6 Social Complexity II: Laws

6.1 Introduction and Motivation

In science, laws describe and theories explain. Laws provide understanding of "how" social complexity occurs; theories answer questions of "why" it occurs. Laws are like mappings between variables; theories are causal stories that account for observed social complexity. Which patterns of social complexity have empirical validity as universal laws that hold cross-culturally and across domains of social science research? How is social complexity explained in terms of existing theories?

This chapter develops the analysis of social complexity by presenting theoretical and empirical laws that describe emergence and subsequent dynamics. The main emphasis in this chapter is on formal description for understanding social complexity. The next chapter progresses toward explanatory theories of social complexity. Understanding of basic patterns in laws of social complexity is necessary for developing viable computational models.

6.2 History and First Pioneers

The history of laws of social complexity dates to the early twentieth century, when pioneers such as Vilfredo Pareto, Max O. Lorenz, Corrado Gini, and Felix Auerbach demonstrated the first power laws in human and social domains of science, half a century before power laws entered physics. These early discoveries were soon followed by social power laws discovered by Alfred Lotka, George K. Zipf, Lewis F. Richardson, Herbert A. Simon, and Manus I. Midlarksy. Most recent work on these and other non-equilibrium distributional models focuses on discovering additional domains (e.g., the Internet) as well as replicating earlier discoveries with newly available and better data.

By contrast, research on structural laws of social complexity is more recent, beginning in the Cold War years with the pioneering work of Albert Wohlstetter, William Riker, Martin Landau, Jeffrey L. Pressman, Aaron Wildavsky, Elinor Ostrom, and John W. Kingdon. Research on both types of laws of social complexity is

C. Cioffi-Revilla, *Introduction to Computational Social Science*,
Texts in Computer Science, DOI 10.1007/978-1-4471-5661-1_6,

still active and promises new discoveries as CSS researchers expand the domains of universal patterns.

1896 Economist Vilfredo Pareto [1848–1923] pioneers power laws through his comparative research on income and wealth in his classic textbook, *Cours d'economie politique*.

1905 Max Otto Lorenz [1876–1959] publishes his seminal paper on the curve named after him in the *Journal of the American Statistical Association*, while still a doctoral student at the University of Wisconsin.

1912 Sociologist Corrado Gini [1884–1965] proposes his classic coefficient of inequality in *Mutabilitá e Variabilitá*.

1913 Physicist Felix Auerbach [1856–1933] discovers the rank-size law of human settlement sizes, published in *Das Gesetz der Bevölkerungskonzentration* (The Law of Population Concentration), rediscovered years later by Zipf.

1926 Statistician Alfred Lotka [1880–1949] publishes his discovery of the inverse-square law in the "The Frequency Distribution of Scientific Productivity," *Journal of the Washington Academy of Sciences*.

1935 Linguist George Kingsley Zipf [1902–1950] publishes his first papers on the rank-size distribution of settlements.

1941 Meteorologist Lewis Fry Richardson [1881–1953] discovers the scaling power-law of conflicts, inaugurating the modern scientific study of war through a series of papers in 1941, 1945, and 1948. His first monograph dates to 1919, on "The Mathematical Psychology of War."

1955 Herbert A. Simon publishes his classic paper "On a Class of Skew Distributions" in the journal *Biometrika*, followed in 1958 by his first paper on the power-law distribution of business firms in the *American Economic Review*.

1958 Gutenburg-Richter Law for earthquakes is discovered, arguably the first true power law in the physical sciences.

1959 Albert Wohlstetter publishes his classic paper on Deterrence Theory, "The Delicate Balance of Terror," based on the Conjunctive Principle examined in this chapter and the next, in the influential policy journal *Foreign Affairs*.

1960 Richardson's *Statistics of Deadly Quarrels* is published posthumously.

1962 William H. Riker formalizes the Theory of Political Coalitions and demonstrates the Conjunctive Law for minimal-winning coalitions.

1969 Martin Landau explicitly identifies conjunctive redundancy in his seminal paper published in the *Public Administrative Review*, followed in 1972 by his classic *Political Theory and Political Science: Studies in the Methodology of Political Inquiry*.

1973 Jeffrey L. Pressman and Aaron Wildavsky publish the classic *Implementation: How Great Expectations in Washington Are Dashed in Oakland*, based on the Conjunctive Law.

