}_A \mid A \in \mathcal{F}\}.$$

Bemerkungen:

- Die Erzeuger $\mathcal{E}_0$ bis $\mathcal{E}_3$ sind *nicht* durchschnittstabil, wohl aber $\mathcal{E}_4$ und $\mathcal{E}_5$. Beispielsweise gilt für disjunkte $K_1, K_2 \in \mathcal{K}$ nur

$$\mathcal{F}_{K_1} \cap \mathcal{F}_{K_2} = \{F \in \mathcal{F} \mid F \cap K_1 \neq \emptyset \text{ und } F \cap K_2 \neq \emptyset\} \subseteq \mathcal{F}_{K_1 \cup K_2}$$

ohne Gleichheit. Insbesondere ist $\mathcal{F}_{K_1} \cap \mathcal{F}_{K_2} \neq \mathcal{F}_K$ für jedes $K \in \mathcal{K}$. Denn entweder ist K zu K_1, K_2 disjunkt - in diesem Fall gilt $K \in \mathcal{F}_K$, aber $K \notin \mathcal{F}_{K_1} \cap \mathcal{F}_{K_2}$ - oder es gilt $K \cap (K_1 \cup K_2) \neq \emptyset$, o.B.d.A. $F := K \cap K_1 \neq \emptyset$. In diesem Fall gilt $F \in \mathcal{F}_K$, aber $F \cap K_2 = \emptyset$ und somit $F \notin \mathcal{F}_{K_1} \cap \mathcal{F}_{K_2}$. $\mathcal{E}_0$ ist also nicht durchschnittstabil (analog argumentiert man für $\mathcal{E}_1$ bis $\mathcal{E}_3$).

Andererseits gilt beispielsweise für alle $O_1, O_2 \in \mathcal{O}$

$$\mathcal{G}_{O_1} \cap \mathcal{G}_{O_2} = \{F \in \mathcal{F} \mid F \subseteq O_1 \text{ und } F \subseteq O_2\} = \{F \in \mathcal{F} \mid F \subseteq O_1 \cap O_2\} = \mathcal{G}_{O_1 \cap O_2},$$

d.h. $\mathcal{E}_4$ ist durchschnittstabil (analog argumentiert man für $\mathcal{E}_5$).

- Es lässt sich zeigen, dass auch die Systeme $\mathcal{G}_k$ und $\mathcal{H}_k := \{F \in \mathcal{F} \mid F \supseteq K\}$ für $K \in \mathcal{K}$ (durchschnittstabile) Erzeuger von $\mathfrak{F}$ sind, ebenso $\mathcal{H}_O$ für $O \in \mathcal{O}$ und $\mathcal{H}_A$ für $A \in \mathcal{F}$.

Lemma 8: Es sei $\{A_n\}_{n \in \mathbb{N}}$ eine Mengenfolge mit $A_n \subseteq \mathfrak{X}$ für alle $n \in \mathbb{N}$.

a) Es gilt ohne weitere Einschränkungen: $\mathcal{G}_{\bigcap_{n \in \mathbb{N}} A_n} = \bigcap_{n \in \mathbb{N}} \mathcal{G}_{A_n}$.

b) Unter der Voraussetzung $A_n \uparrow A = \bigcup_{k \in \mathbb{N}} A_k$ für eine Menge $A \subseteq \mathfrak{X}$, wobei die A_n^c kompakt sind für alle $n \in \mathbb{N}$ gilt auch: $\mathcal{G}_{\bigcup_{n \in \mathbb{N}} A_n} = \bigcup_{n \in \mathbb{N}} \mathcal{G}_{A_n}$.

Satz 12: Es bezeichne $\mathcal{H}_K := \{F \in \mathcal{F} \mid F \supseteq K\}$ für $K \in \mathcal{K}$. Dann gilt $\mathcal{H}_K \in \mathfrak{F}$.

Bemerkung: Aus dem letzten Satz bzw. der Bemerkung am Ende von Lemma 7 ergibt sich insbesondere die interessante Schlussfolgerung, dass für jede kompakte bzw. allgemeiner abgeschlossene Menge $F \in \mathcal{F}$ wegen $\{F\} = \mathcal{G}_F \cap \mathcal{H}_F$ auch $\{F\} \in \mathfrak{F}$ liegt.

Definition 6: Es sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Ein Zufallselement $\Xi: (\Omega, \mathcal{A}) \rightarrow (\mathcal{F}, \mathfrak{F})$ heißt *zufällige abgeschlossene Menge (ZAM)*. Die durch

$$T_{\Xi}(K) := P(\Xi \cap K \neq \emptyset) = P(\Xi \in \mathcal{F}_K) = P^{\Xi}(\mathcal{F}_K), \quad K \in \mathcal{K}$$

auf $\mathcal{K}$ definierte Abbildung heißt das zu Ξ gehörige (oder von Ξ induzierte) *Kapazitätsfunktional*.

Lemma 9: Das Kapazitätsfunktional einer ZAM Ξ besitzt folgende Eigenschaften:

- a) $0 \leq T_{\Xi} \leq 1$ mit $T_{\Xi}(\emptyset) = 0$
- b) Ist $\{K_n\}_{n \in \mathbb{N}} \subseteq \mathcal{K}$ eine monoton fallende Folge kompakter Mengen mit $K_n \downarrow K \in \mathcal{K}$, so gilt auch $T_{\Xi}(K_n) \downarrow T_{\Xi}(K)$ für $n \rightarrow \infty$.
- c) Die Abbildungen S_n seien für $n \in \mathbb{Z}^+$ induktiv wie folgt definiert:

$$S_0(K) := 1 - T_{\Xi}(K) = P(\Xi \cap K = \emptyset) \text{ für } K \in \mathcal{K}$$

$$S_n(K; K_1, \dots, K_n) := S_{n-1}(K; K_1, \dots, K_{n-1}) - S_{n-1}(K \cup K_n; K_1, \dots, K_{n-1})$$

für $K, K_1, \dots, K_n \in \mathcal{K}$, $n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} S_n(K; K_1, \dots, K_n) &= -T_{\Xi}(K) + \sum_{i=1}^n T_{\Xi}(K \cup K_i) - \sum_{1 \leq i < j \leq n} \sum T_{\Xi}(K \cup K_i \cup K_j) \pm \dots \\ &\quad \dots + (-1)^{n+1} T_{\Xi}\left(K \cup \bigcup_{i=1}^n K_i\right) \geq 0 \end{aligned}$$

für $K, K_1, \dots, K_n \in \mathcal{K}$, $n \in \mathbb{N}$.

Bemerkung: Es gilt auch

$$\begin{aligned} S_n(K; K_1, \dots, K_n) &= P\left(\{\Xi \cap K = \emptyset\} \cap \bigcap_{i=1}^n \{\Xi \cap K_i \neq \emptyset\}\right) \\ &= 1 - P\left(\{\Xi \cap K \neq \emptyset\} \cup \bigcup_{i=1}^n \{\Xi \cap K_i = \emptyset\}\right) \end{aligned}$$

für alle $n \in \mathbb{N}$ ($\rightarrow$ Siebformel).

Satz 13 (Choquet): Ist T ein Funktional auf $\mathcal{K}$, das die Eigenschaften a) bis c) von Lemma 9 erfüllt, so existiert eine eindeutig bestimmte Verteilung Q auf $\mathfrak{F}$ mit

$$T(K) = Q(\mathcal{F}_K) \text{ für alle } K \in \mathcal{K}.$$

Der Beweis ist sehr aufwändig. Das wesentliche Problem besteht hier darin, dass der Erzeuger $\mathcal{E}_0 = \{\mathcal{F}_K \mid K \in \mathcal{K}\}$ nicht durchschnittstabil, also insbesondere kein Semiring ist, so dass der Maßfortsetzungssatz von Carathéodory nicht direkt angewendet werden kann.

Lemma 10: Die folgenden Abbildungen sind messbar:

a) $U: \mathbb{R}^d \ni \mathbf{x} \mapsto \{\mathbf{x}\} \in \mathcal{F}^d$

b) $V_i: \mathbb{R}^d \ni \mathbf{x} \mapsto \{x_i\} \in \mathcal{F}^1$ für jedes feste $i \in \{1, \dots, d\}$.

Bemerkung: Lemma 10 b) impliziert, dass für einen d -dimensionalen Zufallsvektor $\mathbf{X}$ die Projektion $\Xi_i = \{X_i\}$ eine ZAM in $\mathbb{R}^1$ ist.

II.2. Messbare Operationen mit zufälligen abgeschlossenen Mengen

In diesem Abschnitt zeigen wir, dass die üblichen Mengenoperationen wie Vereinigung, Durchschnitt, kartesische Produkte usw. messbare Abbildungen sind, so dass solche Operationen, angewendet auf zufällige abgeschlossene Mengen, auch wieder zufällige abgeschlossene Mengen ergeben.

Lemma 11: Die Mengenoperationen „Vereinigung“ und „Durchschnitt“ auf $\mathcal{F}$ sind $\mathfrak{F}$ -messbar.

Ähnlich zeigt man, dass auch endliche Vereinigungen und abzählbare Durchschnitte auf $\mathcal{F}$ messbare Mengenoperationen sind.

Bemerkungen:

- Es ergibt sich für Lemma 10 b) sofort folgende Verallgemeinerung: Die Abbildung V mit $V: \mathbb{R}^d \ni \mathbf{x} \mapsto \{x_1, \dots, x_d\} = \bigcup_{i=1}^d \{x_i\} \in \mathcal{F}^1$ ist messbar, so dass für einen d -dimensionalen Zufallsvektor $\mathbf{X}$ die Menge der Koordinatenwerte $\{X_1, \dots, X_d\}$ eine ZAM ist (ggf. nach Streichung von Vielfachheiten).
- Endliche (auch zufällige) Vereinigungen und abzählbare (auch zufällige) Durchschnitte von ZAM sind wieder ZAM.
- Die durch $W: \mathcal{F}^d \rightarrow \mathcal{F}^d: F \mapsto \overline{F^c}$ und $R: \mathcal{F}^d \rightarrow \mathcal{F}^d: F \mapsto \partial F$ definierten Abbildungen (Abschluss des Komplements von F bzw. Rand von F) sind $\mathfrak{F}^d$ - $\mathfrak{F}^d$ -messbar.

Lemma 12: Die Abbildung $R: \mathbb{R}^n \rightarrow \mathbb{R}^m$ sei stetig, für $n, m \in \mathbb{N}$. Dann ist die Urbildabbildung $U: \mathcal{F}^m \rightarrow \mathcal{F}^n: F \mapsto R^{-1}(F)$ für $F \in \mathcal{F}$ $\mathfrak{F}^m$ - $\mathfrak{F}^n$ -messbar.

Bemerkung: Mit Ξ ist dann auch $R^{-1}(\Xi)$ eine ZAM.

Lemma 13: Die Mengenoperation „kartesisches Produkt“ für abgeschlossene Mengen ist messbar.

Bemerkung: mit zwei n - bzw. m -dimensionalen ZAM Ξ_1 und Ξ_2 ist also auch $\Xi_1 \times \Xi_2$ eine ZAM.

Lemma 14: Die Abbildung $S: \mathcal{F}^d \times \mathbb{R}^d \rightarrow \mathbb{R}: (F, x) \mapsto \mathbb{1}_F(x)$ ist $\mathfrak{F}^d \otimes \mathcal{B}^d$ -messbar.

Satz 15: Es sei μ ein σ -endliches Maß auf der Borel- σ -Algebra $\mathcal{B}^d$ und für $B \in \mathcal{B}^d$ die Abbildung T_B definiert durch:

$$T_B: \mathcal{F}^d \rightarrow \mathbb{R}^+: F \mapsto \mu_B(F) := \mu(F \cap B) \quad [\text{Spurmaß}].$$

Dann ist T_B $\mathfrak{F}^d$ -messbar für jedes $B \in \mathcal{B}^d$.

Satz 16 („Robbin’s Theorem“): Es sei μ ein σ -endliches Maß auf $\mathcal{B}^d$ und Ξ eine zufällige abgeschlossene Menge. Dann gilt:

- a) $\mu_{\Xi}(B) = \mu(\Xi \cap B) = \mu_B(\Xi)$ ist eine nicht-negative Zufallsvariable für alle $B \in \mathcal{B}^d$.
- b) $E[\mu_{\Xi}(B)] = E[\mu(\Xi \cap B)] = \int_B T_{\Xi}(\{x\}) \mu(dx)$ für alle $B \in \mathcal{B}^d$.

Lemma 17: Es sei $F \in \mathcal{F}^d$ mit $\mathbf{0} \in F$. Dann ist die *Skalierungs-Abbildung*

$$\rho: \mathbb{R}^+ \rightarrow \mathcal{F}^d: r \mapsto r \cdot F := \{r \cdot x \mid x \in F\}$$

$\mathcal{B}^1$ - $\mathfrak{F}^d$ -messbar.

Bemerkung: Lemma 17 impliziert, dass das Objekt $\Xi := R \cdot F$ eine ZAM bildet, wenn R eine nicht-negative reelle Zufallsvariable ist. Diese Tatsache führt uns später zu einer mathematischen Formulierung eines *Minimal-Areals* mit Anwendungen in der Biologie und Ökologie.

Ist speziell $F = K(1)$ der Einheitskreis in $\mathfrak{X} = \mathbb{R}^2$, so ist also $\Xi = R \cdot F = K(R)$ mit dem Kapazitätsfunktional

$$T_{\Xi}(\{x\}) = P(x \in \Xi) = P(R \geq \|x\|_2) = 1 - F_R(\|x\|_2) \text{ für alle } x \in \mathbb{R}^2,$$

wobei hier zusätzlich angenommen sei, dass die Verteilungsfunktion F_R von R stetig ist. Eine formale Anwendung von Satz 16 ergibt dann mit dem σ -endlichen Lebesgue-Maß $\mu = m^2$ nach Umrechnung in Polarkoordinaten

$$\begin{aligned} E[\mu(\Xi)] &= \int_{\mathbb{R}^2} T_{\Xi}(\{x\}) \mu(dx) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(1 - F_R(\sqrt{x_1^2 + x_2^2})\right) dx_1 dx_2 \\ &= \int_0^{2\pi} \int_0^{\infty} r \cdot (1 - F_R(r)) dr d\varphi = \pi \int_0^{\infty} 2r \cdot (1 - F_R(r)) dr = \pi \int_0^{\infty} \underset{r=\sqrt{z}}{\left(1 - F_R(\sqrt{z})\right)} dz \\ &= \pi \int_0^{\infty} (1 - F_{R^2}(z)) dz = \pi \cdot E(R^2), \end{aligned}$$

was wegen $\mu(\Xi) = \pi R^2$ zu erwarten war.

Definition 8: Eine d -dimensionale ZAM Ξ heißt *stationär*, wenn $P^\Xi = P^{\Xi+\mathbf{a}}$ gilt für alle $\mathbf{a} \in \mathbb{R}^d$, und *isotrop*, wenn $P^\Xi = P^{D(\Xi)}$ gilt für alle Drehungsabbildungen D auf $\mathbb{R}^d$. Die ZAM Ξ heißt *bewegungsinvariant*, wenn sie stationär und isotrop ist.

Bemerkung: Die Begriffe *stationär* und *isotrop* sind für ZAM sinnvoll, weil $x \mapsto x - \mathbf{a}$ und $x \mapsto D^{-1}(x)$ stetige (sogar bijektive) Abbildungen von $\mathbb{R}^d$ auf sich sind und damit nach Lemma 12 für jede ZAM Ξ die durch die entsprechenden Urbilder induzierten Objekte $\Xi + \mathbf{a}$ und $D(\Xi)$ ebenfalls zufällige abgeschlossene Mengen sind. Die Stationarität einer ZAM ist dabei keine einfache Eigenschaft; nicht-triviale Beispiele werden wir erst in einem späteren Abschnitt kennen lernen.

Lemma 18: Eine d -dimensionale ZAM Ξ ist genau dann stationär, wenn mit $K \in \mathcal{K}^d$ $T_\Xi(K) = T_\Xi(K - \mathbf{a})$ gilt für alle $\mathbf{a} \in \mathbb{R}^d$, und *isotrop*, wenn $T_\Xi(K) = T_\Xi(D^{-1}(K))$ gilt für alle Drehungsabbildungen D auf $\mathbb{R}^d$.

Definition 9: Es sei Ξ eine d -dimensionale stationäre ZAM. Dann heißt die Größe

$$p_{\Xi} := E \left[m^d \left(\Xi \cap [0,1]^d \right) \right]$$

Bedeckungsgrad von Ξ . Die Abbildung

$$C_{\Xi} : \mathbb{R}^d \rightarrow \mathbb{R} : h \mapsto P \left(\{\mathbf{0} \in \Xi\} \cap \{h \in \Xi\} \right)$$

heißt *Kovarianzfunktion* von Ξ .

Bemerkung: Der Bedeckungsgrad einer ZAM ist ein charakteristischer Parameter; ihm kommt vor allem in biologischen und ökologischen Anwendungen eine wesentliche Bedeutung zu.

Zur Beachtung: es gilt $C_{\Xi}(h) = P \left(\{\mathbf{0} \in \Xi\} \cap \{h \in \Xi\} \right) = P^{\Xi} \left(\mathcal{F}_{\{\mathbf{0}\}}^d \cap \mathcal{F}_{\{h\}}^d \right)$ für alle $h \in \mathbb{R}^d$.

Lemma 19: Es sei Ξ eine d -dimensionale stationäre ZAM. Dann gilt

$$\rho_{\Xi} = T_{\Xi}(\{\mathbf{0}\}) = P(\mathbf{0} \in \Xi) = C_{\Xi}(\mathbf{0}).$$

Ist die ZAM Ξ zusätzlich isotrop, dann gilt noch

$$C_{\Xi}(h) = C_{\Xi}(k) \text{ für alle } h, k \in \mathbb{R}^d \text{ mit } \|h\|_2 = \|k\|_2,$$

d.h. der Ausdruck $C_{\Xi}(h)$ hängt funktional nur von $\|h\|_2$ ab; Schreibweise:

$$C_{\Xi}(h) = c_{\Xi}(\|h\|_2) \text{ für alle } h \in \mathbb{R}^d.$$

Statistische Auswertung des Bedeckungsgrades einer stationären ZAM:

Eine methodisch recht einfache Vorgehensweise besteht darin, vor Ausführung des Experiments eine feste endliche Punktmenge $\mathfrak{X}_0 = \{x_1, \dots, x_N\} \subset \mathbb{R}^d$ mit $N \in \mathbb{N}$ als „Testmenge“ auszuwählen und als Schätzer für p_Ξ den Mittelwert

$$\hat{p}_\Xi := \frac{1}{N} \sum_{k=1}^N \mathbb{1}_\Xi(x_k) = \frac{1}{N} \#(\Xi \cap \mathfrak{X}_0)$$

anzusetzen. Dieser Schätzer ist erwartungstreu wegen

$$E(\hat{p}_\Xi) = \frac{1}{N} \sum_{k=1}^N E[\mathbb{1}_\Xi(x_k)] = \frac{1}{N} \sum_{k=1}^N P(x_k \in \Xi) = P(\mathbf{0} \in \Xi) = p_\Xi$$

und besitzt unter Isotropie die Schätzvarianz

$$\text{Var}(\hat{p}_\Xi) = \frac{1}{N} p_\Xi - p_\Xi^2 + \frac{2}{N^2} \sum_{1 \leq i < j \leq N} c_\Xi(\|x_i - x_j\|_2).$$

II.3. Das Minimal-Areal-Problem der Statistischen Ökologie

1. Was ist ein Minimal-Areal ?

Das Problem des „richtigen“ Stichprobenumfangs oder der „richtigen“ Probenflächengröße ist seit jeher ein brisantes Thema in Biologie und Ökologie. Hierzu zwei Zitate von Fachleuten:

Optimum size of a plot for vegetation has always intrigued plant ecologists. Qualitative ecologists emphasize sampling for recurring plant assemblages and are interested in a plot size on which the species composition of the community is adequately represented. Such a plot is also referred to as „*minimal area*“.

The selection of an appropriate plot size for measurement is a subjective decision and is based primarily on the size and spacing of individuals of a species. Some ecologists have tried to reach a plot size compromise by taking a sufficient number of smaller plots so that the sum of the areas of smaller plots is equal to, or larger than, the minimal area. Such an objective is achieved either by placing the smaller plots side by side until the size and shape of the minimal area is attained or by random placement of the plots in the community to be sampled.

[Aus: CH. D. BONHAM (1989): Measurements for Terrestrial Vegetation]

II.3. Das Minimal-Areal-Problem der Statistischen Ökologie

1. Was ist ein Minimal-Areal ?

Das Problem des „richtigen“ Stichprobenumfangs oder der „richtigen“ Probenflächengröße ist seit jeher ein brisantes Thema in Biologie und Ökologie. Hierzu zwei Zitate von Fachleuten:

Wie groß soll die Aufnahmefläche gewählt werden? - Sie soll einerseits möglichst alle vorkommenden Arten enthalten, andererseits sollte ihre Mindestgröße nicht zu hoch angesetzt werden, weil man sonst sehr leicht mit den Forderungen nach Einheitlichkeit und Quasi-Homogenität der Verteilung in Konflikt gerät.

Man kann versuchen, diese Frage objektiv zu entscheiden, indem man die Probefläche allmählich vergrößert und die Artenzahl gegen die Flächengröße aufträgt. Wenn man dies im Gelände tatsächlich durchführt, so erhält man eine Art Sättigungskurve, die allerdings entgegen der Erwartung nicht vollständig in eine Waagerechte übergeht, sondern bei starker Erweiterung des so ermittelten „*Minimalareals*“ immer noch Anstiegstendenzen erkennen läßt. Die Form der Aufnahmefläche spielt dabei keine Rolle.

[Nach M. MÜHLENBERG (1993): Freilandökologie]

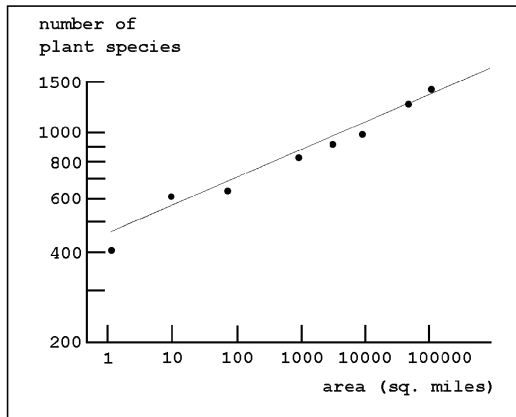
II.3. Das Minimal-Areal-Problem der Statistischen Ökologie

1. Was ist ein Minimal-Areal ?

Offensichtlich enthalten beide „Definitionen“ eines Minimal-Areals eine gewisse Art von Willkür: für Bonham ist seine Bestimmung gar eine subjektive Angelegenheit, das Kriterium ist die „Adäquanz“ der Artenrepräsentation; für Mühlenberg gibt es immerhin eine Art mathematisches Kriterium, nach dem seine Größe bestimmt werden kann, indem man den Verlauf der Art-Areal-Kurve zur Entscheidung heranzieht. Beiden gemeinsam scheint jedoch unübersehbar der „dynamische“ Charakter des Minimalareals zu sein: man vergrößert das Probegebiet solange, bis der gewünschte Grad an „Zufriedenheit“ über das Ergebnis erreicht ist. Wir werden später sehen, daß genau diese Idee für eine sinnvolle Mathematisierung des Minimal-Areal-Begriffs verwendet werden kann, die alle wesentlichen Vorstellungen der Substanzwissenschaften dazu in sich trägt.

2. Flächen-Arten-Beziehungen

Ein bereits seit längerer Zeit bekannter Effekt ist die empirisch immer wieder belegbare Tatsache, daß die in Freilandversuchen ermittelte Anzahl gefundener Arten, in einem doppelt-logarithmischen Maßstab gegen die Probenflächengröße aufgetragen, den Eindruck einer linearen Beziehung vermittelt.



Die linke Graphik stammt aus entsprechenden Untersuchungen von Williams (1964) zum Pflanzenvorkommen in England; die Probenflächengröße erstreckt sich dabei in einer Skala von wenigen hundert Quadratmetern bis hin zur Gesamtgröße von England.

Wir wollen nun in einem ersten Schritt zeigen, daß ein solcher Effekt auch durch ein geeignetes - sogar recht einfaches - stochastisches Verbreitungsmodell erklärt werden kann. Dazu machen wir folgende idealisierende Annahmen über die insgesamt s in der Ebene vorhandenen Arten:

- Das Verbreitungsmuster jeder Art i wird durch einen räumlich homogenen Poisson-Prozeß ξ_i mit Parameter $\lambda_i > 0$ beschrieben.
- Die Verbreitungsmuster der s vorhandenen Arten sind stochastisch unabhängig.

Definieren wir nun für Borel-Mengen $A \in \mathcal{B}^2$ die Größen

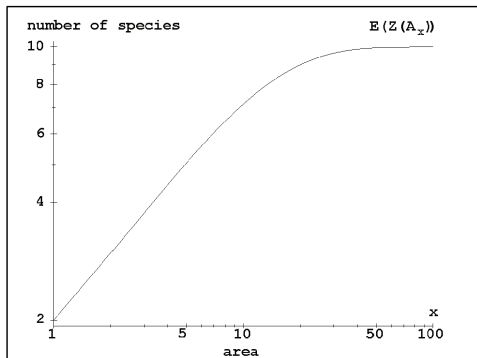
$$Z_i(A) := \begin{cases} 1, & \text{wenn } \xi_i(A) > 0 \\ 0, & \text{sonst} \end{cases} \quad \text{mit} \quad Z(A) := \sum_{i=1}^s Z_i(A),$$

so gibt $Z(A)$ die Anzahl der in der Menge A gefundenen Arten an. Der Erwartungswert dieser Zufallsgröße läßt sich auf Grund unserer Annahmen leicht bestimmen zu

$$E(Z(A_x)) = \sum_{i=1}^s (1 - e^{-\lambda_i x}) =: S(x)$$

für jede Borel-Menge $A_x \in \mathcal{B}^2$ mit Lebesgue-Maß $m^2(A_x) = x > 0$, unabhängig von der Form der Menge A_x . Die Annahme eines homogenen Poisson-Prozesses im Verbreitungsmodell entspricht dabei der Quasi-Homogenität bei Mühlenberg; daher ist es nicht verwunderlich, daß sowohl bei Mühlenberg als auch in unserem Modell die *Form* des Aufnahmegebiets A keine Rolle spielt.

Die Funktion S heißt in der Ökologie *Art-Areal-Kurve* (engl.: species-area-curve).



Die linke Graphik zeigt die erwartete Anzahl gefundener Arten im Poisson-Modell in Abhängigkeit von der Flächen-größe für einen in Öko-systemen „typischen“ Parametersatz der Form

$$\lambda_1 = \dots = \lambda_7 = 0,1$$

$$\lambda_8 = 0,15$$

$$\lambda_9 = 0,3$$

$$\lambda_{10} = 2,7.$$

Man erkennt hier sehr deutlich sowohl den im unteren Bereich linear erscheinenden Verlauf der Art-Areal-Kurve als auch den von Mühlensberg in diesem Zusammenhang angedeuteten „Sättigungseffekt“.

3. Eine mathematische Definition des Minimal-Areals

Nach den obigen Ausführungen können wir jetzt einen Versuch unternehmen, ein Minimal-Areal mathematisch zu definieren, wobei die Ideen von Bonham und Mühlenberg angemessene Berücksichtigung finden sollen.

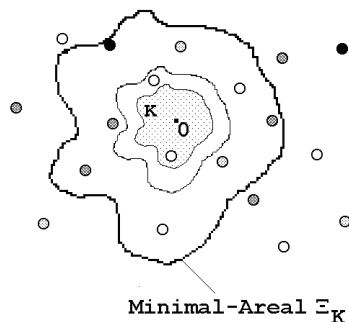
Definition 10: Es sei $K \in \mathcal{K}^2$ eine kompakte, sternförmige Menge mit $0 \in K$ und $m^2(K) = 1$. Setze, mit den Bezeichnungen von oben,

$$\rho_K = \inf \{r > 0 \mid Z(rK) = s\}, \quad \Xi_K = \rho_K K.$$

Die zufällige abgeschlossene Menge Ξ_K heißt *Minimal-Areal* vom Typ K (vgl. auch Lemma 17).

Lemma 20: Bezeichnet die Größe $W_K = m^2(\Xi_K)$ das (zufällige) Lebesgue-Maß des Minimal-Areals Ξ_K vom Typ K , so gilt für die zugehörige Verteilungsfunktion:

$$P(W_K \leq x) = P\left(\bigcap_{i=1}^s \{Z_i(A_x) = 1\}\right) = \prod_{i=1}^s (1 - e^{-\lambda_i x}), \quad x > 0.$$



Beispiel für ein
Minimal-Areal mit vier
Arten

Die obige Definition reflektiert gut den schon erwähnten „dynamischen“ Charakter eines Minimal-Areals: man vergrößere (oder verkleinere, je nachdem) das Probengebiet K solange, bis *erstmalig* – bei Erreichen von ρ_K – alle s vorhandenen Arten erfaßt sind. Hierdurch wird einerseits einer gewissen Art von „Minimalität“ Rechnung getragen, auf der anderen Seite wird aber auch die grundsätzlich stochastische Natur des Minimal-Areals deutlich.

Die Sternförmigkeit des Gebiets K spielt dabei aus mathematischer Sicht keine Rolle; sie ist lediglich deshalb in die Definition mit aufgenommen worden, weil in der Praxis etwa im Rahmen einer Freilanduntersuchung bereits beprobte Teilgebiete grundsätzlich weiterverwendet werden und ein nicht-stern-förmiges Gebiet bei Vergrößerung u.U. „Löcher“ erzeugt, in denen schon Proben genommen wurden (man denke etwa an einen echten Kreisring K , der bei Vergrößerung nach außen „wandert“). Die Bedingung $0 \in K$ dient entsprechend lediglich einer Festlegung des aktuellen „Standorts“ als Nullpunkt innerhalb des Ausgangsgebiets; Kompaktheit bezieht sich hier im wesentlichen auf die räumliche Beschränktheit des Probengebiets (ansonsten würde aus mathematischer Sicht auch die Abgeschlossenheit von K genügen).

Eine elegante Art, Momente der Minimal-Areal-Fläche W_K zu berechnen, liegt in der Bestimmung der zugehörigen momenterzeugenden Funktion ψ_{W_K} .

Satz 17: Für das Minimal-Areal Ξ_K vom Typ K gilt:

$$\psi_{W_K}(t) = E(e^{tW_K}) = \sum_{k=1}^s (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq s} \frac{\lambda_{i_1} + \dots + \lambda_{i_k}}{\lambda_{i_1} + \dots + \lambda_{i_k} - t} \quad \text{für } t < \min\{\lambda_1, \dots, \lambda_s\},$$

woraus sich für die Momente ergibt:

$$E(W_K^n) = n! \sum_{k=1}^s (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq s} \frac{1}{(\lambda_{i_1} + \dots + \lambda_{i_k})^n} \quad \text{für alle } n \in \mathbb{N}.$$

4. Anwendungen in der Ökologie

Die folgende Tabelle enthält Artenspektren aus Monitoring-Untersuchungen im Niederländischen Wattenmeer, die mit einer Probenfläche von 0,43 m² erhoben wurden (drei Standorte B,C,J in den Monaten 03 (März) und 08 (August)). Die Angaben sind hochgerechnete Anzahlen je m².

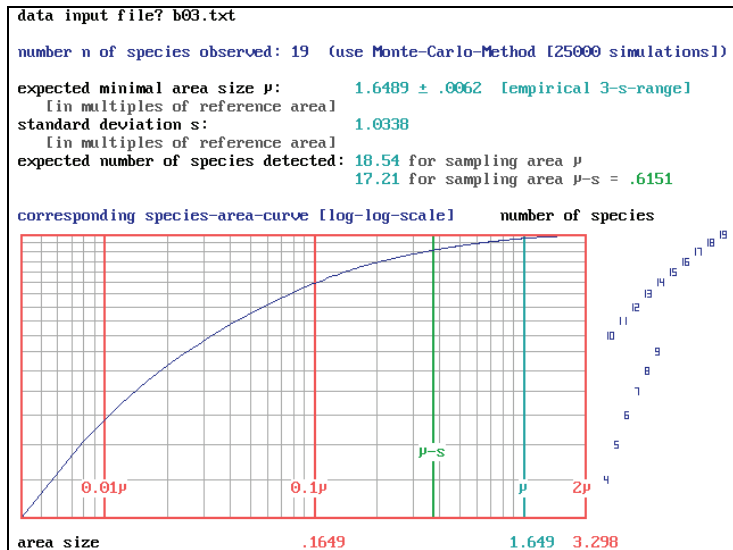
Species	B03	C03	J03	B08	C08	J08
1 Sagartia troglodytes	0	0	0	0	0	2
2 Nemertini sp.	0	0	6	0	5	2
3 Littorina littorea	6	0	0	0	0	0
4 Hydrobia ulvae	0	31	0	0	1190	0
5 Mytilus edulis	129	16	0	3224	0	2
6 Cerastoderma edule	72	9	0	56	64	0
7 Tellina tenuis	0	0	82	0	0	63
8 Macoma balthica	100	76	27	108	95	32
9 Scrobicularia plana	10	0	0	5	0	0
10 Ensis americanus	1	0	0	4	18	52
11 Mya arenaria	9	12	0	27	171	0
12 Harthothoe sarsi	0	0	1	5	2	0
13 Eteone longa	19	13	10	7	24	0
14 Anaitides mucosa	13	3	8	0	26	0
15 Nereis diversicolor	50	42	0	164	81	0
16 Nereis virens	0	0.3	0	0	0	0
17 Nereis longissima	2	0	0	0	0	2
18 Nephtys hombergii	4	1	22	2	0	17
19 Scoloplos armiger	21	148	19	14	731	95
20 Scoleopsis foliosa	0	2	6	0	2	7
21 Marenzelleria viridis	1	1	0	0	0	0
22 Magelona papillicornis	0	0	9	0	0	21
23 Tharyx marioni	0	0	0	0	0	2
24 Heteromastus filiformis	94	98	0	279	100	0
25 Arenicola marina	48	42	7	40	52	7
26 Lanice conchilega	2	0	4	24	0	21
27 Gammarus locusta	9	0	0	19	0	0
28 Bathyporeia sarsi	0	0	11	0	0	7
29 Urothoe poseidonis	0	0	1	0	0	0
30 Corophium arenarium	0	2	0	0	50	0
31 Crangon crangon	0	2	3	17	52	21
32 Carcinus maenas	9	0	0	38	18	2
33 Asterias rubens	0	0	0	0	0	2

$$E(W_K) = E\left(W_K^2(\Xi_K)\right) \quad \mathbf{1,66} \quad \mathbf{3,78} \quad \mathbf{1,55} \quad \mathbf{0,66} \quad \mathbf{0,77} \quad \mathbf{1,29}$$

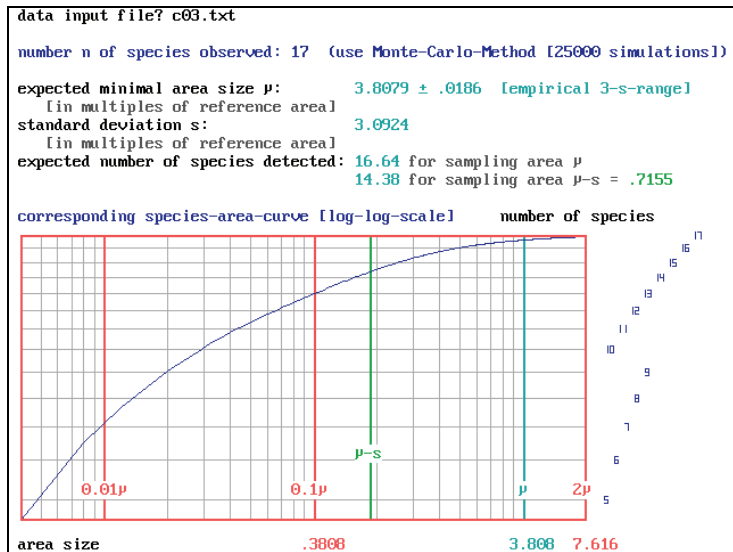
Nimmt man idealisierend an, daß sich die verschiedenen Arten über der Probenfläche „rein zufällig“ und ohne gegenseitige Beeinflussung verteilen, so kann man das obige Minimal-Areal-Modell zumindest näherungsweise zur Anwendung bringen. Die erhobenen Abundanzen können dann als erwartungstreue Schätzer für die benötigten Poisson-Parameter betrachtet werden. Man erhält damit für den Erwartungswert der Minimal-Areal-Fläche die obigen Zahlen (in m^2).

