

Log-Normalverteilungen, Potenz- und Paretoverteilungen, $1/f$ -Rauschen, und zugrunde liegende Mechanismen¹

Inhaltsverzeichnis

1	Einführung	2
1.1	Die Log-Normalverteilung	2
1.2	Pareto- und Potenzverteilungen	3
1.3	Skaleninvarianz und Potenzfunktionen	4
1.4	$1/f$ -Rauschen und damit verbundene Verteilungen	9
2	Mechanismen und Verteilungen	10
2.1	Log-Normalverteilung	10
2.1.1	Prozesse erster Ordnung und die Log-Normalverteilung	10
2.1.2	Gibrats Gesetz und Shockleys Modell	11
2.1.3	Das obere Ende: Der Montroll-Schlesinger-Ansatz	14
2.1.4	Thermodynamische Modelle	17
2.2	Potenz-(Pareto-)Verteilungen	18
2.2.1	Die Kombination von Exponentialfunktionen	19
2.2.2	Rauschen in Halbleitern und die Fluten des Nils	20
2.2.3	Selbst-organisierte Kritikalität	21

¹Korrektur 19. 05. 2024

1 Einführung

Die Häufigkeit von Wörtern, die Häufigkeiten spontanen Verhaltens, die Intensität von Kriegen und die Häufigkeit terroristischer Attacken, die Wartezeiten zwischen epileptischen Anfällen oder schizophrenen Schüben, die Verteilung von Einkommen, von wissenschaftlichen Publikationen, und von *spike bursts* in neuronalen Populationen haben gemeinsam, dass sie, wenn schon nicht exakt, so doch oft in guter Näherung $\propto x^{-\alpha}$ verteilt sind, wobei α eine für den jeweiligen Fall zu bestimmende Konstante und x die entsprechende Variable ist. Es sollen einige Mechanismen beschrieben werden, die diesen Sachverhalt erklären können.

1.1 Die Log-Normalverteilung

Definition 1.1 *Es sei X eine zufällige Veränderliche mit $X = e^Y$, wobei Y eine normalverteilte zufällige Veränderliche mit Erwartungswert μ und Varianz σ^2 ist. Dann ist $\log X$ normalverteilt mit diesen Parametern:*

$$f(\log x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad (1)$$

und X heißt log-normalverteilt.

Der Erwartungswert und die Varianz sind durch

$$\mathbb{E}(X) = e^{\mu + \sigma^2/2} \quad (2)$$

$$\text{Var}(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) \quad (3)$$

$$\text{Median} = \text{med}(X) = e^\mu, \quad (4)$$

dh μ ist der Median von $\log y$. Die *multiplikative Standardabweichung* ist durch

$$\sigma^* = e^\sigma \quad (5)$$

definiert. μ ist ein Skalenparameter: wird X in verschiedenen Einheiten gemessen, verändert sich μ entsprechend, während σ^* invariant bleibt. Der Variationskoeffizient ist durch

$$cv = [\exp(\sigma^2) - 1]^{1/2} \quad (6)$$

definiert, – er hängt nur von σ^2 ab. Die asymptotisch effizientesten Schätzungen für μ und σ^* sind

$$\bar{x} = \exp\left(\frac{1}{n} \sum_{i=1}^n \log x_i\right) = \left(\prod_{i=1}^n x_i\right)^{1/n} \quad (7)$$

$$s^* = \exp\left[\left(\frac{1}{n-1} \sum_{i=1}^n \left(\log\left(\frac{x_i}{\bar{x}}\right)\right)^2\right)^{1/2}\right] \quad (8)$$

Anmerkung Bekanntlich ist die Summe zweier unabhängiger normalverteilter Variablen wieder normalverteilt. Sind X_1 und X_2 zwei log-normalverteilte Variablen, so ist das Produkt $X = X_1 X_2$ wieder log-normalverteilt, denn

$$X = X_1 X_2 = e^{Y_1} e^{Y_2} = e^{Y_1 + Y_2} = e^Y, \quad Y \sim N(\mu, \sigma^2) \quad (9)$$

mit $\mu = \mu_1 + \mu_2$, $\sigma^2 = \sigma_1^2 + \sigma_2^2$. \square

1.2 Pareto- und Potenzverteilungen

Definition 1.2 Es sei X eine zufällige Veränderliche mit der Dichtefunktion

$$f(x) = \begin{cases} \frac{k}{b} \left(\frac{x-a}{b}\right)^{-k-1} = \frac{k}{b} \left(\frac{b}{x-a}\right)^{k+1}, & \text{für } x > a+b \\ 0, & \text{sonst} \end{cases} \quad (10)$$

mit $a \in \mathbb{R}$, $b, k \in \mathbb{R}_+$. Dann heißt f Pareto-Dichte².

Im Folgenden wird i. a. $a = 0$ sein. Dann ist b gleich dem minimal möglichen Wert x_{min} für x . Für die Verteilungsfunktion findet man

$$F(x) = \int_{a+b}^x f(u) du = 1 - \left(\frac{x-a}{b}\right)^{-k} = 1 - \left(\frac{b}{x-a}\right)^k. \quad (11)$$

Anmerkung: Die Dichtefunktion (10) wird oft auch Pareto-Verteilung genannt. Andererseits ist die Verteilungsfunktion $F(x)$ als Integral der Dichtefunktion definiert; deswegen wird $f(x)$ hier explizit als Dichtefunktion eingeführt. \square

Für $x \downarrow a+b$ folgt $F(x) \rightarrow 0$, und für $x \uparrow \infty$ folgt $b/(x-a) \rightarrow 0$ und damit $F(x) \rightarrow 1$.

Es sei $a = 0$; dann sind Erwartungswert und Varianz durch

$$\mathbb{E}(X) = \begin{cases} b \frac{k}{k-1}, & k > 1 \\ \infty, & k \leq 1 \end{cases} \quad (12)$$

$$Var(X) = \begin{cases} b \left(\frac{k}{k-1} - \frac{k^2}{(k-1)^2} \right), & k > 2 \\ \infty, & k \leq 2 \end{cases} \quad (13)$$

Dass ein endlicher Erwartungswert nur für $k > 1$ und eine endliche Varianz nur für $x > 2$ existiert erklärt die Ausdrucksweise 'heavy tailed' für Verteilungen vom Pareto-Typ; zwar wird die Dichtefunktion für kleinere Werte von x schnell

²Vilfredo Federico Pareto (1848 – 1923), italienischer Ingenieur, Ökonom und Soziologe. Pareto betrachtete u.a. die Verteilung von Einkommen bzw. Vermögen in einer Gesellschaft, die der nach ihm benannten Verteilung folgt.

kleiner, für größere Werte von x konvergiert sie langsam relativ zu anderen Dichtungen gegen Null.

Anmerkung Eine ähnliche Verteilung ist die Potenzverteilung:

Definition 1.3 *Es sei*

$$f(x) = \begin{cases} \frac{c}{b} \left(\frac{x-a}{b}\right)^{c-1}, & \text{für } a \leq x \leq a+b \\ 0, & \text{sonst} \end{cases} \quad (14)$$

Dann heißt f Potenzverteilung.

Die Verteilungsfunktion ist durch

$$F(x) = \int_a^x f(u) du = \left(\frac{x-a}{b}\right)^c \quad (15)$$

gegeben. Für $x \rightarrow a$ folgt $F(x) \rightarrow 0$, und für $x \rightarrow a+b$ ergibt sich $F(x) \rightarrow 1$. In den hier behandelten Anwendungen ist aber die Pareto-Verteilung von größerem Interesse.

Rang-Häufigkeiten-Darstellung f sei durch die Pareto-Dichte gegeben, und es werde $R(x) = P(X > x)$ betrachtet. $R(x)$ ist der Anteil der Fälle, die größer als x sind und wird gelegentlich als *Rang* bezeichnet. Aus (11) folgt sofort

$$R(x) = \left(\frac{b}{x-a}\right)^c. \quad (16)$$

Offenbar ergibt die log-log-Darstellung wieder eine Gerade:

$$\log R(x) = c(\log b - \log(x-a)). \quad (17)$$

Man kann nun den x -Werten einen Rang zuordnen: der größte Wert bekommt den Rangplatz 1, der zweitgrößte bekommt den Rangplatz 2, und der kleinste Wert bekommt schließlich den Rangplatz n . Trägt man $R(x)$ gegen x bzw gegen den Rangplatz $i(x)$, $i = 1, 2, \dots, n$ auf, so erhält man die Rang-Häufigkeits-Darstellung, – wobei der Ausdruck 'Häufigkeitsdarstellung' ein wenig irreführend ist, da es sich nicht um Häufigkeiten handeln muß; der Name hat historische Gründe. Benutzt man logarithmische Skalen, so erhält man wegen (17) eine Gerade.