1978 Gabriel Almond and Bingham Powell publish their influential input-output model of a complex polity, where policies in the outcome space follow a sequential conjunctive law.

1984 John W. Kingdon publishes his classic *Agendas*, *Alternatives*, *and Public Policies*, demonstrating the sequential conjunctive law for policy-making processes in complex polities.

1985 Elinor Ostrom [1933–2012] and colleagues from Indiana University (Vincent Ostrom, Roger Parks, Harvey Starr), the University of Illinois (Claudio Cioffi-Revilla, Richard L. Merritt, Robert Muncaster, and Dina A. Zinnes), and the University of Iowa (Robert Boynton) establish the Triple-I Seminar on Complex Systems.

Since 1990 Power laws are replicated in numerous domains of social science research, such as elections, budgetary processes, finance, terrorism, and the Internet.

1999 Cioffi-Revilla discovers that civil wars scale across the global system, demonstrating long-range spatio-temporal correlations.

2003 Economist Christian Kleiber and statistician Samuel Kotz [1930–2010] publish *Statistical Size Distributions in Economics and Actuarial Sciences*, the first comprehensive treatise on the Pareto Law and related distributions of social complexity.

2003 The same year Cioffi-Revilla and Midlarsky demonstrate that a uniform distribution can be critically misjudged as a power law (Type II error) when diagnostic bending in the lower and upper tails is ignored. In the same paper they demonstrate power law scaling for the deadliest wars.

6.3 Laws of Social Complexity: Descriptions

In this section we examine descriptive laws of social complexity. These are grouped into two main categories, structural and distributional, each of which consists of a variety of models. The *comparative statics* of these laws are interesting, because most equations are nonlinear in nature. This often results in non-intuitive or counterintuitive consequences on the emergent behavior of social complexity. Both share two additional, scientifically deep properties: they are *related* to one another, as well as being *universal* across domains of social complexity.

6.3.1 Structural Laws: Serial, Parallel, and Hybrid Complexity

The **structure of social complexity** refers to the way systems and processes are organized across social domains, including coupled socio-techno-natural systems and components within them, as we have already seen in the case of near-decomposability. Figures 6.1 and 6.2 illustrate isomorphic examples of structural configurations found in social systems and processes, which can often (not always!) be expressed in terms of networks or trees, respectively. A salient feature of structural laws of social complexity is that they have dual isomorphic representation as logic and probabilistic formalism, which facilitates computational modeling. Here we examine more closely the character of causal structures and how they generate emergent social complexity.

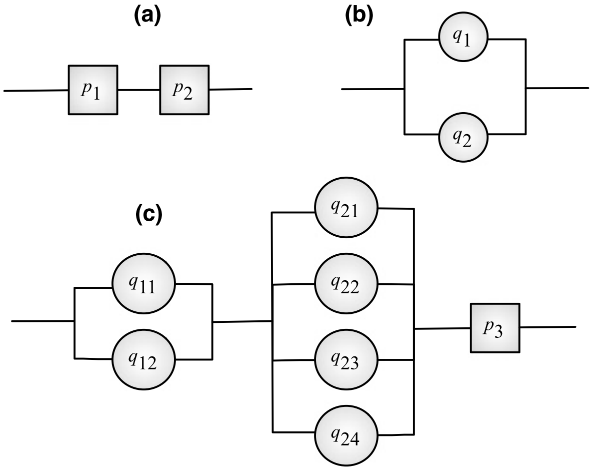

Fig. 6.1 *Structural patterns of social complexity by causal necessity and sufficiency.* (**a**) Serial complexity by causal conjunction; (**b**) parallel complexity by causal disjunction; and (**c**) a case of hybrid serial-parallel complexity with some parallelized disjunctive components within an overall serialized 3-conjunctive structure