Man sieht hier aber, daß die Erwartungswerte der Minimal-Areal-Flächen in jedem Fall die ursprüngliche Probenflächengröße von $0,43 \text{ m}^2$ weit übersteigen; das bedeutet, daß insbesondere die selteneren Arten bzw. solche mit niedriger Besiedlungsdichte häufig in der Probenfläche nicht gefunden werden, auch wenn sie – in großräumigerem Maßstab – tatsächlich vorhanden sind. Die zahlreichen Nullen in der Tabelle sind also nicht unbedingt ein Indikator dafür, daß die entsprechenden Arten nicht vorhanden wären – man hat sie u.U. nur einfach nicht entdeckt!

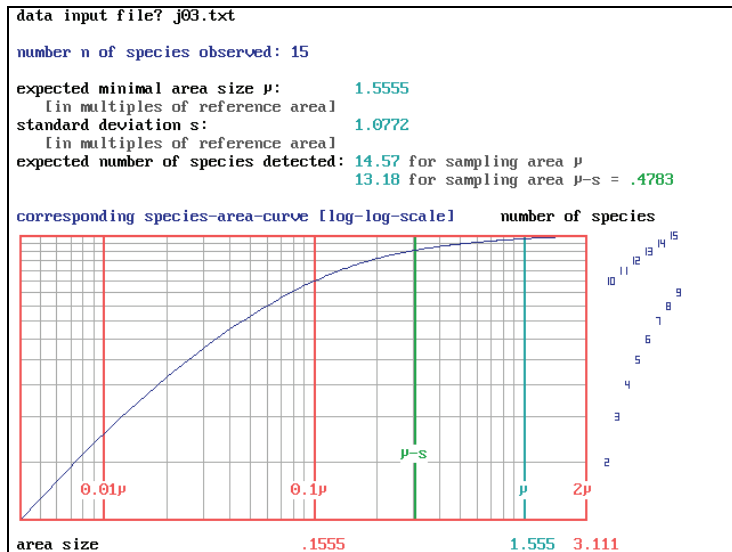
Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.



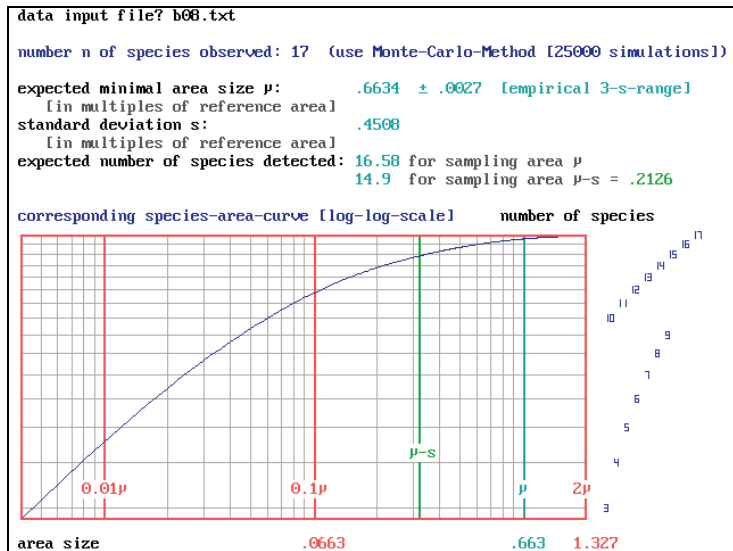
Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.



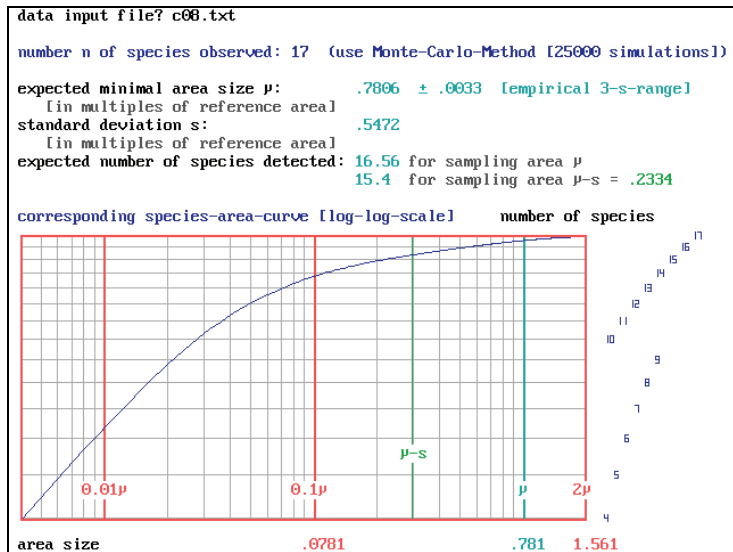
Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.



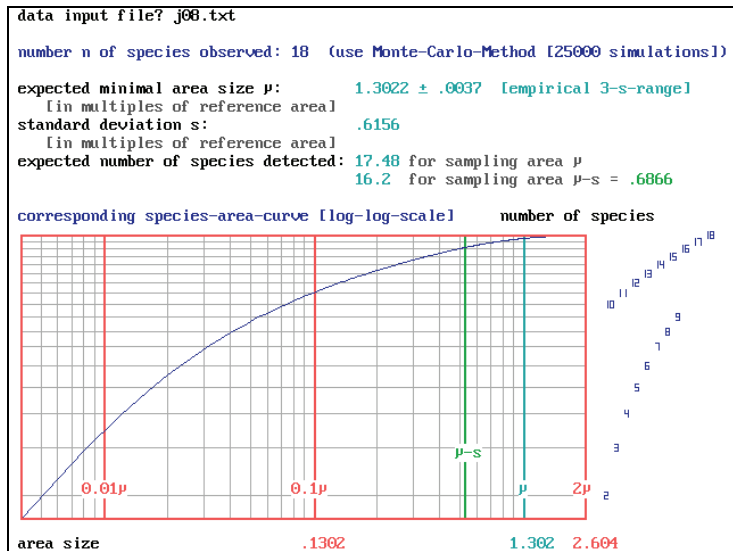
Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.

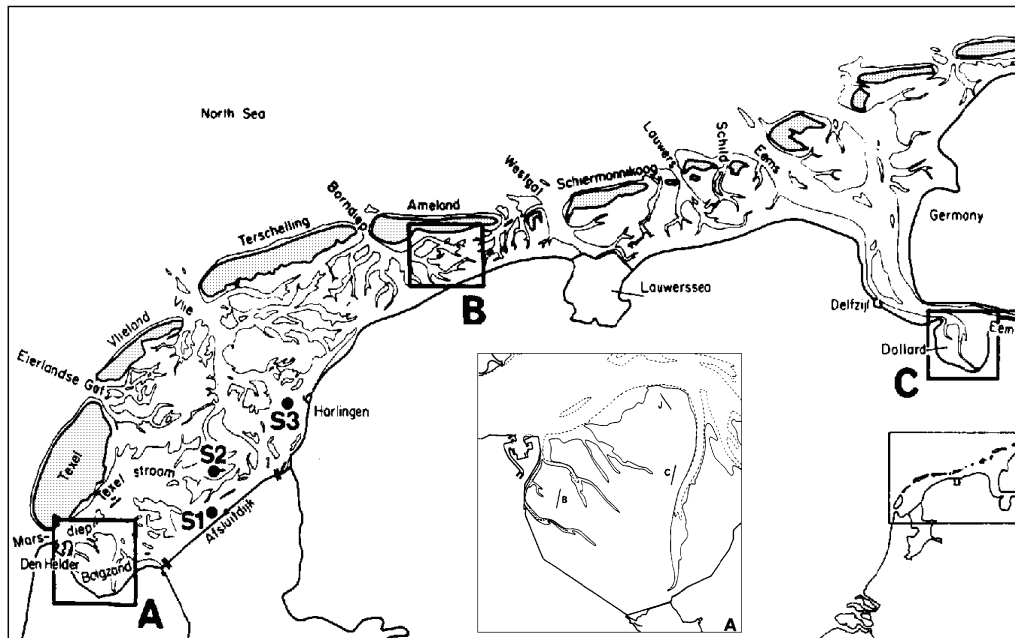


Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.



Die folgenden Graphiken zeigen die mit einem Programm berechneten Minimal-Areal-Auswertungen für die 6 Proben aus dem Wattenmeer.





Karte des Probengebiets

II. 4. Boole-Modelle

Eine besonders interessante und zugleich verhältnismäßig einfache Konstruktion stationärer / isotroper zufälliger abgeschlossener Mengen erhält man durch das Ersetzen der „Punkte“ eines stationären / isotropen Punktprozesses durch geeignete „einfache“ zufällige abgeschlossene Mengen (z.B. Kugeln, Würfel usw.) Die so entstehenden Boole-Modelle gehen auf Matheron und andere Autoren gegen Ende der 1960er Jahre zurück; sie werden oft auch als *Keim-Korn-Modelle* bezeichnet. Insbesondere lassen sich damit geeignete Punktprozesse auch mit zufälligen abgeschlossenen Mengen identifizieren.

Lemma 21: Es sei N eine endliche Zufallsvariable mit Werten in $\mathbb{Z}^+$, unabhängig von $\{X_n\}_{n \in \mathbb{N}}$, einer Familie stochastisch unabhängiger d -dimensionaler Zufallsvektoren mit derselben atomlosen Verteilung Q . Dann ist $\Xi := \bigcup_{k=1}^N \{X_k\} = \{X_1, \dots, X_N\}$ eine zufällige abgeschlossene Menge (mit der Konvention $\Xi = \emptyset$, falls $N = 0$) mit dem Kapazitätsfunktional

$$T_{\Xi}(K) = E\left[1 - (1 - Q(K))^N\right] = 1 - \varphi_N(1 - Q(K)) \text{ für alle } K \in \mathcal{K}^d,$$

wobei φ_N die wahrscheinlichkeitserzeugende Funktion der Zufallsvariablen N bezeichne.

Bemerkung: Ist $N = n$ eine konstante Zufallsvariable für ein $n \in \mathbb{N}$, so gilt offensichtlich

$$T_{\Xi}(K) = 1 - (1 - Q(K))^n \text{ für alle } K \in \mathcal{K}^d.$$

Mit Hilfe von Lemma 21 ist es möglich, einen fundamentalen Zusammenhang zwischen Punktprozessen und zufälligen abgeschlossenen Mengen herzustellen. Wir beginnen mit einer Beschreibung der Identifikation zwischen Poisson-Punktprozessen und ZAM.

Satz 18: Es sei $\xi = \sum_{n=1}^N \varepsilon_{X_n}$ ein kanonischer Poisson-Punktprozess mit den Parametern $\lambda > 0$ und einer atomlosen Verteilung Q im Sinne von Satz 3. Dann ist $\Xi := \bigcup_{k=1}^N \{X_k\} = \{X_1, \dots, X_N\}$ eine zufällige abgeschlossene Menge mit dem Kapazitätsfunktional

$$T_{\Xi}(K) = 1 - e^{-\lambda Q(K)} \text{ für alle } K \in \mathcal{K}^d.$$

Ist umgekehrt $\Xi = \bigcup_{k=1}^N \{X_k\} = \{X_1, \dots, X_N\}$ eine ZAM gemäß der Konstruktion von Lemma 21 mit $P^N = \mathcal{P}(\lambda)$, so ist $\xi := \sum_{n=1}^N \varepsilon_{X_n}$ ein kanonischer Poisson-Punktprozess mit den Parametern $\lambda > 0$ und Q im Sinne von Satz 3.

Satz 19: Es sei ξ ein homogener Poisson-Punktprozess mit Parameter $\lambda > 0$ im Sinne von Definition 5 und $I = \{\mathbf{j}_n \mid n \in \mathbb{N}\}$ eine Abzählung der Menge $I = \mathbb{Z}^d$. Dann gibt es Familien $\{N_n\}_{n \in \mathbb{N}}$ und $\{X_{ij}\}_{i,j \in \mathbb{N}}$ stochastisch unabhängiger Zufallsvariablen mit den Eigenschaften:

- $P^{N_n} = \mathcal{P}(\lambda)$ für alle $n \in \mathbb{N}$
- $P^{X_{ij}} = \mathcal{U}(I_i)$ für alle $i, j \in \mathbb{N}$ mit dem Standard-Intervall $I_i = I_{\mathbf{j}_i, \mathbf{j}_i + \mathbf{1}} \in \mathcal{B}^d$ (wobei $\mathbf{j}_i + \mathbf{1}$ durch koordinatenweise Addition der 1 aus $\mathbf{j}_i$ hervorgeht)

so dass ξ die Darstellung

$$\xi = \sum_{i=1}^{\infty} \sum_{j=1}^{N_i} \varepsilon_{X_{ij}}$$

besitzt. Dann ist $\Xi := \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{N_i} \{X_{ij}\}$ eine zufällige abgeschlossene Menge mit dem Kapazitätsfunktional

$$T_{\Xi}(K) = 1 - e^{-\lambda m^d(K)} \text{ für alle } K \in \mathcal{K}^d.$$

Bemerkungen:

- Da das Lebesgue-Maß m^d auf $\mathcal{B}^d$ σ -endlich ist, kann Robbin's Theorem (Satz 16) angewendet werden mit dem Ergebnis

$$E[\mu_{\Xi}(B)] = E[\mu(\Xi \cap B)] = \int_B T_{\Xi}(\{x\}) m^d(dx) = \int_B 0 m^d(dx) = 0$$

für alle $B \in \mathcal{B}^d$ wegen

$$T_{\Xi}(\{x\}) = 1 - \exp(-\lambda m^d(\{x\})) = 1 - 1 = 0 \text{ für alle } x \in \mathbb{R}^d.$$

Der Bedeckungsgrad p_{Ξ} eines homogenen Poisson-Prozesses Ξ in ZAM-Form ist also erwartungsgemäß Null, da (auch abzählbar unendlich viele) isolierte Punkte des $\mathbb{R}^d$ das Lebesgue-Maß Null besitzen.

- Der in Satz 19 konstruierte homogene Poisson-Prozess Ξ in ZAM-Form ist offensichtlich auch mit diesem Sprachgebrauch stationär und isotrop im Sinne von Definition 8. Damit ergibt sich das vorige Ergebnis auch direkt aus Lemma 19.

Die Idee eines Boole-Modells besteht nun - vereinfachend ausgedrückt - darin, die „Punkte“ einer solchen ZAM geeignet zu „vergrößern“, indem diese durch Objekte mit positivem Lebesgue-Maß ausgetauscht werden, was einen positiven Bedeckungsgrad im Falle der Stationarität zur Folge hat.

Definition 11 (Boole-Modell): Es sei Ξ_0 eine *abzählbare* zufällige abgeschlossene Menge und $\{\Xi_n\}_{n \in \mathbb{N}}$ eine auch von Ξ_0 stochastisch unabhängige Familie von zufälligen *beschränkten* abgeschlossenen (d.h. zufälligen kompakten) Mengen mit demselben Kapazitätsfunktional $T = T_{\Xi_n}$ für alle $n \in \mathbb{N}$. Dann heißt die zufällige abgeschlossene Menge

$$\Xi := \bigcup_{x_n \in \Xi_0} \{x_n + \Xi_n\}$$

Boole-Modell mit den *Keimen* $x \in \Xi_0$ und den *Körnern* Ξ_n . Ist insbesondere Ξ_0 ein Poisson-Punktprozess in ZAM-Form, so spricht man auch von einem Poisson-Boole-Modell.

Bemerkungen:

- Häufig wählt man als Körner d -dimensionale abgeschlossene Kugeln mit Mittelpunkt $\mathbf{0} \in \mathbb{R}^d$ oder andere geeignete kompakte Mengen wie abgeschlossene Standard-Intervalle oder abgeschlossene Ellipsoide, die die $\mathbf{0} \in \mathbb{R}^d$ als Element enthalten.
- Im einfachen Poisson-Fall des Satzes 18 lässt sich das Boole-Modell auch so schreiben:

$$\Xi = \bigcup_{k=1}^N \{X_k + \Xi_k\}.$$

Lemma 22: Die Mengenoperationen „Spiegelung“ und „Minkowski-Addition“ auf $\mathcal{F}^d$ sind $\mathfrak{F}^d$ -messbar.

Lemma 23: Es sei X ein d -dimensionaler Zufallsvektor mit der Verteilung Q und Ξ_1 eine von X unabhängige zufällige abgeschlossene Menge. Dann besitzt die zufällige abgeschlossene Menge $\Xi := X + \Xi_1$ das Kapazitätsfunktional

$$T_{\Xi}(K) = E\left[Q\left(\check{\Xi}_1 + K\right)\right] \text{ für alle } K \in \mathcal{K}^d.$$

Satz 20: Es sei Ξ ein Poisson-Boole-Modell mit einem Poisson-Punktprozess Ξ_0 in ZAM-Form mit den Parametern $\lambda > 0$ und atomlosem Q . Dann gilt für das zugehörige Kapazitätsfunktional:

$$T_{\Xi}(K) = 1 - \exp\left(-\lambda \cdot E\left[Q\left(\check{\Xi}_1 + K\right)\right]\right) \text{ für alle } K \in \mathcal{K}^d.$$

Satz 21: Es sei Ξ ein Poisson-Boole-Modell mit einem homogenen Poisson-Prozess Ξ_0 in ZAM-Form mit Parameter $\lambda > 0$. Dann ist Ξ stationär, und es gilt:

$$T_{\Xi}(K) = 1 - \exp\left(-\lambda \cdot E\left[m^d(\check{\Xi}_1 + K)\right]\right) = 1 - \exp\left(-\lambda \cdot E\left[m^d(\Xi_1 + \check{K})\right]\right)$$

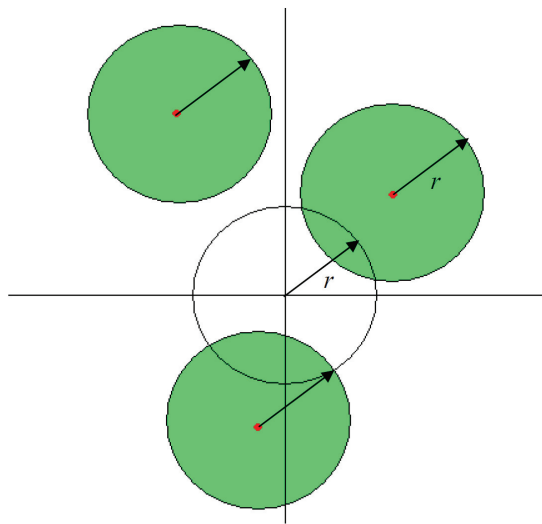
für alle $K \in \mathcal{K}^d$. Für den Bedeckungsgrad p_{Ξ} ergibt sich hieraus noch

$$p_{\Xi} = T_{\Xi}(\{\mathbf{0}\}) = 1 - \exp\left(-\lambda \cdot E\left[m^d(\Xi_1)\right]\right).$$

Bemerkung: Im speziellen Fall, dass die Körner $\{\Xi_n\}_{n \in \mathbb{N}}$ alle deterministische abgeschlossene Kugeln $K(r)$ um den Nullpunkt mit Radius $r > 0$ sind, folgt die Formel für den Bedeckungsgrad in Satz 21 auch direkt aus der Überlegung

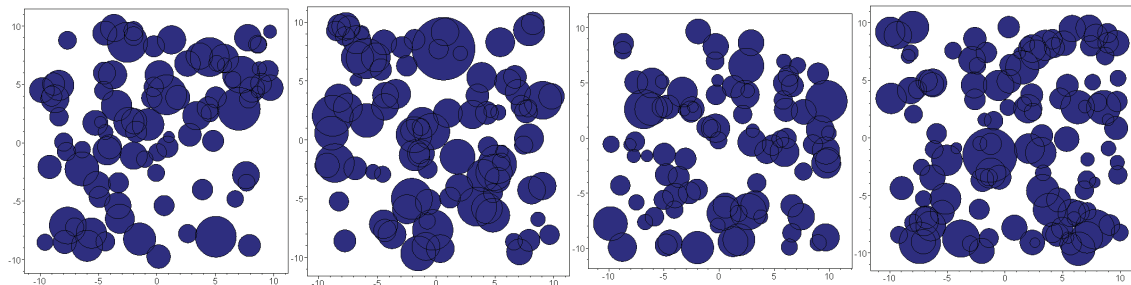
$$1 - p_{\Xi} = P(\mathbf{0} \notin \Xi) = P(\xi(K(r)) = 0) = \exp\left(-\lambda \cdot m^d(K(r))\right),$$

wobei ξ einen homogenen Poisson-Punktprozess mit Parameter $\lambda > 0$ bezeichne.



Veranschaulichung für $d = 2$; die roten Punkte repräsentieren ξ
dargestellt ist der Fall $\{\mathbf{0} \notin \Xi\} = \{\xi(K(r)) = 0\}$

Die nachfolgenden Graphiken zeigen vier Simulationen von *zweidimensionalen* homogenen Poisson-Boole-Modellen mit Parameter $\lambda = 0,25$ und zufälligen abgeschlossenen Kreisen als Körnern. Als Verteilung für den Radius R wurde eine Lognormal-Verteilung mit Parametern $\mu = 0$ und $\sigma^2 = 0,1$ gewählt. Die Fenster für den zu Grunde liegenden homogenen Poisson-Prozess wurden auf den Bereich $[-10,10] \times [-10,10]$ beschränkt.

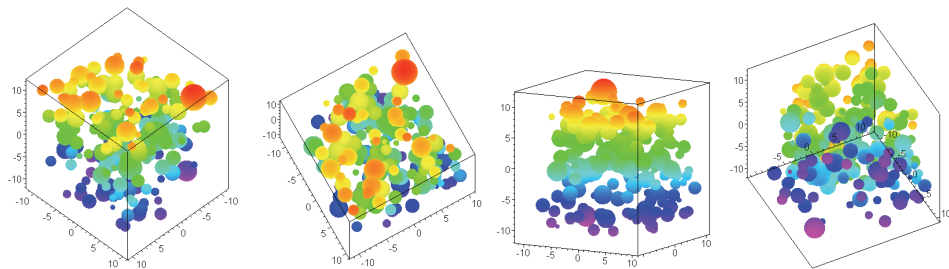


Für den Überdeckungsgrad p_{Ξ} erhält man hier

$$p_{\Xi} = T_{\Xi}(\{\mathbf{0}\}) = 1 - \exp\left(-\lambda \cdot E\left[m^d(\Xi_1)\right]\right) = 1 - \exp\left(-\frac{1}{4}E(R^2)\pi\right) = 0,6168\dots$$

wegen $E(R^2) = e^{2(\mu+\sigma^2)} = e^{0,2} = 1,2214\dots$

Die nachfolgenden Graphiken zeigen eine Simulation eines *dreidimensionalen* homogenen Poisson-Boole-Modells mit Parameter $\lambda = 0,25$ und zufälligen abgeschlossenen Kugeln als Körnern aus vier unterschiedlichen Perspektiven. Als Verteilung für den Radius R wurde wieder eine Lognormal-Verteilung mit Parametern $\mu = 0$ und $\sigma^2 = 0,1$ gewählt. Die Fenster für den zu Grunde liegenden homogenen Poisson-Prozess wurden auf den Bereich $[-10,10] \times [-10,10] \times [-10,10]$ beschränkt.



Für den Überdeckungsgrad p_{Ξ} erhält man hier

$$p_{\Xi} = T_{\Xi}(\{\mathbf{0}\}) = 1 - \exp\left(-\lambda \cdot E\left[m^d(\Xi_1)\right]\right) = 1 - \exp\left(-E(R^3) \frac{\pi}{3}\right) = 0,8064... \text{ wegen}$$

$$E(R^3) = e^{3\mu + \frac{9}{2}\sigma^2} = e^{0,45} = 1,5683...$$

Literatur:

C.D. Bonham (1989): *Measurements for Terrestrial Vegetation*. Wiley, N.Y.

I. Molchanov (2005): *Theory of Random Sets*. Springer, London.

M. Mühlenberg (1993) : *Freilandökologie*. Quelle & Meyer, Heidelberg.

D. Stoyan, W.S. Kendall and J. Mecke (1987): *Stochastic Geometry and Its Applications*. Wiley, New York.

D. Stoyan und H. Stoyan (1992): *Fraktale – Formen – Punktfelder*. Methoden der Geometrie-Statistik. Akademie-Verlag Berlin.

VI. Berufung nach Hamburg (1995-2000)



Nachfolge Prof. Dr. Christian Hipp

(Versicherungsmathematik)

Lehraufgaben:

Versicherungsmathematik, Risikotheorie

Unterstützung der Arbeitsgruppe Stochastik

Mitwirkung an der Ausbildung von Lehramtstudenten

Projekt:

PZVA – Prognose Zukünftiger Versorgungs-Ausgaben (im Öffentlichen Dienst des Landes Hamburg)

Laufzeit: 1999 - 2009

The zero utility principle for scale families of risk distributions

Dietmar Pfeifer and Bernd Heidergott (Hamburg)

1. Introduction

A premium calculation principle is a general rule that assigns a premium H to a given risk X . Intuitively, H is what the insurance carrier charges (apart from an expense allowance) for taking over the risk X . As a policy of premium calculation, the insurance carrier might base his decision on a utility function u , which is assumed to be a monotonically increasing, differentiable and concave function on $\mathbf{R}$. The principle requires that the utility $u(0)$ before assuming responsibility for the claims be equal to the expected utility $E[u(H - X)]$ after taking over such responsibility in exchange for the premium H . Mathematically X is a random variable, and H depends on X through its distribution function and the utility function, as well. In the case $u(0) = 0$ we obtain the zero utility principle given as

$$E[u(H - X)] = 0.$$

So far this principle has been mainly of theoretical interest. If we take the utility function as exponential, then the zero utility principle equals the exponential principle and has an explicit solution. For twice differentiable non-exponential utility functions we have, moreover, the variance principle as a first approximation, see Heilmann (1987). Since this is the only approach known so far to deal with general utility functions, many authors define a utility function as being twice differentiable, see Govaerts and de Vylder (1979), Gerber (1985), Reich (1986), Kremer (1986) or Heilmann and Schröter (1987). In this paper, we show that for certain classes of not necessarily twice differentiable utility functions – which will be characterized as being *scale invariant* – and for certain classes of risk distributions where the expectation plays the role of the “natural” parameter we find explicit representations for the zero utility premium. Further, comparisons with other premium calculation principles are given.

A simple method to estimate parametric claim size distributions from grouped data

Joachim Brix and Dietmar Pfeifer (Hamburg)

1. Introduction

Reinsurance brokers and companies are frequently faced with the problem that claim size data obtained for actuarial analysis are usually processed in grouped form, and mostly even only available for the larger claim size layers. The statistical estimation of appropriate claim size distributions for the total portfolio – say with the aim of forecasting probable maximum losses as upper quantiles of that distribution – is then a difficult task which cannot be performed with the usual elementary statistical tools, although some useful recommendations can be found in the literature such as moment and modified maximum likelihood methods (cf. e.g. [3], section 4.3.A), modified minimum-distance methods (cf. e.g. [3], section 3.3 and section 4.3.A), linear regression methods in the particular case of Pareto distributions (cf. [4], section 3.3.3(c)), or particular methods in the case of lognormal distributions (cf. [2], section 1.4.3). For a similar discussion with respect to extreme value distributions, see [1].

In this paper, we want to show that such an analysis can, however, be more simply performed for most parametric classes of claim size distributions using certain non-linear regression techniques for densities that are nowadays implemented in several statistical software packages, such as STATISTICA. The powerfulness of this method will be demonstrated using both artificial as well as real data from fire, windstorm and health care losses.