1.3 Skaleninvarianz und Potenzfunktionen

Dynamische Prozesse können auf "schnellen" oder "langsamen" (Zeit-)Skalen ablaufen, dh die Veränderungen der Variablen können etwa im Millisekunden-, im Sekunden-, Minuten- oder Stundenbereich ablaufen, oder in noch größeren Zeiträumen. Ein Prozess wird *skalenfrei* genannt, wenn

”no characteristic scales dominate the dynamics of the underlying process. Scale-free behavior can be revealed with scaling analysis, which quantifies the fluctuations of a parameter as a function of the scale at which the parameter is evaluated. Scale-free behavior reflects a tendency of complex systems to develop correlations that decay more slowly and extend over larger distances in time and space than the mechanisms of the underlying process would suggest.” Linkenkaer-Hansen et al (2001). p. 1370

Die 'longe-range'-Korrelationen bilden sich dabei durch lokale Interaktionen und erstrecken sich über das ganze System. Dies bedeutet, dass das System ein Skalenverhalten entwickeln wird, dass durch eine Potenzfunktion abgebildet wird, und darüber hinaus arbeitet der Prozess in einem 'kritischen Zustand' (self-organized criticality (SOC) - Bak, Tang und Wiesenfeldt (1987)).

Skaleninvarianz Ein wesentliches Merkmal der Potenzfunktion ist die *Skaleninvarianz*. Es sei $f(x)$ eine Funktion von x . Der Verlauf von f hängt von der Skala ab, auf der x gemessen wird, bzw in deren Einheit x in f eingeht. Die Skala kann verändert werden, indem x mit einem geeigneten Faktor a multipliziert wird: Geht man von Metern zu Zentimetern über, so wird x mit 100 multipliziert. Es sei insbesondere $f(x) = x^k$, und es werde x in ax skaliert: Man hat $g(y) = g(ax) = (ax)^k = a^k x^k = a^k f(x)$. Die Funktion g auf der neuen Skala verhält sich wie die alte Funktion f auf der x -Skala, mit dem einzigen Unterschied, dass die Funktionswerte nun alle mit a^k multipliziert werden. In diesem Sinne ist f *skaleninvariant*. Umgekehrt gelte $f(ax) = a^k f(x)$, und gesucht ist f . Für $x = 1$ erhält man $f(a) = f(1)a^k$. Benennt man a um in x , so hat man $f(x) = cx^k$, $c = f(1)$.

Eine allgemeine Definition der Skaleninvarianz erweist sich als nützlich:

Definition 1.4 *Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Abbildung, die der Bedingung*

$$f(x) = h(a)f(ax) \tag{18}$$

genügt. Dann heißt f skaleninvariant.

Geht man also von der x -Skala zur Skala ax mit $a > 0$ über, so müssen die f -Werte nur mit einem nur vom Skalenfaktor a abhängenden Faktor $h = h(a)$ multipliziert werden, damit $hf(ax)$ die gleichen Werte wie die ursprüngliche Funktion f annimmt. Die Potenzfunktion erfüllt offenbar diese Bedingung: ist $f(x) = x^k$, so hat man $f(ax) = (ax)^k = a^k x^k$ und

$$a^{-k}f(ax) = x^k = f(x), \tag{19}$$

also $h(a) = a^{-k}$.

Beispiel 1.1 Häufigkeit von Erdbeben³ Richter (1935) definierte die Stärke y (magnitude) eines Erdbebens als Logarithmus zur Basis 10 der Verhältnisses der Amplitude x der Wellen, die durch einen Seismographen gemessen werden, zu einer beliebig gewählten, kleinen Standardamplitude. Also hat ein ein Erdbeben der Stärke 5 eine zehn mal größere Amplitude als ein Erdbeben der Stärke 4; dies bedeutet einen 31.6-mal größere freigesetzte Energie.

Man kann nun die Häufigkeit $n(y)$ von Erdbeben zu ihrer Stärke y in Beziehung setzen. Die Daten werden sehr gut durch eine Potenzfunktion der Form

$$\log_{10} n(y) = a - by \quad (20)$$

beschrieben, wobei $n(y)$ die Anzahl der Erdbeben ist, deren Stärke größer als y ist. Man findet $.5 < b < 1.0$ in Abhängigkeit von der Region, in der Erdbeben auftreten. Weiter ist $n_{ges} = 10^a$ die Gesamtzahl der Erdbeben, d.h. $a = \log_{10} n_{ges}$. Gleichung (20) impliziert

$$n(y) = 10^{a-by}. \quad (21)$$

Es ist $n_0 = 10^0 = 1$, $n_1 = 10^1 = 10$, $n_2 = 10^2 = 100$, $n_3 = 10^3 = 1000$, $n_4 = 10^4 = 10000$, etc., so dass $\log_{10} n_k = k$. Trägt man k gegen die dazu korrespondierende Stärke x auf, so erhält man eine Gerade, die der Gleichung (20) entspricht.

(21) hat – zunächst – nicht die Form der Potenzfunktion (19). Berücksichtigt man aber die Definition der Stärke y , nämlich $y(x) = \log_{10} x/x_0$, wobei x_0 die beliebig wählbare Standardamplitude ist, so findet man

$$\log_{10} n(y) = a - by = a - b \log_{10} x/x_0,$$

so dass

$$\log_{10} n(y) = \log_{10} \alpha - b \log_{10} x/x_0 = \log_{10} \left(\frac{\alpha x_0}{x^b} \right)$$

oder

$$n(x) = \frac{\alpha x_0}{x^b} = \beta x^{-b}. \quad (22)$$

□

Die Frage ist nun, ob die Potenzfunktion die einzige Funktion ist.

Satz 1.1 $f(x) = x^k$ ist die einzige Funktion, die der Gleichung (18) genügt.

Beweis: Dazu werde (18) in der Form

$$f(ax) = g(a)f(x) \quad (23)$$

³Gutenberg, Richter, C.F.: Seismicity of the Earth and Associated Phenomena, Princeton, N.J. 1954

mit $h(a) = 1/g(a)$ geschrieben. Durch entsprechende Normierung kann man stets erreichen, dass $f(1) = 1$, und für $x = 1$ folgt dann $f(a) = g(a)$, so dass f und g identisch sein müssen und man

$$f(ax) = f(a)f(x) \quad (24)$$

schreiben kann. Mit $y = ax$ ist

$$\frac{df(y)}{da} = \frac{xd f(y)}{dy} = x f'(ax) = f'(a)f(x),$$

und für $a = 1$ erhält man

$$x f'(x) = f'(1)f(x), \quad \text{d.h. } f'(1) \frac{1}{x} = \frac{f'(x)}{f(x)}.$$

Dann folgt wiederum

$$f'(1) \int \frac{1}{x} dx = \int \frac{f'(x)}{f(x)} dx,$$

und wegen $\int (1/x) dx = \log x + c$, $\int (f'(x)/f(x)) dx = \log f(x) + c_0$ hat man nun

$$f'(1) \log x = \log f(x) + c_0 - c.$$

Es sollte aber $f(1) = 1$ gelten, und $\log f(1) = \log 1 = 0$, so dass $f'(1) \log 1 = \log 1 + c_0 - c = 0$ und damit $c_0 - c = 0$. Damit erhält man $f'(1) \log x = \log f(x)$ oder $\log x^{f'(1)} = \log f(x)$, so dass

$$f(x) = x^k, \quad k = f'(1).$$

□

Hat man also $f_0(t) = t^k$ und $f_1(\tau) = \tau^k$ mit $\tau = at$, $a > 0$ so folgt

$$f_1(\tau) = f_0(t) \quad \text{für } t = \tau/a, \quad (25)$$

denn für alle $\tau = at$ gilt ja $\tau^k = (at)^k$, so dass für $t = \tau/a$ folgt, dass $t^k = (\tau/a)^k$.