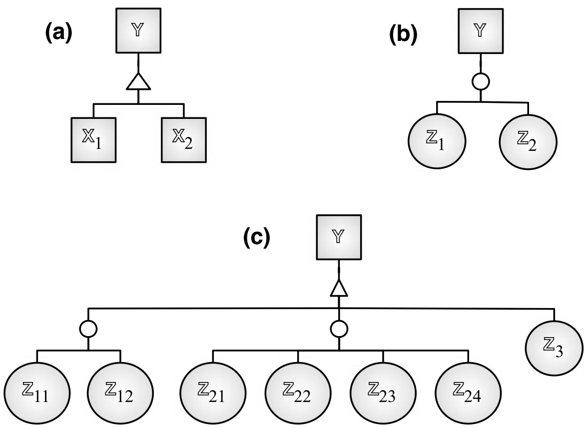

Fig. 6.2 *Structural patterns of social complexity by logic conjunction and disjunction.* (**a**) Serial complexity by causal conjunction; (**b**) parallel complexity by causal disjunction; and (**c**) a case of hybrid serial-parallel complexity with some parallelized disjunctive components within an overall serialized 3-conjunctive structure

6.3.1.1 Serial Complexity by Conjunction

The fundamental structure of complexity in social systems and processes is generated by **compound events**, which emerge from the conjunction of causal events. For example, in the standard model of a polity, the occurrence of successful governance is an emergent compound event generated by a sequential process that begins with (1) an issue collectively affecting a significant sector of society; followed by (2) pressure groups placing demands on government to act; followed by (3) decision-makers doing something to relieve societal stress by enacting policies; and, finally, (4) the public issue being mitigated.

The example just seen is that of a **serial** system (Figs. 6.1(a) and 6.2(a) with 4 components rather than just 2), which is based on *necessary* causal events occurring as a *conjunction* (by Boolean logic AND operator) and emergent overall probability Y_s given by its associated indicator structure function $\Psi_\cap$ according to the following set of related equations:

$$\begin{aligned}\mathbb{Y}_s &= \Psi_\cap(\mathbb{X}_1, \mathbb{X}_2, \mathbb{X}_3, \ldots, \mathbb{X}_n) &(6.1)\\ &\Leftarrow \mathbb{X}_1 \wedge \mathbb{X}_2 \wedge \mathbb{X}_3 \wedge \cdots \wedge \mathbb{X}_n &(6.2)\\ Y_s &= p_1 \cdot p_2 \cdot p_3 \cdots p_n = \prod_{i=1}^{n} p_i &(6.3)\\ &= P^{\Theta}, &(6.4)\end{aligned}$$

where $\mathbb{Y}_s$ denotes the compound event for overall conjunction with necessary causal conditions, $\mathbb{X}_i$ are the n causal events, the symbol $\wedge$ denotes conjunction (Boolean AND), p_i are the probabilities of the causal events, P is their probability when they are all the same, and $\Theta = 1, 2, 3, \ldots, n$ denotes the number of causal events.

An important variation of serial conjunction is when necessary conditions occur in sequence, called **sequential conjunction**, equivalent to Boolean logic SEQAND. Note that probabilities are conditional for sequential causal events. In this case Eqs. (6.1)–(6.4) are simply edited to take into account conditional probabilities, which still require multiplication.

Regardless of whether causal probabilities are conditional or unconditional, overall probability P_s is always *decreased* when social complexity is serialized. **Hypoprobability**, defined by the inequality $Y_s < \min p_i$, is a fundamental property of serial social complexity. It means that serially structured social systems have an overall probability of performing that is smaller than that of the most poorly performing component. Accordingly, the popular aphorism of a chain being as strong as its weakest link ($P = \min p_i$) is objectively wrong, because it overestimates overall serial probability.[1]

6.3.1.2 Parallel Complexity by Disjunction

By contrast, at other times a social system or process may operate according to concurrent activities, as when policy is based on a set of multiple public programs. For example, anti-inflationary policies used by governments are often based on a mix of (1) price controls, (2) subsidies of various kinds (for food, housing, medicines), and (3) other programs that are implemented simultaneously. This example is represented in Figs. 6.1(b) and 6.2(b) with three as opposed to just two causal component events.