Blätter der DGVM XXII (1996)

Blätter der DGVM XXIV (2000)



```

DOSBox 0.74, Cpu speed: 3000 cycles, frameskip: 0, Program: LEBEN
LEBEN U1.10

Ein Programm zur Berechnung von Versicherungsprämien und zur graphischen
Darstellung von Sterbetafeln, Lebenserwartungen und Deckungskapitalien

Autor: Prof. Dr. Dietmar Pfeifer, Fachbereich Mathematik
Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg
e-mail: dietmar.pfeifer@uni-oldenburg.de

Systemanforderungen: 80386+Coprocessor oder höher; Standard-VGA

Kurzanleitung: F6: Weiter mit beliebiger Taste

```

```

DOSBox 0.74, Cpu speed: 3000 cycles, frameskip: 0, Program: LEBEN
File mit Sterbewahrscheinlichkeiten (ohne .qx: 'q' für quit)? stm86
Zinssatz i (in %)? 1.5

Alter x? 26
Laufzeit w? 50 [ maximal omega-x = 75 ]

a_x = .497406 [Todesfall, lebenslange Deckung]
p_x = .014626

a1_xm = .298253 [Todesfall, temporäre Deckung]
p1_xm = .009006

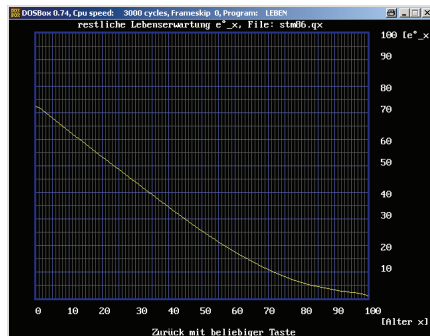
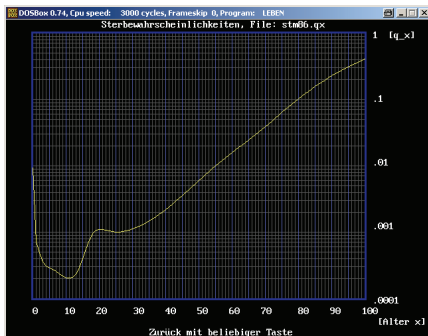
a_xm1 = .233456 [Erlebensfall]
p_xm1 = .007244

a_xm = .523709 [gemischte Versicherung]
p_xm = .01625

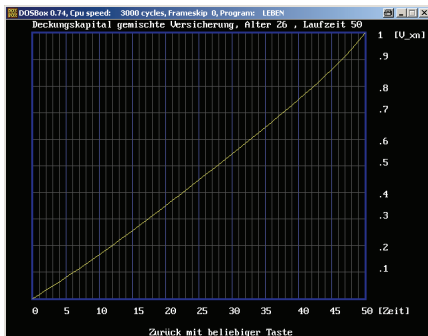
Optionen:
's': Speichern der Ergebnisse in Datei [File].[Nummer]
'c': neu, 'x': neues Alter, 'n': neue Laufzeit, 'l': neuer Zinssatz

Graphik:
F1: Sterbewahrscheinlichkeiten, F2: restliche Lebenserwartung
F3: Deckungskapital gemischte, F4: temporäre Todesfallversicherung
F5: Tabellen, F6: Kurzanleitung

```



Eigene Software zur Lebensversicherungsmathematik



DOSBox 0.74, Cpu speed: 3000 cycles, frameskip: 0, Program: LEBEN

Sterbewahrscheinlichkeiten, tabellarisch

q_0 = .0093	q_1 = .0097	q_2 = .0095	q_3 = .0094	q_4 = .0093
q_5 = .0093	q_6 = .0093	q_7 = .0093	q_8 = .0092	q_9 = .0093
q_10 = .0092	q_11 = .0092	q_12 = .0092	q_13 = .0092	q_14 = .0093
q_15 = .0094	q_16 = .0096	q_17 = .0098	q_18 = .0091	q_19 = .0091
q_20 = .0091	q_21 = .0091	q_22 = .0091	q_23 = .0091	q_24 = .0091
q_25 = .0091	q_26 = .0091	q_27 = .0091	q_28 = .0091	q_29 = .0091
q_30 = .0091	q_31 = .0092	q_32 = .0092	q_33 = .0093	q_34 = .0094
q_35 = .0095	q_36 = .0096	q_37 = .0097	q_38 = .0098	q_39 = .0092
q_40 = .0092	q_41 = .0094	q_42 = .0097	q_43 = .0099	q_44 = .0092
q_45 = .0096	q_46 = .0094	q_47 = .0094	q_48 = .0099	q_49 = .0094
q_50 = .009	q_51 = .0097	q_52 = .0094	q_53 = .0092	q_54 = .0091
q_55 = .01	q_56 = .011	q_57 = .0121	q_58 = .0133	q_59 = .0145
q_60 = .0159	q_61 = .0173	q_62 = .0189	q_63 = .0206	q_64 = .0224
q_65 = .0245	q_66 = .0267	q_67 = .0292	q_68 = .0321	q_69 = .0352
q_70 = .0388	q_71 = .0429	q_72 = .0473	q_73 = .0523	q_74 = .0578
q_75 = .0638	q_76 = .0704	q_77 = .0776	q_78 = .0853	q_79 = .0937
q_80 = .1026	q_81 = .1122	q_82 = .1224	q_83 = .1332	q_84 = .1447
q_85 = .1568	q_86 = .1695	q_87 = .183	q_88 = .197	q_89 = .2117
q_90 = .2269	q_91 = .2429	q_92 = .2592	q_93 = .2761	q_94 = .2934
q_95 = .3112	q_96 = .3293	q_97 = .3478	q_98 = .3665	q_99 = .3853
q_100 = .4043	q_101 = 1			

Weiter mit beliebiger Taste

DOSBox 0.74, Cpu speed: 3000 cycles, frameskip: 0, Program: LEBEN

restliche Lebenserwartung, tabellarisch

e^*_0 = 72.21	e^*_1 = 71.88	e^*_2 = 70.93	e^*_3 = 69.97	e^*_4 = 68.99
e^*_5 = 68.01	e^*_6 = 67.83	e^*_7 = 66.65	e^*_8 = 65.67	e^*_9 = 64.68
e^*_10 = 63.1	e^*_11 = 62.11	e^*_12 = 61.12	e^*_13 = 60.14	e^*_14 = 59.15
e^*_15 = 58.17	e^*_16 = 57.19	e^*_17 = 56.22	e^*_18 = 55.26	e^*_19 = 54.32
e^*_20 = 53.37	e^*_21 = 52.43	e^*_22 = 51.49	e^*_23 = 50.54	e^*_24 = 49.6
e^*_25 = 48.65	e^*_26 = 47.69	e^*_27 = 46.74	e^*_28 = 45.79	e^*_29 = 44.83
e^*_30 = 43.88	e^*_31 = 42.93	e^*_32 = 41.98	e^*_33 = 41.03	e^*_34 = 40.08
e^*_35 = 39.14	e^*_36 = 38.2	e^*_37 = 37.26	e^*_38 = 36.32	e^*_39 = 35.38
e^*_40 = 34.45	e^*_41 = 33.53	e^*_42 = 32.61	e^*_43 = 31.69	e^*_44 = 30.79
e^*_45 = 29.88	e^*_46 = 28.99	e^*_47 = 28.1	e^*_48 = 27.23	e^*_49 = 26.36
e^*_50 = 25.5	e^*_51 = 24.65	e^*_52 = 23.81	e^*_53 = 22.98	e^*_54 = 22.17
e^*_55 = 21.37	e^*_56 = 20.58	e^*_57 = 19.8	e^*_58 = 19.04	e^*_59 = 18.29
e^*_60 = 17.55	e^*_61 = 16.83	e^*_62 = 16.11	e^*_63 = 15.41	e^*_64 = 14.73
e^*_65 = 14.65	e^*_66 = 13.99	e^*_67 = 13.25	e^*_68 = 12.11	e^*_69 = 11.5
e^*_70 = 10.9	e^*_71 = 10.32	e^*_72 = 9.76	e^*_73 = 9.22	e^*_74 = 8.7
e^*_75 = 8.2	e^*_76 = 7.73	e^*_77 = 7.28	e^*_78 = 6.85	e^*_79 = 6.44
e^*_80 = 6.65	e^*_81 = 5.69	e^*_82 = 5.34	e^*_83 = 5.62	e^*_84 = 4.71
e^*_85 = 4.43	e^*_86 = 4.16	e^*_87 = 3.9	e^*_88 = 3.67	e^*_89 = 3.44
e^*_90 = 3.23	e^*_91 = 2.94	e^*_92 = 2.85	e^*_93 = 2.67	e^*_94 = 2.5
e^*_95 = 2.33	e^*_96 = 2.16	e^*_97 = 1.97	e^*_98 = 1.75	e^*_99 = 1.48
e^*_100 = 1.1				

Weiter mit beliebiger Taste

Eigene Software zur Lebensversicherungsmathematik

Mitteilung des Senats an die Bürgerschaft

Stellungnahme des Senats zu dem Ersuchen der Bürgerschaft vom 13./14./15. Dezember 1999 (Drucksache 16/3543) – Versorgungsausgaben (I) –

Gegenstand des Ersuchens

Die Bürgerschaft hat am 13. Dezember 1999 das folgende Ersuchen Drucksache 16/3543 – Versorgungsausgaben (I) – beschlossen:

- „Die Bürgerschaft ersucht den Senat, mit Vorlage des Haushaltsplan-Entwurfs 2001 zu berichten,
1. wie sich die aktuellen Prognoseberechnungen für die Versorgungsausgaben der Freien und Hansestadt Hamburg darstellen,
 2. welche quantitativen Auswirkungen das zum 1. Januar 1999 in Kraft getretene Versorgungsreformgesetz auf die hamburgischen Versorgungsausgaben hat.“

Gliederung der Stellungnahme des Senats

A.

Zusammenfassung der Prognoseergebnisse

1. Gesamtentwicklung
2. Entwicklung im Beamtenversorgungsbereich
3. Entwicklung im Zusatzversorgungsbereich der Arbeitnehmerinnen und Arbeitnehmer

B.

Zu Ziffer 1 des Ersuchens:

Aktuelle Prognoseberechnungen für die Versorgungsausgaben der Freien und Hansestadt Hamburg

1. Forschungsprojekt von Prof. Dr. Dietmar Pfeifer: „Entwicklung eines Prognoseinstruments zur Berechnung zukünftiger Versorgungsausgaben der Freien und Hansestadt Hamburg“
2. Beschreibung des Prognoseinstruments (Grundmodell)
 - 2.1 Zugrunde liegender Datenbestand
 - 2.2 Mathematische Rechenmethode und angenommene Prämissen

3. Prognose zukünftiger Versorgungsausgaben und der erwarteten Zahl der Anspruchsberechtigten (Gutachten Prof. Dr. Pfeifer)

- 3.1 Entwicklung der erwarteten Berechtigtenzahlen
- 3.2 Ergebnisse der Prognoseberechnungen zu den erwarteten Kostenentwicklungen
4. Zusammenfassung und Bewertung der Prognoseergebnisse
5. Bewertung von Möglichkeiten zur Reduzierung des Kostenanstiegs
- 5.1 Beamtenversorgungsbereich
- 5.2 Zusatzversorgungsbereich der Arbeitnehmerinnen und Arbeitnehmer

C.

Zu Ziffer 2 des Ersuchens:
Quantitative Auswirkungen des zum 1. Januar 1999 in Kraft getretenen Versorgungsreformgesetzes auf die hamburgischen Versorgungsausgaben

D.

Schlussbewertung
Anlage (Mathematische Rechenformeln)

Anlage
zur Bürgerschaftsdrucksache 16/4727

Mathematische Rechenformeln (Prof. Dr. Pfeifer)

Für eine Person mit Lebensalter x ergeben sich die erwarteten Kosten EK_x zu einem bestimmten Prognosezeitpunkt aus

$$EK_x = E(K_x) = \sum_{z \in Z} E(K_x(z)) P(Z = z),$$

wobei $E()$ den Erwartungswertoperator bezeichnet, z die Menge der möglichen Zustandsentwicklungen bis zum Prognosezeitpunkt, $P(Z=z)$ die Wahrscheinlichkeit dafür, dass die Person bis zum Prognosezeitpunkt die Zustandsentwicklung $z \in Z$ realisiert hat, K_x die (zufallsabhängigen) Gesamtkosten und $K_x(z)$ die Kosten, die die Person zum Prognosezeitpunkt verursacht, wenn sie die Zustandsentwicklung $z \in Z$ realisiert hat.

Die Wahrscheinlichkeiten $P(Z=z)$ lassen sich durch die Ausscheide- bzw. Invalidisierungswahrscheinlichkeiten der genannten Tabellen ausdrücken.

So gilt im Passivbestand:

für eine x -jährige Person, die unverheiratet ist, den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpx hat:

$$P(Z = z) = \prod_{i=x}^{zpx} (1 - q_i)$$

für eine x -jährige Person, die mit einer y -jährigen Person verheiratet ist, vor dem Prognosezeitpunkt zur Zeit t verstirbt und dessen Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t \prod_{j=y}^{zpy} (1 - \bar{q}_j) \prod_{i=x}^{t-1} (1 - q_i)$$

Hierbei bezeichnen q_i bzw. $\bar{q}_j$ die komplementär-geschlechtsspezifischen Sterbewahrscheinlichkeiten für eine i -jährige Person bzw. dessen Ehepartner.

Im Aktivbestand gilt entsprechend (unter Berücksichtigung von Invalidisierungen):

Für eine x -jährige aktive Person, die den Prognosezeitpunkt als gesunder Passiver erlebt und zum Prognosezeitpunkt das Alter zpx hat:

$$P(Z = z) = \prod_{i=x}^{zpx} (1 - q_i)(1 - q_i^{inv})$$

für eine x -jährige aktive verheiratete Person, die den Prognosezeitpunkt als gesunder Passiver nicht erlebt, zur Zeit t verstirbt,

dessen jetzt y -jährige(r) Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t \prod_{j=y}^{zpy} (1 - \bar{q}_j) \prod_{i=x}^{t-1} (1 - q_i)(1 - q_i^{inv})$$

für eine x -jährige aktive Person, die vor dem Prognosezeitpunkt und vor der Altersgrenze zur Zeit t Invalide wird und den Prognosezeitpunkt mit dem Alter zpx erlebt:

$$P(Z = z) = q_t^{inv} \prod_{i=x}^t (1 - q_i) \prod_{j=t}^{t-1} (1 - q_j^{inv}) \prod_{k=t+1}^{zpx} (1 - q_k^{inv})$$

für eine x -jährige aktive verheiratete Person, die vor dem Prognosezeitpunkt und vor der Altersgrenze zur Zeit s Invalide wird, zur Zeit t verstirbt und dessen jetzt y -jährige(r) Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t^{inv} q_s \prod_{i=x}^{t-1} (1 - q_i) \prod_{k=s+1}^{t-1} (1 - q_k^{inv}) \prod_{j=y}^{zpy} (1 - \bar{q}_j)$$

Hierbei bezeichnen q_i^{inv} bzw. $q_i^{inv,d}$ die Wahrscheinlichkeiten, dass die Person im Jahr i Invalide wird bzw. als Invalide(r) im Jahr i stirbt.

Für die Berechnung der erwarteten Berechtigten gilt:

Die erwarteten Anzahlen $E(A)$ bzw. $E(W)$ der originären Berechtigten bzw. deren Hinterbliebenen zum Prognosezeitpunkt lassen sich entsprechend berechnen. Hier gilt:

$$E(A) = \sum_{n \in N} \sum_{z \in Z_A(A)} P(Z = z), \quad E(W) = \sum_{n \in N} \sum_{z \in Z_W(W)} P(Z = z),$$

wobei N die Menge aller zu berücksichtigenden Personen und $z_{A(A)}$ bzw. $z_{W(W)}$ die Menge aller Zustandsentwicklungen bezeichnet, die im Prognosezeitpunkt für Person n zum Status als persönlich Berechtigte(r) bzw. Witwe(r) führen.

Zusätzlich zu den genannten Prognosewerten geben die Programme der Gruppe PZVA auch die zugehörigen Standardabweichungen für die Versorgungskosten bzw. die Empfängerzahlen aus. Diese stellen ein Maß für die aufgrund der biometrischen Rechnungsgrundlagen vorhandenen – unvermeidbaren – statistischen Schwankungen in den Prognosewerten dar.

Erläuterung des Forschungsvorhabens

Prognose **z**ukünftiger **V**ersorgung**a**usgaben (**PZVA**)

im Auftrag des Senats der Freien und Hansestadt Hamburg

1. Allgemeine Grundlagen und Ziele

- Prognose der **erwarteten Versorgungskosten** sowie der erwarteten **Zahl originärer Versorgungsempfänger** bzw. deren Hinterbliebenen (Witwen/ Witwer, Invalidisierungen) für die Jahre ab 2000 auf versicherungsmathematischer Basis (Prognosezeitraum: jeweils 20 Jahre)
- Verwendung **personenbezogener, anonymisierter Daten** der aktiven und passiven Angehörigen des Öffentlichen Dienstes der Freien und Hansestadt Hamburg (**Individualmethode**)
- **Differenzierungsmöglichkeiten** z.B. zwischen Beamten und Angestellten / Arbeitern einerseits und ausgewählten Einzelplänen / Kapiteln / Berufsgruppen etc. andererseits
- Entwicklung **bedarfsspezifischer Auswerte-Software**
- **Option:** Berücksichtigung spezifischer Parameter zur Abschätzung der Auswirkung von **demographischen Entwicklungen** und **politischen Maßnahmen zur Kostendämpfung** (z.B. Berücksichtigung von Alterstrends, Wahlmöglichkeit von Antragsaltersgrenzen, Änderungen von Versorgungssätzen / Lebensarbeitszeit, etc.)

2. Mathematische Rechnungsgrundlagen

Für eine im Bestand befindliche Person mit **Lebensalter** x ergibt sich für die **erwarteten Kosten** EK_x zu einem bestimmten Prognosezeitpunkt

$$EK_x = E(K_x) = \sum_{z \in \mathfrak{Z}} K_x(z) P(Z = z)$$

wobei

$E(\dots)$ den **Erwartungswertoperator** bezeichnet,

$\mathfrak{Z}$ die Menge der möglichen **Zustandsentwicklungen** bis zum Prognosezeitpunkt, Z die **Zufallsvariable** "Zustandsentwicklung",

$P(Z=z)$ die **Wahrscheinlichkeit** dafür, daß die Person bis zum Prognosezeitpunkt die Zustandsentwicklung $z \in \mathfrak{Z}$ realisiert hat,

$K_x(z)$ die **Kosten**, die die Person zum Prognosezeitpunkt verursacht, wenn sie die Zustandsentwicklung $z \in \mathfrak{Z}$ realisiert hat und

K_x die hieraus resultierenden zufallsabhängigen **Gesamtkosten**.

Mögliche Zustandsentwicklungen $z \in \mathfrak{Z}$ und deren Wahrscheinlichkeiten im **Passiv-Bereich**:

für eine x -jährige Person, die unverheiratet ist, den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpx hat:

$$P(Z = z) = \prod_{i=x}^{zpx} (1 - q_i)$$

für eine x -jährige Person, die mit einer y -jährigen Person verheiratet ist, vor dem Prognosezeitpunkt zur Zeit t verstirbt und dessen Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t \prod_{j=y}^{zpy} (1 - \bar{q}_j) \prod_{i=x}^{t-1} (1 - q_i)$$

Hierbei bezeichnen q_i bzw. $\bar{q}_i$ die komplementär-geschlechtsspezifischen Sterbewahrscheinlichkeiten für eine i -jährige Person bzw. dessen Ehepartner.

Mögliche Zustandsentwicklungen $z \in \mathfrak{Z}$ und deren Wahrscheinlichkeiten im **Aktiv-Bereich (I)**:

für eine x -jährige aktive Person, die den Prognosezeitpunkt als gesunder Passiver erlebt und zum Prognosezeitpunkt das Alter zpx hat:

$$P(Z = z) = \prod_{i=x}^{zpx} (1 - q_i)(1 - q_i^{inv})$$

für eine x -jährige aktive verheiratete Person, die den Prognosezeitpunkt als gesunder Passiver nicht erlebt, zur Zeit t (vor oder nach Altersgrenze, 2 Fälle) verstirbt, dessen jetzt y -jährige(r) Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t \prod_{j=y}^{zpy} (1 - \bar{q}_j) \prod_{i=x}^{t-1} (1 - q_i)(1 - q_i^{inv})$$

Hierbei bezeichnen q_i^{inv} die geschlechtsspezifischen Invalidisierungswahrscheinlichkeiten für eine i -jährige Person bzw. dessen Ehepartner.

Mögliche Zustandsentwicklungen $z \in \mathfrak{Z}$ und deren Wahrscheinlichkeiten im **Aktiv-Bereich (II)**:

für eine x -jährige aktive Person, die vor dem Prognosezeitpunkt und vor der Altersgrenze zur Zeit t Invalide wird und den Prognosezeitpunkt mit dem Alter zpx erlebt:

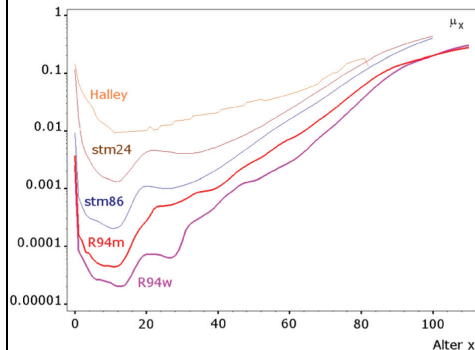
$$P(Z = z) = q_t^{inv} \prod_{i=x}^t (1 - q_i) \prod_{j=x}^{t-1} (1 - q_j^{inv}) \prod_{k=t+1}^{zpx} (1 - q_k^{inv,t})$$

für eine x -jährige aktive verheiratete Person, die vor dem Prognosezeitpunkt und vor der Altersgrenze zur Zeit s Invalide wird, zur Zeit t verstirbt und dessen jetzt y -jährige(r) Witwe(r) den Prognosezeitpunkt erlebt und zum Prognosezeitpunkt das Alter zpy hat:

$$P(Z = z) = q_t^{inv,s} q_s^{inv} \prod_{i=x}^{s-1} (1 - q_i) (1 - q_i^{inv}) \prod_{k=s+1}^{t-1} (1 - q_k^{inv,s}) \prod_{j=y}^{zpy} (1 - \bar{q}_j)$$

Hierbei bezeichnet $q_k^{inv,t}$ die Wahrscheinlichkeit, dass die Person im Jahr t Invalide wird und dann im Jahr k [als Invalide(r)] stirbt.

3. Demographie: Sterbetafeln



jährliche Sterbewahrscheinlichkeiten

Sterbetafel Halley (Breslau, 1696)

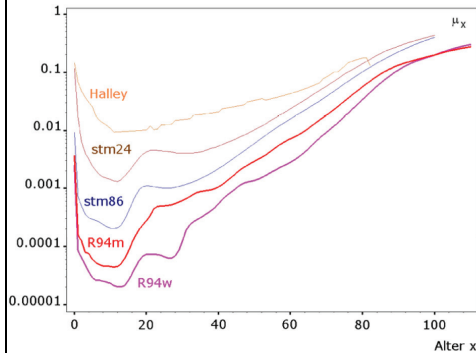
Sterbetafel männlich (1924)

Sterbetafel männlich (1986)

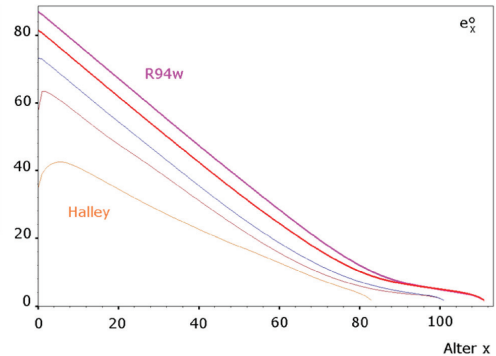
DAV-Sterbetafeln Typ R

männlich und weiblich (1994)

3. Demographie: Sterbetafeln



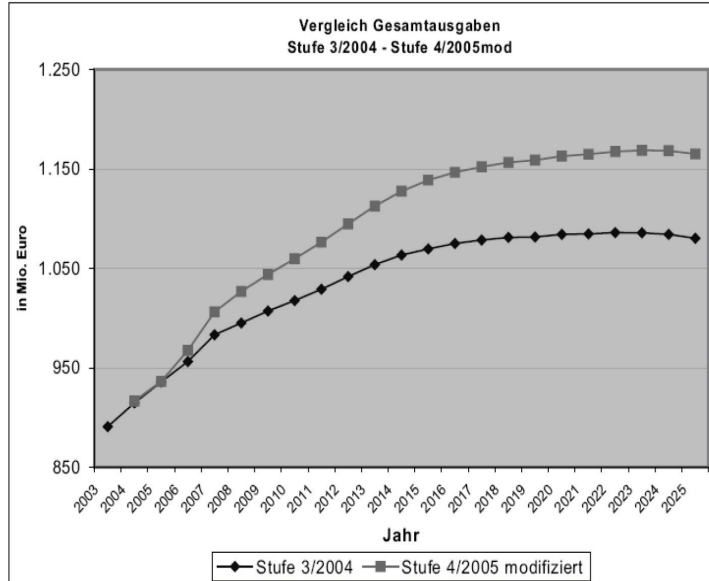
jährliche Sterbewahrscheinlichkeiten

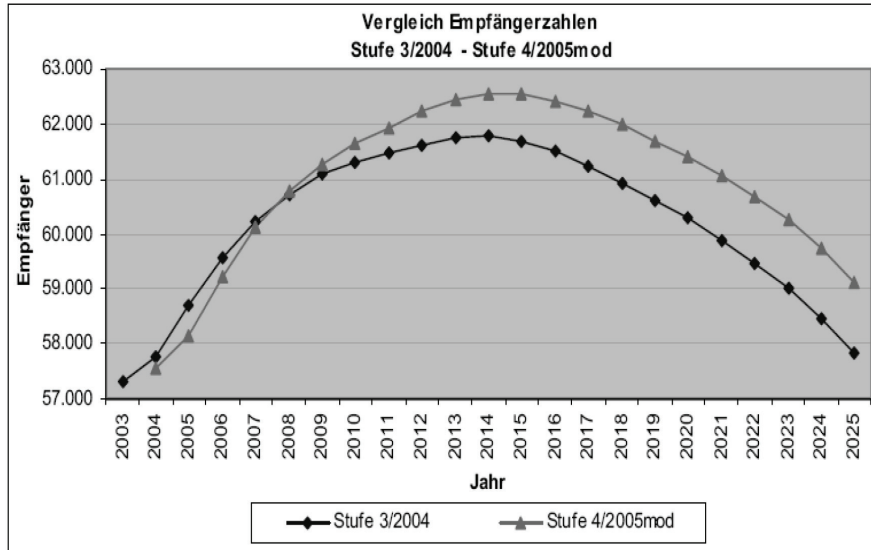


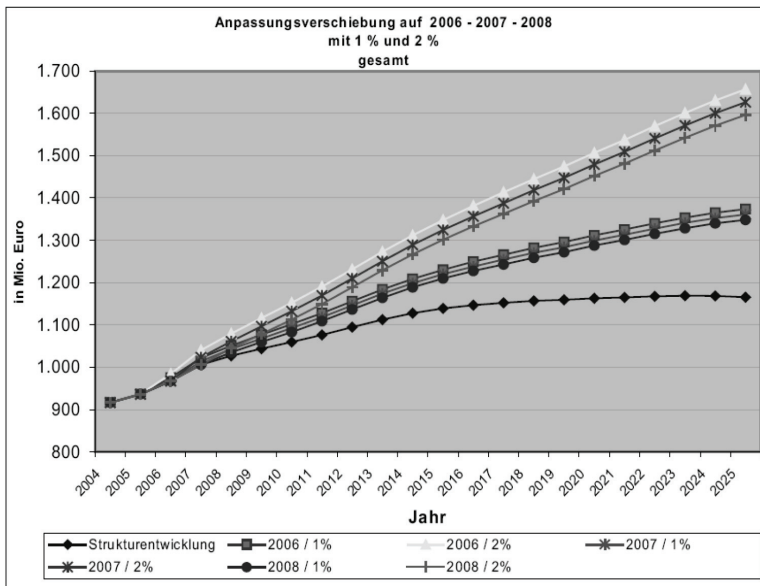
restliche Lebenserwartung

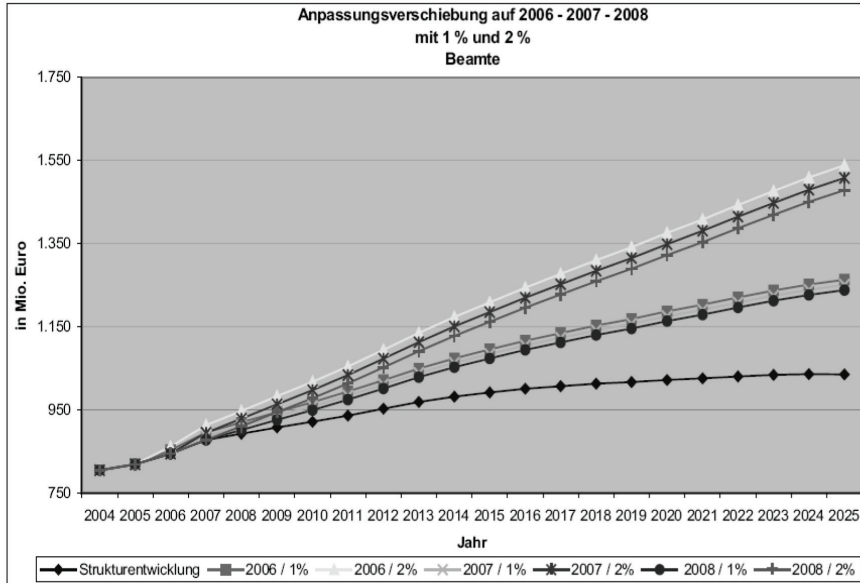
4. Beispielrechnungen:

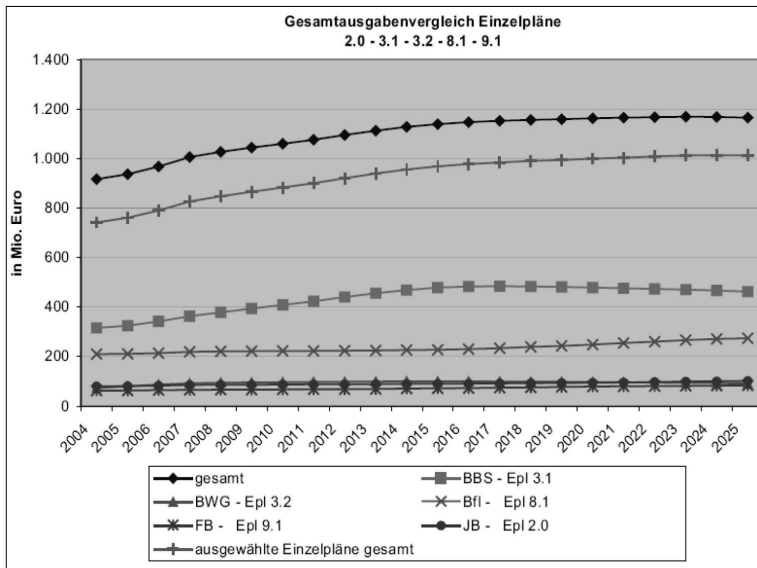
- Strukturentwicklungsvergleich Stufe3 /Stufe 4 [18/3741]
- Anpassungsverschiebungen [18/3741]
- Einzelplanvergleich [18/3741]
- Strukturentwicklungsvergleich Stufe4 /Stufe 5 [18/4736]
- Anpassungsverschiebungen [18/4736]
- Einzelplanvergleich [18/4736]
- Vergleich bisheriger Prognosen

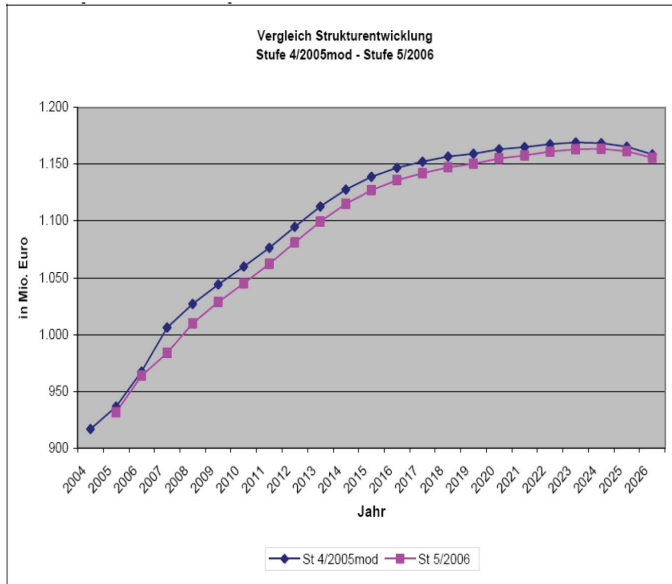


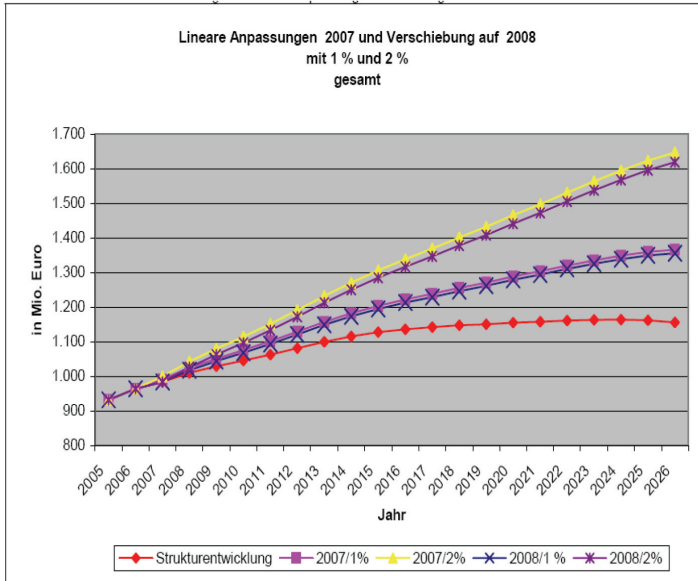


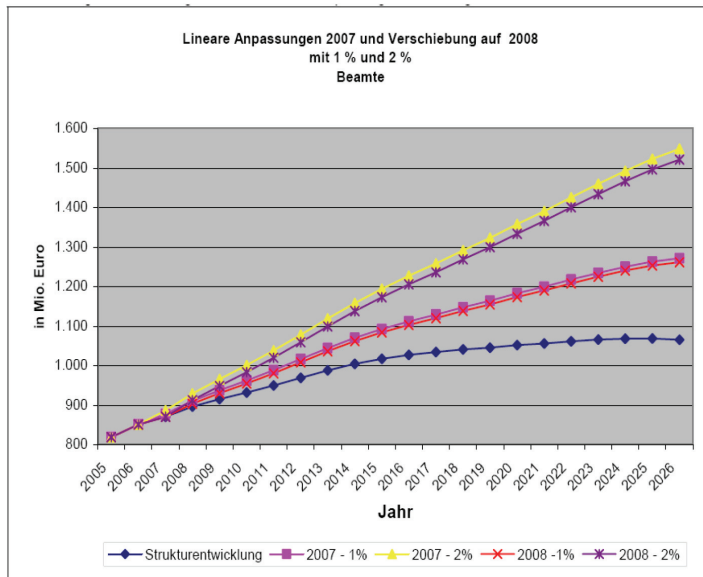


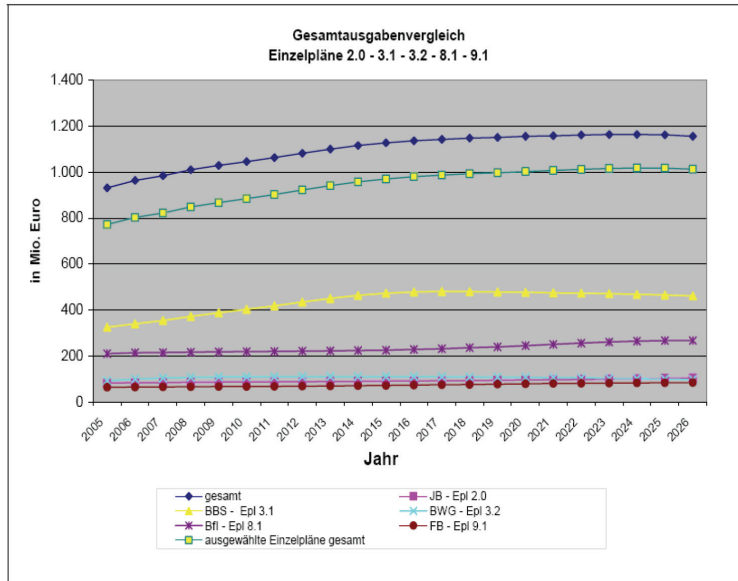


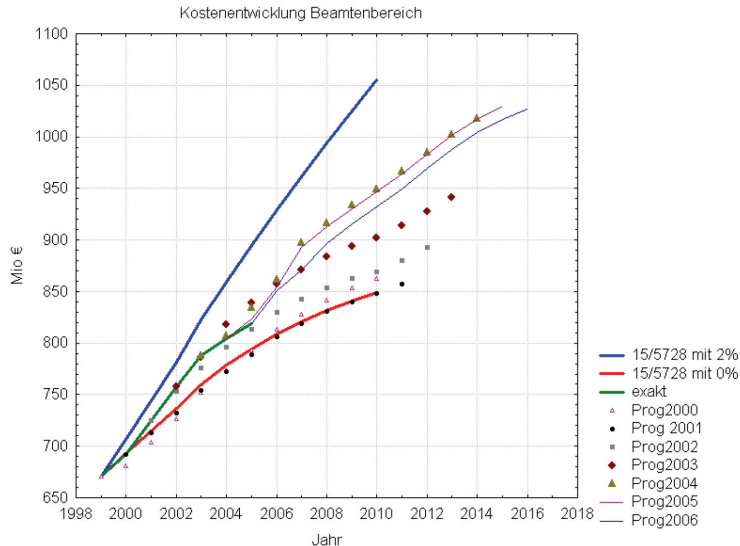


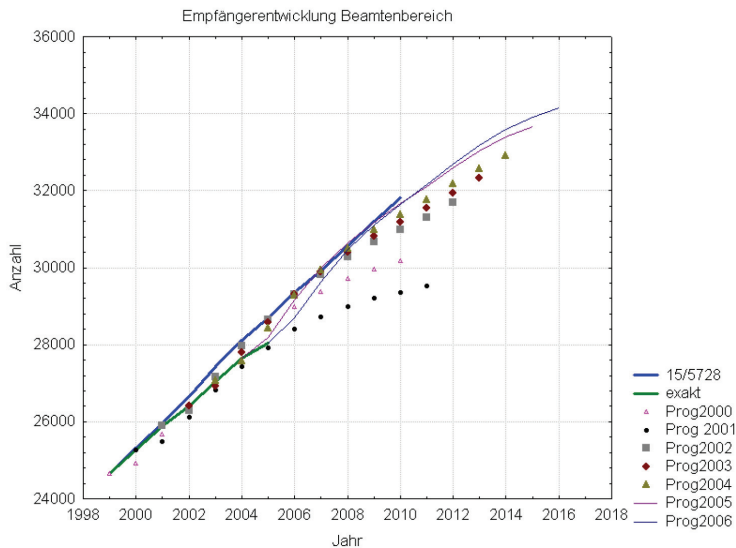












PSUEDO-POISSON APPROXIMATION FOR MARKOV CHAINS

K. A. Borovkov

Statistics Department, University of Melbourne, Australia

D. Pfeifer

Fachbereich Mathematik, Universität Hamburg, Germany

We consider the problem of approximating the distribution of a Markov chain with 'rare' transitions in an arbitrary state space by that of the corresponding pseudo-Poisson process. Sharp estimates for both first- and second-order approximations are obtained. The remarkable fact is that the convergence rate in this setup can be better than that in the ordinary Poisson theorem: the ergodicity of the embedded 'routing' Markov chain improves essentially the degree of approximation. This is of particular importance if the accumulated transition intensity of the chain is of a moderate size so that neither the usual estimates from the Poisson theorem nor the existence of a stationary distribution alone provide good approximation results. On the other hand, the estimates also improve the known results in the ordinary Poisson theorem.