hier Abb. 1

Beispiel 1.2 Fallgesetz Das Fallgesetz

$$s(t) = \frac{g}{2} t^2, \quad g = 9.81 \dots \quad (26)$$

ist offenbar skaleninvariant. t wird in Sekunden, s in Metern gemessen. Der Übergang zu einer anderen Zeitskala (Minuten, Stunden, Jahre, etc) bedeutet die Multiplikation von t mit einer geeignet gewählten Zahl a , und man hat

$$s(at) = \frac{g}{2} (at)^2 = \frac{ga^2}{2} t^2 \Rightarrow \frac{1}{a^2} s(at) = s(t),$$

und dies entspricht (18) mit $h(a) = a^{-2}$. □

Beispiel 1.3 Webers Gesetz Webers Gesetz (das in der behaupteten Allgemeinheit falsch ist) besagt, dass

$$\frac{\Delta x}{x} = c \quad (27)$$

c eine Konstante, x eine Stimulusintensität und Δx der Zuwachs an Intensität, der nötig ist, damit $x + \Delta x$ gerade von x unterschieden werden kann. Das sich die Maßeinheit von x in (27) herauskürzt, liegt nahe, dass hier ein Fall von Skaleninvarianz vorliegt. Wird etwa die Länge eines Gegenstandes beurteilt und in Metern [m] gemessen, so folgt für $c = .2$, dass für $x = 1$ [m] der Wert von Δx gleich .2 Meter, oder 20 cm ist. Wird in Zentimeter [cm] gemessen, so erhält man für den gleichen c -Wert für $x = 10$ cm den Wert $\Delta x = 10 \times .2 = 2$ cm. Für $x = 1$ cm folgt $\Delta x = .2 \times 1 = .2$ cm, etc. \square

Log-log-Darstellung Für $f(x) = ax^k$ folgt

$$\log f(x) = k \log x + \log a, \quad (28)$$

dh in log-log-Darstellung ergibt sich eine Gerade. Die Steigung hängt nicht von der Maßeinheit ab, – diese bestimmt nur die additive Konstante, dh die *Form* der Beziehung zwischen f und x bleibt invariant.

Andere Aspekte der Skaleninvarianz

1.4 $1/f$ -Rauschen und damit verbundene Verteilungen

Stellt man sein Radio nicht auf einen Sender ein, so hört man ein Rauschen. Auf die physikalischen Ursachen des Rauschens soll hier nicht im Detail eingegangen werden, es genügt, festzustellen, dass die Lautsprechermembran mehr oder weniger zufällig angeregt wird. Man kann – zumindest im Prinzip – das Geräusch während der Dauer eines Zeitintervalls durch eine Funktion $\xi(t)$ darstellen, die den durch die Membran erzeugten Schwingungen der Luft repräsentiert. Der Verlauf von ξ ist sehr irregulär und in diesem Sinne "zufällig". ξ ist demnach eine zufällige Funktion. Betrachtet man die Luftschwankungen während verschiedener Zeitabschnitte, aber unter sonst identischen Bedingungen, so wird man verschiedene Verläufe ξ finden. Man kann alle diese Verläufe zu einer Menge $X = \{\xi\}$ zusammenfassen. Die "Beobachtung" des Rauschens während eines Zeitintervalls $[0, T]$ entspricht dann der zufälligen Wahl einer Funktion $\xi \in X$, und ξ selbst zeigt dann wiederum einen irregulären, "zufälligen" Verlauf. Die Art dieser Verläufe wird durch ein Wahrscheinlichkeitsgesetz \mathbb{P} bestimmt; $X_t = \{X, \mathbb{P}\}$, also die Menge X zusammen mit dem W-Maß \mathbb{P} , bezeichnet man als *stochastischen Prozeß*.

Viele Funktionen können als Überlagerung von Sinusfunktionen oder Kosinusfunktionen $A \sin(\omega x + \varphi)$ dargestellt werden. Dazu müssen, für eine gegebene Funktion, die Amplituden $A = A(\omega)$ und die Phase $\varphi = \varphi(\omega)$ richtig gewählt werden. Das mathematische Verfahren, mit dem man diese Darstellung bestimmt, ist als Fourier-Analyse⁴ bekannt. Das Spektrum wird durch eine komplexe Funktion $F(\omega) = F_1(\omega) + iF_2(\omega)$ dargestellt, $i = \sqrt{-1}$, wobei F_1 und F_2 reelle Funktionen sind. F ist eine kompakte Darstellung zweier Funktionen, nämlich der Amplituden- und der Phasenfunktion, wobei

$$A(\omega) = (F_1^2(\omega) + F_2^2(\omega))^{1/2}, \quad \varphi(\omega) = \tan^{-1} \left(\frac{F_2(\omega)}{F_1(\omega)} \right) \quad (29)$$

Das Power-Spektrum ist

$$S(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} |A(\omega)|^2, \quad \omega = 2\pi f \quad (30)$$

$|A(\omega)|^2$ heißt auch *Energie-Spektrum* von X , oder die *Energie-Spektraldichte*. $S(\omega)$ gibt also die Energie an, die durch die Frequenz $\omega = 2\pi f$ geliefert wird, wobei f hier für eine Frequenz steht.

Autokorrelation eines Prozesses; Power spectrum als Fourier-Transformierte der Autokorrelation \rightarrow Power-Funktion.

Die Beziehung zwischen $1/f$ -Rauschen und Verteilungen mit "long tails" wird in Montroll und Schlesinger (1982) diskutiert; in Abschnitt ?? wird darauf näher eingegangen.

⁴Nach Jean Baptiste Joseph Fourier (1768 – 1830), französischer Mathematiker und Physiker, der dieses Verfahren im Zusammenhang mit Fragen der Wärmeleitung zuerst entwickelte.

2 Mechanismen und Verteilungen

2.1 Log-Normalverteilung

Hinweis: as time-to-failure-distribution: multiplicative degradation model, multiplicative shocks – small, but proportional to already existing degradation, $y_i = (1 + \varepsilon_i)y_{i-1}$, with ε_i small, iid random perturbation⁵, Schneeball, der beschneiten hang hinunterrollt und dabei Schnee mitnimmt. Je mehr Schnee schon dran ist, dh je dicker er ist, desto mehr Schnee kann er aufnehmen.

$$y_n = y_0 \prod_{i=1}^n (1 + \varepsilon_i)$$

Logarithmieren und Central-Limit-Theorem; Beispiele: chemische Reaktionen bis neue Komponente gebildet wird (= "Failure"), Diffusion bzw Wanderung von Ionen, Crack growth or propagation

2.1.1 Prozesse erster Ordnung und die Log-Normalverteilung

Ein bestimmtes Gift werde einem Organismus zur Zeit $t = 0$ mit der Konzentration C_0 zugeführt, und der Organismus baue das Gift ab oder zertöre es mit einer Rate, die proportional zur Konzentration zur Zeit t ist. Die Konzentration zur Zeit $C(t)$ ist dann

$$C(t) = C_0 e^{-kt}, \quad k > 0. \quad (31)$$

Die Rate k wird in Einheiten reziprok zur Zeit angegeben, dh kt ist dimensionslos. Die Wirkung des Gifts bestehe darin, eine metabolische oder physiologische Funktion zu hemmen, und die Hemmung dauere so lange, bis die Konzentration ein kritisches Niveau C_L angenommen hat. Das Überleben des Organismus hänge von der Dauer der Inhibition ab: dauert diese länger als t_L , so stirbt der Organismus, für eine geringere Dauer lebt der Organismus. Die minimale lethale Dosis für einen gegebenen Organismus sei

$$D_L = C_L e^{kt_L} \quad \text{oder} \quad \log D_L = kt_L + \log C_L. \quad (32)$$

Für verschiedene Organismen sei der Wert von k verschieden, so dass k als zufällige Veränderliche betrachtet werden kann, wenn Organismen in eine Strichprobe aufgenommen werden. Insbesondere sei k normalverteilt. Dann ist D_i log-normalverteilt.

Der Punkt bei diesem Beispiel ist, dass die Reaktion mit einem Alles-oder-Nichts-Prozess gekoppelt ist; üblicherweise kehrt eine chemische Reaktion zu ihrer normalen Rate zurück, wenn die Inhibition aufgehoben wird. Das hier

⁵iid = independent, identically distributed

betrachtete Beispiel wird also für bestimmte Organismen oder Zellen zutreffen, für die ein solcher Prozess gilt.

Man kann ebenfalls einen Wachstumsprozess betrachten, wie er sich zB aus selbst-reproduzierenden Organismen oder allgemein Elementen ergibt (autokatalytischer Prozess);

$$N = N_0 e^{\lambda T}, \quad (33)$$

wobei N die am Ende erreichte Anzahl und N_0 der Startwert ist. λ ist die Wachstumsrate, und T ist die Dauer. Ist λ oder T normalverteilt, so ist N log-normalverteilt. Koch (1966) nennt die embryonale Zelldifferentiation als Beispiel. Dabei gibt es Perioden, in denen verschiedene Zelltypen unbeschränkt wachsen. Wenn die Wachstumsraten der Zelltypen zufällig fluktuieren, wird der Anteil der Zelltypen am Ende einer Periode log-normalverteilt sein.