This is an example of a **parallel** system, which is based on *sufficient* causal events occurring as a *disjunction* (by Boolean logic OR operator) and emergent overall probability Y_p given by its associated indicator structure function $\Psi_\cup$ and the following set of related equations:

$$\begin{aligned}\mathbb{Y}_p &= \Psi_\cup(\mathbb{Z}_1, \mathbb{Z}_2, \mathbb{Z}_3, \ldots, \mathbb{Z}_m) &(6.5)\\ &\Leftarrow \mathbb{Z}_1 \vee \mathbb{Z}_2 \vee \mathbb{Z}_3 \vee \cdots \vee \mathbb{Z}_m &(6.6)\end{aligned}$$

[1] The correct aphorism should be that a chain is *weaker* than its *weakest* link, which is an even worse condition than being *as weak as* the weakest link.

$$Y_p = 1 - (1 - q_1) \cdot (1 - q_2) \cdot (1 - q_3) \cdots (1 - q_m) = 1 - \prod_{j=1}^{m} (1 - q_j) \quad (6.7)$$

$$= 1 - (1 - Q)^{\Gamma}, \quad (6.8)$$

where notation follows the same conventions as for Eqs. (6.1)–(6.4). By De Morgan's Law, it can be easily demonstrated that parallelization equations (6.5)–(6.8) follow from serialization equations (6.1)–(6.4).

An important variation of parallel disjunction occurs when sufficient conditions are mutually exclusive, called exclusive disjunction, equivalent to the Boolean logic XOR operator and the common language phrase "either." In this case the probabilities of causal events must add up to 1, so the parallel complexity equations we just presented now become

$$\mathbb{P}_p = \Psi(\mathbb{Y}_1, \mathbb{Y}_2, \mathbb{Y}_3, \ldots, \mathbb{Y}_m) \quad (6.9)$$

$$\Leftarrow \mathbb{Y}_1 \veebar \mathbb{Y}_2 \veebar \mathbb{Y}_3 \veebar \cdots \veebar \mathbb{Y}_m \quad (6.10)$$

$$P_p = q_1 + q_2 + q_3 + \cdots + q_m = \sum_{j=1}^{m} q_j \quad (6.11)$$

$$= mq. \quad (6.12)$$

There is a symmetrical result for hypoprobability. Regardless of whether causal disjunctive probabilities are inclusive (OR) or exclusive (XOR), overall probability P_p is always *increased* when social complexity is based on a parallel structure—which is also common at the second- and higher-order of causation. **Hyperprobability**, defined by the inequality $Y_p > \max q_j$, is the fundamental property of parallel social complexity. It means that parallel structured social systems have an overall probability of performance that is greater than that of the best performing component.[2]

6.3.1.3 Hybrid Structural Complexity

Most social systems and processes in the real world operate through some combination of serial and parallel structure, especially those that are complex artifacts or complex policies. Examples of this kind of structural complexity are shown in Figs. 6.1(c) and 6.2(c), which show first-order 3-conjunction that embeds 2- and 3-disjunctions of the second-order.

The following two kinds of symmetrical patterns (serial-parallel and parallel-serial) serve as building blocks for modeling far more complex social forms, to *any* desirable degree of structural complexity.

[2]Popular culture is silent about an analog of the serial chain metaphor for the case of a parallel structure. If it existed, it should say: a parallelized system is stronger than its strongest component.

A **serial-parallel system** has first-order Θ-degree serialization, second-order Γ-degree parallelization, and overall probability equation given by

$$Y_{sp} = \left[1 - (1 - Q)^{\Gamma}\right]^{\Theta}. \tag{6.13}$$

This is the kind of structural complexity shown earlier in Figs. 6.1(c) and 6.2(c). In this instance, we may have a 3-stage social *process* where the first and second stages are carried out by two and four parallel activities, respectively. Alternatively, the same structure may represent a social *system* that requires three operating components to undertake action (e.g., legislative, executive, judicial branches of government), the first of which relies on two parallel components (say, a senate and an assembly), and the second relies on four agencies (e.g., such as for policies on security, economics, health, and infrastructure).

The symmetrical opposite is a **parallel-serial system**, which has first-order parallelization, second-order serialization, and overall probability equation

$$Y_{ps} = 1 - \left(1 - P^{\Theta}\right)^{\Gamma}. \tag{6.14}$$

The origin of chiefdoms (sociogenesis) provides an excellent example of hybrid social complexity. Within the overall formative process, a first-order structure of the compound event $\mathbb{P}$ ("the potential for sociogenesis occurs") is given by the following conjunction of necessary causal events:

$$\mathbb{P} = \Psi(\mathbb{X}_{kin}, \mathbb{X}_{com}, \mathbb{X}_{norm}, \ldots, \mathbb{X}_{ca}), \tag{6.15}$$
$$\Leftarrow \langle \mathbb{X}_{kin} \wedge \mathbb{X}_{com} \wedge \mathbb{X}_{norm} \wedge \cdots \wedge \mathbb{X}_{ca} \rangle, \tag{6.16}$$