AMS 1991 Subject Classifications: 60F05, 60J10, 60J25.

Poisson approximation * Markov chain * pseudo-Poisson process

ON ASYMPTOTIC BEHAVIOR OF WEIGHTED SAMPLE QUANTILES

K. BOROVKOV¹, H. DEHLING² and D. PFEIFER³

¹ Dept. of Statistics, University of Melbourne
Parkville 3052, Australia

² Department of Mathematics, University of Groningen
P. O. Box 800, 9700 Groningen, The Netherlands

³ Institut für Mathematische Stochastik, Universität Hamburg
Bundesstr. 55, 20146 Hamburg, Germany

We study the limit behavior of L -statistics with the 'delta type' weight sequences, of which the O -statistics and usual kernel quantile estimators are special cases. Deviations of these statistics from the sample quantiles are shown to have (after proper normalization) Gaussian limit distributions. We evaluate the main term of the asymptotics of the covariance of this deviation with the corresponding sample quantile. This gives also the asymptotic relative deficiency of sample quantiles w.r.t. our L -statistics.

Key words: sample quantile, quantile estimator, L -statistic, quantile process, relative efficiency, second order properties.

Stochastic Processes and
Their Applications 61 (1996)

Mathematical Methods of
Statistics 5 (1996)

ON IMPROVEMENTS OF THE ORDER OF APPROXIMATION IN THE POISSON LIMIT THEOREM

K. BOROVKOV,^{*} *University of Melbourne*

D. PFEIFER,^{**} *Universität Hamburg*

Abstract

In this paper we consider improvements in the rate of approximation for the distribution of sums of independent Bernoulli random variables via convolutions of Poisson measures with signed measures of specific type. As a special case, the distribution of the number of records in an i.i.d. sequence of length n is investigated. For this particular example, it is shown that the usual rate of Poisson approximation of order $O(1/\log n)$ can be lowered to $O(1/n^2)$. The general case is discussed in terms of operator semigroups.

Journal of Applied Probability 33 (1996)

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

ТОМ 45
ВЫПУСК

2

© 2000 г.

BOROVKOV K. A.*, PFEIFER D.**
**ESTIMATES FOR THE SYRACUSE PROBLEM
VIA A PROBABILISTIC MODEL**

В статье с помощью простой стохастической модели для «сиракузской проблемы» (известной так же как « $(3x+1)$ -проблема») получены оценки «среднего поведения» траекторий исходной детерминированной динамической системы. Использование этой модели оправдывается не только некоторым сходством между управляющими правилами систем, но и качественной оценкой скорости аппроксимации (теорема 2). Из модели мы выводим явные формулы для асимптотических плотностей некоторых множеств, характеризующих исходную последовательность. Также получены аппроксимации асимптотических распределений для «моментов остановки» (времени до поглощения в единственном известном цикле $\{1, 2\}$) исходной системы и приведены численные иллюстрации результатов.

Ключевые слова и фразы: «сиракузская задача», итерация дискретных функций, динамическая система, случайное блуждание.

1. Introduction. The Syracuse problem, also known as $3x+1$ problem, Collatz' problem, Kakutani's problem, Ulam's problem, and Hasse's algorithm, has been drawing attention of many mathematicians for more than 40 years. The problem belongs to the class of questions which can be explained to anybody, but which are very difficult to answer. For its history, details, references and prizes for its solution see [6] and [7].

The Syracuse problem is concerned with the behaviour of a simply defined discrete time dynamical system. In general, such a system is given by the relation

$$x_{n+1} = f(x_n), \quad n \geq 0, \quad f: \mathcal{X} \rightarrow \mathcal{X}, \quad (1)$$

which defines the sequence $\{x_n\}_{n \geq 0}$ in the phase space $\mathcal{X}$ of the system, starting with some initial point $x_0 \in \mathcal{X}$. Such systems arise in a very natural way in many fields of

* Department of Mathematics & Statistics, The University of Melbourne, Parkville 3052, Australia; e-mail: kostya@ms.unimelb.edu.au

** Institut für Mathematische Stochastik, Universität Hamburg, Bundesstraße 55, D-20146 Hamburg, Germany.

VII. Zweite Berufung nach Oldenburg (2000-2016)



Nachfolge Prof. Dr. Ortwin Emrich
(Angewandte Wahrscheinlichkeitstheorie)

Lehraufgaben:

Stochastik und verwandte Gebiete (Grundkurs Stochastik, Spieltheorie, Räumliche Statistik, Geostatistik, Risikotheorie, Monte-Carlo-Methoden, Stochastische Prozesse, Semigroups and Probability, Diskrete Stochastische Finanzmathematik, Quantitative Risk Management)

Unterstützung der Arbeitsgruppe Stochastik

Mitwirkung an der Ausbildung von Lehramtstudenten

Aufbau des neuen Schwerpunkts Versicherungs- und Finanzmathematik

Zur Mathematik derivativer Finanzinstrumente: Anregungen für den Stochastik-Unterricht

Dietmar Pfeifer, Carl von Ossietzky Universität Oldenburg

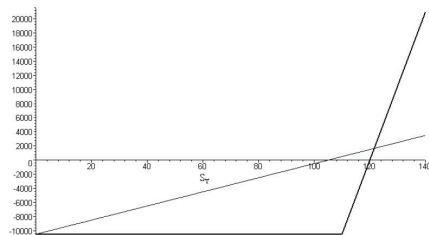
Zusammenfassung: Spätestens seit der Verleihung des Nobelpreises für Ökonomie im Jahr 1997 an Myron Scholes und Robert Merton ist die zusammen mit Fischer Black¹ seit 1972 entwickelte Theorie der mathematischen Bewertung von Finanzderivaten in aller Munde. Hiervon zeugen auch die mehr als ein Dutzend allein seit 1996 erschienenen Monographien und Handbücher zu diesem Thema. In diesem Artikel soll aufgezeigt werden, daß und wie elementare Grundlagen der wesentlich mit Methoden der Stochastik arbeitenden modernen Finanzmathematik in der Schule vermittelt werden können.

1. Einleitung

Es begann wie so oft bei bahnbrechenden wissenschaftlichen Neuerungen: die Arbeit mit der fundamentalen Formel zu einer "objektiven" Bewertung von Aktien-Optionen, die zunächst von Fischer Black und Myron Scholes, später auch von Robert Merton seit Anfang der 70er Jahre entwickelt wurde, hatte damals in den führenden Journalen im Bereich der Ökonomie aufgrund allgemeinen Unverständnisses keine Chance auf Publikation; die Arbeit erschien schließlich im relativ unbekannten *Journal of Political Economy*. Erst mit der von J.C. Cox, A. Ross und M. Rubinstein 1979 vorgestellten Vereinfachung des Modells durch Binomialbäume erlangte das mathematische Konzept größere Akzeptanz; seitdem hat sich die Theorie rasant entwickelt und als inzwischen eigenständiger Bereich unter dem Stichwort *Stochastische Finanzmathematik* in der Mathematischen Stochastik fest etabliert.

Wenn die allgemeine Theorie auch recht anspruchsvoll und für den schulischen Gebrauch im allgemeinen sicher zu schwierig ist, lassen sich die wesentlichen Grundbegriffe aber doch mit elementaren stochastischen Methoden leicht vermitteln. Im folgenden soll aufgezeigt werden, wie dies sogar im Rahmen eines Grundkurses zur Stochastik geleistet werden kann. Wir beschränken uns dabei im wesentlichen auf die sogenannten *Europäischen Aktien-Optionen im Ein-Perioden-Modell*.

Die folgende Graphik zeigt die Hebelwirkung von Optionsgeschäften noch einmal aus einer anderen Sichtweise.



Gewinn / Verlust aus Optionen (dicke Linie) und aus Aktien (dünne Linie) als Funktion des Kurswerts S_T ; $C_0 = 9,52$, $X = 110$, $K = 10000$, $i = 5\%$

4. Kombinationen von Optionsgeschäften

In diesem Abschnitt wollen wir den "spekulativen" Aspekt von Optionsgeschäften ein wenig weiter vertiefen, indem wir die Auswirkungen bestimmter Kombinationen, die in der Praxis häufig anzutreffen sind, untersuchen, wie z.B. gleichzeitiger Kauf und Verkauf gewisser Optionen. Ausgangspunkt unserer Rechnungen ist dabei wieder das obige Beispiel von Eberlein, d.h. bei der Berechnung der Optionspreise gehen wir von der Annahme aus, daß die zukünftigen Kurse nur zwei Werte annehmen können; wir untersuchen allerdings die Auswirkungen der getätigten Geschäfte für eine wesentlich größere Bandbreite möglicher zukünftiger Kurse.

Ein erster wichtiger Typ von Kombinationsgeschäften ist der sog. "Spread"; darunter versteht man den gleichzeitigen Kauf und Verkauf je einer Option desselben Typs zu unterschiedlichen Ausübungspreisen. Die Auswirkungen dieser Handels-Strategie bestehen im wesentlichen in einer Reduzierung des Verlustrisikos bei gleichem Kapitaleinsatz, allerdings werden die Gewinn-Chancen damit ebenfalls geringer. Die folgenden beiden Graphiken zeigen die Gewinne aus einem sog. *Bull-Call-Spread* und einem sog. *Bear-Call-Spread* als Funktion des Kurswerts zur Zeit T ; im ersten Fall ist der Ausübungspreis X_1 der gekauften Option niedriger als der

Stochastik in der Schule 20 (2000)

Projekte:

bmb+f –Projekt e-stat (www.emilea.de)

Laufzeit: 2001 – 2006 (in Oldenburg)

Die Entwicklung der multimedialen, internetbasierten und interaktiven Lehr- und Lernumgebung EMILeA-stat durch das Projekt e-stat wird seit April 2001 vom Bundesministerium für Bildung und Forschung (bmb+f) im Programm "Neue Medien in der Bildung" gefördert. Die Bezeichnung EMILeA-stat ist gleichzeitig eine eingetragene Marke.

An dem bundesweiten Vorhaben sind Arbeitsgruppen unter der Leitung von Prof. Dr. Joachim Buhmann (Universität Bonn, seit WS 2003/2004 ETH Zürich), Prof. Dr. Erhard Cramer (Universität Oldenburg, seit WS 2004/2005 RWTH Aachen), Prof. Dr. Ursula Gather (Universität Dortmund), Prof. Dr. Heinz Lothar Grob (Universität Münster), Prof. Dr. Wolfgang Härdle (Humboldt-Universität zu Berlin), Prof. Dr. Udo Kamps (Universität Oldenburg, seit SS 2004 RWTH Aachen), Prof. Dr. Claus Möbus (Universität Oldenburg), Prof. Dr. Dietmar Pfeifer (Universität Oldenburg), Prof. Dr. Kristina Reiss (Universität Oldenburg, seit WS 2002/2003 Universität Augsburg), Prof. Dr. Bernd Rönz (Humboldt-Universität zu Berlin), Prof. Dr. Hans Gerhard Strohe (Universität Potsdam), Prof. Dr. Karl-Heinz Waldmann (Universität Karlsruhe) und Prof. Dr. Claus Weihs (Universität Dortmund) beteiligt.

Graphik: q_x Graphik: e_x^0 Graphik: ${}_kV_{x:\overline{n}|}$ Graphik: ${}_kV_{x:\overline{n}|}^1$ Tabelle: q_x Tabelle: e_x^0 Tabelle: ${}_kV_{x:\overline{n}|}$ Tabelle: ${}_kV_{x:\overline{n}|}^1$

Sterbewahrscheinlichkeit



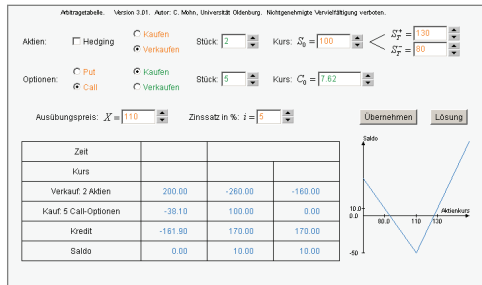
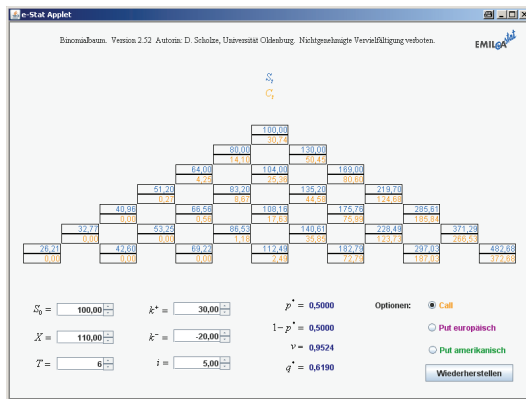
Parameter

 $x =$ $n =$ $i =$

PKV95

Prämien und Deckungskapital

 $v = 0,970874$ $d = 0,029126$ $A_x = 0,196465$ $P_x = 0,007121$ $A_{x:\overline{n}|}^1 = 0,027963$ $P_{x:\overline{n}|}^1 = 0,001408$ $A_{x:\overline{n}|}^{\frac{1}{i}} = 0,393444$ $P_{x:\overline{n}|}^{\frac{1}{i}} = 0,019806$ $A_{x:\overline{n}|} = 0,421407$ $P_{x:\overline{n}|} = 0,021213$



Microsoft Excel - Call.xls											
Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster ?											
A17											
	A	B	C	D	E	F	G	H	I	J	K
1			T = 0		T = 1						
2			Start		Price Increase		Price Decrease				
3		Units	Value	Account	Value	Account	Value	Account			
4	Stocks	-2	100	200,00	130	-260,00	80	-160,00			
5	Call-Options	5	7,62	-38,10	20	100,00	0	0,00			
6	Money	161,90	1	-161,90	1,05	170,00	1,05	170,00			
7		Balance		0,00		10,00		10,00			
8											
9											
10	So	100									
11	S+	130	p	0,4							
12	S-	80	1-p	0,6							
13											
14	X	110									
15											
16	Interest (in %)	5	v	0,9524							
17											
18	Co	7,62	E(vC)	7,62							
19											
20	Hedge Ratio	0,4									
21											
22											
23											
24											
25											
26											

EMIL@stat

Microsoft Excel - CRR.xls												
Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster ?												
A30												
A	B	C	D	E	F	G	H	I	J	K	L	
1 So	100								482,68			
2 k+	30	p*	0,5						372,68			
3 k-	-20	1-p*	0,5					371,29				
4								266,53				
5 X	110	q*	0,6190				285,61		297,03			
6							185,84		187,03			
7 Interest (in %)	5	v	0,9524			219,70		228,49				
8						124,68		123,73				
9					169,00		175,76		182,79			
10					80,60		75,99		72,79			
11				130,00		135,20		140,61				
12				50,45		44,58		35,85				
13	S _t		100,00		104,00		108,16		112,49			
14	C _t		30,74		25,36		17,63		2,49			t = 6T
15				80,00		83,20		86,53				
16				14,10		8,67		1,18				
17					64,00		66,56		69,22			
18					4,25		0,56		0,00			
19						51,20		53,25				
20						0,27		0,00				
21							40,96		42,60			
22							0,00		0,00			
23								32,77				
24								0,00				
25									26,21			
26									0,00			
t = 0												
Call-Option Put-Option europäisch Put-Option amerikanisch												

Aktien, Derivate, Arbitrage:

Eine Einführung in die moderne Finanzmathematik

Prof. Dr. Dietmar Pfeifer

Institut für Mathematik

Schwerpunkt Versicherungs- und Finanzmathematik



Kursverläufe des DAX:



Tagesgang 5.1.2011

Kursverläufe des DAX:



Tagesgang 5.1.2011

Wochengang

Kursverläufe des DAX:



Tagesgang 5.1.2011

Wochengang

Jahresgang

(Selbstähnlichkeit auf allen Zeitskalen!)

Möglichkeiten, auf schwankende Kurse zu reagieren:

Möglichkeiten, auf schwankende Kurse zu reagieren:

➤ **spekulativ (in der Hoffnung auf schnelle Gewinne)**

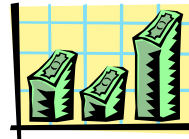


Möglichkeiten, auf schwankende Kurse zu reagieren:

➤ **spekulativ (in der Hoffnung auf schnelle Gewinne)**



➤ **konservativ (Portfolio-Hedging)**

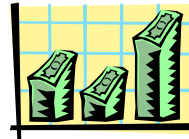


Möglichkeiten, auf schwankende Kurse zu reagieren:

➤ **spekulativ (in der Hoffnung auf schnelle Gewinne)**



➤ **konservativ (Portfolio-Hedging)**



➤ **verfügbare Finanzinstrumente:**



Derivative (Optionen, Futures, ...)

Option:

Option:

Recht, eine bestimmte Menge eines bestimmten Gutes zu einem im voraus festgesetzten Preis

Option:

Recht, eine bestimmte Menge eines bestimmten Gutes zu einem im voraus festgesetzten Preis

➤ *innerhalb einer bestimmten Frist* (amerikanische Option)

Option:

Recht, eine bestimmte Menge eines bestimmten Gutes zu einem im voraus festgesetzten Preis

➤ *innerhalb einer bestimmten Frist* (amerikanische Option)

oder

➤ *zu einem bestimmten Zeitpunkt* (europäische Option)

Option:

Recht, eine bestimmte Menge eines bestimmten Gutes zu einem im voraus festgesetzten Preis

➤ *innerhalb einer bestimmten Frist* (amerikanische Option)

oder

➤ *zu einem bestimmten Zeitpunkt* (europäische Option)

o zu kaufen (Call-Option)

Option:

Recht, eine bestimmte Menge eines bestimmten Gutes zu einem im voraus festgesetzten Preis

➤ *innerhalb einer bestimmten Frist* (amerikanische Option)

oder

➤ *zu einem bestimmten Zeitpunkt* (europäische Option)

o zu kaufen (Call-Option)

oder

o zu verkaufen (Put-Option).

Bezeichnungen:

Bezeichnungen:

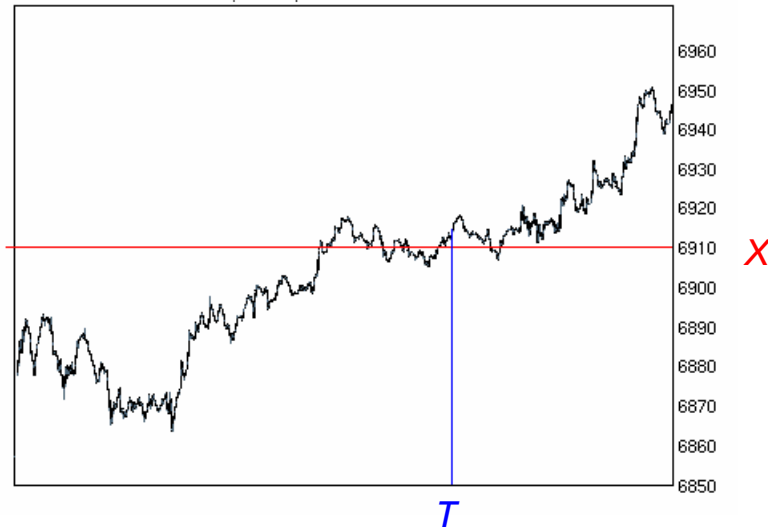
T :	Laufzeit; Verfalltag
X :	Ausübungspreis, Basispreis
S_t :	Kurswert zur Zeit t , mit $0 \leq t \leq T$
i :	risikoloser Zins
$r = 1 + i$:	Zinsfaktor
$v = 1/r$:	Diskontfaktor

Bezeichnungen:

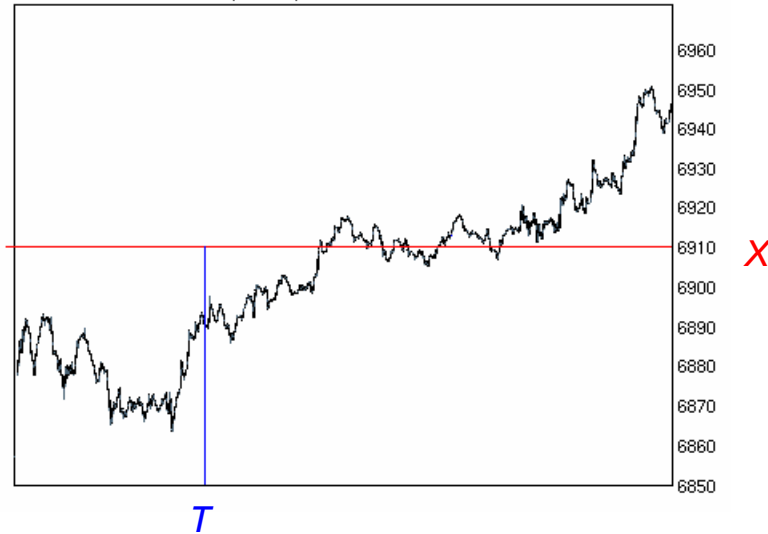
T :	Laufzeit; Verfalltag
X :	Ausübungspreis, Basispreis
S_t :	Kurswert zur Zeit t , mit $0 \leq t \leq T$
i :	risikoloser Zins
$r = 1 + i$:	Zinsfaktor
$v = 1/r$:	Diskontfaktor

Call-Option:	$S_T > X$:	Option <i>ausüben</i>
	$S_T \leq X$:	Option <i>nicht ausüben</i>
Put-Option:	$S_T < X$:	Option <i>ausüben</i>
	$S_T \geq X$:	Option <i>nicht ausüben</i>

Veranschaulichung:



Veranschaulichung:



Ziel: Bestimmung des „richtigen“ Optionswertes

Ziel: Bestimmung des „richtigen“ Optionswertes

z.B. C_t : Wert der Call-Option zur Zeit $t \in [0, T]$

Ziel: Bestimmung des „richtigen“ Optionswertes

z.B. C_t : Wert der Call-Option zur Zeit $t \in [0, T]$

Es gilt am Verfalltag $t = T$:

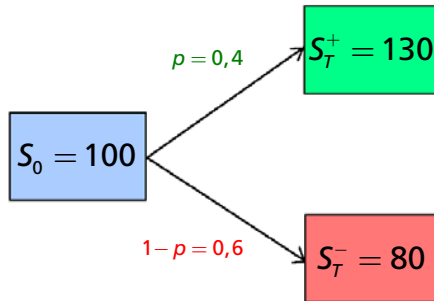
Ziel: Bestimmung des „richtigen“ Optionswertes

z.B. C_t : Wert der Call-Option zur Zeit $t \in [0, T]$

Es gilt am Verfalltag $t = T$:

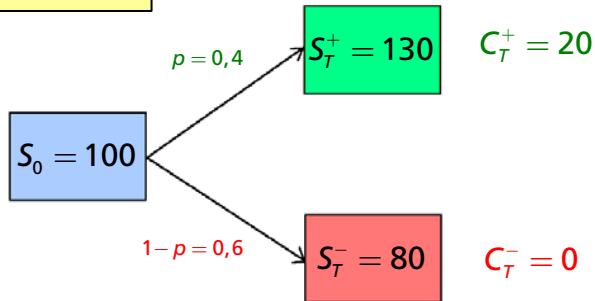
$$C_T = \max(S_T - X; 0) = (S_T - X)^+$$

Beispiel:



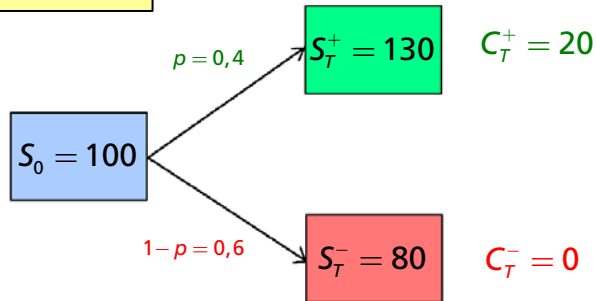
Beispiel:

$$i = 5\% \quad X = 110$$



Beispiel:

$$i = 5\% \quad X = 110$$

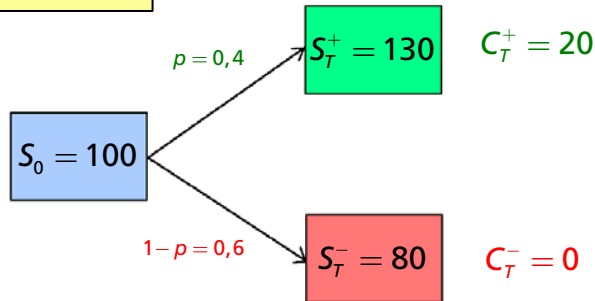


Callpreis nach Erwartungswert-Prinzip ($\rightarrow$ faires Spiel):

$$C_0 = E[v \cdot C_T]$$

Beispiel:

$$i = 5\% \quad X = 110$$

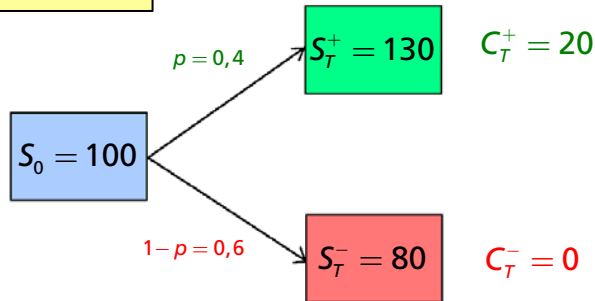


Callpreis nach Erwartungswert-Prinzip ($\rightarrow$ faires Spiel):

$$C_0 = E[v \cdot C_T] = v \cdot p \cdot C_T^+ = v \cdot p \cdot (S_T^+ - X)$$

Beispiel:

$$i = 5\% \quad X = 110$$

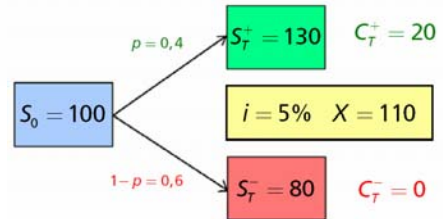


Callpreis nach Erwartungswert-Prinzip ($\rightarrow$ faires Spiel):

$$C_0 = E[v \cdot C_T] = v \cdot p \cdot C_T^+ = v \cdot p \cdot (S_T^+ - X) = \frac{0,4 \cdot 20}{1,05} = \underline{\underline{7,62}}$$

Analyse:

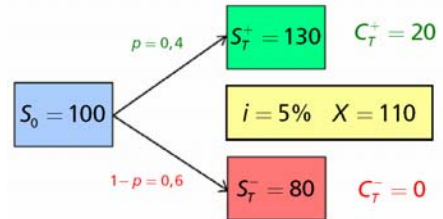
$$C_0 = 7,62$$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$

Analyse:

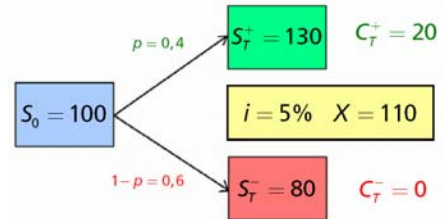
$$C_0 = 7,62$$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Leerverkauf 2 Aktien	+200,00	-260,00	-160,00

Analyse:

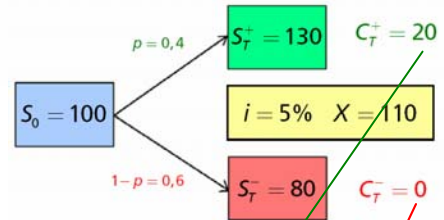
$$C_0 = 7,62$$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Leerverkauf 2 Aktien	+200,00	-260,00	-160,00
Kauf 5 Call-Optionen	-38,10	+100,00	0,00

Analyse:

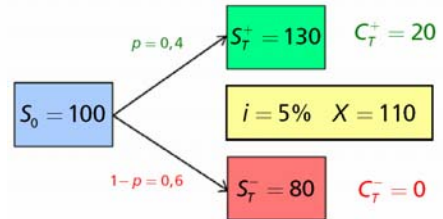
$$C_0 = 7,62$$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Leerverkauf 2 Aktien	+200,00	-260,00	-160,00
Kauf 5 Call-Optionen	-38,10	+100,00	0,00

Analyse:

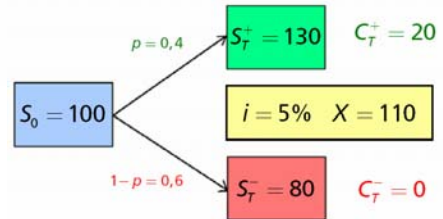
$$C_0 = 7,62$$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Leerverkauf 2 Aktien	+200,00	-260,00	-160,00
Kauf 5 Call-Optionen	-38,10	+100,00	0,00
Kredit vergeben	-161,90	+170,00	+170,00
Saldo	0,00	+10,00	+10,00

Analyse:

$$C_0 = 7,62$$

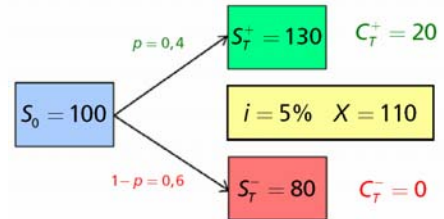


$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$

Arbitrage-Möglichkeit!

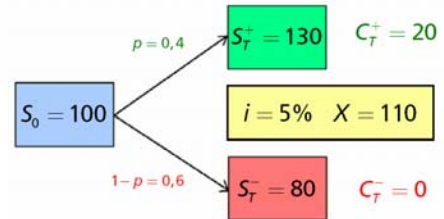
Kredit vergeben	-161,90	+170,00	+170,00
Saldo	0,00	+10,00	+10,00

Analyse: höherer Preis $C_0 = 12$



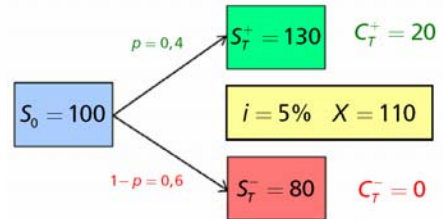
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$

Analyse: höherer Preis $C_0 = 12$



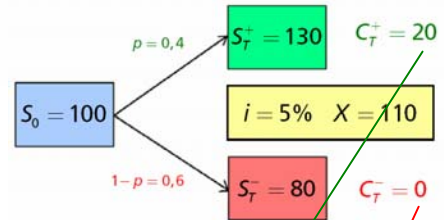
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf 2 Aktien	-200,00	+260,00	+160,00

Analyse: höherer Preis $C_0 = 12$



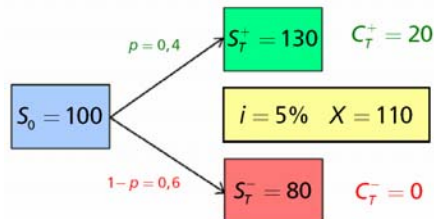
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf 2 Aktien	-200,00	+260,00	+160,00
Verkauf 5 Call-Optionen	+60,00	-100,00	0,00

Analyse: höherer Preis $C_0 = 12$



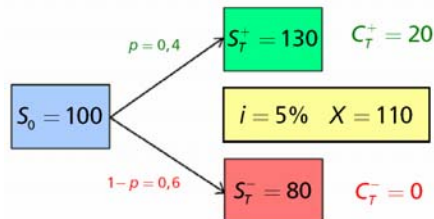
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf 2 Aktien	-200,00	+260,00	+160,00
Verkauf 5 Call-Optionen	+60,00	-100,00	0,00

Analyse: höherer Preis $C_0 = 12$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf 2 Aktien	-200,00	+260,00	+160,00
Verkauf 5 Call-Optionen	+60,00	-100,00	0,00
Kredit aufnehmen	+140,00	-147,00	-147,00
Saldo	0,00	+13,00	+13,00

Analyse: höherer Preis $C_0 = 12$

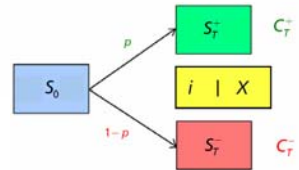


$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$

Arbitrage-Möglichkeit!