2.1.2 Gibrats Gesetz und Shockleys Modell

Gibrat (1931) untersuchte das Wachstum von Firmen und stellte die Hypothese auf, dass sie proportional zu ihrer Größe wachsen, wobei die Wachstumsrate zufällig verteilt ist. Demnach soll gelten

$$x_t - x_{t-1} = \varepsilon_t x_{t-1}, \quad (34)$$

wobei x_{t-1} und x_t die Firmengrößen zu den Zeiten $t-1$ und t sind und ε_t die zufällige Wachstumsrate zur Zeit t ist. Es folgt

$$x_t = x_{t-1} + \varepsilon_t x_{t-1} = (1 + \varepsilon_t) x_{t-1}$$

und daraus durch sukzessives Einsetzen schließlich

$$x_t = x_0 (1 + \varepsilon_1)(1 + \varepsilon_2) \cdots (1 + \varepsilon_t). \quad (35)$$

Ist das Zeitintervall $[t-1, t)$ hinreichend klein, so dass $\log(1 + \varepsilon_t) \approx \varepsilon_t$, so folgt

$$\log x_t = \log x_0 + \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_t \quad (36)$$

und unter den Bedingungen des Zentralen Grenzwertsatzes ist für hinreichend großen Wert für t und $\mu = \mathbb{E}(\varepsilon_t)$, $\sigma^2 = \text{Var}(\varepsilon_t)$

$$\log x_t - \log x_0 = \log \left(\frac{x_t}{x_0} \right) \sim N(t\mu, t\sigma^2). \quad (37)$$

Dies bedeutet, dass die Firmengrößen für hinreichend großen t -Wert log-normalverteilt sind.

Natürlich gilt dieses Resultat nicht nur für Firmengrößen, sondern für jede Größe, für die der Ansatz (34) gemacht werden kann. Unter anderem sind

die Vermögen in einer Population (in einem Land) in guter Näherung log-normalverteilt, – bis auf große Vermögen, für die die Pareto-Verteilung gilt. Darauf wird weiter unten noch ausführlicher eingegangen. Der Grund für die Log-Normalverteilung ist, dass "normale" Einkommen bzw Vermögen in guter Näherung proportional zum vorhandenen Vermögen wachsen, was zunächst eine Exponentialverteilung der Vermögen impliziert. Da aber die Zinssätze zufällig und in guter Näherung unabhängig vom Vermögen variieren, ergibt sich eine Log-Normalverteilung der Einkommen.

Gibrats Gesetz kann auch auf die Größen und Wachstumsraten von Städten angewendet werden. Ein proportionaler Wachstumsprozess erzeugt die log-Normalverteilung für Städtegrößen, – nur für das obere Ende wird die Größe approximativ durch das Zipfsche Gesetz beschrieben, einfach weil das obere Ende der log-Normalverteilung und das Zipfsche Gesetz empirisch nicht gut unterscheidbar sind. Allgemein gilt, dass Prozesse die durch Gibrats Gesetz beschrieben werden können, in Abhängigkeit von spezifischen Annahmen über den stochastischen Prozess gegen die log-Normalverteilung oder gegen eine Potenzfunktion konvergieren.

Autokatalytische Prozesse Allgemein kann man einen autokatalytischen Prozess der Form

$$\frac{dX}{dt} = \pm kX, \quad (38)$$

betrachten. Es folgt

$$\frac{dX}{X} = kdt.$$

Nun ist $\int dX/X = \log X + c$, dh

$$\frac{1}{X} = \frac{d \log X}{dX},$$

so dass unter Berücksichtigung von (38)

$$\frac{d \log X}{dt} = \frac{d \log X}{dX} \frac{dX}{dt} = \pm \frac{1}{X} kX = \pm k \Rightarrow \log X(t) = \log X_0 \pm kt \quad (39)$$

Die Konstante k sei nicht wirklich konstant, sondern variere zufällig, so dass zur Zeit t

$$(\log X)_{t+1} = (\log X)_t \pm k_t, \quad (40)$$

wobei k_t der zufällige Wert von k zur Zeit t sei. Man hat wieder

$$(\log X)_{t+1} = (\log X)_0 \pm \sum_{s=0}^t k_s, \quad (41)$$

und der Zentrale Grenzwertsatz liefert eine Log-Normalverteilung, auch wenn die k_t nicht Gauß-verteilt sind.

Shockleys Modell Shockley⁶ (1957) ging davon aus, dass es mindestens acht Fähigkeiten oder Eigenschaften bedarf, um einen wissenschaftlichen Artikel zu veröffentlichen: (1) die Fähigkeit, ein geeignetes Problem zur Bearbeitung auszusuchen, (2) die Kompetenz, das Problem zu bearbeiten, (3) die Fähigkeit, ein relevantes Resultat als solches zu erkennen, (4) die Fähigkeit, einen geeigneten Schlußpunkt zu setzen und mit dem Verfassen des Manuskripts zu beginnen, (5) die Fähigkeit, die Resultate und Schlußfolgerungen adäquat zu präsentieren, (6) die Fähigkeit, die Kritiken anderer zu integrieren, die von anderen Wissenschaftlern gemacht werden, (7) den festen Willen, das Manuskript fertigzustellen und einzureichen, (8) die Fähigkeit, positiv auf die Kritiken der Gutachter zu reagieren. Es sei p_i die Wahrscheinlichkeit, über die i -te dieser Komponenten zu verfügen, und weiter seien die Komponenten stochastisch unabhängig voneinander. Die Wahrscheinlichkeit, dass ein Wissenschaftler einen Artikel publiziert, ist dann durch

$$P = p_1 p_2 \cdots p_8 \quad (42)$$

gegeben. Dann ist aber

$$\log P = \log p_1 + \log p_2 + \cdots + \log p_8. \quad (43)$$

Unter allgemeinen Bedingungen kann man nun den Zentralen Grenzwertsatz anwenden, dh $\log P$ sollte in guter Näherung normalverteilt sein. Damit ist aber P log-normalverteilt, und dieser Befund korrespondiert zu den Shockleyschen Daten. Die inhaltlichen Punkte (1) bis (8) sind dabei gar nicht so wichtig. Interessant ist die Anwendung des Zentralen Grenzwertsatzes für die Verteilung von $\log P$. Zwar weiß man schon aus (9), dass ein Produkt log-normalverteilt ist, wenn nur die Faktoren log-normalverteilt sind, aber bezöge man sich nur auf (9), so müßte man begründen, warum eine individuelle Wahrscheinlichkeit p_i log-normalverteilt sein soll: p_i kann nicht log-normalverteilt sein, da p_i auf $(0, 1)$ beschränkt ist, eine log-normalverteilte Variable aber auf $[0, \infty)$ definiert ist.

Die Publikation eines wissenschaftlichen Artikels ist nur ein Beispiel, das für alle komplexen Prozesse mit vielen Komponenten steht. Es werde nun ein Prozess mit n Komponenten betrachtet. Jede der Komponenten sei mit einer

⁶William Bradford Shockley Jr. (1910 – 1989), Physiker; erfand zusammen mit John Bardeen und Walter Houser Brattain den Transistor. Die drei bekamen dafür 1956 den Nobel-Preis für Physik. Shockley kommerzialisierte seine Erfindung, aber seine Mitarbeiter rebellierten und machten sich ihrerseits selbstständig. Dies war der Startpunkt für das Silicon Valley. Er war auch Anhänger eugenischer Ideen und schlug vor, das genetisch Benachteiligten eine finanzielle Belohnung bekommen sollten, wenn sie sich freiwillig sterilisieren ließen, – generell hielt er Schwarze für genetisch bedingt weniger intelligent als Weiße, und da sie sich stärker vermehrten als Weiße, befürchtete Shockley eine "retrogression in human evolution", vergl. den Nachruf in der New York Times <http://www.nytimes.com/learning/general/onthisday/bday/0213.html>.

Varianz σ_j^2 assoziiert, und es sei

$$\bar{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n \sigma_j^2 \quad (44)$$

Die Varianz der Log-Normalverteilung sollte dann in der Größenordnung $n\bar{\sigma}^2$ sein. Je größer der Wert von n , desto größer ist der Bereich, über dem die Log-Normalverteilung eine $1/f$ -Verteilung approximiert.