where $\mathbb{X}_i$ denote various necessary conditions for chiefdom formation, such as existence of kinship knowledge $\mathbb{X}_{kin}$, communicative ability $\mathbb{X}_{com}$, normative knowledge $\mathbb{X}_{norm}$, and collective action ability $\mathbb{X}_{ca}$, among others as examined in the next chapter. Thus, the first-order probability equation is simply

$$P = X_{kin} \cdot X_{com} \cdot X_{norm} \cdots X_{ca} = \prod_{i=kin}^{ca} X_i \tag{6.17}$$
$$= X^{\Theta}, \tag{6.18}$$

consistent with earlier notation. In turn, collective action ability is satisfied through a variety of Γ strategies (e.g., providing incentives, exercising authority, among others), not in just one unique way.[3] Accordingly, the second-order probability equation in terms of Γ strategies is:

$$P = X^{\Theta-1} X_{ca} \tag{6.19}$$
$$= X^{\Theta-1} \cdot \left[1 - (1 - Q)^{\Gamma}\right], \tag{6.20}$$

[3] We will examine collective action theory more closely in the next chapter.

where Q now represents the probability of individual collective action strategies being known.

A more contemporary example consists of modeling the probability of crisis management policies in issue domains such as humanitarian disasters, financial crises, or cybersecurity. First-order complexity is typically serial,

$$P = X_1 \cdot X_2 \cdot X_3 \cdots X_n \tag{6.21}$$

$$= \prod_{i=1}^{n} X_i \tag{6.22}$$

$$= \prod_{i=1}^{n} \left[1 - \prod_{j=1}^{m} (1 - Z_j) \right]_i, \tag{6.23}$$

because n requirements (e.g., accurate intelligence, available capacity, implementation plans, among others) must occur in conjunction. In the case of humanitarian disaster response, supply chain management is also a prominent serialized structure, as are lines of communication. In the case of financial crisis management, passage of legislation and other regulatory procedures have similar serialized structures. However, second-order complexity is often parallelized, as each requirement is ensured through m different approaches or strategies. Alternative locations are often used for dropping humanitarian relief in affected zones, whereas financial crisis policies employ multiple interventions, rather than a single act of government.

From a computational perspective, hybrid social complexity is modeled with code that makes extensive use of functions as subprograms. For example, separate functions can be defined for computing each structural component. This also results in a program being more modular, which is almost always a desirable feature and a real necessity when dealing with algorithmic complex.

6.3.2 Distributional Laws: Scaling and Non-equilibrium Complexity

Social complexity is also characterized by statistical and probability distributions, specifically by **non-equilibrium distributions** and power laws. As suggested earlier in this chapter by the historical overview of milestones and pioneers, over the past century power laws have been shown to exist across multiple domains of social complexity. In almost all cases these distributions are about *size* variables, not durations, which is a intriguing feature that remains somewhat of a scientific mystery. To better appreciate and understand this area of CSS it is best to begin by defining a power law.

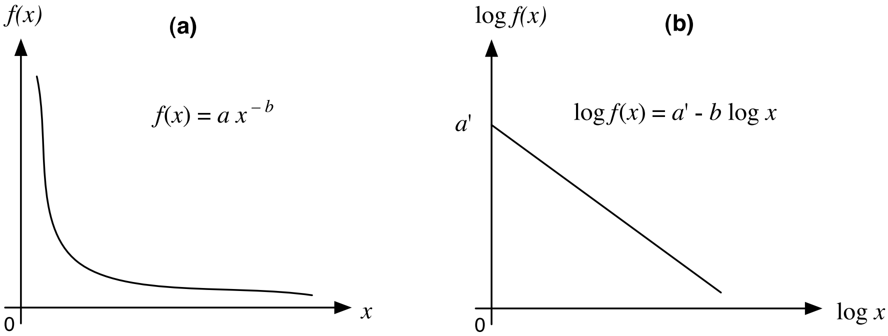

Fig. 6.3 The power law in (**a**) untransformed hyperbolic form and (**b**) linearized or log-linear form in log-log space

Definition 6.1 (Power Law) Let X be a real variable with a set of values $x \in \Re$. A power law is a function of x that is inversely proportional to x itself. Formally,

$$\begin{aligned} f(x) &\propto x^b \\ &= ax^b, \end{aligned} \tag{6.24}$$

where $a > 0$ and $b > 0$.