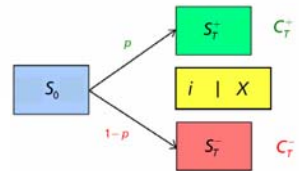
Kredit aufnehmen	+140,00	-147,00	-147,00
Saldo	0,00	+13,00	+13,00

Analyse: richtiger Preis $C_0 = ?$



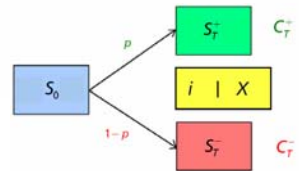
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$-n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$

Analyse: richtiger Preis $C_0 = ?$



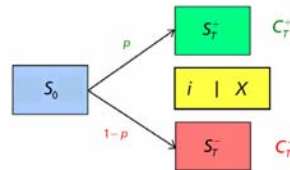
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$-n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Verkauf m Call-Optionen	$+m \cdot C_0$	$-m \cdot (S_T^+ - X)$	$0,00$

Analyse: richtiger Preis $C_0 = ?$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$-n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Verkauf m Call-Optionen	$+m \cdot C_0$	$-m \cdot (S_T^+ - X)$	0,00
Kredit aufnehmen	$+(n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$
Saldo	0,00	0,00	0,00

Analyse: richtiger Preis $C_0 = ?$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$-n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Verkauf m Call-Optionen	$+m \cdot C_0$	$-m \cdot (S_T^+ - X)$	$0,00$
Kredit aufnehmen	$+(n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$
Saldo	$0,00$	$0,00$	$0,00$

Notwendige Bedingung: $n \cdot S_T^- = n \cdot S_T^+ - m \cdot (S_T^+ - X)$

- 11d -

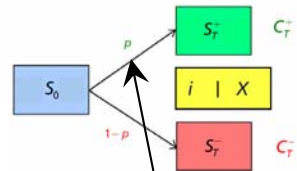
$\Rightarrow$ Hedge Ratio

$$h_{\text{Call}} = \frac{n}{m} = \frac{S_T^+ - X}{S_T^+ - S_T^-}$$

und

$$C_0 = h_{\text{Call}} \cdot (S_0 - v \cdot S_T^-)$$

Analyse: richtiger Preis $C_0 = ?$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$n \cdot S_0$	$n \cdot S_T^+$	$n \cdot S_T^-$

Lösung ist unabhängig von p !

Kredit aufnehmen	$+(n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$	$-r \cdot (n \cdot S_0 - m \cdot C_0)$
Saldo	0,00	0,00	0,00

Notwendige Bedingung: $n \cdot S_T^- = n \cdot S_T^+ - m \cdot (S_T^+ - X)$

$\Rightarrow$ Hedge Ratio

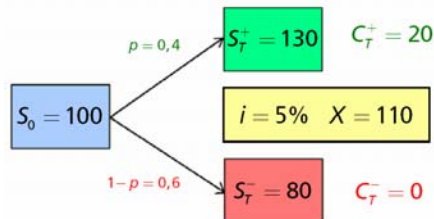
$$h_{\text{Call}} = \frac{n}{m} = \frac{S_T^+ - X}{S_T^+ - S_T^-}$$

und

$$C_0 = h_{\text{Call}} \cdot (S_0 - v \cdot S_T^-)$$

Probe:

$$C_0 = 9,52$$

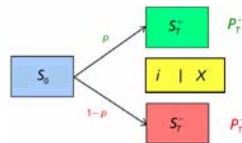


$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf 2 Aktien	-200,00	+260,00	+160,00
Verkauf 5 Call-Optionen	+47,60	-100,00	0,00
Kredit aufnehmen	-152,40	-160,00	-160,00
Saldo	0,00	0,00	0,00

$$h_{Call} = \frac{S_T^+ - X}{S_T^+ - S_T^-} = \frac{20}{50}$$

$$C_0 = h_{Call} \cdot (S_0 - v \cdot S_T^-) = \frac{2}{5} \cdot \frac{105 - 80}{1,05} = 9,52$$

Analyse für Put-Preis: richtiger Preis $P_0 = ?$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Kauf n Aktien	$-n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Kauf m Put-Optionen	$-m \cdot P_0$	0,00	$+m \cdot (X - S_T^-)$
Kredit aufnehmen	$+(n \cdot S_0 + m \cdot P_0)$	$-r \cdot (n \cdot S_0 + m \cdot P_0)$	$-r \cdot (n \cdot S_0 + m \cdot P_0)$
Saldo	0,00	0,00	0,00

Notwendige Bedingung: $n \cdot S_T^+ = n \cdot S_T^- + m \cdot (X - S_T^-)$

- 13a -

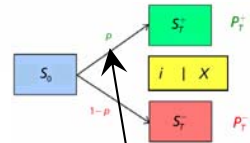
$\Rightarrow$ Hedge Ratio

$$h_{\text{Put}} = \frac{n}{m} = \frac{X - S_T^-}{S_T^+ - S_T^-}$$

und

$$P_0 = h_{\text{Put}} \cdot (v \cdot S_T^+ - S_0)$$

Analyse für Put-Preis: richtiger Preis $P_0 = ?$



$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$

Lösung ist unabhängig von p !

Kredit aufnehmen	$+(n \cdot S_0 + m \cdot P_0)$	$-r \cdot (n \cdot S_0 + m \cdot P_0)$	$-r \cdot (n \cdot S_0 + m \cdot P_0)$
Saldo	0,00	0,00	0,00

Notwendige Bedingung: $n \cdot S_T^+ = n \cdot S_T^- + m \cdot (X - S_T^-)$

$\Rightarrow$ Hedge Ratio

$$h_{Put} = \frac{n}{m} = \frac{X - S_T^-}{S_T^+ - S_T^-}$$

und

$$P_0 = h_{Put} \cdot (v \cdot S_T^+ - S_0)$$

Beobachtung:

Beobachtung:

➤ Für die Hedge Ratios gilt:

$$h_{\text{Call}} + h_{\text{Put}} = \frac{S_T^+ - X}{S_T^+ - S_T^-} + \frac{X - S_T^-}{S_T^+ - S_T^-} = 1$$

Beobachtung:

➤ Für die Hedge Ratios gilt:

$$h_{\text{Call}} + h_{\text{Put}} = \frac{S_T^+ - X}{S_T^+ - S_T^-} + \frac{X - S_T^-}{S_T^+ - S_T^-} = 1$$

➤ Optionspreise sind genau dann positiv, wenn $S_T^- < rS_0 < S_T^+$ gilt,
d.h. wenn positive Kursbewegungen potenziell mehr Gewinn erzielen als die risikolose Geldanlage

Beobachtung:

- Für die Hedge Ratios gilt:

$$h_{\text{Call}} + h_{\text{Put}} = \frac{S_T^+ - X}{S_T^+ - S_T^-} + \frac{X - S_T^-}{S_T^+ - S_T^-} = 1$$

- Optionspreise sind genau dann positiv, wenn $S_T^- < rS_0 < S_T^+$ gilt, d.h. wenn positive Kursbewegungen potenziell mehr Gewinn erzielen als die risikolose Geldanlage
- Bei Arbitrage-freien (fairen) Optionspreisen ist durch **keine** mengenmäßige Kombination von Derivaten eine Arbitragemöglichkeit gegeben

Beobachtung:

- Für die Hedge Ratios gilt:

$$h_{\text{Call}} + h_{\text{Put}} = \frac{S_T^+ - X}{S_T^+ - S_T^-} + \frac{X - S_T^-}{S_T^+ - S_T^-} = 1$$

- Optionspreise sind genau dann positiv, wenn $S_T^- < rS_0 < S_T^+$ gilt, d.h. wenn positive Kursbewegungen potenziell mehr Gewinn erzielen als die risikolose Geldanlage
- Bei Arbitrage-freien (fairen) Optionspreisen ist durch **keine** mengenmäßige Kombination von Derivaten eine Arbitragemöglichkeit gegeben
- Die Arbitrage-freien (fairen) Optionspreise hängen nicht von **p** ab!

Aber:

„Äquivalentes“ Erwartungswertprinzip: Im Fall von $S_T^- < rS_0 < S_T^+$ gilt

$$0 < p^* = \frac{rS_0 - S_T^-}{S_T^+ - S_T^-} = 1 - q^* < 1$$

Aber:

„Äquivalentes“ Erwartungswertprinzip: Im Fall von $S_T^- < rS_0 < S_T^+$ gilt

$$0 < p^* = \frac{rS_0 - S_T^-}{S_T^+ - S_T^-} = 1 - q^* < 1$$

mit

$$C_0 = h_{\text{Call}} \cdot (S_0 - v \cdot S_T^-) = E^*[v \cdot C_T] = v \cdot p^* \cdot C_T^+$$

Aber:

„Äquivalentes“ Erwartungswertprinzip: Im Fall von $S_T^- < rS_0 < S_T^+$ gilt

$$0 < p^* = \frac{rS_0 - S_T^-}{S_T^+ - S_T^-} = 1 - q^* < 1$$

mit

$$C_0 = h_{Call} \cdot (S_0 - v \cdot S_T^-) = E^*[v \cdot C_T] = v \cdot p^* \cdot C_T^+$$

$$P_0 = h_{Put} \cdot (v \cdot S_T^+ - S_0) = E^*[v \cdot P_T] = v \cdot q^* \cdot P_T^-$$

Aber:

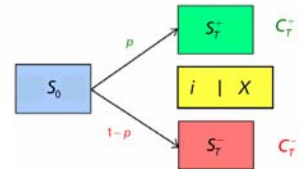
„Äquivalentes“ Erwartungswertprinzip: Im Fall von $S_T^- < rS_0 < S_T^+$ gilt

$$0 < p^* = \frac{rS_0 - S_T^-}{S_T^+ - S_T^-} = 1 - q^* < 1$$

mit dem Arbitrage-freien Preis

$$D_0 = E^*[v \cdot D_T] = v \cdot (p^* \cdot D_T^+ + q^* \cdot D_T^-)$$

für jedes beliebige Derivat D !

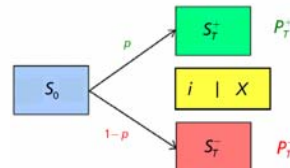


Hedging mit Call-Optionen:

$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Besitz n Aktien	$+n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Verkauf m Call-Optionen	$+m \cdot C_0$	$-m \cdot (S_T^+ - X)$	0,00
Kredit vergeben	$-m \cdot C_0$	$+r \cdot m \cdot C_0$	$+r \cdot m \cdot C_0$
Saldo	$+n \cdot S_0$	$+r \cdot n \cdot S_0$	$+r \cdot n \cdot S_0$

mit der Hedge Ratio $h_{\text{Call}} = n / m = (S_T^+ - X) / (S_T^+ - S_T^-)$ wegen

$$n \cdot S_T^- + r \cdot m \cdot C_0 = n \cdot S_T^- + m \cdot h_{\text{Call}} \cdot (rS_0 - S_T^-) = n \cdot S_T^- + n \cdot (rS_0 - S_T^-) = r \cdot n \cdot S_0$$



Hedging mit Put-Optionen:

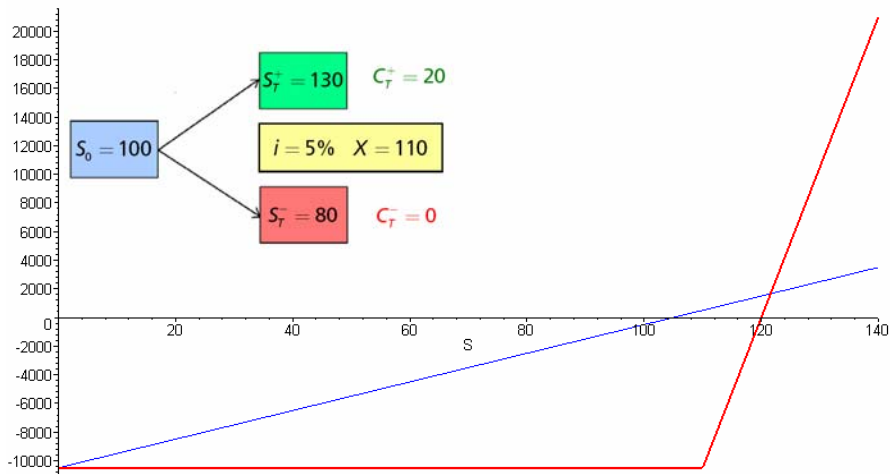
$t = 0$		$t = T$	
Aktion	Kontobewegung	$S_T^+ > X$	$S_T^- < X$
Besitz n Aktien	$+n \cdot S_0$	$+n \cdot S_T^+$	$+n \cdot S_T^-$
Kauf m Put-Optionen	$-m \cdot P_0$	0,00	$+m \cdot (X - S_T^-)$
Kredit aufnehmen	$+m \cdot P_0$	$-r \cdot m \cdot P_0$	$-r \cdot m \cdot P_0$
Saldo	$+n \cdot S_0$	$+r \cdot n \cdot S_0$	$+r \cdot n \cdot S_0$

mit der Hedge Ratio $h_{Put} = n / m = (X - S_T^+) / (S_T^+ - S_T^-)$ wegen

$$n \cdot S_T^+ - r \cdot m \cdot P_0 = n \cdot S_T^+ - m \cdot h_{Put} \cdot (S_T^+ - rS_0) = n \cdot S_T^+ - n \cdot (S_T^+ - rS_0) = r \cdot n \cdot S_0$$

Hebelwirkung von Optionsgeschäften:

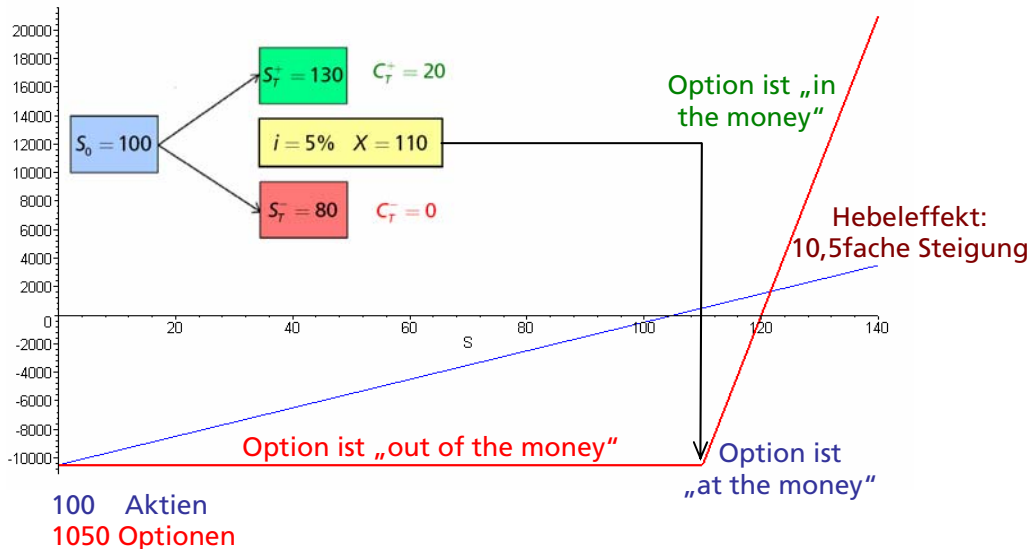
Gewinn aus **Call-Optionen** und **Aktien** als Funktion des Kurses bei Kapitaleinsatz 10.000



100 Aktien
1050 Optionen

Hebelwirkung von Optionsgeschäften:

Gewinn aus **Call-Optionen** und **Aktien** als Funktion des Kurses bei Kapitaleinsatz 10.000



Kursverläufe des DAX:



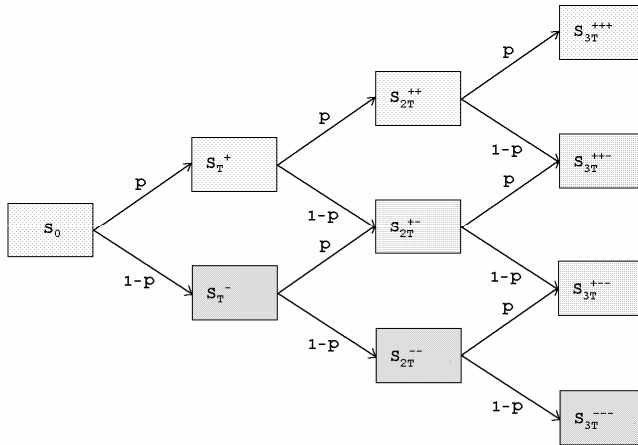
Tagesgang 5.1.2011

Wochengang

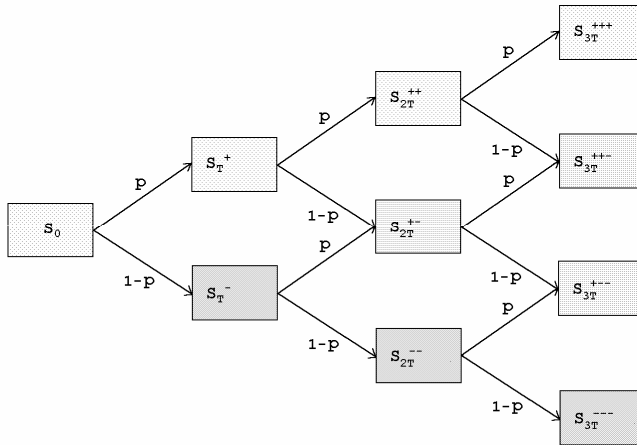
Jahresgang

(Selbstähnlichkeit auf allen Zeitskalen!)

Mehr-Stufen-Modell von Cox-Ross-Rubinstein:



Mehr-Stufen-Modell von Cox-Ross-Rubinstein:

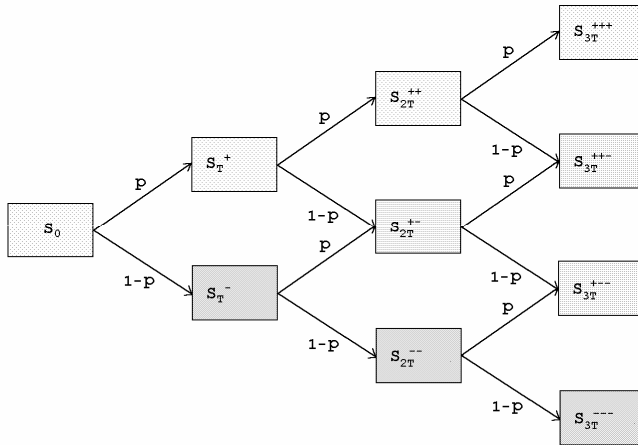


Annahmen:

konstante Kursverhältnisse:

$$k^{\pm} = \frac{S_T^{\pm}}{S_0} - 1 \quad (\text{Volatilitäten})$$

Mehr-Stufen-Modell von Cox-Ross-Rubinstein:



Annahmen:

konstante Kursverhältnisse:

$$k^{\pm} = \frac{S_T^{\pm}}{S_0} - 1 \quad (\text{Volatilitäten})$$

stochastisch unabhängige
Auf-/ Abwärtsbewegungen:

$$S_{nT} = (1 + k^+)^N \cdot (1 + k^-)^{n-N}$$

$$P(N = j) = \binom{n}{j} p^j (1-p)^{n-j}, \quad j = 0, \dots, n.$$

Arbitrage-freie Berechnung von Optionspreisen:

$$\begin{aligned}
 C_0 &= E^*(\mathbf{v}^n C_{nT}) = \mathbf{v}^n E^*(S_{nT} - X)^+ \\
 &= \mathbf{v}^n \sum_{j=0}^n \binom{n}{j} p^{*j} q^{*n-j} [S_0(1+k^+)^j(1+k^-)^{n-j} - X]^+
 \end{aligned}$$

mit

$$p^* = \frac{rS_0 - (1+k^-)S_0}{(1+k^+)S_0 - (1+k^-)S_0} = \frac{i_T - k^-}{k^+ - k^-}$$

und $i_T = (1+i)^T - 1$ (Zins für Periodenlänge T)

Die Formel von Black und Scholes:

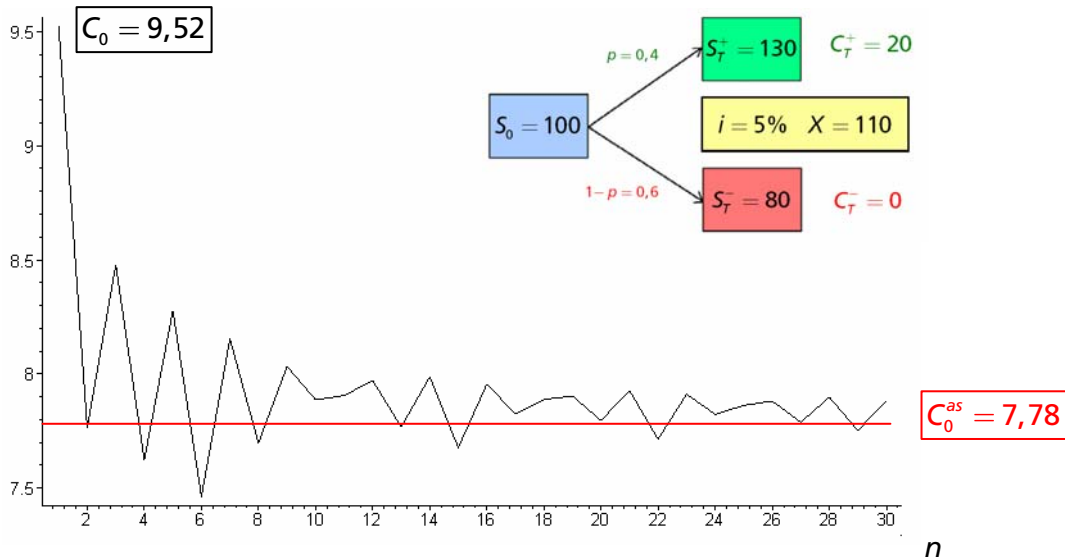
- Grenzfall für wachsende n mit angepassten Volatilitäten:

$$k_n^+ = \sigma \sqrt{\frac{T}{n}}, \quad k_n^- = -\sigma \sqrt{\frac{T}{n}}$$

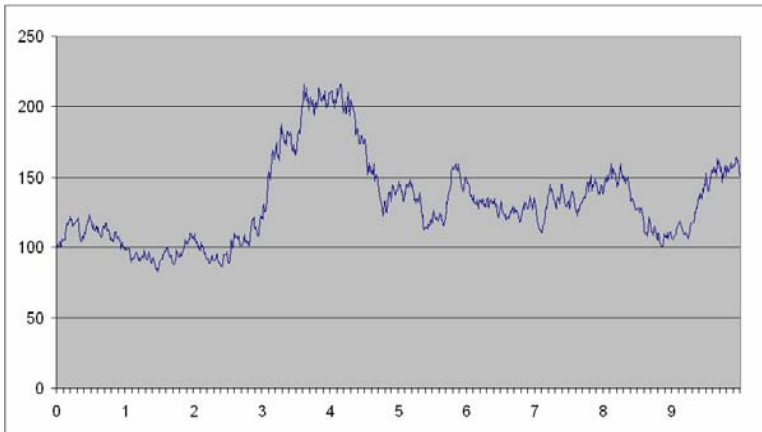
- Approximation der Binomial- durch die Normalverteilung:

$$C_0^{as} = S_0 \Phi \left(\frac{\ln \left(\frac{r^T S_0}{X} \right) + \frac{\sigma^2}{2} T}{\sigma \sqrt{T}} \right) - v^T X \Phi \left(\frac{\ln \left(\frac{r^T S_0}{X} \right) - \frac{\sigma^2}{2} T}{\sigma \sqrt{T}} \right)$$

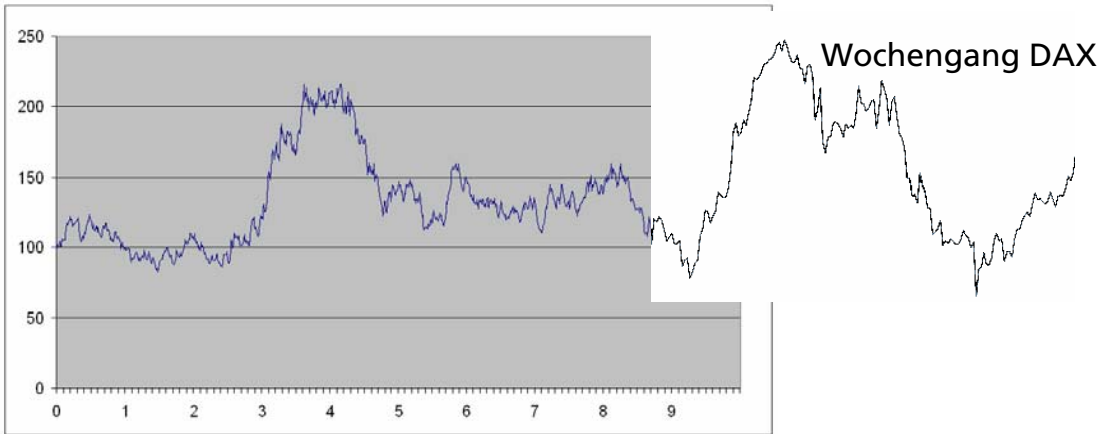
Die Formel von Black und Scholes:



Simulation des Grenzprozesses (geometrische Brownsche Bewegung)



Simulation des Grenzprozesses (geometrische Brownsche Bewegung)



Solche Funktionen sind überall stetig, aber nirgends differenzierbar!

Ausblick:

➤ Preisfindung für

- nicht-arbitragefreie (unvollständige) Märkte
- allgemeinere Finanzprodukte (z.B. CDO's)
- Derivate an Warenterminbörsen
- Derivate an Energiebörsen
- Realloptionen

➤ Modellierung von Zinsstrukturkurven

➤ Asset-Liability-Management

Aktien, Derivate, Arbitrage:

Eine Einführung in die moderne Finanzmathematik

Prof. Dr. Dietmar Pfeifer

Institut für Mathematik

Schwerpunkt Versicherungs- und Finanzmathematik



Weitere Projekte:

1. EFRE-Projekt „Entwicklung eines Solvency II-kompatiblen Risikomanagementsystems für kleine Versicherungsvereine auf Gegenseitigkeit“

(zusammen mit mit Angelika May und Thomas Kneib)

Förderlinie: 2.2.1 Kooperationsprojekte mit KMU

Laufzeit: 2009 - 2010

Beteiligte Unternehmen:

5 in der Region Weser-Ems ansässige VVaG, 1 VVaG im Gebiet Lüneburg

2. EFRE-Projekt Entwicklung und Konzeption eines Master-Weiterbildungsstudiengangs „Risikomanagement für Banken und Versicherungen“

(zusammen mit Angelika May und Jörg Prokop)

Förderlinie: 4.2 Kooperationsprojekte mit KMU

Laufzeit: 2010 - 2013

Risikomanagement mit Diplom

Uni Oldenburg startet im Sommersemester mit neuem Studiengang

Die Bedeutung des Risikomanagements in der Finanz-Branche wächst und damit der Bedarf an hochqualifizierten Fachkräften. Die Universität Oldenburg hat deshalb einen neuen Master-Studiengang entwickelt, der sich sowohl an Mitarbeiter von Versicherungen als auch von Banken wendet. „Risikomanagement für Banken und Versicherungen“ heißt der Studiengang, der im Sommersemester beginnt. Er ist auf eine Dauer von sechs Semestern in Teilzeit angelegt und richtet sich an Fachkräfte von Banken, Finanzdienstleistungs- und Versicherungsunternehmen. Wer das komplette Masterstudium nicht absolvieren möchte, kann an einem Programm mit individueller Schwerpunktsetzung teilnehmen und einzelne versicherungs- bzw. bankspezifische Module als eigenständige Weiterbildungseinheit wählen.

Initiiert und konzipiert haben den neuen Studiengang Dietmar Pfeifer und Jörg Prokop von der Universität Oldenburg sowie die Finanzmathematikerin Angelika May, die ähn-



Professoren der **Uni Oldenburg** reagieren auf wachsende Nachfrage nach Risikomanagern.

Foto: Uni Oldenburg

liche Weiterbildungs-Projekte bereits an der Universität Heidelberg und im Bonner Forschungszentrum Caesar entwickelt hat. Der Studiengang legt Wert auf Praxisnähe. Zu den Bausteinen gehören Betriebslehre, Informationsmanagement, Risikomodelle und Operationelle Risiken sowie Themenblöcke rund um Solvency II.

VW

Aus: Versicherungswirtschaft, Heft 9, Mai 2012

ON AN ESTIMATION PROBLEM FOR TYPE I CENSORED SPATIAL POISSON PROCESSES

J. HURT, P. LACHOUT, * *Charles University, Prague*

D. PFEIFER, ** *Universität Hamburg*

Abstract

In this paper we consider the problem of estimating the intensity of a spatial homogeneous Poisson process if a part of the observations [quadrat counts] is censored. The actual problem has occurred during a court case when one of the authors was a referee for the defense.

TYPE I CENSORING; SPATIAL POISSON PROCESS; INTENSITY ESTIMATION

AMS CLASSIFICATION: PRIMARY: 62F12; SECONDARY: 62N25, 90B25

Kybernetika 37 (2001)

Wahldebakel: BUSH + GORE + CHAOS = RUEGE?

Anmerkungen zu einer mathematischen Denksportaufgabe

DIETMAR PFEIFER, OLDENBURG UND MICHAEL NAATZ, BERLIN

Zusammenfassung: Stochastische Such- und Optimierungsalgorithmen lassen sich oft effizienter zur Lösung konkreter mathematischer Aufgaben einsetzen als „klassische“ numerische oder enumerative Verfahren. Dies soll an dem Beispiel einer Denksportaufgabe aus dem Stern-Magazin, dem die obige Rätsel-Gleichung entstammt, veranschaulicht werden.

1. Das Problem und seine Herkunft

In der Ausgabe 2 des Stern-Magazins vom 4.1.2001 findet sich unter der Rubrik Rätsel auf Seite 71 unter dem Stichwort *Wahldebakel* die folgende mathematische Denksportaufgabe:

„Das Chaos bei der Bush-und/or-Gore-Wahl verdient sicherlich eine Rüge, so sehen es zumindest Vertreter der Politmathematik. Ihre Formel

$$BUSH + GORE + CHAOS = RUEGE$$

lässt sich in eine schlüssige Gleichung verwandeln, indem ihre Buchstaben durch Ziffern – gleiche Buchstaben durch gleiche, verschiedene durch verschiedene Ziffern – ersetzt werden. Unter den sechs möglichen Lösungen wird diejenige gesucht, bei der das CHAOS am größten ist.“

– wobei im doppeldeutigen Sinne hier die numerisch größte Lösung gemeint ist. Durch geschicktes Überlegen lassen sich natürlich auf Anhieb einige Beziehungen zwischen den zehn Unbekannten herstellen: so gilt etwa $H + S = 10$. Die Beziehung zwischen C und R ist dagegen nicht so einfach wegen der drei Summanden auf der linken Seite; hier ist sowohl $R = C + 1$ als auch $R = C + 2$ denkbar, und beide Möglichkeiten kommen auch in der Lösungsmenge vor. Ansonsten ist man wohl mehr oder weniger auf eine geeignete Rate-Strategie angewiesen.

Wenn man darauf verzichten will, bleibt eigentlich zunächst nur die Alternative einer systematischen Abzählung aller denkbaren Fälle (hier: Permutationen der Ziffern 0 bis 9 ohne Wiederholung), das sind im vorliegenden Fall maximal $10! = 3\,628\,800$ Möglichkeiten (wegen $B > 0$, $G > 0$, $C > 0$ und $R > 0$ reduzieren sich diese aber faktisch noch auf $4096! = 1\,451\,520$ Möglichkeiten auf insgesamt $2\,177\,280$ Möglichkeiten). Die Programmierung

geeigneter Algorithmen ist aus schulischer Sicht allerdings aufwendig, auch wenn hier die schnellsten Verfahren nur wenig mehr als die theoretisch minimale Größenordnung von $10!$ Schritten erfordern (vgl. z.B. Priese/Ruskey (1994) oder Canfield / Williamson (1995)). Erstmalig biete hier nun die Stochastik einen interessanten und effizienten Ausweg aus diesem Dilemma, indem man die oben angedeutete vollständige Enumeration aller sinnvollen Fälle durch eine zufällige Enumeration ersetzt. Dabei macht man sich die Tatsache zunutze, dass jede Permutation $\{0(0), \dots, 9(9)\}$ der zehn Ziffern $\{0, \dots, 9\}$ durch eine Verkettung von paarweisen Vertauschungen erreicht werden kann. Dies sei exemplarisch an dem Beispiel der Permutation $\{2, 4, 1, 7, 5, 3, 0, 6, 8, 9\}$ veranschaulicht:

$[0, \underline{1}, 2, 3, 4, 5, 6, 7, \underline{8}, 9]$
 $[0, \underline{2}, 1, 3, 4, 5, 6, 7, \underline{8}, 9]$
 $[2, 0, 1, 3, 4, 5, 6, 7, \underline{8}, 9]$
 $[2, 0, 1, \underline{4}, 3, 5, 6, 7, \underline{8}, 9]$
 $[2, 0, \underline{4}, 1, 3, 5, 6, 7, \underline{8}, 9]$
 $[2, \underline{4}, 0, 1, 3, 5, 6, 7, \underline{8}, 9]$
 $[2, \underline{4}, 3, 0, 1, 5, 6, 7, \underline{8}, 9]$
 $[2, \underline{4}, 1, 0, 3, 5, 7, \underline{6}, 8, 9]$
 $[2, \underline{4}, 1, 0, 3, 7, \underline{5}, 6, 8, 9]$
 $[2, \underline{4}, \underline{1}, 0, 7, 3, 5, \underline{6}, 8, 9]$
 $[2, \underline{4}, 1, \underline{7}, 0, 3, 5, \underline{6}, 8, 9]$
 $[2, \underline{4}, 1, \underline{7}, 0, \underline{6}, 3, \underline{5}, 8, 9]$
 $[2, 4, 1, \underline{7}, \underline{2}, 0, 3, \underline{6}, \underline{5}, 8, 9]$
 $[2, \underline{4}, 1, \underline{7}, \underline{2}, \underline{0}, 3, \underline{6}, \underline{5}, 8, 9]$

Dabei sind die im nächsten Schritt zu vertauschenden Ziffern jeweils fett markiert, die schon richtig positionierten Ziffern unterstrichen. Das Prinzip ist offensichtlich: man „schiebt“ die jeweils nächste Ziffer aus dem noch nicht richtig angeordneten „Rest“ von rechts nach links durch sukzessives Vertauschen mit dem jeweiligen linken Nachbarn an die richtige Position. Ein solcher Vertauschungsschritt

$$BUSH + GORE + CHAOS = RUEGE$$

$6728 + 1350 + 48932 = 57010$ (1)
 $5346 + 1792 + 86074 = 93212$ (2)
 $9537 + 8041 + 27603 = 45181$ (3)
 $4173 + 2860 + 53987 = 61020$ (4)
 $2364 + 7195 + 84016 = 93575$ (5)
 $2819 + 6045 + 39701 = 48565$ (6)
 $8419 + 7056 + 39201 = 54676$ (7)
 $8637 + 1259 + 47023 = 56919$ (8)
 $3764 + 9025 + 14806 = 27595$ (9)
 $5073 + 1492 + 83647 = 90212$ (10)

Mit einem handelsüblichen PC, Ausstattung Pentium II mit 233 MHz Taktfrequenz, benötigte das (kompilierte) Programm dabei in keinem Fall mehr als 5 Minuten bis zum Auffinden dieser Lösungen.