2.1.3 Das obere Ende: Der Montroll-Schlesinger-Ansatz

Pareto (1897) argumentierte, dass der Anteil der Personen mit einem Einkommen größer als x in guter Näherung proportional zu $1/x^k$ sei,

$$P(X \geq x) = Ax^{-k}, \quad x > 0, \quad k \geq 0 \quad (45)$$

A eine Normierungskonstante. - bei Einkommensverteilungen findet man oft $k \approx 1.5$ (für $k \leq 1$) existiert kein Erwartungswert). Dieser Ausdruck korrespondiert zu (11) mit $a = 0$ und $b = 1$. Der Befund (45) gilt für unterschiedliche Länder, die Verteilungen unterscheiden sich nur durch Parameterwerte. Allerdings beschreibt (45) nicht die gesamte Einkommensverteilung in einer Population; spätere Untersuchungen zeigen, dass diese Beziehung nur für die oberen Einkommen gilt, - der Hauptteil der Einkommen ist log-normalverteilt. Eine Beziehung der Form (45) gilt auch für die Anzahl von Publikationen, die von Wissenschaftlern geschrieben werden, - aber im oberen Bereich. Lotka (1926) fand, dass die Anzahl N von Chemikern, die mehr als n Artikel geschrieben haben, durch die Beziehung $P(\nu > n) = A(1/n^k)$ gegeben ist, wobei $k \approx 2$ sei. Shockleys Beobachtungen, die zur Log-Normalverteilung korrespondieren, beziehen sich möglicherweise auf eine andere Population von Wissenschaftlern.

Es sei $F_{ln}(x)$ die Verteilungsfunktion (also $P(X \leq x)$) der Log-Normalverteilung und $F_p(x)$ die Verteilungsfunktion der Pareto-Verteilung. Generell verhalten sich $P(X > x) = 1 - F_{ln}(x)$ und $P(X > k) = 1 - F_p(x)$ von einem von den Parametern der Verteilungen abhängigen x_0 -Wert an sehr ähnlich (vergl. Abb. LN-Pareto-Vergleich).

Abbildung verteilungsvergleich hier, ebenso Montroll-Schlesinger-Abb

Neuere Daten zeigen, dass für Einkommensverteilungen die Log-Normalverteilung im Bereich zwischen dem 5-ten und dem 97-ten Perzentil einen exzellenten Fit hat und die obersten 2 - 3 % davon abweichend durch eine Pareto-Verteilung beschrieben werden. Dies legt nahe, dass für den 5 - 97 Perzentilbereich das Shockley-Modell eine gute Erklärung für die Einkommensverteilung ist, während die oberen 2 - 3% Möglichkeiten zur Einkommensvermehrung haben, die von denen der übrigen abweichen und für die das Shockley-Modell

nicht mehr gilt. Eine Möglichkeit der Erklärung besteht darin, für den unteren Perzentilbereich eine Log-Normalverteilung mit einem Parametersatz (I) zu postulieren, und für den oberen eine Log-Normalverteilung mit einem Parametersatz (II), wobei für diese zweite Verteilung nur das obere Ende in den Daten aufscheint. Alternativ dazu könnte man annehmen, dass für den oberen Teil Mechanismen verantwortlich sind, die für Verteilungen vom Potenztyp typisch sind, über die im Folgenden noch geredet werden wird.

Montroll & Schlesinger (1982, 1983) verfolgen einen etwas anderen Ansatz. Sie gehen ebenfalls davon aus, dass für die oberen 2 - 3% (oder gar nur für die oberen 1%) andere Einkommensgesetze gelten als für die unteren Bereiche. Die Reichen haben eine Möglichkeit der Vermehrung (*amplification*) ihres Einkommens, die den Beziehern unterer Einkommen verwehrt ist. Montroll et al. zitieren die Beatles: auf dem Höhepunkt ihrer Popularität wurde jede der von ihnen neu herausgebrachten Schallplatten sofort millionenfach verkauft, wodurch ihr Einkommen jedesmal einen beträchtlichen Zuwachs erhielt, der mit einem Gehalt entlohnten Menschen nicht zukommt. Personen, die ihr Geld durch Investitionen vermehren, verfügen ebenfalls über eine beträchtliche *leverage*. Läuft die Wirtschaft gut, ist Geld leichter verfügbar (u.a. niedrige Zinsen), das wiederum investiert werden kann. Insbesondere können etwa Aktien durch eine Anzahlung von etwa 10% gekauft werden. Der Wert der Aktien erhöht sich um ein Mehrfaches, so dass die noch ausstehende Restzahlung klein im Vergleich zum Gewinn ist. Montroll et al zitieren J. P. Morgan⁷, der mit der Anhäufung seines Vermögens in der beschriebenen Art begann.

Montroll et al. liefern ein Modell der Amplifikation, dass das Pareto-Ende der Verteilung beschreibt. Es sei \bar{x} der Erwartungswert der Log-Normalverteilung ausschließlich des Pareto-Endes. Betrachtet wird nun die Verteilung von $Y = X/\bar{x}$; offensichtlich ist Y frei von Maßeinheiten. Weiter sei λ die Wahrscheinlichkeit einer auf \bar{x} wirkenden Amplifikation $\bar{x} \rightarrow N\bar{x}$, und Y sei nun mit diesem neuen Parameter $N\bar{x}$ verteilt. Die Verteilung $g(y) = g(X/\bar{x})$ geht dann über in die Verteilung $g(X/N\bar{x})$. Mit der Wahrscheinlichkeit λ^2 schließt sich nun eine weitere Amplifikation an, mit $N\bar{x} \rightarrow N^2\bar{x}$, etc. Die Verteilung von Y ergibt sich dann gemäß

$$G(Y) = (1 - \lambda) \left[g(y) + \frac{\lambda}{N} g\left(\frac{\lambda}{N}\right) + \frac{\lambda^2}{N^2} g\left(\frac{Y}{N^2}\right) + \dots \right]. \quad (46)$$

Der Wert von λ definiert den Bereich, für den die Anfangsverteilung gilt. Ersetzt man Y durch Y/N , so erhält man

$$G(Y) = \frac{\lambda}{N} G(Y/N) + (1 - \lambda)g(Y). \quad (47)$$

⁷1837–1913, der von dem berühmten Mathematiker Felix Klein eine Privatdozentenstelle in Göttingen angeboten bekommen hatte. Morgan hatte aber von seinem Vater bereits eine Million Dollars bekommen. Statt die Stelle zu akzeptieren investierte er die Million in der beschriebenen Weise, – der Rest ist bekannt (<http://de.wikipedia.org/wiki/J.-P.-Morgan>)

Für $\lambda \rightarrow 0$ gibt es keine Amplifikation und $G(Y) = g(y)$. Mit größer werdendem Wert von λ folgt $g(Y) \rightarrow 0$. Wenn $g(Y) \rightarrow 0$ schneller als $G(y) \rightarrow 0$ folgt

$$G(Y) = \frac{\lambda}{N} G\left(\frac{\lambda}{N}\right). \quad (48)$$

Setzt man $G(Y) = AY^{-1-\mu}$, so erhält man

$$\mu = [\log(1/\lambda)] / \log N. \quad (49)$$

Für die Daten in Abb. 1 findet man $\mu = 1.63$, und für $\lambda = .01$ ergibt sich eine Amplifikation von 16.8. Dieser Wert korrespondiert zu dem Befund, dass sich für eine kleinere Firma die beste Amplifikation bei 15 bis 20 Angestellten ergibt.