In purely mathematical terms, a power law refers to any equation of the form

$$y = ax^b, \tag{6.25}$$

where constants a and b can assume any value, such that $f(x)$ in Eq. (6.24) can be either increasing ($b > 0$), decreasing ($b < 0$), or constant ($b = 0$) in x, as well as positive ($a > 0$) or negative ($a < 0$). However, within the context of social complexity theory the term "power law" always implies a negative exponent ($b < 0$) and a positive function ($a > 0$), which in algebraic terms makes Eq. (6.25) the same as a hyperbolic function that is asymptotic in both Cartesian axes, as in Fig. 6.3(a).

For reasons that will become apparent in Sect. 6.3.2.1, the general functional equation (6.25) can be and often is *linearized* by applying a base-10 logarithmic transformation to both sides of the equation, which yields

$$\log f(x) = a' + b \log x, \tag{6.26}$$

where $a' = \log a$ and b now represent an *intercept* and a *slope*, respectively (Fig. 6.3(b)), in log-log space. Note that the slope b is an *elasticity* in log-log space, since

$$\eta_{y,x} = \frac{\partial \log y}{\partial \log x} = \frac{\partial y}{\partial x}\frac{x}{y}.$$

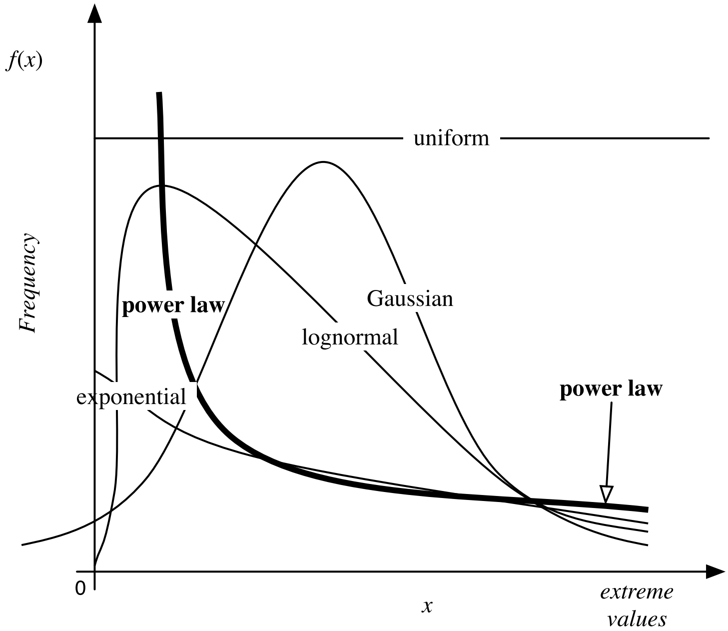

Fig. 6.4 The power law and other distribution models

The log-linear form of Eq. (6.26) is useful from an empirical perspective, because values of x can be plotted on log-log space to examine the form of the distribution, although strictly speaking the term "power law" refers to Eq. (6.25) (with $a > 0$ and $b < 0$), not Eq. (6.26) in log-linear form. For reasons shown below, Eq. (6.25) is the more theoretically relevant equation.

Social scientists familiar with regression analysis will readily recognize Eq. (6.26) as a log-linear regression equation, where both dependent (y) and independent (x) variables have been log-transformed using base 10. In power law analysis the main purpose of log-linearization is *not* to be able to apply ordinary least square (OLS) methods, but to observe how linear the resulting empirical x-y scattergram is and how constant the value of an observed slope $\hat{b}$ is.

Each form of a power law—linear or non-linear, in log-log or linear Cartesian space, respectively—highlights different properties of social complexity, similar to the way in which different forms of the same game in a game-theoretic model (i.e., normal or extensive forms) highlight different features of strategic interaction, or different probability functions (density, cumulative, intensity) provide different views on the uncertainty properties of the same random variable. In addition, each power-law function can also be related to other probability functions, as we shall examine.