[Die Wahrscheinlichkeit, dass es außer den gefundenen noch weitere Lösungen gibt, ist mit fast zweifach insgesamt über 200 Millionen durchgeführten Suchschritten nach (2) kleiner als 10^{-11} ...]

4. Diskussion

Das angesprochene Problem eignet sich m.E. sehr gut für eine Diskussion im Zusammenhang mit dem Computereinsatz in der Schule. Im Gegensatz zu bekannten Computer-Algorithmen-Systemen wie DERIVE, MATHEMATICA oder MAPLE müssen sich die Schüler hier ein problemorientiertes Lösungskonzept selbst überlegen, da keines der genannten Programme über geeignete Module verfügt, die hier direkt einsetzbar wären. [Eine theoretisch denkbare Auflistung aller $10!$ Permutationen in einem Feld oder einer Datei ist mit diesen Systemen zwar technisch möglich, aufgrund des riesigen Umfangs und der damit verbundenen Rechenzeit aber nicht praktikabel.] Hinzu kommt der denkportale Elternteil, eine Lösung der auf den ersten Blick vielleicht einfach erscheinenden Aufgabe zu finden und dabei auch die der Rätsel-Redaktion eines großen Wochenmagazins Fehler nachzuweisen! Die benötigten Kenntnisse aus der Stochastik gehen dabei über elementare Kombinatorik bzw. die diskrete Gleichverteilung nicht hinaus, zeigen aber durchaus moderne Anwendungen der Stochastik auf, die vor allem in der Informatik und der Numerik zunehmend an Bedeutung gewinnen (Stichwort: Simulated Annealing, vgl. Mathar / Pfeifer (1990), Kapitel 3.3).

Eine kompilierte Fassung des Programms kann bei Interesse auch direkt von den Autoren per e-mail bezogen werden.

Literatur

- Athen, H.; Griesel, H.; Postel, H. (1995): Mathematik heute. Leistungskurs Stochastik. Schroedel Schulbuchverlag, Hannover.
 Baierlein, M.; Bath, F.; Greifneger, U.; Krumbacher, G. (1984): Anschauliche Analysis II. Leistungskurs. Ehrenwirth Verlag, München.
 Barth, F.; Haller, R. (1998): Stochastik Leistungskurs. Oldenbourg Verlag, München.
 Canfield, E.R.; Williamson, S.G. (1995): A Loop-Free Algorithm for Generating the Linear Extensions of a Poset. order 12 No. 1, 57-75.
 Engel, A. (1987): Stochastik. Ernst Klett Schulbuchverlag, Stuttgart.
 Engel, A. (1991): Mathematisches Experimentieren mit dem PC. Ernst Klett Schulbuchverlag, Stuttgart.
 Kregel, U. (2000): Einführung in die Wahrscheinlichkeitstheorie und Statistik. Für Studium, Berufspraxis und Lehramt. 5. Aufl., Vieweg, Braunschweig und Wiesbaden.
 Mathar, R.; Pfeifer, D. (1990): Stochastik für Informatiker. Leitfaden und Monographien der Informatik. Teubner, Stuttgart.
 Pfanzagl, J. (1991): Elementare Wahrscheinlichkeitsrechnung. 2. Aufl., Walter de Gruyter, Berlin.
 Priese, G.; Ruskey, F. (1994): Generating Linear Extensions Fast. SIAM J. Comp. 23, 373-386.
 Stern-Magazin Nr. 2 vom 4.1.2001, Gruner+Jahr AG&Co., Hamburg.

Anschriften der Verfasser:

- Prof. Dr. Dietmar Pfeifer
 Carl von Ossietzky Universität Oldenburg
 Fachbereich Mathematik
 Postfach 2503
 26111 Oldenburg
 pfeifer@mathematik.uni-oldenburg.de
 Dipl.-Math. Michael Naatz
 Technische Universität Berlin
 Fakultät II – Mathematik und Naturwissenschaften
 Institut für Mathematik
 Straße des 17. Juni 136
 10623 Berlin
 naatz@math.tu-berlin.de

Stochastik in der Schule 21 (2001)

SIZE SELECTION AND COMPETITION FOR MUSSELS, *MYTILUS EDULIS*, BY OYSTERCATCHERS, *HAEMATOPUS OSTRALEGUS*, HERRING GULLS, *LARUS ARGENTATUS*, AND COMMON EIDERS, *SOMATERIA MOLLISSIMA*

Gudrun Hilgerloh¹ and Dietmar Pfeifer²

¹Forschungszentrum Terramar, Schleusenstr. 1, D-26382 Wilhelmshaven, Germany.
Present address: Institut für Zoologie, Abt. V, Johannes Gutenberg-Universität Mainz,
Becherweg 13, D-55099 Mainz

²Fachbereich Mathematik, Universität Oldenburg, Postfach 2503, D-26111 Oldenburg

ABSTRACT

Investigations on size selection of blue mussels, *Mytilus edulis*, and exploitation competition by ventral hammering oystercatchers, *Haematopus ostralegus*, herring gulls, *Larus argentatus*, and common eiders, *Somateria mollissima*, were carried out on various blue mussel beds on the tidal flats of Spiekeroog (Niedersachsen/Germany). The size-class was considered to be the ideal one, which would be selected if all size-classes were available. On one mussel bed the mussel sizes selected by oystercatchers were the most abundant mussel sizes, the median being 51 mm. It was concluded that the most abundant size-class corresponded to the ideal size-class which was much bigger than predicted for ventral hammering oystercatchers. If all mussels available were smaller than the ideal size, they selected the largest ones. Unlike oystercatchers, eiders did not select the largest mussels present when they were all smaller than the ideal size. Instead, they selected mussels that were only slightly longer than the median of the available mussels. As eiders dive for several mussels when they are feeding on small mussels, they are time-restricted and less selective than if diving for single large mussels. For herring gulls, the ideal mussel size seemed to be 20 mm. Mussel sizes selected by different bird species on the same mussel bed differed significantly. Each species chose other size-classes if the ideal size was not available. The species never competed for the same size-class on a mussel bed.

Keywords: size selection, competition, blue mussel (*Mytilus edulis*), eider (*Somateria mollissima*), oystercatcher (*Haematopus ostralegus*), herring gull (*Larus argentatus*).

INTRODUCTION

On the tidal flats of the Wadden Sea (southern North Sea), the most significant avian predators on blue mussels are oystercatchers (*Haematopus ostralegus*), herring gulls (*Larus argen-*

tus) and common eiders (*Somateria mollissima*) (Hilgerloh 1997). Responding to different constraints, each species differ in the most taken size-class. Herring gulls, which are studied here, and eiders swallow the whole mussel while oystercatchers open the mussel and eat only the flesh. Thus, the gape of the bill and the diameter of the oesophagus could limit the size of mussels that could be swallowed by herring gulls and eiders. The thickness of the shell and the energetic value can be critical for all three species.

Mussel populations may show a large variation of size structure and abundance within and between years, which has consequences for their exploitation by different predators. Birds are known to feed on prey yielding the highest benefit for them which thus is the optimal prey (Krebs & Kacelnik 1991). The optimal size can vary very much between sites depending on the available mussels. There is, however, an ideal size that would be taken, if all mussel sizes were available at a site. Since mussel sizes and characteristics correlate with the size variation between the sites, the ideal size is not always available. At any given situation the birds feed on the optimal size.

The ideal size is expected to differ between the species according to the species specific constraints. If the ideal size is not available or too scarce, the bird predator can cope with smaller sizes. There is, however, a minimum size for each predator species. They would hardly derive any benefit from the prey below the minimum size, which accordingly they would not take (Krebs & Kacelnik 1991).

Ophelia 56 (2002)

ON THE DISTANCE BETWEEN THE DISTRIBUTIONS OF RANDOM SUMS

BERO ROOS,* *Universität Hamburg*

DIETMAR PFEIFER,** *Universität Oldenburg*

Abstract

In this paper, we consider the total variation distance between the distributions of two random sums S_M and S_N with different random summation indices M and N . We derive upper bounds, some of which are sharp. Further, bounds with so-called magic factors are possible. Better results are possible when M and N are stochastically or stop-loss ordered. It turns out that the solution of this approximation problem strongly depends on how many of the first moments of M and N coincide. As approximations, we therefore choose suitable finite signed measures, which coincide with the distribution of the approximating random sum S_N , if M and N have the same first moments.

Keywords: Binomial distribution; discrete Taylor formula; finite signed measures; Poisson approximation; random sums; sharpness results; stochastic order; stop-loss order; total variation distance.

AMS 2000 Subject Classification: Primary 60E05

Secondary 62E17; 60E15; 60F05

Journal of Applied Probability 40 (2003)

Möglichkeiten und Grenzen der mathematischen Schadenmodellierung^{*}

Von Dietmar Pfeifer, Oldenburg

Inhaltsübersicht

1. Theoretische Überlegungen
 - 1.1 Generelle Aspekte der Modellbildung
 - 1.2 Univariate Verteilungsmodelle
 - 1.3 Multivariate Verteilungsmodelle
2. Praktische Überlegungen
 - 2.1 US-Hurricanes
 - 2.2 Sturmschäden in Europa
 - 2.3 Abhängige Gefahren
3. Resumee und Ausblick

Zeitschrift für die gesamte Versicherungswissenschaft, Heft 4 (2003)

Modeling dependence in finance and insurance: the copula approach

Dietmar Pfeifer and Johana Nešlehová (Oldenburg)

1. Introduction

The concept of “copulas” is not really a new one in the mathematical world. It goes essentially back to problems posed by Hoeffding and Fréchet more than 60 years ago concerning the maximal and minimal possible correlation for bivariate distributions when the marginals are fixed, or, more generally, upper and lower bounds for the joint distribution under this condition. A related question originally posed by the Italian geometrist Pompilj concerns the minimum of the L^α -distance of random variables X and Y , i.e. the minimum value of $E|X - Y|^\alpha$ when the marginal distributions F_X and F_Y are given. It was actually this approach which lead to the concept of “minimal metrics” in the important field of probabilistic metric spaces and the thorough investigation of copulas, see e.g. Schweizer and Sklar [24] or Zolotarev [27] for an extended exposition. The word “copula” however was seemingly first coined by Sklar in his famous 1959 paper [26]. (For a nice survey over the historic development of copulas we refer to the monograph of Dall’Aglio et al. [5].)

Blätter der DGVFM Band XXVI, Heft 2 (2003)



Maximum Likelihood Estimators in a Statistical Model of Natural Catastrophe Claims with Trend

ALEXANDER KUKUSH

Kiev National Taras Shevchenko University, Volodymyrska 60, 01033 Kiev, Ukraine
E-mail: alexander_kukush@univ.kiev.ua

YURI CHERNIKOV

Celenia Software, Rybalska 22, 01011 Kiev, Ukraine
E-mail: chernikov@ukr.net

DIETMAR PFEIFER

University Oldenburg, Fachbereich Mathematik, P.O. Box 2503, D-26111 Oldenburg, Germany
E-mail: pfeifer@mathematik.uni-oldenburg.de

[Received October 1, 2002; Revised April 27, 2005; Accepted May 18, 2005]

Abstract. A statistical model to analyse stochastically increasing claims arising out of natural catastrophes is presented. Based on record values, the exponential trends over time can be identified. A more specific three-parameter model involving such a trend is also proposed. Observed claims are modeled as a stochastically increasing sequence of Fréchet distributed random variables. Consistency and asymptotic normality of the joint maximum likelihood estimator are shown. Possible applications in forecasting of claims are indicated. In particular claims data from U.S. hurricanes and Japanese taifuns are discussed.

Key words. catastrophe claims, Fréchet distribution, maximum likelihood estimator, Nevzorov's record model, trend

**MODELING AND GENERATING DEPENDENT RISK PROCESSES
FOR IRM AND DFA**

BY

DIETMAR PFEIFER

Institut für Mathematik

Carl von Ossietzky Universität, Oldenburg

and AON Jauch&Hübener, Hamburg, Germany

AND

JOHANNA NEŠLEHOVÁ

RiskLab

ETH-Zentrum, Zürich, Switzerland

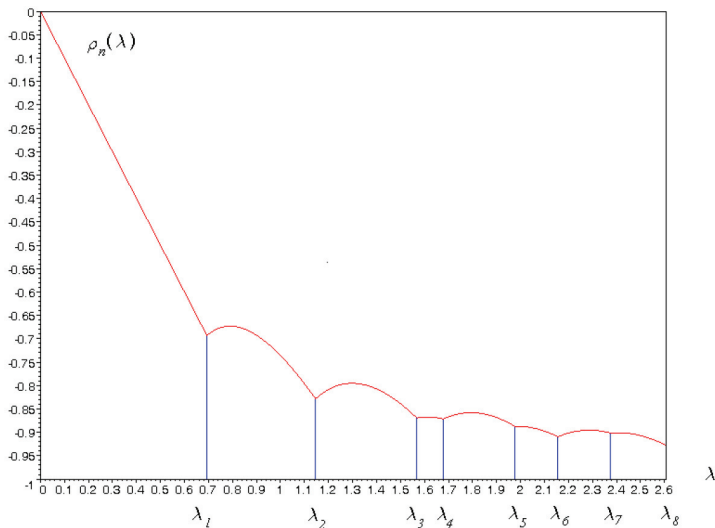
ABSTRACT

Modern Integrated Risk Management (IRM) and Dynamic Financial Analysis (DFA) rely in great part on an appropriate modeling of the stochastic behavior of the various risky assets and processes that influence the performance of the company under consideration. A major challenge here is a more substantial and realistic description and modeling of the various complex dependence structures between such risks showing up on all scales. In this presentation, we propose some approaches towards modeling and generating (simulating) dependent risk processes in the framework of collective risk theory, in particular w.r.t. dependent claim number processes of Poisson type (homogeneous and non-homogeneous), and compound Poisson processes.

ASTIN Bulletin 34 (2004)

n	$\rho_n(\lambda)$	λ_n	is the positive solution of
1	$-\lambda$	0.6931	$1 - 2e^{-\lambda} = 0$
2	$\frac{1 - 2e^{-\lambda}}{\lambda} - \lambda$	1.1461	$1 - (2 + \lambda)e^{-\lambda} = 0$
3	$\frac{3 - (6 + 2\lambda)e^{-\lambda}}{\lambda} - \lambda$	1.5681	$2 - (4 + 2\lambda + \lambda^2)e^{-\lambda} = 0$
4	$\frac{5 - (10 + 4\lambda + \lambda^2)e^{-\lambda}}{\lambda} - \lambda$	1.6783	$1 - (2 + 2\lambda)e^{-\lambda} = 0$
5	$\frac{6 - (12 + 6\lambda + \lambda^2)e^{-\lambda}}{\lambda} - \lambda$	1.9761	$2 - \left(4 + 2\lambda + \lambda^2 + \frac{1}{3}\lambda^3\right)e^{-\lambda} = 0$
6	$\frac{8 - \left(16 + 8\lambda + 2\lambda^2 + \frac{1}{3}\lambda^3\right)e^{-\lambda}}{\lambda} - \lambda$	2.1559	$2 - (4 + 4\lambda + \lambda^2)e^{-\lambda} = 0$
7	$\frac{10 - \left(20 + 12\lambda + 3\lambda^2 + \frac{1}{3}\lambda^3\right)e^{-\lambda}}{\lambda} - \lambda$	2.3761	$2 - \left(4 + 2\lambda + \lambda^2 + \frac{1}{3}\lambda^3 + \frac{1}{12}\lambda^4\right)e^{-\lambda} = 0$
8	$\frac{12 - \left(24 + 14\lambda + 4\lambda^2 + \frac{2}{3}\lambda^3 + \frac{1}{12}\lambda^4\right)e^{-\lambda}}{\lambda} - \lambda$	2.6082	$2 - \left(4 + 4\lambda + \lambda^2 + \frac{1}{3}\lambda^3\right)e^{-\lambda} = 0$

Tabellarische Darstellung der Minimalkorrelation zwischen Poisson-verteilten Risiken mit gleichem Parameter λ



Graphische Darstellung der Minimalkorrelation zwischen Poisson-verteilten Risiken mit gleichem Parameter λ

Strichlisten bei Laplace-Experimenten – zum Paradox der ungleichmäßigen Verteilung

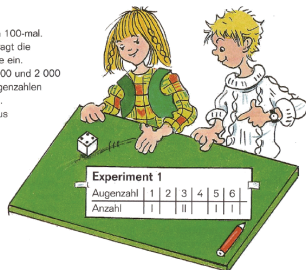
DIETMAR PFEIFER, OLDENBURG

Zusammenfassung: Das Themengebiet „Zufall und Wahrscheinlichkeit“ wird auf elementarem Niveau inzwischen bereits in der Grundschule behandelt. Dabei stehen experimentelle Aufgaben wie wiederholtes Werfen eines Würfels oder das Ziehen aus Dosen mit Kugeln unterschiedlicher Farbe im Vordergrund, meist in Verbindung mit der Erfassung der Ergebnisse in Strichlisten. Bei Laplace-Experimenten wie dem wiederholten Würfelwurf herrscht dabei – motiviert durch das Gesetz der großen Zahlen – intuitiv die Vorstellung, dass sich die Zählergebnisse in den Strichlisten einigermaßen „gleichmäßig“ verteilen. Dies ist aber paradoxerweise nicht so, wie in dem vorliegenden Beitrag gezeigt werden soll.

1. Das Problem und seine Herkunft

In dem letzten Kapitel Zufall und Wahrscheinlichkeit des neuen Unterrichtswerks **Mathebaum 4** findet sich im einführenden Abschnitt **Zufallsexperimente** die folgende Aufgabe:

- Maren und Lukas würfeln 100-mal. Macht es genauso und tragt die Ergebnisse in eine Tabelle ein.
- Vermutet, wie oft bei 1 000 und 2 000 Würfen die einzelnen Augenzahlen ungefähr erreicht werden.
- Addiert die Ergebnisse aus euren Gruppen und prüft die Vermutung.



Betrachtet man die einzelnen Augenzahl-Listen separat, so ist klar, dass die Anzahl der Striche in jeder solchen Liste einer $\mathcal{B}(n, p)$ -Binomialverteilung genügt mit $n=100$ und $p=1/6$, woraus sich eine erwartete Strich-Anzahl von $n \cdot p = 100/6 = 16,6$ ergibt. Bekanntlich liegt der maximale Wert der zugehörigen Elementarwahrscheinlichkeiten bei $[(n+1)p]$ (abgerundeter Wert), wenn $(n+1)p$ nicht ganzzahlig ist (vgl. BARTH und HALLER (1998), Satz 245.1). In dem hier betrachteten konkreten Beispiel liegt dieser Wert bei 16. Es ist also nahe liegend, zu vermuten, dass sich die Striche in den Augenzahl-Listen um diesen Wert herum einpendeln werden. Dies ist auch der Hintergrund des Frageteils b). Allerdings werden die Schüler bei der Durchführung des Experiments enttäuscht werden, denn diese idealisierte gleichmäßige Verteilung wird sich bei 100 vorgegebenen Wurf-Wiederholungen mit sehr hoher Wahrscheinlichkeit gerade nicht einstellen. Rückfragen bei der Lehrkraft werden hier vermutlich keine Klarheit bringen und eher für Verwirrung auf allen Seiten sorgen.

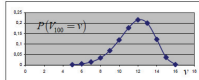
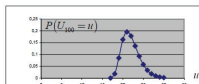
für das Maximum $U_n = \max_{1 \leq i \leq n} \{Z_{n,i}\} = S_{n,n}$ der Strichliste mit Wert u bzw. durch

$$K_n(u) = \left\{ (i_1, \dots, i_n) \mid i_1 = \dots = i_n = u, \sum_{j=1}^n i_j = n \right\}$$

für das Minimum $V_n = \min_{1 \leq i \leq n} \{Z_{n,i}\} = S_{n,1}$ der Strichliste mit Wert v . Die folgenden Tabellen listen einige der zugehörigen Ergebnisse auf.

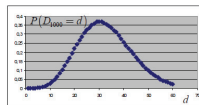
u	$P(U_{100}=u)$	v	$P(V_{100}=v)$
17	0.00042094	5	0.0017449
18	0.01834261	6	0.0055238
19	0.08565795	7	0.0148055
20	0.16341641	8	0.0342032
21	0.19574612	9	0.0668687
22	0.17826556	10	0.1196188
23	0.13644896	11	0.1775653
24	0.09282839	12	0.2160462
25	0.05795118	13	0.2000018
26	0.03380292	14	0.1223051
27	0.01860631	15	0.03659704
28	0.0097157	16	0.00233859
29	0.00482646		
30	0.0022848		

Als maximaler Wert der Strichliste ergibt sich „am häufigsten“ 21, als minimaler Wert „am häufigsten“ 12.



4. Diskussion

Das angesprochene Problem zeigt, dass man bei einer intuitiven Argumentation, die bei der Einführung eines neuen Stoffgebiets sicher angemessen und auch notwendig ist, gerade in der Stochastik aufpassen muss, weil sie hier manchmal zu falschen Vorstellungen führen kann. Der vorliegende Aufsatz soll deshalb einen Beitrag dazu leisten, sich auch kritisch mit Lehrinhalten auseinander zu setzen, gerade dann, wenn Schüler aufgrund ihres Alters selbst noch nicht in der Lage sind, sich die Unterschiede zwischen realer Beobachtung und intuitiver Vorstellung zu erklären. Bei der konkreten Aufgabe dürfte es auch schwierig sein, den Frageteil b) oben richtig zu beantworten. Die nachfolgende Graphik zeigt die empirische Verteilung der Spannweite bei einer Würfelwurfzahl von $n=1000$. Hierfür wurden insgesamt 1 Milliarde (!) Würfelwürfe mit dem PC simuliert, das entspricht der Simulation von 1 Million Strichlisten.



Die intuitiv erwartete „gleichmäßigere“ Verteilung der Werte der Strichliste (wegen der höheren Modalzahl von $n=1000$) zeigt sich mit einem Modalwert von etwa 30 für die Spannweite D_{1000} auch hier noch nicht sehr deutlich.

Literatur

- Barth, F.; Haller, R. (1998): Stochastik-Leistungskurs. Oldenbourg Verlag, München.
- Henze, N. (2003): Stochastik für Einsteiger. Vieweg Verlag, Braunschweig.
- Höbner, G. et al. (2004): Mathebaum 4. Schroedel Verlag, Braunschweig.

Stochastik in der Schule 26 (2006)



Applied Section

Solvency II: stability problems with the SCR aggregation formula

DIETMAR PFEIFER* and DOREEN STRASSBURGER

Institute of Mathematics, Carl von Ossietzky University, Oldenburg, Germany

(Accepted 4 October 2007)

One of the central issues in the Solvency II process will be an appropriate calculation of the Solvency Capital Requirement (SCR). This is the economic capital that an insurance company must hold in order to guarantee a one-year ruin probability of at most 0.5%. In the so-called standard formula, the overall SCR is calculated from individual SCRs in a particular way that imitates the calculation of the standard deviation for a sum of normally distributed risks (SCR aggregation formula). However, in order to cope with skewness in the individual risk distributions, this formula must be calibrated accordingly in order to maintain the prescribed level of confidence. In this paper, we want to show that the methods proposed and discussed thus far still show stability problems within the general setup.

Keywords: Solvency II; SCR; Skewness; Symmetry; Calibration; Copulas

CHAPMAN & HALL/CRC FINANCE

Handbook of Solvency for Actuaries and Risk Managers Theory and Practice

Arne Sandström

 **CRC Press**
Taylor & Francis Group
A CHAPMAN & HALL BOOK

15.3.5.1 General Calibration Problems

The general problem in using the baseline approach for aggregating the different risk modules as proposed for the European SCR, as discussed above, is discussed in Pfeifer and Strassburger (2008). They show that it is not sufficient for general purposes with a calibration of the standard SCR for the skewness. For the class of risk distributions they consider,

the square root formula tends to underestimate, and in some cases also overestimate, the true SCR. This is a general drawback of the baseline formula outside the world of elliptical risk distributions.

In their view, the only reasonable “all purpose” calibration is to apply the maximum value of one to all correlations, that is, having $C_{SCR} = \sum_{i=1}^d C_i$.

In the next chapter we will introduce another way of calibrating the CR formula using the above-mentioned Cornish–Fisher expansion.

26.4.1 Standard Formula Dependence

Only for a restricted class of distributions the aggregation with linear correlation coefficients produces the correct result. In the literature, numerous examples can be found where linear correlations do not well reflect the dependence between distributions and the use of linear correlations lead to wrong or even absurd aggregation results; see, for example, Embrecht et al. (2002) and Pfeifer and Strassburger (2008), but also Chapter 15. Two main reasons for this aggregation problem are

- The dependence between the distributions is not linear; for example, there are tail dependencies
- The shape of the marginal distributions is significantly different from the normal distribution; for example, the distributions are skewed

Both characteristics are shared by many of the main risks and subrisks in the Solvency II model. A discussion on the aggregation technique is done in CEIOPS (2009e12, 2010a09). They have observed that under the financial crisis, mainly 2007–2009, market parameters such as credit spreads, property prices, equity prices, and currency exchange rates, simultaneously showed strong adverse changes. These changes are not observed under normal conditions. It also became apparent that a change in one parameter had a reinforcing effect on the deterioration of other parameters.

The underlying distributions are not normal, but usually skewed and often truncated by reinsurance or hedging.



Sendung zum Lotto-Jackpot von NDR 3 am 31.1.2009, 19:30 Uhr

<http://www>

otto.mov

Ein Reservierungsverfahren für die Rechtsschutzversicherung nach Art der Lebensversicherung

Dietmar Pfeifer · Soeren Henniges ·
Doreen Straßburger · Alice Winkel

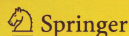
Online publiziert: 28. August 2012
© Springer-Verlag 2012

Zusammenfassung In der vorliegenden Arbeit wird eine dem spezifischen Geschäftsmodell der Rechtsschutzversicherung angepasste Verfahrensweise zur Berechnung von Reserven für bekannte gemeldete und unbekannte, nachgemeldete Schadenfälle vorgestellt. Ein wesentliches Merkmal ist hierbei die Trennung von Fallzahlen und durchschnittlichen Fallkosten, die eine genauere Modellierung des Abwicklungsverhaltens ermöglichen und die spezifischen Besonderheiten der Rechtsschutzversicherung besser abbilden. Im Unterschied zu anderen Reservierungsverfahren für Fallzahlen und Regulierungskosten, die in der Praxis häufig ohne Überprüfung der Voraussetzungen zur Anwendbarkeit verwendet werden, wird hier ein der Lebensversicherung nachgebildetes Modell gewählt, bei dem die Zeit bis zur Schließung eines Falls durch eine „Lebensdauerverteilung“ modelliert wird. Die Trennung in die genannten Modellkomponenten erlaubt auch eine einfache, den gesetzlichen Forderungen nachkommende Berücksichtigung von Teuerungseffekten. Schließlich lässt sich das vorgestellte Verfahren auch in einfacher, aber angemessener Weise auf die Bestimmung der Nachmeldereserve übertragen.

Claudia Becker
Roland Fried
Sonja Kuhn *Editors*

Robustness and Complex Data Structures

Festschrift in Honour of Ursula Gather



Chapter 18 Correlation, Tail Dependence and Diversification

Dietmar Pfeifer

18.1 Introduction

What is frequently abbreviated as Solvency II is perhaps the most challenging legislative adventure in the European Union (besides Basel II/III for the banking sector) in the last decade. It is a fundamentally new, risk driven approach towards a harmonization of financial regulation for insurance and reinsurance companies writing business in the European Union. One of the major aims of the Solvency II framework is a customer protection limiting the yearly ruin probability of the company to at most 0.5 % by requiring sufficient economic capital. The calculation of this so called Solvency Capital Requirement (SCR) is based on a complicated mathematical and statistical framework derived from an economic balance sheet approach (for more details, see, e.g., Buckham et al. 2011; Cruz 2009; Doff 2007 or Sandström 2006). An essential aspect in the SCR calculation here is the notion of diversification, which aims at a reduction of the overall capital requirement by “distributing” risk in an appropriate way. There are several definitions and explanations of this term, some of which are presented in the sequel.

“Although it is an old idea, the measurement and allocation of diversification in portfolios of asset and/or liability risks is a difficult problem, which has so far found many answers. The diversification effect of a portfolio of risks is the difference between the sum of the risk measures of stand-alone risks in the portfolio and the risk measure of all risks in the portfolio taken together, which is typically non-negative, at least for positive dependent risks.”

[Hürlimann (2009a, p. 325)]

“Diversification arises when different activities complement each other, in the field of both return and risk. [...] The diversification effect is calculated by using correlation factors. Correlations are statistical measures assessing the extend to which events could occur simultaneously. [...] A correlation factor of 1 implies that certain events will always occur

D. Pfeifer (✉)
Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany
e-mail: dietmar.pfeifer@uni-oldenburg.de

C. Becker et al. (eds.), *Robustness and Complex Data Structures*,
DOI 10.1007/978-3-642-35494-6_18, © Springer-Verlag Berlin Heidelberg 2013

301

Table 18.1 Joint distribution of risks

$P(X = x, Y = y)$	x			$P(Y = y)$	$P(Y \leq y)$
	0	50	100		
y	β	$0.440 - \beta$	0.000	0.440	0.440
40	$0.554 - \beta$	β	0.001	0.555	0.995
50	0.000	0.001	0.004	0.005	1.000
$P(X = x)$	0.554	0.441	0.005		
$P(X \leq x)$	0.554	0.995	1.000		

Table 18.2 Moments and correlations

$E(X)$	$E(Y)$	$\sigma(X)$	$\sigma(Y)$	$\rho(\beta) = \rho(X, Y)$
22.550	22.450	25.377	19.912	-0.9494 + 3.9579 β

Table 18.3 Distribution of aggregate risk

s	0	40	50	90	100	140	150
$P(S = s)$	β	$0.554 - \beta$	$0.440 - \beta$	β	0.001	0.001	0.004
$P(S \leq s)$	β	0.554	$0.994 - \beta$	0.994	0.995	0.996	1.000

Example 18.3 Let the joint distribution of the non-negative risks X and Y be given by Table 18.1, with $0 \leq \beta \leq 0.440$, giving $\text{VaR}_0(X) = 50$, $\text{VaR}_0(Y) = 40$.

For the moments of X and Y , we obtain the values in Table 18.2 (with σ denoting the standard deviation). This shows that the range of possible risk correlations is the interval $[-0.9494; 0.7921]$, with a zero correlation being attained for $\beta = 0.2399$.

Table 18.3 shows the distribution of the aggregated risk $S = X + Y$.

We thus obtain a risk concentration due to $\text{VaR}_0(S) = 100 > 90 = \text{VaR}_0(X) + \text{VaR}_0(Y)$, independent of the parameter β and hence also independent of the possible correlations between X and Y .

A closer look to the joint distribution of X and Y shows that the reason for this perhaps unexpected result is the fact that although one can have a “diversification effect” in the central body of the distribution, where a fraction of little less than $1 - \alpha$ of the risk pairs are located, the essential “concentration effect”, however, is caused by a joint occurrence of very high losses, with a fraction of α of all risk pairs.

The following result is related to the consideration of “worst VaR scenarios” as in McNeil et al. (2005, Sect. 6.2).

Proposition 18.3 Let X and Y be non-negative risks with cumulative distribution functions F_X and F_Y , respectively, which are continuous and strictly increasing on

their support. Denote, for a fixed $\alpha \in (0, 1)$,

$$Q^*(\alpha, \delta) := \min \{ Q_X(u) + Q_Y(2 - \alpha - \delta - u) \mid 1 - \alpha - \delta \leq u \leq 1 \} \\ \text{for } 0 \leq \delta < 1 - \alpha.$$

Then there exists a sufficiently small $\epsilon \in (0, 1 - \alpha)$ with the property

$$Q^*(\alpha, \epsilon) > Q_X(1 - \alpha) + Q_Y(1 - \alpha) = \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y). \quad (18.6)$$

Assume further that the random vector (U, V) has a copula C as joint distribution function with the properties

$$V < 1 - \alpha - \epsilon \iff U < 1 - \alpha - \epsilon \quad \text{and} \\ V = 2 - \alpha - \epsilon - U \iff U \geq 1 - \alpha - \epsilon. \quad (18.7)$$

If we define

$$X^* := Q_X(U), \quad Y^* := Q_Y(V), \quad S^* := X^* + Y^*,$$

then the random vector (X^*, Y^*) has the same marginal distributions as (X, Y) , and it holds

$$\text{VaR}_\alpha(X^* + Y^*) \geq Q^*(\alpha, \epsilon) > \text{VaR}_\alpha(X^*) + \text{VaR}_\alpha(Y^*) = \text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y),$$

i.e., there is a risk concentration effect. Moreover, under the assumption (18.7), the correlation $\rho(X^*, Y^*)$ is minimal if $V = 1 - \alpha - \epsilon - U$ for $U < 1 - \alpha - \epsilon$ (lower extremal copula $\underline{C}$) and maximal if $V = U$ for $U < 1 - \alpha - \epsilon$ (upper extremal copula $\bar{C}$).

Proof By assumption, the (non-negative) quantile functions Q_X and Q_Y are continuous and strictly increasing over the interval $[0, 1]$ (with a possibly infinite value at the point 1), so that

$$Q^*(\alpha, 0) \\ = \min \{ Q_X(u) + Q_Y(2 - \alpha - u) \mid 1 - \alpha \leq u \leq 1 \} > Q_X(1 - \alpha) + Q_Y(1 - \alpha),$$

the minimum being actually attained. Since by the continuity assumptions above, $Q^*(\alpha, \epsilon)$ is continuous in ϵ and decreasing when ϵ is increasing, relation (18.6) follows.