Um den Übergang von einer Log-Normalverteilung zu einer Pareto-Verteilung zu modellieren, werde zunächst noch einmal die Log-normalverteilung betrachtet:

$$\begin{aligned} g(\log x) &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \\ &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \log m)^2}{2\sigma^2}\right), \\ &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x/m))^2}{2\sigma^2}\right), \quad m = e^\mu. \end{aligned} \quad (50)$$

Montroll et al betrachten x/m und drücken x als Vielfaches f von m aus:

$$x = fm. \quad (51)$$

Weiter gelte

$$f = e^n, \quad (52)$$

und es gelte $\sigma^2 \rightarrow \infty$, – dh σ^2 sei groß gegen x . Dann hat man

$$\log g(f) = -n - \frac{1}{2} \left(\frac{n}{\sigma}\right)^2 - \frac{1}{2} \log(2\pi\sigma^2). \quad (53)$$

Für einen großen Wert von σ^2 kann man nun den größten Wert von n schätzen, für den (53) innerhalb bestimmter Genauigkeitsgrenzen linear in n ist, wobei der konstante Term vernachlässigt werden kann. Betrachtet man den Quotienten des mittleren und des ersten Terms in (53), so sei

$$\frac{(n/\sigma)^2/2}{n} \leq \theta|n| \quad \text{oder} \quad |n| \leq 2\theta\sigma^2. \quad (54)$$

Ist $\theta = .1$, $\sigma = 5$ so folgt für alle $|n| \leq 5$, dass $g(f)$ eine $1/f$ -Verteilung mit 10%-iger Genauigkeit approximiert. Dies bedeutet 11 Werte $n \in \mathbb{N}$, oder 11 e -Faktoren (vergl. (52)) oder 4 Größenordnungen. Es läßt sich zeigen, dass allgemein $g(f)$ eine $1/f$ -Verteilung für $(4\theta\sigma^2 + 1)$ e -Faktoren innerhalb eines relativen Fehlers θ approximiert. Je größer σ^2 , desto größer ist der Bereich, in dem die $1/f$ -Verteilung approximiert wird.

2.1.4 Thermodynamische Modelle

Viele Untersuchungen legen nahe, dass ca 90% des Einkommens in einer Population in sehr guter Näherung durch eine Log-Normalverteilung beschrieben werden können

Die Verteilung von Einkommen in einer Gesellschaft scheint auf den ersten Blick log-normal verteilt zu sein; Souma (2001) spricht von einer universellen Struktur der Verteilung des persönlichen Einkommens, – die aber nicht ausschließlich durch eine Log-Normalverteilung gegeben sei, sondern an ihrem oberen Ende durch eine Verteilung, die durch eine Potenzfunktion beschrieben wird. Andere Autoren argumentieren, es sei weniger das Einkommen, sondern der Verbrauch (consumption), der durch eine Log-Normalverteilung charakterisiert wird (Battistin, Blundell & Lewbell, 2007).

Das Einkommen beruht nicht zuletzt auf der Interaktion mit anderen Mitgliedern der Gesellschaft. Die Frage ist, wieweit hierbei Zufälligkeiten eine Rolle spielen. Im Extremfall könnte es sein, dass die beobachtete Verteilung durch stochastische Prinzipien, wie sie in der Statistischen Physik (Thermodynamik) diskutiert werden, bestimmt wird. Derartige Anwendungen der Statistischen Physik werden seit einiger Zeit verfolgt, man spricht dementsprechend von *econophysics*.

Die Physik hatte bereits früh zur Theorienbildung in der Ökonomie beigetragen; bereits Pierre-Simon Laplace und der belgische Astronom Alphonse Quetelet waren der Ansicht, dass bestimmte physikalische Gesetzmäßigkeiten auch in sozialen Systemen aufzufinden seien, ebenso John Stuart Mill und sogar Karl Marx. Das Leitbild war zunächst die newtonsche Physik. Erst die Ökonomen Francis Edgeworth und Alfred Marshall übernahmen die Ideen von Maxwell und Boltzmann aus der statistischen Physik, insbesondere die Idee, dass ökonomische Systeme wie Gase gegen bestimmte Gleichgewichtszustände konvergierten. Diese Annahme wurde zur Grundannahme der neoklassischen Ökonomie, die heute noch von der Mehrheit der Ökonomen gelehrt wird. Es waren wiederum Physiker, die darauf hinwiesen, dass (nicht-lineare) Systeme, wie sie in der Ökonomie betrachtet werden müssten, nicht nur mehr als ein Gleichgewicht haben können und von einem Equilibriumzustand in einen anderen gelangen können, sondern dass ökonomische Systeme im Allgemeinen weit entfernt von einem solchen Zustand operieren. Ökonomen nehmen diesen Sachverhalt offenbar ungern zu Kenntnis ("Equilibrium thinking still has them in its unshakeable thrall", so der kritische Ökonom Steven Keen, zitiert in Ball (2006)). Ball (2006) gibt einen kurzen Überblick über die Diskussionen über die *econophysics*, worauf hier nicht weiter eingegangen werden kann. Hier soll nur gezeigt werden, dass sehr einfache Ansätze bereits zu log-normalen Verteilungen und deren Übergang zu Potenzverteilungen führen können.

Zur Illustration wird kurz der Ansatz von (Chatterjee, Sinha & Chakra-

barti, 2007) vorgestellt. Gegeben seien N "Spieler", und zur Zeit $t = 0$ habe jeder Spieler eine Einheit "Geld". Das Gesamtkapital ist dann $M = N$, und N sei groß. Weiter bleibe M konstant für alle folgenden Zeitpunkte. Zu je einem Zeitpunkt werden zwei Spieler zufällig gewählt. Sie teilen das Kapital, das sie bis zu diesem Zeitpunkt akkumuliert haben, wiederum zufällig zwischen sich auf, – Schulden dürfen sie allerdings nicht machen. Dadurch wird die anfangs gegebene Gleichverteilung aufgehoben und die Frage ist, ob es eine Verteilung von Kapital gibt, gegen die Verteilungen zu den verschiedenen Zeitpunkten streben. Die Frage wurde in etwas anderem Zusammenhang in der statistischen Physik (Thermodynamik) diskutiert und beantwortet: Die Wahrscheinlichkeit, dass ein Spieler über den Betrag m verfügt, beträgt

$$P(m) = A \exp\left(-\frac{m}{T}\right), \quad T = \frac{M}{N} \quad (55)$$

T ist der durchschnittliche Betrag, über den ein Mitglied der Gemeinde der Spieler verfügt. $P(m)$ ist als Boltzmann- oder Gibbs-Verteilung bekannt⁸.

Nun werde angenommen, dass jeder Spieler einen bestimmten Betrag λ von seinem bisher angesammelten Betrag spart und den Rest wie zuvor unter den Mitspielern verteilt. Jeder Spieler spare den gleichen Betrag. Die Frage ist nun, wie die *steady-state* (Gleichgewichts-) Verteilung nach einer großen Anzahl von Schritten aussieht. Es wird eine Gamma-Verteilung

$$P(m) = Am^a \exp\left(-\frac{m}{\lambda T}\right), \quad a = \frac{3\lambda}{1-\lambda}, \quad 0 \leq \lambda \leq 1, \quad (56)$$

A eine Normierungskonstante.

Nun sei λ nicht identisch für alle Spieler. Es sei $\rho(\lambda)$ die Wahrscheinlichkeitsverteilung für die Neigung zum Sparen, wenn λ vorliegt, und $\rho(\lambda) \rightarrow \rho_0 > 0$ für $\lambda \rightarrow 1$. Dann läßt sich zeigen, dass $P(m) \rightarrow m^{-(1+\nu)}$ mit $\nu = 1$ für $m \rightarrow \infty$ (Chatterjee, Chakrabarti & Manna (2004), für $\rho(\lambda) \rightarrow 0$ für $\lambda < \lambda_0$ strebt die Verteilung $P(m)$ allerdings gegen eine Gamma-Verteilung für kleine und mittlere Werte von m , die allerdings durch eine Log-Normalverteilung approximiert werden kann; diese Verteilung wird Pareto-artig für $m \rightarrow \infty$, $\nu = 1$.

2.2 Potenz-(Pareto-)Verteilungen

Potenzverteilungen können durch verschiedene Mechanismen zustandekommen. Im Folgenden werden verschiedene Mechanismen vorgestellt.

⁸Die Boltzmann-Verteilung ist durch $N_i/N = g_i \exp(-E_i/(k_B T))$ definiert, N_i die Anzahl der Partikel mit Energie E_i , $N = \sum_i N_i$, T die Temperatur, k_B die Boltzmann-Konstante, $Z(T)$ eine Partitionsfunktion, vgl. Fließbach (1995)

2.2.1 Die Kombination von Exponentialfunktionen

Eine Möglichkeit, Potenz-Verteilungen zu erhalten, ergibt sich aus der Kombination von exponentialverteilten Größen. Gesucht werde die Verteilung einer zufälligen Veränderlichen X , die über eine Exponentialfunktion mit einer Größe Y verbunden ist: $X = \exp(\alpha Y)$, α eine Konstante, und Y sei exponentialverteilt. Es gelte also $f(y) = \lambda \exp(-\lambda y)$ und somit $F(y) = 1 - \exp(-\lambda y)$. Es ist

$$\begin{aligned} P(X \leq x) &= P(\exp(\alpha Y) \leq x) = P\left(Y \leq \frac{1}{\alpha} \log x\right) \\ &= 1 - \exp(-\lambda(\log x^{1/\alpha})) = 1 - x^{-\lambda/\alpha}, \end{aligned} \quad (57)$$

also eine Potenzverteilung mit der Dichtefunktion

$$f(x) = \frac{dP(X \leq x)}{dx} = \frac{\lambda}{\alpha} x^{-\lambda/\alpha - 1}. \quad (58)$$

Die Beziehung $X = \exp(\alpha Y)$ impliziert ja

$$\frac{dX}{dY} = \alpha X,$$

dh die Veränderung von $X(Y)$ ist proportional zu $X(Y)$ selbst. Y variiert zufällig. Betrachtet man die Größen zu verschiedenen Zeitpunkten t , so ist $X_t = e^{Yt}$.