Figure 6.4 shows a power law in the context of other distributions. Compared to the so-called normal, Gaussian, or bell-shaped distribution, a power law distribution has *many* small values, *some* (fewer) medium-range values, and a few *rare* extreme

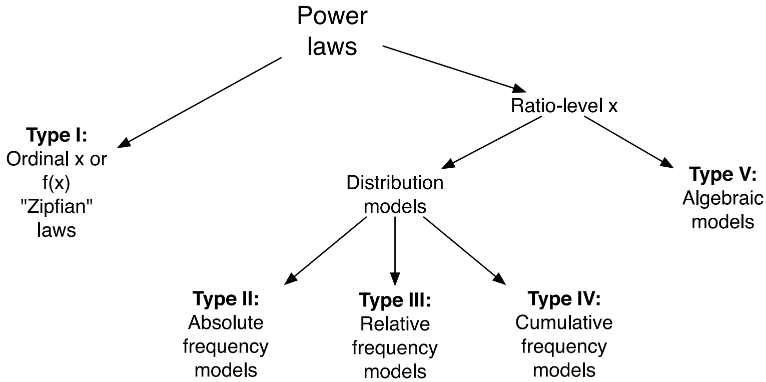

Fig. 6.5 Taxonomy of power law models according to types of dependent variables

values. By contrast, in a Gaussian distribution both smallest and largest values are extremely rare (with vanishingly small probability) and mid-range values are *the norm*.

Crucially, in terms of understanding complexity, extreme events are many times more "normal" in a power law distribution than in a Gaussian distribution. There are also other significant differences with respect to other major types of distributions, such as exponential, uniform, and lognormal, as examined in the next sections.

6.3.2.1 Systematics of Social Power Laws

It would appear from the preceding formalization that power law models are all analytically or formally similar (Eq. (6.25)), in the same sense that all hyperbolas are similar, in that they would differ only by the numerical value of the coefficients a and b. However, that is not the case, because the term on the *left* side of a power law—the function $f(x)$ that is inversely proportional to a given variable x—often denotes widely different quantities when examined in different disciplines and different empirical domains. In addition, as in the case of Zipf's Law, the independent variable can sometimes assume rank-ordinal values, such that the independent variable is not ratio-level.

Given such confusing practices in the literature, it is useful to identify and systematize the most common types of power laws, because the (seemingly) simple form of the linear log-log plots that are commonly reported in publications often conceal interesting subtle differences that stem from quite different quantities being plotted in vertical and horizontal axes, i.e., dependent and independent variables. Similarities and differences among various types of power laws of social complexity are meaningful and should be understood. The taxonomy shown in Fig. 6.5 spans five types of power laws across various social and natural phenomena.

As illustrated in Fig. 6.5, power law models are a class composed of two distinct—albeit related—subclasses or sets of models according to the level of measurement of the independent variable x (ordinal or ratio).[4] In turn, ratio-level power laws comprise several subtypes, as explained in the next sections. In spite of these differences, it must be stressed that all power law models are mathematical representations of extreme skewed variability that are **scale-free**, in the sense discussed below.

6.3.2.2 Type I: Rank-Size or Zipfian Models

The first (and oldest) type of power law model is **Zipf's Law** of harmonic sizes, also known as a **Rank-Size Law** (geography, linguistics) or **rank-size rule** (anthropological archaeology). Given an ordered set of values $\langle x_1, x_2, x_3, \ldots, x_n \rangle$ of a variable X, where the subscript i denotes rank from highest ($i = 1$ or first) to lowest ($i = n$ or last), the power law for values of X with respect to rank i of each value $x_i \in X$ is given by the equation

$$x_i = \frac{a}{i^b} \quad \text{(Type I power law)}, \tag{6.27}$$

where $a = x_1$ (the largest value) and $b \approx 1$. Note that from Eq. (6.27) it also follows that for this type of distribution the product of any value $x_i \in X$ times its rank i always equals (or approximates) the constant a (the largest value x_1). Therefore, the largest value determines all other values of the distribution. Such a decreasing series of values is also known as a **harmonic series**, wherein the second largest value is $1/2$ the size of the largest, the third largest value is $1/3$ the size of the largest, ..., and the last (the nth value) is $1/n$ the size of the largest. From Eq. (6.27) it also follows that

$$\log x_i = a' - \log i, \tag{6.28}$$

which is commonly used for analyzing empirical data with log-log plots. By definition, therefore, this type of power law has elasticity equal to 1.