The copula construction above now implies that

$$P(S^* \leq s) \leq 1 - \alpha - \epsilon \\ \text{for } s \leq Q_X(1 - \alpha - \epsilon) + Q_Y(1 - \alpha - \epsilon) = \text{VaR}_{\alpha+\epsilon}(X) + \text{VaR}_{\alpha+\epsilon}(Y), \\ P(S^* \leq s) = 1 - \alpha - \epsilon \quad \text{for } \text{VaR}_{\alpha+\epsilon}(X) + \text{VaR}_{\alpha+\epsilon}(Y) < s < Q^*(\alpha, \epsilon), \\ P(S^* \leq s) \geq 1 - \alpha - \epsilon \quad \text{for } s \geq Q^*(\alpha, \epsilon). \quad (18.8)$$

Table 18.4 Examples of risk measures and correlations for various values of σ

σ	$\text{VaR}_\alpha(X)$ $= \text{VaR}_\alpha(Y)$	$\text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y)$	$\text{VaR}_\alpha(X^* + Y^*)$	$\rho_{\min}(X^*, Y^*)$	$\rho_{\max}(X^*, Y^*)$
0.1	1.2873	2.5746	2.6205	-0.8719	0.9976
0.2	1.6408	3.2816	3.3994	-0.8212	0.9969
0.3	2.0704	4.1408	4.3661	-0.7503	0.9951
0.4	2.5866	5.1732	5.5520	-0.6620	0.9920
0.5	3.1992	6.3984	6.9901	-0.5598	0.9873
0.6	3.9177	7.8354	8.7134	-0.4480	0.9802
0.7	4.7497	9.4994	10.7537	-0.3310	0.9700
0.8	5.7011	11.4022	13.1401	-0.2136	0.9556
0.9	6.7750	13.5500	15.8969	-0.1002	0.9362
1.0	7.9712	15.9424	19.0412	0.0050	0.9108
1.5	15.4675	30.9350	40.4257	0.3127	0.6839
2.0	23.3748	46.7496	66.8923	0.2723	0.3794
2.5	27.5107	55.0214	86.2673	0.1399	0.1637
3.0	25.2162	50.4324	86.7034	0.0565	0.0611

Relation (18.8) in turn implies that $\text{VaR}_\alpha(S^*) = \text{VaR}_\alpha(X^* + Y^*) \geq Q^*(\alpha, \epsilon)$ which proves the first part of Proposition 18.3, due to relation (18.6).

The remainder part follows from Theorem 5.25 in McNeil et al. (2005) when looking at the conditional distribution of (X^*, Y^*) given the event $\{U < 1 - \alpha - \epsilon\}$. $\square$

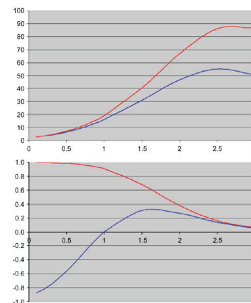
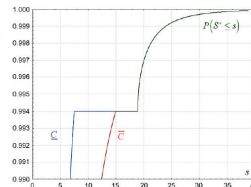
Note that both types of copulas that provide the extreme values for the correlations, $\underline{C}$ and $\bar{C}$, are of the type “shuffles of M ”, see Nelsen (2006, Sect. 3.2.3).

In the following example, we shall show some consequences of Proposition 18.3 in the case of lognormally distributed risks, which are of special importance for Pillar One under Solvency II, see, e.g., Hürfmann (2009a, 2009b).

Example 18.4 To keep things simple and comparable with Solvency II specifications, we shall assume that X and Y follow the same lognormal distribution $\mathcal{LN}(\mu, \sigma)$ with $\mu \in \mathbb{R}$, $\sigma > 0$ and $E(X) = E(Y) = 1$ which corresponds to the case $\mu = -\sigma^2/2$. Table 18.4 shows all relevant numerical results for the extreme copulas $\underline{C}$ and $\bar{C}$ in Proposition 18.3, especially the maximal range of correlations induced by them. According to the Solvency II standard, we choose $\alpha = 0.005$ (and $\epsilon = 0.001$, which will be sufficient here).

Note that the bottom graph in Fig. 18.1 resembles the graph in Fig. 5.8 in McNeil et al. (2005).

The graph in Fig. 18.2 shows parts of the two cumulative distribution functions under the extreme copulas $\underline{C}$ and $\bar{C}$ for $S^* := X^* + Y^*$ in the case $\sigma = 1$. Note that especially for smaller values of σ (which is typical for the calculation of the SCR in

Fig. 18.1 Top: Graph of $\text{VaR}_\alpha(X^* + Y^*)$ (red) and $\text{VaR}_\alpha(X) + \text{VaR}_\alpha(Y)$ (blue) as functions of σ ; bottom: graph of $\rho_{\min}(X^*, Y^*)$ (red) and $\rho_{\max}(X^*, Y^*)$ (blue) as functions of σ **Fig. 18.2** Graph of cumulative distribution functions for extreme copulas for $S^* = X^* + Y^*$ with $\sigma = 1$ 

the non-life risk module of Solvency II) the range of possible negative and positive correlations between the risks is quite large, with the same significant discrepancy between the Value at Risk of the aggregated risks and the sum of individual Values at Risk.

Note also that any correlation value ρ of $\rho(X^*, Y^*)$ between $\rho_{\min}(X^*, Y^*)$ and $\rho_{\max}(X^*, Y^*)$ can be achieved by a proper mixture of the extreme copulas $\underline{C}$ and $\bar{C}$.



FMA AUFSICHTSKONFERENZ 2012

**Regulierung im Wandel:
Von der quantitativen
zur qualitativen Aufsicht –
Erfahrungen und
Herausforderungen**

RAFF/SMREKAR/KALB/PFEIFER/BRAUMÜLLER

WORKSHOP 2
**SOLVENCY II ALS TOOL DER
UNTERNEHMENSSTEUERUNG**

DI MANFRED RAFF, VORSTANDSMITGLIED SPARKASSEN
VERSICHERUNG AG VIG, **MAG. THOMAS SMREKAR**,
PARTNER, KPMG, **DR. STEPHAN KALB**, SENIOR DIRECTOR
FITCH RATINGS, **PROF. DR. DIETMAR PFEIFER**, UNIVER-
SITÄT OLDENBURG, **DR. PETER BRAUMÜLLER**, BEREICHS-
LEITER VERSICHERUNGS AUFSICHT UND PENSIONS KASSEN-
AUFSICHT, FMA

PETER BRAUMÜLLER: Herzlich willkommen zu unserem Workshop zum Thema Solvency II als Tool der Unternehmenssteuerung. Wir haben am Vormittag schon einiges gehört, das im weiteren Sinne durchaus dazu passt. Ich habe vernommen, dass sich in den Unternehmen sehr viele Personen mit anderen Themen beschäftigen müssen und dass sie offensichtlich schon Schwierigkeiten haben, genügend Mitarbeiter zu finden, die das Risikomanagement und im weiteren Sinne die Führung des Unternehmens wahrnehmen können. Ebenfalls haben wir vernommen, wie wichtig Ports sind, wie wichtig die Führung und Steuerung des Unternehmens für den Erfolg oder Misserfolg eines Unternehmens sind. Eine zusätzliche Erkenntnis – die wir im Übrigen vor einigen Jahren im Bereich der Aufsichtsbehörden im Rahmen einer Untersuchung gewonnen haben – konnten wir in Erfahrung bringen: Vieles wurde in breiterem Sinn auch durch Managementfehler verursacht oder verstärkt. Und darüber hinaus habe ich noch mitbekommen, es sollte „The company wants a model“ und nicht „The model wants the company“ heißen. Ich glaube, dass alle Punkte sehr gut zu unserem Thema hier passen werden. Wir wollen uns Solvency II als neues Aufsichtssystem vor dem Hintergrund anschauen, inwieweit dieses neue System kompatibel mit



Workshop 2: Dr. Stephan Kalb, Dr. Manfred Rapp, Dr. Peter Braumüller, Mag. Thomas Smeke und Prof. Dr. Dietmar Pfeifer (v.l.)

Fragen ist, die sich in der Unternehmenssteuerung stellen. Und wir haben das Vergnügen, dieses aus verschiedenster Perspektive tun zu können. Wir werden es aus der Sicht des Versicherers, aus der Sicht der Abschlussprüfung, aus der Sicht der Ratingagentur und last but not least auch aus wissenschaftlicher Perspektive beleuchten und haben dann hoffentlich noch etwas Zeit für eine Diskussion. Ich darf das Wort an Dr. Rapp übermitteln.

MANFRED RAPP: Das Thema lautet Solvency II als Tool der Unternehmenssteuerung. Ich habe mich zu Beginn gefragt, ob ich hinter den Titel ein Satzzeichen setzen soll, das vielleicht für die weiteren Ausführungen indikativ ist. So habe ich es mit einem Rufzeichen probiert, was im Schach ein Zeichen für einen guten Zug ist – das hat mir nicht gefallen. Ein Punkt passt irgendwie gar nicht, mir war am aller nächsten, dahinter ein großes Fragezeichen zu setzen. Und im Schach würden wir wahrscheinlich noch ein zweites dazusetzen als Zeichen für einen schlechten Zug. Nur dass Sie vielleicht vorweg erkennen können, wo ich stehe, und ich fürchte, es wird sich vom Tenor her auch nicht wahnsinnig viel ändern. Als analytisch denkender Mensch habe ich mir gedacht, wenn Solvency II ein Tool für die Unternehmenssteuerung werden soll, dann stelle ich gewisse Anforderungen daran. Und ich habe ein wenig Brainstorming betrieben – auch mit Kollegen. Welche Anforderungen würde man denn als Verantwortlicher für ein Unternehmen an ein Tool stellen, das man dann auch in der Praxis verwendet? Und

wir den ganzen Tag schon gehört, es ist alles viel zu kompliziert. Es ist kompliziert, es ist intransparent, es wurde vorher gesagt, wir haben am Anfang acht Seiten Dokumentationen gehabt, hinterher 400. Wenn Sie die ganzen Annexe dazunehmen, waren es, glaube ich, über 3.000 bedruckte Seiten Papier, um diese ganzen Technical Specifications durchzulesen. Das ist nicht transparent, es ist auch keine stringente Notation da. Da werden dann teilweise Formeln auf den Markt geworfen, da frage ich mich auch, was ist damit eigentlich gemeint? Das versteht man gar nicht mehr richtig. Es werden Parameter gesetzt aus politischer Motivation heraus, z. B. irgendwelche Faktoren auf die Rückstellungen, die dann aus irgendwelchen anderen Dingen kommen, z. B. aus dem Markt, die man in den Unternehmen gar nicht abbilden kann. Leider gibt es die unternehmensspezifischen Parameter nur noch auf Antrag. Das ist ein sehr komplizierter Prozess, um das im Endeffekt hinzubekommen. Wir haben fehlende Kosten-Nutzen-Überlegungen. Auch das ist heute schon mehrmals gesagt worden. Wenn ich beispielsweise das Rückversicherungsausfallrisiko angucke und was in QIS 4 und QIS 5 x-mal gerechnet werden musste. Immer wurde ein Teil herausgenommen, dann wurde neu gerechnet. Im Endeffekt haben Sie dann noch 0,5 Prozent des SCR verändert. Das macht eigentlich keinen Sinn. Eine Erhöhung der Komplexität produziert nicht automatisch eine höhere Präzision.

Mein Fazit ist schlussendlich: Ein risikoorientierter Bewertungsansatz ist prinzipiell möglich mit einem Standardansatz, aber der muss sehr viel einfacher sein. Die aktuelle Konzeption der Säule I ist viel zu kompliziert, sie ist mathematisch inkonsistent, die Auswirkungen sind nicht abschätzbar, die Berücksichtigung von Korrelationen ist mathematisch nicht gerechtfertigt, die korrekte Repräsentanz des Unternehmens oder der Risikolage des Unter-



DIETMAR PFEIFER: „Alle noch so komplexen mathematischen Modelle zur Beschreibung realer ökonomischer Prozesse sind letztendlich falsch. Also wäre es doch vernünftig, zur Erhöhung der Praktikabilität eher ein einfaches falsches Modell als ein komplexes falsches Modell auszuwählen.“

Piotr Jaworski
Fabrizio Durante
Wolfgang Karl Härdle *Editors*

Copulae in Mathematical and Quantitative Finance

Proceedings of the Workshop Held in
Cracow, 10-11 July 2012



Chapter 8 Singular Mixture Copulas

Dominic Lauterbach and Dietmar Pfeifer

Abstract We present a new family of copulas—the *Singular Mixture Copulas*. We begin with the construction of singular copulas whose supports lie on the graphs of two given quantile functions. These copulas are then mixed with respect to a continuous distribution resulting in a nonsingular parametric copula.

The Singular Mixture Copulas we construct have a Lebesgue density and in special cases even a closed form representation. Moreover, they have positive lower and upper tail dependence. Because Singular Mixture Copulas are mixtures of “simple” singular copulas, they can be simulated easily.

8.1 Introduction

Copulas provide an effective and versatile tool for modeling multivariate stochastic dependence. Since their introduction by Sklar in 1959 (see [11]) there have been intense developments in both the copula theory and their applications, see, e.g., [1, 5–7, 9, 10, 12].

In [10] several geometric methods of constructing copulas are presented. One approach deals with the construction of singular copulas whose supports lie in a given set. Another approach mixes an infinite family of copulas with respect to a mixing distribution. We present a new family of copulas—the *Singular Mixture Copulas*. These copulas result from a combination of the above-mentioned methods. In Sect. 8.2 we construct singular copulas whose supports lie on the graphs of two given quantile functions. These copulas are then mixed with respect to a continuous distribution resulting in an absolutely continuous parametric copula (Sect. 8.3).

D. Lauterbach (✉) · D. Pfeifer
Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany
e-mail: dominic.lauterbach@uni-oldenburg.de; dietmar.pfeifer@uni-oldenburg.de

P. Jaworski et al. (eds.), *Copulae in Mathematical and Quantitative Finance*,
Lecture Notes in Statistics 213, DOI 10.1007/978-3-642-35407-6_8,
© Springer-Verlag Berlin Heidelberg 2013

165

As mixing distribution we particularly use a generalized beta distribution, i.e., a linear transformation of a beta distribution. Section 8.4 summarizes the results and gives an outlook on some extensions of this approach.

8.2 Singular Copulas

Let F be a continuous distribution function on $[0, 1]$ and let α be some constant in $]0, 1[$. Then there exists a continuous function G such that

$$\alpha F(x) + (1 - \alpha)G(x) = x \quad (8.1)$$

for all $x \in [0, 1]$. The function G is given by

$$G(x) = \frac{x - \alpha F(x)}{1 - \alpha}. \quad (8.2)$$

In general, G is not necessarily a distribution function. However, we are interested in exactly this case.

Let us assume for a moment that G is also a distribution function. Let X be a random variable with a continuous uniform distribution on $[0, 1]$, and let I be a random variable, independent of X , with a binomial $B(1, \alpha)$ -distribution. Define the random variable Y via

$$Y := I \cdot F^{-1}(X) + (1 - I) \cdot G^{-1}(X). \quad (8.3)$$

Easy calculations show that Y follows a continuous uniform distribution on $[0, 1]$. As a consequence the distribution function of (X, Y) is a certain singular copula. So with two distribution functions F and G satisfying (8.1) we can construct singular copulas. Those copulas are given by

$$\begin{aligned} C_{X,Y}(x, y) &= \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x, I \cdot F^{-1}(X) + (1 - I) \cdot G^{-1}(X) \leq y) \\ &= \mathbb{P}(I = 1) \mathbb{P}(X \leq x, X \leq F(y)) + \mathbb{P}(I = 0) \mathbb{P}(X \leq x, X \leq G(y)) \\ &= \alpha \min(x, F(y)) + (1 - \alpha) \min(x, G(y)). \end{aligned}$$

As mentioned above, G is not necessarily a distribution function, so we have to make assumptions on F to guarantee that G is also a distribution function.

Lemma 8.1. Let F be a differentiable distribution function on $[0, 1]$. Then the function G given by (8.2) is a differentiable distribution function on $[0, 1]$ if and only if $F'(x) \leq \frac{1}{\alpha}$ for all $x \in [0, 1]$.

Proof. From $F(0) = 0$ and $F(1) = 1$ it follows immediately that $G(0) = 0$ and $G(1) = 1$. From (8.2) we have

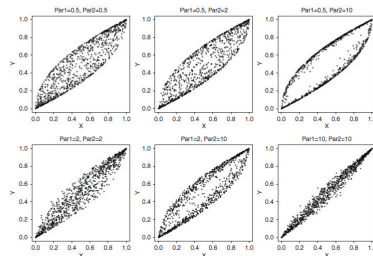


Fig. 8.3 Scatter plots of simulated points from a Singular Mixture Copula with generalized beta mixing distribution for $\alpha = 0.5$ and different shape parameters

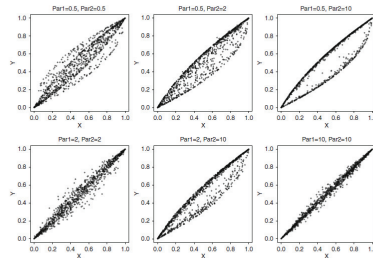


Fig. 8.4 Scatter plots of simulated points from a Singular Mixture Copula with generalized beta mixing distribution for $\alpha = 0.3$ and different shape parameters

From Bernstein polynomials to Bernstein copulas

C. Cottin
Department of Engineering and Mathematics
FH Bielefeld University of Applied Sciences
D-33615 Bielefeld, Germany
claudia.cottin@fh-bielefeld.de

D. Pfeifer
Department of Mathematics
School of Mathematics and Science
Carl von Ossietzky University Oldenburg
D-26111 Oldenburg, Germany
dietmar.pfeifer@uni-oldenburg.de

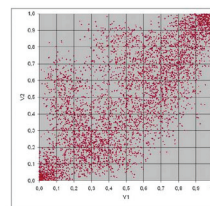
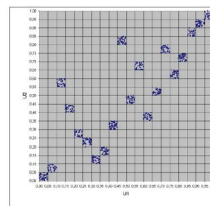
Dedicated to the 65th birthday of Professor Heiner Gonska

Abstract

In this paper we review Bernstein and checkerboard copulas for arbitrary dimensions and general grid resolutions in connection with discrete random vectors possessing uniform margins, and point out the relation to tensor product Bernstein operators. We further suggest a pragmatic and effective way to fit the dependence structure of multivariate data to Bernstein copulas via rook copulas, a subclass of checkerboard copulas, which is based on the multivariate empirical distribution.

2010 AMS Subject Classification: 41A10, 41A63, 60E05, 62E17, 62G30, 62P05.

Key Words and Phrases: Bernstein polynomials, Bernstein copulas, checkerboard copulas, construction of copulas, risk management applications



5000 simulated random vectors
following the rook copula (left) and the Bernstein copula (right)

A generalization of the procedure to arbitrary dimensions, replacing the rectangle $I_{[0,1]^2}$ by a general cube, is obvious. ♦

Note that according to a fundamental theorem in statistics, the empirical distribution function of a multivariate observation converges uniformly to the true cdf when the sample size increases. Likewise, the empirical copula based on the extracted marginal ranks converges uniformly to the true underlying copula. This implies that with an increasing number of observed data, the rook copulas as well as the Bernstein copulas with the discrete skeletons derived from the marginal ranks converge to the true underlying copula as well, since in both cases the grid constant m corresponds to the sample size.

References

- [1] G.A. ANASTASSIOU, S.G. GAL (2000): *Approximation Theory*. Moduli of Continuity and Global Smoothness Preservation. Birkhäuser, Basel.
- [2] T. BOUEZMARNI, J.V.K. ROMBOUTS, A. TAAMOUTI (2008): *Asymptotic properties of the Bernstein density copula for dependent data*. CORE discussion paper 2008/45, Leuven University, Belgium.
- [3] C. COTTIN, S. DÖHLER (2013): *Risikoanalyse*. Modellierung, Beurteilung und Management von Risiken mit Praxisbeispielen. 2. Aufl., Springer Spektrum, Heidelberg.
- [4] C. COTTIN, H.H. GONSKA (1993): *Simultaneous approximation and global smoothness preservation*. Rendiconti del Circolo Matematico di Palermo (2), Suppl. 33, 259 – 279.

Journal of Applied Functional Analysis (2014)



Deutscher Bundestag
Finanzausschuss
Die Vorsitzende

Herrn Prof. Dr. Dietmar Pfeifer
Fakultät V - Mathematik und
Naturwissenschaften Oldenburg
Postfach 25 03
26111 Oldenburg

Berlin, 17. November 2014
Geschäftszeichen: PA 7-18/2956
Anlagen: 6

Ingrid Arndt-Brauer, MdB
Platz der Republik 1
11011 Berlin
Telefon: +49 30 227-32468
Telefon: +49 30 227-37072
Fax: +49 30 227-36944 - 36206
finanz.ausschuss@bundestag.de
rene.probst@bundestag.de

Dienstgebäude:
Paul-Löbe-Haus
Konrad-Adenauer-Straße 1
10557 Berlin

Öffentliche Anhörung zu dem Gesetzentwurf der Bundesregierung „Entwurf eines Gesetzes zur Modernisierung der Finanzaufsicht über Versicherungen“ (Stichwort: Solvency II) - BT-Drucksache 18/2956 -

Sehr geehrter Professor Pfeifer,

der Finanzausschuss des Deutschen Bundestages hat in seiner 24. Sitzung am Mittwoch, dem 12. November 2014 beschlossen, eine öffentliche Anhörung zu dem o. g. Gesetzentwurf durchzuführen. Hierzu möchte ich Sie recht herzlich einladen.

Die öffentliche Anhörung findet am

**Mittwoch, dem 3. Dezember 2014, von 13.00 bis 15.00 Uhr
in Berlin im Marie-Elisabeth-Lüders-Haus
Anhörungssaal 3 101**

statt. Der Zugang zu dem Anhörungssaal erfolgt über den Eingang des Marie-Elisabeth-Lüders-Hauses in der Adele-Schreiber-Krieger-Str. 1/Schiffbauerdamm (Anlage 1). Bitte weisen Sie sich dort mit Ihrem Personalausweis aus.

Zum Verlauf der öffentlichen Anhörung darf ich Sie darüber informieren, dass der Finanzausschuss auf die Erarbeitung eines den Sachverständigen vorab zuzuleitenden Fragenkatalogs verzichtet hat. Es sollen zudem keine generellen einführenden Statements abgegeben werden. Es ist vielmehr beabsichtigt, Ihnen, den Experten, unmittelbar Fragen zu der Materie zu stellen. Der Gesetzentwurf auf BT-Drucksache 18/2956 ist als Anlage 2, die Liste der eingeladenen Sachverständigen als Anlage 3 beigelegt.

CARL
VON
OSSIEZKY
universität
OLDENBURG
FAKULTÄT V
MATHEMATIK- UND
NATURWISSENSCHAFTEN

Stellungnahme

zum Regierungsentwurf eines Gesetzes zur
Modernisierung der Finanzaufsicht über Versicherungen

Inhalt

1. Fachinhaltliche Kritik
 - 1.1. Grundsätzliche Anmerkungen
 - 1.2. Anmerkungen zu den Begriffsbestimmungen
 - 1.3. Anmerkungen zum Formelwerk für das SCR
- 1.4. Fazit
2. Sozialökonomische Kritik
3. Literatur

INSTITUT FÜR MATHEMATIK

Prof. Dr. Dietmar Pfeifer

Professur für Angewandte
Wahrscheinlichkeitstheorie

*Schwerpunkt Versicherungswissenschaften
und Finanzmathematik*

TELEFONDURCHWAHL

(0441) 7 98 - 32 29

EMAIL

dietmar.pfeifer@uni-oldenburg.de

OLDENBURG

25.11.2014



POSTANSCHRIFT

D-26111 Oldenburg

PAKETANSCHRIFT

Assenländer Heerstraße 114 - 118

D-26129 Oldenburg

TELEFON

(0441) 7980

INTERNET

www.uni-oldenburg.de

BANKVERBINDUNG

Landesbank für Ost- und Oldenburg (LZO)

Konto 1 988 112

BLZ 280 501 00

BIC BRLADE21HAN

IBAN DE 4528 0501 0001 0198 8112

1. Fachinhaltliche Kritik

Der vorliegende Gesetzentwurf entspricht einer Umsetzung der Richtlinie 2009/138/EG zur Harmonisierung der Versicherungsaufsicht in der Europäischen Union in nationales Recht. Ein wesentliches Ziel ist dabei, „sicherzustellen, dass es höchstens in einem von 200 Fällen zur Insolvenz kommen kann oder diese Unternehmen mit einer Wahrscheinlichkeit von 99,5 % in den kommenden zwölf Monaten weiterhin in der Lage sein werden, ihren Verpflichtungen gegenüber den Versicherungsnehmern und Begünstigten nachzukommen“ (Richtlinie 2009/138/EG, Erwägungsgrund (64), S. L 335/7). Dieser Erwägungsgrund wird in Artikel 99 (3) der Richtlinie konkretisiert und findet sich im vorliegenden Gesetzentwurf in leicht abgeänderter Form in § 97 (2) wieder:

„Die Solvabilitätskapitalanforderung muss so kalibriert werden, dass alle quantifizierbaren Risiken, denen ein Versicherungsunternehmen ausgesetzt ist, widerspiegelt werden. Dabei sind sowohl der aktuelle Geschäftsumfang als auch die in den nächsten zwölf Monaten erwarteten neuen Geschäfte zugrunde zu legen. In Bezug auf den aktuellen Geschäftsumfang deckt die Solvabilitätskapitalanforderung nur unerwartete Verluste ab. Sie entspricht dem Value-at-Risk der Basis eigenmittel eines Versicherungsunternehmens zu einem Konfidenzniveau von 99,5 Prozent über einen Zeitraum von einem Jahr.“

Sowohl die Richtlinie als auch der deutsche Gesetzentwurf heben in der Berechnung des Solvenzkapitals auf eine Risikobewertung durch Wahrscheinlichkeiten unter Bezug auf das Risikomaß „Value-at-Risk“ ab. Durch diesen konkreten Bezug zur Mathematik ist eine gewisse Objektivität bei der Bewertung von Risiken und der Berechnung des Solvenzkapitals beabsichtigt.

Hierzu ist allgemein anzumerken, dass die Bewertung ökonomischer Prozesse durch „Wahrscheinlichkeiten“ nicht unkritisch ist. Anders als bei physikalischen Vorgängen, deren Wiederholbarkeit unter gleichartigen Bedingungen prinzipiell möglich ist und daher „Wahrscheinlichkeiten“ unter Bezug auf das Gesetz der großen Zahlen als empirisch bestimmbar angesehen werden, sind diese Voraussetzungen bei ökonomischen Vorgängen in der Regel nicht gegeben. Eine „Wahrscheinlichkeit“ für das Fortbestehen eines Versicherungsunternehmens im folgenden Geschäftsjahr kann daher objektiv gar nicht ermittelt werden (vgl. hierzu z.B. N. Bouleau (2009), Kapitel XIV). Die in der Leitlinie und im vorliegenden Gesetzentwurf angestellten Überlegungen zur Bestimmung eines Solvenzkapitals mit einem „garantierten“ Konfidenzniveau von 99,5 % können sich daher nur auf theoretische Modellrechnungen beziehen, von denen man annimmt, dass diese - zumindest näherungsweise - auf die ökonomische Wirklichkeit über-

tragen werden können. Außerdem ist - anders als im Wortlaut des Gesetzentwurfs - das Solvenzkapital im wesentlichen die Differenz zwischen dem zugrunde liegenden Risikomaß und dem Erwartungswert der unter Risiko stehenden Bezugsgrößen (versicherungstechnisches Risiko, Marktrisiko usw.), siehe z.B. A. Sandström (2011), Kapitel 5.2, S. 61. Basis eigenmittel besitzen keinen echten Value-at-Risk, von daher ist die gewählte Formulierung in § 97 (2) irreführend. Die Formulierung in § 97 (2) sollte daher besser durch folgende inhaltlich korrekte Alternative ersetzt werden:

„Die Solvabilitätskapitalanforderung muss so kalibriert werden, dass alle quantifizierbaren Risiken, denen ein Versicherungsunternehmen ausgesetzt ist, widerspiegelt werden. Dabei sind sowohl der aktuelle Geschäftsumfang als auch die in den nächsten zwölf Monaten erwarteten neuen Geschäfte zugrunde zu legen. In Bezug auf den aktuellen Geschäftsumfang deckt die Solvabilitätskapitalanforderung nur unerwartete Verluste ab. Sie **orientiert sich an der Differenz des Value-at-Risk zu den Erwartungswerten der unter Risiko stehenden Bezugsgrößen** eines Versicherungsunternehmens zu einem Konfidenzniveau von 99,5 Prozent über einen Zeitraum von einem Jahr.“

1.2. Anmerkungen zu den Begriffsbestimmungen

In dem vorliegenden Gesetzentwurf werden in § 7 Begriffsbestimmungen vorgenommen, die u.a. einige der verwendeten mathematischen Termini erläutern sollen. Die dabei gewählten Formulierungen sind aber in großen Teilen nicht zutreffend bzw. sinntestellend, so dass hier ein dringender Nachbesserungsbedarf besteht.

§ 7 (5): Der hier angesprochene „Zusammenhang“ zwischen einem Diversifikationseffekt und einer nicht vollen Risikokorrelation ist mathematisch nicht begründbar, jedenfalls nicht für das im Gesetzentwurf zugrunde gelegte Risikomaß „Value-at-Risk“. Es gibt hierzu inzwischen zahlreiche wissenschaftliche Publikationen, die diesen weit verbreiteten Standpunkt widerlegen, z.B. D. Pfeifer (2013a, 2013b); vgl. auch L. Rüschendorf (2013). Die Formulierung in § 7 (5) sollte daher besser durch folgende einfachere, aber korrekte Alternative ersetzt werden:

„Diversifikationseffekte: eine Reduzierung des Gefährdungspotenzials von Versicherungsunternehmen und -gruppen durch die Diversifizierung ihrer Geschäftstätigkeit, die sich aus der Tatsache ergibt, dass das negative Resultat eines Risikos durch das günstigere Resultat eines anderen Risikos **potenziell** ausgeglichen werden kann.“

§ 7 (27): Die hier gegebene „Definition“ eines Risikomaßes ist aus mathematischer Sicht unsinnig und vermischt in unzulässiger Weise den mathematischen Kern mit einer ökonomischen Anwendung. Die Formulierung in § 7 (27) sollte daher besser durch folgende inhaltlich korrekte Alternative ersetzt werden:

„Risikomaß: eine mathematische Funktion, die unter einer bestimmten Wahrscheinlichkeitsverteilung eine für ein theoretisches Risiko charakteristische Zahl nach einer gegebenen Axiomatik (Verteilungs-, Skalen- und Translationsinvarianz sowie Monotonie) bestimmt. Ein Risikomaß wird dazu benutzt, um realen Risiken einen monetären Betrag zuzuweisen, der eine ökonomische Einschätzung des realen Risikos für Zwecke der Solvenzkapitalberechnung ermöglicht.“

§ 7 (36): Den Begriff „Wahrscheinlichkeitsverteilungsprognose“ gibt es weder in der mathematischen Fachsprache noch in der Umgangssprache. Er entspricht auch nicht einer angemessenen Übersetzung des Terminus „probability forecast“, der in der englischen Version der Richtlinie 2009/138/EG verwendet wird. Je nach Kontext müsste in dem Gesetzentwurf entweder der Begriff „Wahrscheinlichkeitsverteilung“ oder „unterstellte Wahrscheinlichkeitsverteilung“ verwendet werden. Im Einzelnen wären folgende alternative Formulierungen korrekter:

S. 10: „§ 116: Statistische Qualitätsstandards für unterstellte Wahrscheinlichkeitsverteilungen“

§ 7 (36): „Wahrscheinlichkeitsverteilung: eine mathematische Funktion, die auf theoretischer Ebene auf der Grundlage einer bestimmten Axiomatik Ereignissen eine Zahl zwischen Null und Eins zuweist. Ein wesentliches Element der im Gesetz verwendeten Axiomatik ist die Wahrscheinlichkeitsbewertung einer disjunkten Vereinigung von Ereignissen durch die Summe der individuellen Ereigniswahrscheinlichkeiten (Kolmogoroff'sche Axiomatik).“

S. 86: „§ 116: Statistische Qualitätsstandards für unterstellte Wahrscheinlichkeitsverteilungen“, ebenso § 116 (4), § 116 (5), § 118 (2), § 118 (3) und S. 264 [Zu § 116 ...]

1.3. Anmerkungen zum Formelwerk für das SCR

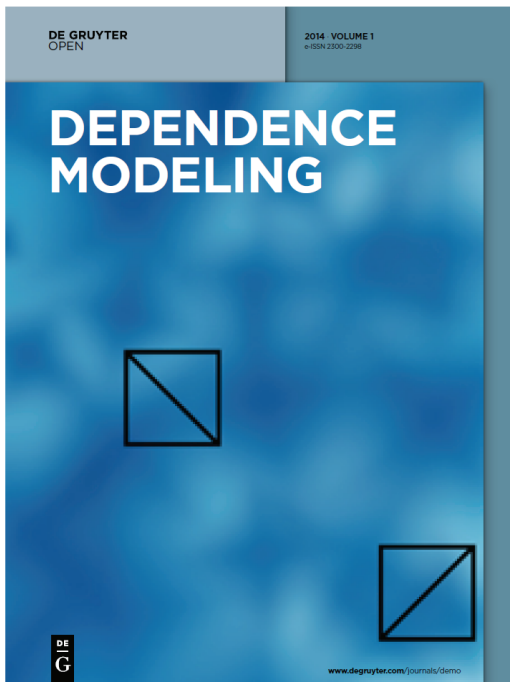
Die im Gesetzentwurf vorgesehene Aggregation der SCR-Werte für einzelne Submodule zu einem Gesamt-SCR auf der Basis von Korrelationen (S. 209 ff.) ist aus mathematischer Sicht nicht zu rechtfertigen, wenn die Risikoverteilungen nicht einer gemeinsamen mehrdimensionalen Normalverteilung entstammen oder das Risikomaß deutlich vom alternativen (aber nicht monotonen) Standardabweichungsprinzip abweicht. Keine der beiden Voraussetzungen wird in der Gesetzesvorlage unterstellt: das Risikomaß „Value-at-Risk“ weicht deutlich vom Standardabweichungsprinzip ab, und die Annahme einer gemeinsamen mehrdimensionalen Normalverteilung für die betrachteten Risiken dürfte in der Praxis wohl ausgeschlossen sein. Insofern ist eine unkritische Anwendung der „Standardformel“ zweifelhaft; hierzu liegen entsprechende wissenschaftliche Publikationen vor (z.B. D. Pfeifer und D. Straßburger (2008) oder A. Sandström (2011), Abschnitt 15.3.5.1, S. 237 f.).