Beispiel 2.1 Verteilung von Größen M_i mit exponentiellem Wachstum Gegeben sei eine Menge von Größen k_i , $i = 1, \dots, n \dots$, die exponentiell wachsen, d.h. $k_i = \exp(-\mu(t - t_i))$, wobei $t \geq t_i$. t_i ist der Zeitpunkt, zu dem das Wachstum begann, so dass für $t = t_i$ folgt, dass $k_i = 1$, d.h. das Wachstum beginnt, sobald eine Einheit existiert. Eine der Größen M_i wird zur Zeit t zufällig gewählt und gesucht ist die Wahrscheinlichkeit $P(k_i < k)$. Es ist

$$\begin{aligned} P(k_i < k) &= P(\exp(-\mu(t - t_i)) < k) \\ &= P(\mu(t - t_i) < \log k) \\ &= P(t - t_i < \frac{1}{\mu} \log k) \\ &= P(t < \log k^{1/\mu}) \\ &= 1 - P(t_i < t - \log k^{1/\mu}) \end{aligned} \quad (59)$$

Es sei $n(t_0)$ die Anzahl der Größen M_i , für die $t_i < t_0$, und $n(t)$ sei die Anzahl der Größen zum Zeitpunkt $t \geq t_0$. Weiter gelte $n(t_0) = \exp(\nu t_0)$ und $n(t) = \exp(\nu t)$. Dann ist

$$P(t_i < t_0) = \frac{n(t_0)}{n(t)} = \frac{\exp(\nu t_0)}{\exp(\nu t)} = \exp(\nu(t_0 - t)). \quad (60)$$

Es sei insbesondere $t_0 = t - \log k^{1/\mu}$. Dann ist

$$\begin{aligned} P(k_i < k) &= 1 - \exp(\nu(t - \log k^{1/\mu} - t)) \\ &= 1 - \exp(-\nu(\log k^{1/\mu})) \\ &= 1 - k^{\mu/k} \end{aligned} \quad (61)$$

Die Wahrscheinlichkeitsdichte für k erhält man durch Differentiation:

$$f(k) = \frac{\nu}{\mu} k^{-(1+\nu/\mu)}. \quad (62)$$

Dies ist gerade die Pareto-Dichte. \square

2.2.2 Rauschen in Halbleitern und die Fluten des Nils

van der Ziel (1950) nahm an, dass das Rauschen in Halbleitern als $1/f$ -Rauschen zu betrachten sei. Um seiner Argumentation folgen zu können, sei daran erinnert, dass die kinetische Energie (also seine Bewegungsenergie) durch $E_{kin} = mv^2/2$ gegeben ist, wobei m die Masse des Körpers und v seine Geschwindigkeit ist. Die Varianz σ^2 als Maß der "Beweglichkeit" einer zufälligen Veränderlichen ist ebenfalls als Summe von Quadraten definiert, gewissermaßen in Analogie zur Energie – je größer die Varianz, desto größer ist die Energie, mit der sie unterschiedliche Werte annimmt und desto größer ist der Bereich dieser Werte. van der Ziel nahm an, dass die Bewegung τ der Elektronen durch eine Energie aktiviert wurde, $\tau \propto \exp(E/kT)$, wobei die Boltzmann-Verteilung ist, E die Energie, T die Temperatur, k eine Proportionalitätskonstante. Innerhalb eines bestimmten Bereiches sollten alle E -Werte mit gleicher Wahrscheinlichkeit auftreten. Über den Zentralen Grenzwertsatz folgert man dann, dass $\log \tau$ log-normalverteilt ist und damit eine $1/f$ -Verteilung approximiert wird. Machlup (1981) argumentiert in diesem Zusammenhang, dass das $1/f$ -Rauschen den Sachverhalt reflektiert, dass die Natur "is sufficiently chaotic to possess . . . a large ensemble of mechanisms with no prejudice about scale." Das Rauschen in einem Halbleiter ist demnach nur Ausdruck der Tatsache, dass die Natur skaleninvariant ist. "Rein zufällige" Prozesse haben eine Autokorrelation $\propto \exp(-t/\tau)$. Das dazu korrespondierende Power Spectrum ist die Fourier-Transformierte dieser Funktion, und die ist durch

$$S(\omega) \propto \frac{\tau}{1 + \omega^2 \tau^2} \quad (63)$$

gegeben. Man habe nun eine große Ansammlung von Zufallsprozessen dieser Art, die sich nur durch die τ -Werte unterscheiden. $\rho(\tau)$ sei die Wahrscheinlichkeitsdichte für die τ -Werte. Das Power Spectrum dieser Mischung von Prozessen hängt nun von der Verteilung $\rho(\tau)$ der τ -Werte – der Korrelationszeiten – ab. Angenommen, es gelte

$$\rho(\tau) \propto \frac{d\tau}{\tau}. \quad (64)$$

Das Power Spectrum ist dann

$$\int_{\tau_1}^{\tau_2} S_{\tau}(\omega)\rho(\tau)d\tau = \int_{\tau_1}^{\tau_2} \frac{\tau}{1 + \omega^2\tau^2} \frac{d\tau}{\tau} = \frac{\tan^{-1}\omega\tau}{\omega} \Big|_{\tau_1}^{\tau_2}. \quad (65)$$

(64) bedeutet Skaleninvarianz, (65) definiert den Bereich (τ_1, τ_2) für die Skaleninvarianz. Ist dieser Bereich hinreichend groß (τ_2/τ_1 ist hinreichend groß), so ist das Spektrum gleich $1/\omega$ in einem hinreichend großen Bereich.

Ist also der betrachtete Prozess multiplikativ im Sinne von (42), so ist die korrespondierende Verteilung die Log-Normalverteilung, die durch eine $1/\omega$ -Verteilung approximiert wird. Log-Normalverteilungen kann man erwarten, wenn sie einem auf Benjamin Franklin zurückgehenden Sprichwort entsprechen: "For the want of a nail a shoe was lost, for want of a shoe the horse was lost, and for want of a horse the rider was lost, etc" (Montroll et al, p. 3381).

Die Fluten des Nils Der Wasserfluss des Nils weist – wie wohl andere Flüsse auch – $1/f$ -Rauschen auf. Das Shockley-Modell kann auch hier angewendet werden. Montroll & Schlesinger betrachten dazu die Stadien, die einen Wassertropfen von der Quelle des Nils bis zur Mündung bringen: zunächst müssen die atmosphärischen Bedingungen gegeben sein, damit sich ein Wassertropfen als Teil des Regens bildet, dann müssen Wind- Temperatur- und Bodenbeschaffenheit (Porosität) in geeigneter Weise zusammenkommen, damit der Tropfen die verschiedenen Stationen des Flusses "überlebt". Die Fließrate des Flusses ist log-normalverteilt und liefert in der beschriebenen Weise ein $1/f$ -Rauschen.

2.2.3 Selbst-organisierte Kritikalität

Ein spezieller, wenn auch oft gefundener Fall liegt vor, wenn $S \propto 1/f^\alpha$; das entsprechende Rauschen ist das $1/f$ -Rauschen. Viele Phänomene, die im Rahmen der statistischen Physik untersucht werden, zeigen ein $1/f$ -Rauschen: man findet es in elektrischen Widerständen, in den Fluktuationen des Fließens des Nils, in Sanduhren, in der Helligkeit von Sternen, etc. Die Frage ist, wie es zu deuten, dh auch: zu erklären ist. Bak, Tang & Wiesenfeldt (1987) den der *selbstorganisierten Kritikalität* (self-organised criticality) vor.

Die Beschreibung soll auf intuitive Betrachtungen beschränkt werden, um einen Ausflug in die Theorie komplexer Systeme zu vermeiden. Bak et al. wählen dazu das Beispiel gekoppelter, gedämpfter Pendel. Ein Pendel werde angestossen. Wegen der Kopplungen pflanzt sich diese Störung durch das Netzwerk von Pendeln fort. Das Netzwerk ist instabil schon für kleine Störungen. Am Ende wird das System aber gegen einen stabilen Zustand streben. Es wird keine spezifische Länge in diesem System geben, so dass es eine skaleninvariante Struktur stabiler Zustände endet. Der hier beschriebene Prozess ist selbstorganisierend, das Rauschen (die fluktuierende Störung) wird eine Autokorrelation vom Typ Potenzfunktion zeigen. Die Zeitdauern der Fluktuationen werden, wegen der Skaleninvarianz, vom Typ $F(t) \propto t^{-\alpha}$ sein, und eine

Störung kann entweder nur ein Pendel in Bewegung setzen, die bei anderen praktisch nicht mehr ankommt, oder sie kann, im anderen Extrem,

Literatur

- [1] Bak, P., Tang, C., Wiesenfeldt, K. (1987) Self-Organized Criticality: An Expansion of $1/f$ Noise. *Physical Review Letters* 59 (4), 381–384
- [2] Barabási, A.L., Albert, R. (1999) Emergence of Scaling in Random Networks. arXiv:cond-mat/9910332v1 [cond-mat.dis-nn] 21 Oct 1999 (<http://arxiv.org/pdf/cond-mat/9910332.pdf>)
- [3] Ball, P. (2006) Culture Crash, *Nature*, 441, 686 – 688
- [4] Barabási, A.L. (2005) The origin of bursts and heavy tails in human dynamics. *Nature*, 435, 207–211
- [5] Battistin, Blundell, R., Newbell, A. (2007) Why is consumption more log-normal than income? Gibrat’s law revisited. *The Insitute of Fiscal Studies*
- [6] Chater, N., Brown, D.A.(1999) Scale-Invariance as a unifying Principle-*Cognition*, 69, B17–B24
- [7] Chatterjee, A., Sinha, S., Chakrabarti, B. (2007) Economic inequality: Is it natural? *Current Science*, 92 (10), 1383 – 1389
- [8] Chatterjee, B. K., Chakrabarti, B.K., Manna, S.S. (2004) Pareto law in a kinetic model of market with random saving propensity. *Pghysica A*, 335, 155 – 163
- [9] Frette, V., Christensen, K., Malthe-Sørensen, A., Feder, J., Jøssang, T., Meakin, P. (1996) Avalanche Dynamics in a pile of rice. *Nature*, 379, 49–52
- [10] Fließbach, T.: Statistische Physik. Heidelberg 1995
- [11] Linkenkaer-Hansen, K., Nikouline, V.V., Ilmoniemi, R. J. (2001) Long-Range temporal correlations and scaling behaviour in human brain oscillations. *The Journal of Neuroscience*, 21 (4), 1370–1377
- [12] Lotka, A. J. (1926) The frequency distribution of scientific productivity. *Journal of gthe Washington Academi of Science*, 16, 317 – 323
- [13] Mega, M.S., Allegrini, P., Grigolini, P., Latora, V., Palatella, L., Rapisarda, A., Vinciguerra, S., (2003) Power-law time distribution of large earthquakes. arXiv.org $\dot{\iota}$ cond-mat $\dot{\iota}$ arXiv:cond-mat/0212529 (*Phys. Rev. Lett.*, 90 (2003))

- [14] Mitzenmacher, M. (2003) A brief history of generative models for power laws and log-normal distributions. *Internet Mathematics*, 1(2), 226–251
- [15] Montroll, E. W., Schlesinger, M. F. (1982) On $1/f$ noise and other distributions with long tails. *Proc. Natl. Acad. Sci* 79, 3380–3383
- [16] Montroll, E. W., Schlesinger, M. F. (1983) Maximum entropy formalisms, fractals, scaling phenomena, and $1/f$ -noise: a tale of tails. *Journal of Statistical Physics*, 32 (2), 209 – 230
- [17] Mora, T., Bialek, W. (2011) Are biological systems poised at criticality? *Journal of Statistical Physics*, 144, 268 – 302
- [18] Newman, M.E.J. (2005) Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics* 46 (5), 323 – 351
- [19] Proekt, A., Banavar, J.R., Maritan, A., Pfaff, D.W. (2012) Scale invariance in the dxynamics of spontaneous behavior. *PNAS*, 109 (26), 10564–10569
- [20] Richter C. F. (1935) An instrumental earthquake magnitude scale. In: *Bulletin of the Seismological Society of America*, 25 (1), 1–32.
- [21] Shockley, W. B. (1957) On the statistics of individual variations of productivity in research laboratories. *Proceedings of the IRE*⁹, 1957, 45(3), 279 – 290
- [22] Simon, H. A. (1955) On s classs of skew distribution functions. *Biometrika* 42 (3/4), 425 – 440
- [23] Souma, W. (2001) Universal structure of the pesonan income distribution. *Fractals*, 9 (4), 463 – 470
- [24] Sutton, J. (1997) Gibrat’s legacy. *Journal of Economic Literature*, 35 (März), 40–59
- [25] Yang, C.B. (2004) The origin of power-law distributions in self-organized criticality. *Journal of Physics A: Mathematical and General*, 37, L523 – L529

⁹IRE = Institute of Radio Engineers

Abbildung 1: Einkommensverteilung USA

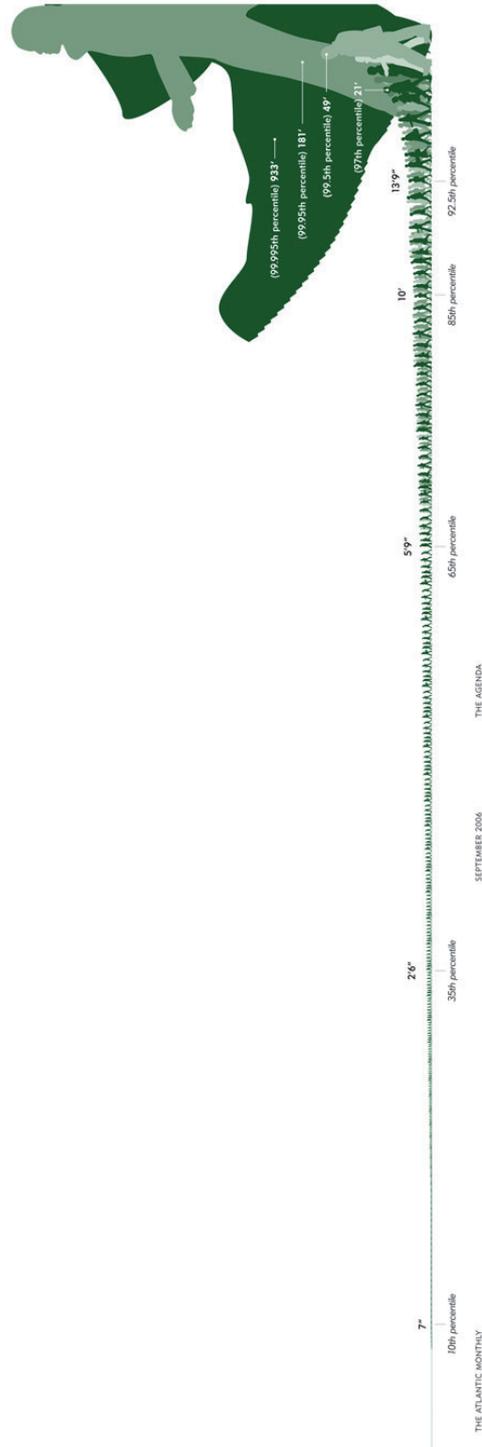


Abbildung 2: Einkommensverteilung USA

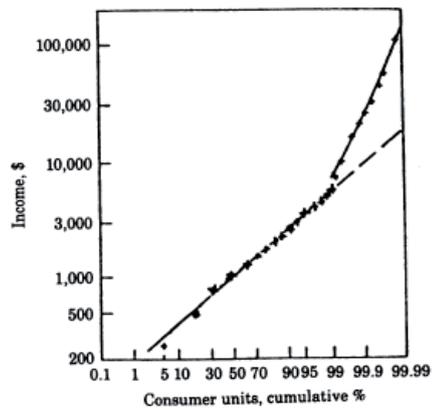


FIG. 1. Distribution of families and single individuals by income level, 1935/1936. Data are from ref. 9. Most of the data follows a log-normal distribution, while the last 1% is governed by a Pareto tail.

Aus Montroll & Sclesinger 1982 Familieneinkommen und Pareto-tail