Es muss daher damit gerechnet werden, dass die ebenfalls im Gesetzentwurf in § 27 verankerte eigene Risikoeinschätzung des Versicherungsunternehmens deutliche Abweichungen gegenüber den Einschätzungen mit der Standardformel aufweist, und zwar aufgrund der genannten Inkonsistenzen innerhalb der Standardformel und weniger aufgrund von gravierenden Abweichungen vom europäischen „Durchschnittsmarkt“, die durch die Struktur des Unternehmens selbst gegeben sein könnten.

1.4. Fazit

Der vorliegende Gesetzentwurf weist zahlreiche formal-inhaltliche Schwächen auf, die ein „garantiertes“ Sicherheitsniveau von 99,5 % für das Fortbestehen eines Versicherungsunternehmens im folgenden Geschäftsjahr nicht mathematisch rechtfertigen. Hier wäre eine etwas neutralere Formulierung ohne Bezug auf konkrete Zahlen „ehrlicher“. Hinzu kommt, dass aufgrund des im deutschen HGB verankerten Vorsichtsprinzips die empirisch beobachtete jährliche „Insolvenzrate“ deutscher Versicherungsunternehmen deutlich unter den tolerierten 0,5 % nach dem neuen Aufsichtsgesetz liegt. Für deutsche Versicherungskunden bedeutet die Umstellung auf Solvency II damit faktisch eine potenzielle Verschlechterung ihrer Sicherheitslage bei steigenden Versicherungsbeiträgen, die zwangsläufig durch die Umlage der anfallenden Umstellungskosten auf die Kunden entstehen.



Dietmar Pfeifer*, Hervé Aououlac Tsatedem[†], Andreas Mändle, and Côme Girschig

New copulas based on general partitions-of-unity and their applications to risk management

DOI 10.1515/demo-2016-0006

Received March 30, 2015; accepted June 20, 2016

Abstract. We construct new multivariate copulas on the basis of a generalized infinite partition-of-unity approach. This approach allows, in contrast to finite partition-of-unity copulas, for tail dependence as well as for asymmetry. A possibility of fitting such copulas to real data from quantitative risk management is also pointed out.

Keywords: copulas, partition-of-unity, tail dependence, asymmetry

MSC: 62H05, 62H12, 62H17, 62H20

1 Introduction

The theory of copulas and their applications has gained much interest in the recent years, especially in the field of quantitative risk management, insurance and finance (see e.g. [13] or [16]). While classical approaches like elliptically contoured copulas and Archimedean copulas are widely explored, other approaches concentrate on non-standard, non-symmetric or data-driven copula constructions (see e.g. [9], [8], [2] or [6] and the papers therein for a survey, especially the contributions related to vine copulas). Statistical and computational aspects of copulas have also been investigated in more detail recently (see e.g. [1] and [12]). In this paper, we want to focus on a particular class of copulas and their generalizations, the so called partition-of-unity copulas (see e.g. [10] or [7]). Whereas in the usual approach, only finite partitions-of-unity are considered, which do not allow for a modelling of tail-dependence, we extend this concept to infinite partitions-of-unity, which allows for tail-dependence as well as for asymmetry, and which can also be used to fit given data to a more realistic copula model. Our investigations resemble in some sense more recent approaches such as [16], [4], [19], [5], or [3]. Whereas in these papers, local modifications of known standard copulas are considered in order to obtain tail dependence or asymmetries, we focus on a closed form representation of completely new copula densities which allows for easy Monte Carlo simulations as well as a data driven modeling of tail dependence and asymmetries. This approach is not restricted to two dimensions in general, but can likewise be used in arbitrary dimensions. However, in order to illustrate our results, we will give examples in the bivariate case only.

To facilitate the readability of the paper, all elaborate proofs are given in an appendix.

*Corresponding Author: Dietmar Pfeifer: Carl von Ossietzky Universität Oldenburg, Germany, E-mail: dietmar.pfeifer@uni-oldenburg.de
Hervé Aououlac Tsatedem[†], Andreas Mändle: Carl von Ossietzky Universität Oldenburg, Germany
Côme Girschig: Ecole Nationale des Ponts et Chaussées, Paris, France

New copulas based on general partitions- of-unity and their applications to risk management

Dietmar Pfeifer¹⁾,
Hervé Awoumlac Tsatedem†,
Côme Girschig²⁾,
Andreas Mändle¹⁾

¹⁾Institut für Mathematik, Universität Oldenburg and
²⁾École des Ponts Paris Tech, Paris

1. Introduction

- Construction of new multivariate copulas on the basis of a generalized infinite partition-of-unity approach
- Allows for tail-dependence as well as for asymmetry
- Can be easily implemented for risk management purposes
- Can be extended to an uncountable infinite partition-of-unity approach

2. Main Results

Let $\mathbb{Z}^+ = \{0, 1, 2, 3, \dots\}$ denote the set of non-negative integers and suppose that $\{\varphi_i(\mathbf{u})\}_{i \in \mathbb{Z}^+}$ and $\{\psi_j(\mathbf{v})\}_{j \in \mathbb{Z}^+}$ are non-negative maps defined on the interval $(0, 1)$ each such that

$$\sum_{i=0}^{\infty} \varphi_i(\mathbf{u}) = \sum_{j=0}^{\infty} \psi_j(\mathbf{v}) = 1$$

and

$$\int_0^1 \varphi_i(\mathbf{u}) d\mathbf{u} = \alpha_i > 0, \quad \int_0^1 \psi_j(\mathbf{v}) d\mathbf{v} = \beta_j > 0 \text{ for } i, j \in \mathbb{Z}^+.$$

2. Main Results

The maps $\varphi_i(\mathbf{u})$ and $\psi_j(\mathbf{v})$ can be thought of representing **discrete distributions** over the non-negative integers $\mathbb{Z}^+$ with parameters $\mathbf{u}$ and $\mathbf{v}$, resp.

The sequences $\{\alpha_i\}_{i \in \mathbb{Z}^+}$ and $\{\beta_j\}_{j \in \mathbb{Z}^+}$ represent the probabilities of the corresponding **mixed distributions**.

2. Main Results

Let $\{p_{ij}\}_{i,j \in \mathbb{Z}^+}$ represent the probabilities of an arbitrary discrete bivariate distribution over $\mathbb{Z}^+ \times \mathbb{Z}^+$ with marginal distributions given by $p_{i.} = \sum_{j=0}^{\infty} p_{ij} = \alpha_i$ and $p_{.j} = \sum_{i=0}^{\infty} p_{ij} = \beta_j$ for $i, j \in \mathbb{Z}^+$. Then

$$c(u, v) := \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{p_{ij}}{\alpha_i \beta_j} \varphi_i(u) \psi_j(v), \quad u, v \in (0, 1)$$

defines the density of a bivariate copula, called *generalized partition-of-unity copula*.

2. Main Results

The function c is in fact is the density of a bivariate copula:

$$\begin{aligned}\int_0^1 c(u, v) dv &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{p_{ij}}{\alpha_i \beta_j} \varphi_i(u) \int_0^1 \psi_j(v) dv = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{p_{ij}}{\alpha_i \beta_j} \beta_j \varphi_i(u) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{p_{ij}}{\alpha_i} \varphi_i(u) = \sum_{i=0}^{\infty} \frac{\varphi_i(u)}{\alpha_i} \sum_{j=0}^{\infty} p_{ij} = \sum_{i=0}^{\infty} \frac{\varphi_i(u)}{\alpha_i} \alpha_i \\ &= \sum_{i=0}^{\infty} \varphi_i(u) = 1,\end{aligned}$$

likewise for $\int_0^1 c(u, v) du$.

2. Main Results

From a „dual“ point of view, we can rewrite the above as

$$c(u,v) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} f_i(u) g_j(v), \quad u,v \in (0,1)$$

where $f_i(\cdot) = \frac{\varphi_i(\cdot)}{\alpha_i}$, $g_j(\cdot) = \frac{\psi_j(\cdot)}{\beta_j}$, $i, j \in \mathbb{Z}^+$ denote the densities induced by $\{\varphi_i(u)\}_{i \in \mathbb{Z}^+}$ and $\{\psi_j(v)\}_{j \in \mathbb{Z}^+}$.

This means that the copula density c can also be seen as a mixture of product densities, which possibly allows for a simple way for a stochastic simulation.

2. Main Results

An extension of this approach to d dimensions with $d > 2$ is obvious: assume that $\{\varphi_{ki}(u)\}_{i \in \mathbb{Z}^+}$ for $k = 1, \dots, d$ represent discrete probabilities with

$$\sum_{i=0}^{\infty} \varphi_{ki}(u) = 1 \text{ for } u \in (0,1)$$

and

$$\int_0^1 \varphi_{ki}(u) du = \alpha_{ki} > 0 \text{ for } i \in \mathbb{Z}^+.$$

2. Main Results

Let $\{p_{\mathbf{i}}\}_{\mathbf{i} \in \mathbb{Z}^{+d}}$ represent the distribution of an arbitrary discrete d -dimensional random vector $\mathbf{Z}$ over $\mathbb{Z}^{+d}$ where, for simplicity, we write $\mathbf{i} = (i_1, \dots, i_d)$, i.e.

$$P(\mathbf{Z} = \mathbf{i}) = p_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^{+d}.$$

Suppose further that for the marginal distributions, there holds

$$P(Z_k = i) = \alpha_{ki}, i \in \mathbb{Z}^+, k = 1, \dots, d.$$

2. Main Results

Then

$$c(\mathbf{u}) := \sum_{\mathbf{i} \in \mathbb{Z}^{+d}} \frac{p_{\mathbf{i}}}{\prod_{k=1}^d \alpha_{k,i_k}} \prod_{k=1}^d \varphi_{k,i_k}(u_k), \quad \mathbf{u} = (u_1, \dots, u_d) \in (0,1)^d$$

defines the density of a d -variate copula, which is also called *generalized partition-of-unity copula*.

Alternatively, we can rewrite this again as

$$c(\mathbf{u}) = \sum_{\mathbf{i} \in \mathbb{Z}^{+d}} p_{\mathbf{i}} \prod_{k=1}^d f_{k,i_k}(u_k), \quad \mathbf{u} = (u_1, \dots, u_d) \in (0,1)^d$$

with densities $f_{ki}(\cdot) = \frac{\varphi_{ki}(\cdot)}{\alpha_{ki}}, \quad i \in \mathbb{Z}^+, k = 1, \dots, d.$

3. The symmetric case (diagonal dominance)

For simplicity, we restrict ourselves to the two-dimensional case in the sequel. The generalization to higher dimensions is obvious.

Let $\varphi_i = \psi_i$ for $i \in \mathbb{Z}^+$ and $\int_0^1 \varphi_i(u) du = \alpha_i > 0$. Define

$$p_{ij} := \begin{cases} \alpha_i, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$c(u, v) := \sum_{i=0}^{\infty} \frac{\varphi_i(u) \varphi_i(v)}{\alpha_i} = \sum_{i=0}^{\infty} \alpha_i f_i(u) f_i(v), \quad u, v \in (0, 1)$$

defines the density of a *generalized partition-of-unity copula with diagonal dominance*.

3. The symmetric case (diagonal dominance)

Example 1 (binomial distributions - Bernstein copula).

For a fixed integer $m \geq 2$, consider the family of binomial distributions given by their point masses

$$\varphi_{m,i}(u) = \begin{cases} \binom{m-1}{i} u^i (1-u)^{m-1-i}, & i = 0, \dots, m-1 \\ 0, & i \geq m. \end{cases}$$

3. The symmetric case (diagonal dominance)

Example 1 (binomial distributions - Bernstein copula).

We have, for $i = 0, \dots, m-1$,

$$\begin{aligned}\alpha_{m,i} &= \int_0^1 \varphi_{m,i}(u) du = \binom{m-1}{i} \int_0^1 u^i (1-u)^{m-1-i} du \\ &= \frac{(m-1)!}{i!(m-1-i)!} \cdot \frac{\Gamma(i+1)\Gamma(m-i)}{\Gamma(m+1)} = \frac{1}{m}\end{aligned}$$

and hence

$$c_m(u,v) = m \sum_{i=0}^{m-1} \binom{m-1}{i}^2 (uv)^i ((1-u)(1-v))^{m-1-i}, \quad u, v \in (0,1),$$

which is the density of a bivariate *Bernstein copula*.

3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

Consider, for fixed $\beta > 0$, the family of negative binomial distributions given by their point masses

$$\varphi_{\beta,i}(u) = \binom{\beta + i - 1}{i} (1-u)^\beta u^i, \quad i \in \mathbb{Z}^+.$$

3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

Here we have, for $i \in \mathbb{Z}^+$,

$$\begin{aligned}\alpha_{\beta,i} &= \int_0^1 \varphi_{\beta,i}(u) du = \binom{\beta+i-1}{i} \int_0^1 u^i (1-u)^\beta du \\ &= \frac{\Gamma(\beta+i)}{i! \Gamma(\beta)} \cdot \frac{\Gamma(i+1) \Gamma(\beta+1)}{\Gamma(\beta+i+2)} = \frac{\beta}{(\beta+i)(\beta+i+1)}\end{aligned}$$

and hence for $u, v \in (0, 1)$,

$$c_\beta(u, v) = (\beta+1)((1-u)(1-v))^\beta \sum_{i=0}^{\infty} \binom{\beta+i-1}{i} \binom{\beta+i+1}{i} (uv)^i.$$

3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

For integer values of β , this expression can be explicitly evaluated as a finite sum.

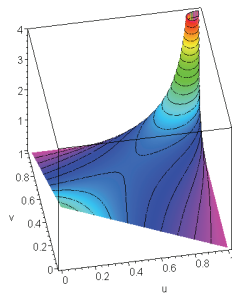
Lemma 1. For $\beta \in \mathbb{N}$ and $u, v \in (0, 1)$, there holds

$$c_{\beta}(u, v) = (\beta + 1) \frac{((1-u)(1-v))^{\beta}}{(1-uv)^{2\beta+1}} \sum_{i=0}^{\beta-1} \binom{\beta-1}{i} \binom{\beta+1}{i} (uv)^i.$$

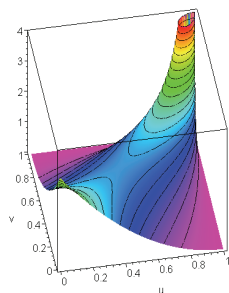
3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

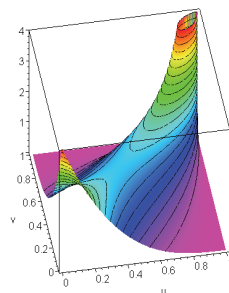
The following graphs show the negative binomial copula densities c_β for $\beta = 1, \dots, 4$.



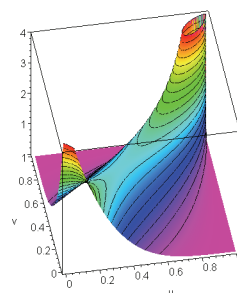
$\beta = 1$



$\beta = 2$



$\beta = 3$



$\beta = 4$

3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

Negative binomial copulas typically show an upper tail dependence, as can be seen from the following exemplary table.

β	1	2	3	4	5	6	7	8	9	10
$\lambda_U(\beta)$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{11}{16}$	$\frac{93}{128}$	$\frac{193}{256}$	$\frac{793}{1024}$	$\frac{1619}{2048}$	$\frac{26333}{32768}$	$\frac{53381}{65536}$	$\frac{215955}{262144}$

3. The symmetric case (diagonal dominance)

Example 2 (negative binomial distributions).

A closed formula for the tail dependence coefficients for integer values of β is given in the following result.

Lemma 2. For $\beta \in \mathbb{N}$, there holds

$$\begin{aligned}\lambda_U(\beta) &= \lim_{t \uparrow 1} \frac{\int_t^1 \int_t^1 c_\beta(u, v) du dv}{1-t} \\ &= \frac{2\Gamma(2\beta)}{\Gamma^2(\beta)} \cdot \int_0^1 \int_0^1 \frac{x^\beta y^\beta}{(x+y)^{2\beta+1}} dx dy = 1 - \frac{\binom{2\beta}{\beta}}{4^\beta}.\end{aligned}$$

3. The symmetric case (diagonal dominance)

Example 3 (Poisson distributions).

Consider the family of Poisson distributions given by their point masses

$$\varphi_{\gamma,i}(u) = (1-u)^{\gamma} \frac{\gamma^i L(u)^i}{i!}, \quad i \in \mathbb{Z}^+$$

where $L(u) = -\ln(1-u) > 0$, $u \in (0,1)$ and $\gamma > 0$.

3. The symmetric case (diagonal dominance)

Example 3 (Poisson distributions).

Here for $i \in \mathbb{Z}^+$, with $z = L(u)$ and $y = (1+\gamma)z$,

$$\begin{aligned}\alpha_{\gamma,i} &= \int_0^1 \varphi_{\gamma,i}(u) du = \int_0^1 (1-u)^\gamma \frac{\gamma^i L(u)^i}{i!} du = \int_0^\infty \frac{\gamma^i z^i}{i!} e^{-(1+\gamma)z} dz \\ &= \frac{\gamma^i}{(1+\gamma)^{i+1}} \int_0^\infty \frac{y^i}{i!} e^{-y} dy = \frac{\gamma^i}{(1+\gamma)^{i+1}} = \left(\frac{\gamma}{1+\gamma} \right)^i \left(1 - \frac{\gamma}{1+\gamma} \right),\end{aligned}$$

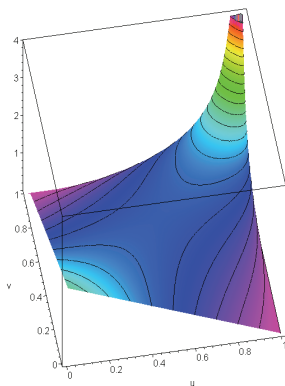
indicating that the $\alpha_{\gamma,i}$ correspond to a geometric distribution with mean γ , and hence for $u, v \in (0,1)$,

$$c_\gamma(u, v) = (1+\gamma)(1-u)^\gamma (1-v)^\gamma \sum_{i=0}^{\infty} \frac{(\gamma(1+\gamma) \ln(1-u) \ln(1-v))^i}{i!^2}.$$

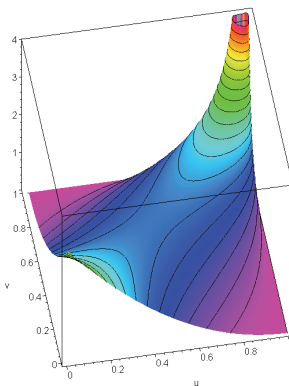
3. The symmetric case (diagonal dominance)

Example 3 (Poisson distributions).

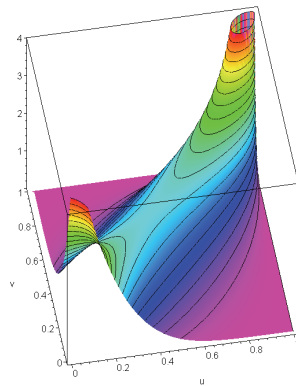
The following graphs show some of these copula densities for different choices of γ .



$\gamma = 1$



$\gamma = 2$

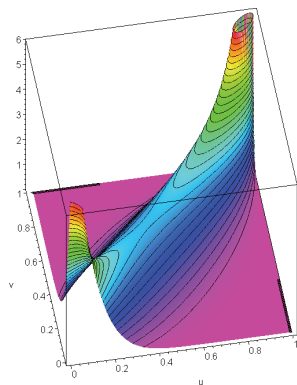


$\gamma = 5$

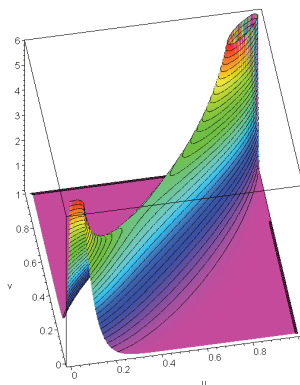
3. The symmetric case (diagonal dominance)

Example 3 (Poisson distributions).

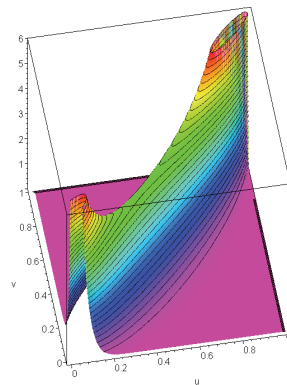
The following graphs show some of these copula densities for different choices of γ .



$\gamma = 10$



$\gamma = 20$



$\gamma = 30$

3. The symmetric case (diagonal dominance)

Example 3 (Poisson distributions).

The corresponding copula C cannot be calculated explicitly. However, in contrast to the visual impression, the coefficient $\lambda_U(\gamma)$ of upper tail dependence is zero here for all $\gamma > 0$, although we have a singularity in the point $(1,1)$ in all cases.

3. The symmetric case (diagonal dominance)

Example 4 (log series distribution).

Consider the family of log series distributions given by their point masses

$$\varphi_i(u) = \frac{u^i}{i \cdot L(u)}, \quad i \in \mathbb{N}$$

where again $L(u) = -\ln(1-u)$, $u \in (0,1)$. Here we get

$$\alpha_i = \int_0^1 \varphi_i(u) du = \frac{\beta_i}{i} = \frac{1}{i} \sum_{j=1}^i \binom{i}{j} (-1)^{j+1} \ln(j+1) \text{ for } i \in \mathbb{N}.$$

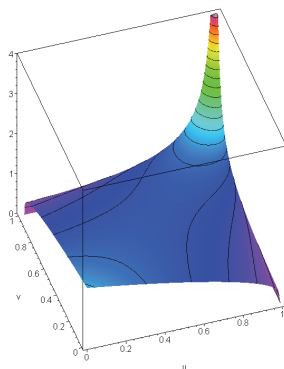
The density of the bivariate log series copula is given by

$$c(u,v) = \sum_{i=1}^{\infty} \frac{1}{\alpha_i} \varphi_i(u) \varphi_i(v) = \frac{1}{\ln(1-u) \ln(1-v)} \sum_{i=1}^{\infty} \frac{(uv)^i}{i \beta_i}, u, v \in (0,1).$$

3. The symmetric case (diagonal dominance)

Example 4 (log series distribution).

The following graph shows the corresponding copula density. The log series copula also has no tail dependence.



plot of $c(u, v)$

4. The asymmetric case

Specifying the probabilities p_{ij} in a non-symmetric way we obtain asymmetric copula densities even if the maps $\varphi_i(\cdot)$ and $\psi_j(\cdot)$ are identical. A very simple approach to this problem is a specification of a suitable non-symmetric $(n+1) \times (n+1)$ -matrix $M_n = [p_{ij}]_{i,j=0,\dots,n}$ for $n \in \mathbb{Z}^+$ with

$$\sum_{k=0}^n p_{ik} = \sum_{k=0}^n p_{ki} = \alpha_i \quad \text{for } i = 0, \dots, n$$

and

$$p_{ij} := \begin{cases} \alpha_i, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \quad \text{for } i, j > n.$$

4. The asymmetric case

Example 5 (negative binomial distributions).

We consider negative binomial distributions with $\beta = 1$.

Then $\alpha_i = \int_0^1 \varphi_{1,i}(u) du = \frac{1}{(1+i)(2+i)}$ for $i \in \mathbb{Z}^+$. With $n = 4$ and

$$M_4 := \frac{1}{60} \begin{bmatrix} 18 & 5 & 5 & 0 & 2 \\ 10 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{bmatrix}$$

we obtain

4. The asymmetric case

Example 5 (negative binomial distributions).

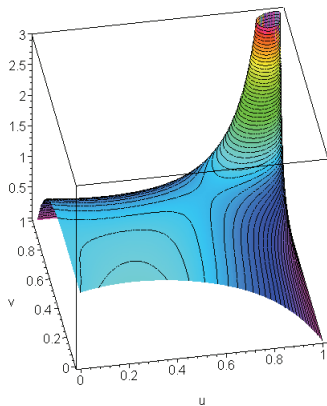
$$c(u,v) = \frac{(1-u)(1-v)}{5(1-uv)^3} H(u,v), \quad u,v \in (0,1)$$

with the polynomial

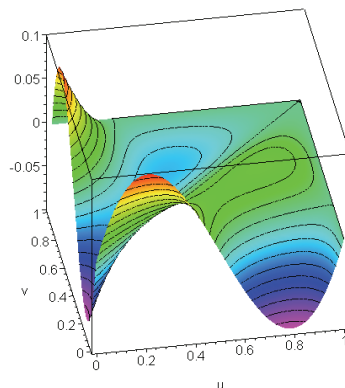
$$\begin{aligned} H(u,v) = & 150u^7v^7 - 450u^6v^6 - 10u^7v^3 + 510u^5v^5 - 10u^3v^7 - \dots \\ & \dots - 30u^5v^4 - 10u^3v^5 + 30u^2v^6 - 300u^4v^4 + 30u^6v^2 - \dots \\ & \dots - 5u^3v^4 + 80u^4v^3 - 30u^5v + 94u^3v^3 + 30u^2v^4 - \dots \\ & \dots - 30uv^5 - 60u^3v^2 + 15u^2v^3 + 10u^4 + 18u^2v^2 - 30uv^3 + \dots \\ & \dots + 10v^4 - 15uv^2 + 10v^2 - 18uv + 10u + 5v + 6. \end{aligned}$$

4. The asymmetric case

Example 5 (negative binomial distributions).



plot of $c(u, v)$



plot of $c(u, v) - c(v, u)$

Coefficient of upper tail dependence: $\lambda_u = \frac{1}{2}$

4. The asymmetric case

Example 6 (different negative binomial distributions).

For $\beta = 1$ and $\beta = 2$, resp. we get

$$\alpha_i = \int_0^1 \varphi_{1,i}(u) du = \frac{1}{(1+i)(2+i)} \quad \text{and}$$

$$\beta_j = \int_0^1 \varphi_{2,j}(v) dv = \frac{2}{(2+j)(3+j)} = 2\alpha_{j+1} \quad \text{for } i, j \in \mathbb{Z}^+.$$

$$\text{Let further } [p_{ij}]_{i,j \in \mathbb{Z}^+} = \begin{bmatrix} \beta_0 & \beta_1 & 0 & 0 & 0 & \dots & \dots & \dots & \dots \\ 0 & 0 & \beta_2 & \beta_3 & 0 & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \beta_4 & \beta_5 & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots & \beta_6 & \beta_7 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

4. The asymmetric case

Example 6 (different negative binomial distributions).

Then $p_{i.} = \sum_{j=0}^{\infty} p_{ij} = \alpha_i$ and $p_{.j} = \sum_{i=0}^{\infty} p_{ij} = \beta_j$ for $i, j \in \mathbb{Z}^+$

since

$$\beta_{2i} + \beta_{2i+1} = \frac{2}{(2+2i)(3+2i)} + \frac{2}{(3+2i)(4+2i)} = \frac{1}{(1+i)(2+i)} = \alpha_i.$$

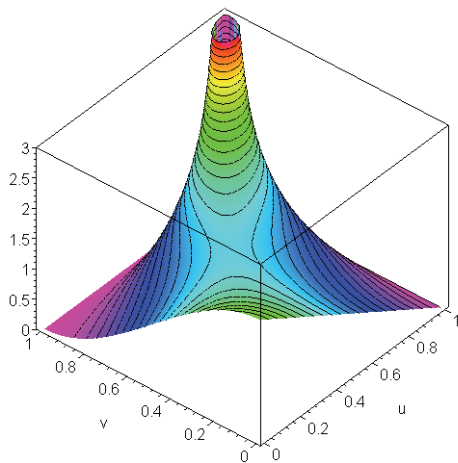
It now follows that

$$c(u, v) = \frac{2(1-u)(1-v)^2(1+2v+5uv^2+4uv^3)}{(1-uv^2)^4}, \quad u, v \in (0, 1)$$

which obviously is asymmetric.

4. The asymmetric case

Example 6 (different negative binomial distributions).



plot of $c(u, v)$

4. The asymmetric case

Example 6 (different negative binomial distributions).

The corresponding copula C is, for $x, y \in (0, 1)$, given by

$$C(x, y) = \frac{xy}{(1 - xy^2)^2} (2 - x - 2xy^3 + xy^4 + x^2y^3 - 2y^2 + y^3)$$

This copula has a coefficient of upper tail dependence of

$$\lambda_u = \frac{5}{9}$$

5. Conclusions and Applications

Remark 1: Negative binomial copulas can easily be simulated through the alternative representation formula (2.5) involving mixed product-Beta distributions. Poisson copulas can be simulated using the transformation $z \mapsto 1 - e^{-z}$ applied to Gamma distributed random variables Z with a random shape parameter α , where $\alpha - 1$ is generated by a geometric distribution, and scale parameter $1 + \gamma$.

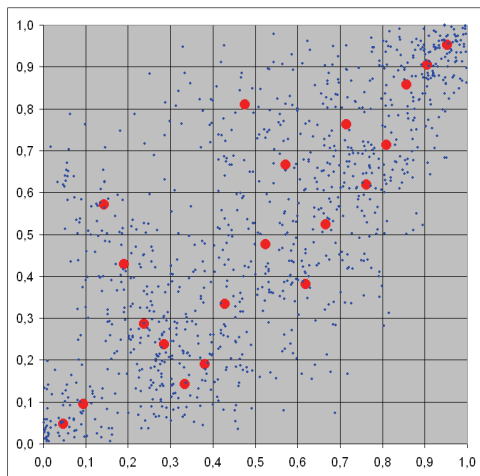
5. Conclusions and Applications

Remark 2: For practical applications in quantitative risk management, it seems reasonable to fit the required probabilities $[p_{ij}]_{i,j \in \mathbb{Z}^+}$ to empirical data via their empirical copula, for instance as was proposed in PFEIFER, STRASSBURGER AND PHILIPPS (2009). In the particular case of Bernstein copulas such a procedure can be very easily implemented, even in higher dimensions (cf. COTTIN AND PFEIFER (2014)).

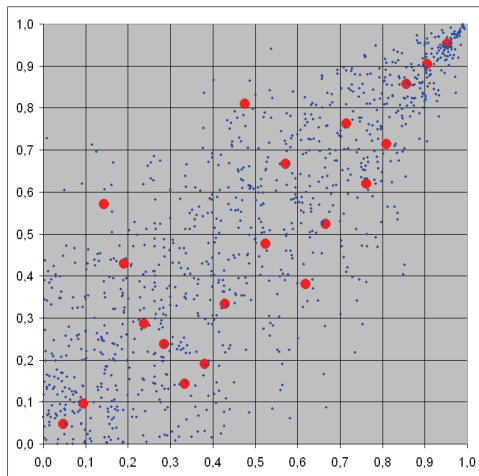
5. Conclusions and Applications

As a practical exercise, we refer to Example 4.2 in COTTIN AND PFEIFER (2014) where the empirical copula from an original data set was fitted to a Bernstein copula. The following two graphs show the scatter plot from the empirical copula (big red dots) superimposed by 1000 simulated points of that Bernstein copula (left) and of a negative binomial copula, with $\beta = 5$.

5. Conclusions and Applications



Bernstein copula fit



negative binomial copula fit

The Bernstein copula represents the local asymmetry of the empirical copula better, but shows no tail dependence, as does the negative binomial copula.

5. Conclusions and Applications

The fit to the negative binomial copula was performed by a numerical match between the theoretical correlation for the negative binomial copula and the correlation of the empirical copula, which is 0.815. Note that the theoretical correlation $\rho(\beta)$ for the negative binomial copula can be explicitly calculated as

$$\begin{aligned}\rho(\beta) &= 12\beta \left(\sum_{i=0}^{\infty} \frac{(i+1)^2}{(\beta+i)(\beta+i+1)(\beta+i+2)^2} \right) - 3 \\ &= 3\beta \left(2(\beta+1)^2 \Psi(1, \beta+2) - 2\beta - 1 \right)\end{aligned}$$

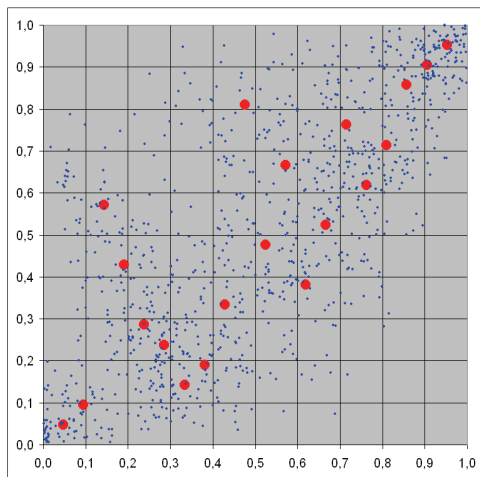
where $\Psi(1, z)$ denotes the first derivative of the digamma function, or $\Psi(1, z) = \frac{d^2}{dz^2} \ln \Gamma(z)$, $z > 0$.

5. Conclusions and Applications

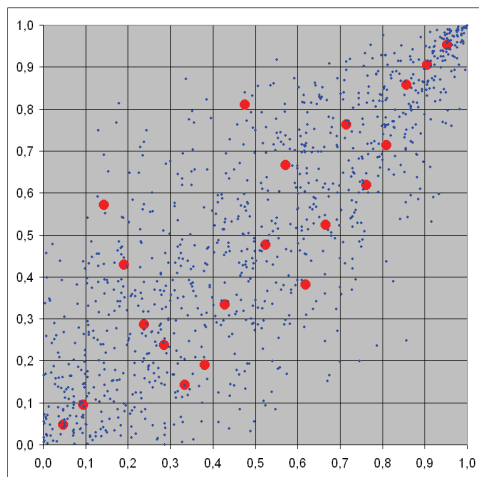
β	1	2	3	4	5	6	7
$\rho(\beta)$	0.4784	0.6529	0.7410	0.7937	0.8288	0.8537	0.8723

For the sake of completeness, we finally show a comparison between the Bernstein copula fit and a Poisson copula fit with parameter $\gamma = 6$. The empirical correlation for the Poisson copula here is 0.814.

5. Conclusions and Applications



Bernstein copula fit



Poisson copula fit

Note that although the empirical plot for the Poisson copula might suggest some tail dependence here this is actually not the case.