Felix Auerbach was the first to discover this type of power law in the harmonic frequency of population concentrations. Perhaps somewhat unfairly, the model is commonly named after the Harvard linguist George Kingsley Zipf [1902–1950] because it was he who popularized it. This type of power law may be of unique interest in the social sciences and the life sciences (laws of so-called "allometry" or proportion), and perhaps they remain undiscovered in the physical sciences.

As shown in Fig. 6.5, the next three types of power laws consider different distributions of values of X in terms of various frequency measures: absolute frequency (Type II), relative frequency (Type III), and cumulative frequency (Type IV). All three distribution types of power laws—which are canonical variations on the common theme of modeling scale-free inequality—occur in both the social sciences and the natural sciences.

[4] Using the **Stevens level of measurement** as a classification criterion is useful for distinguishing formally different mathematical forms that are analyzed through different statistical and mathematical methods (discrete vs. continuous). The same classification might be less useful in physical power laws, where ranks and ordinal variables are not as common as they are in social science.

6.3.2.3 Type II: Absolute Frequency Models

In the second type of power law the *absolute frequency* ϕ of a given value $x \in X$ is inversely proportional to x. Thus,

$$\phi(x) = \frac{a}{x^b} \quad \text{(Type II power law).} \tag{6.29}$$

From Eq. (6.29) it follows that

$$\log \phi(x) = a' - b \log x, \tag{6.30}$$

where $a' = \log a$ is the intercept and b is the slope (exponent in Eq. (6.29)). Recall that b is also in this case the elasticity η of $\log \phi(x)$ with respect to $\log x$.

In the social sciences this type of power law has been frequently reported for variables as diverse as the size of archaeological sites in a given region, personal income, number of Internet routers, network links, and the number of fatalities that have occurred in warfare on all scales in modern history. Lewis Fry Richardson's Law of War Severity, describing the skewed distribution of fatalities generated by conflicts of all magnitudes, is a power law of this type. In the natural sciences, this type of power law has been reported for the size of species, the lifespan of genera, earthquake energy releases, meteor diameters, and the relative sizes of avalanches in Conway's Game of Life (a cellular automata model examined in Chap. 7).

The next two types of power laws are somewhat similar, since they are both based on probability functions, but different in several interesting, crucial details that are easy to overlook.

6.3.2.4 Type III: PDF Models

The third and closely related type of power law is stated in terms of *relative frequency*, which in the statistical limit approximates a **probability density**. Formally, this is the **hyperbolic probability density function** (p.d.f.)

$$p(x) = \frac{a}{x^b} \quad \text{(Type III power law).} \tag{6.31}$$

(In physics, Eq. (6.31) is often called a "distribution function," which is a mathematical misnomer that can cause confusion. The term "distribution function" refers to the cumulative density function $\Phi(x)$, or "mass function," as in the next section.)[5]

[5]For example, Bak (1996), Jensen (1998), and Barabasi (2002) misname these functions repeatedly—c.d.f., p.d.f., and complementary c.d.f.—as if they were synonymous, whereas each function refers to the probability of a different event: $\mathbf{Pr}(X \le x)$, $\mathbf{Pr}(x < X \le x + dx)$, and $\mathbf{Pr}(X > x)$, respectively. The obvious but important point is simply that probability functions that refer to different events should be named differently and consistently.

The log-linear form for the Type III power law is easily derived, from Eq. (6.31), as

$$\log p(x) = a' - b \log x, \tag{6.32}$$

with $a' = \log a$, and, again, b is the elasticity of $\log \phi(x)$ with respect to $\log x$.[6]

This type of power law also has strong empirical support across social domains. It has been reported for the size of firms in terms of employees (Simon's Law), the number of publications by scholars (Lotka's Law), the number of collaborations by movie actors, the size of commodity price fluctuations (Mandelbrot's Law), and other social variables. In the natural and engineering sciences, this same Type-III power law has been reported for the size of species, the connectivity of the US power grid, the size of forest fires (Turcotte's Law), and the size of sandpile avalanches (Bak's Law).

6.3.2.5 Type IV: Log-Survival or Log-CCDF Models

A fourth type of power law is based on the *complementary cumulative density function*, or $1 - \Phi(x) = \mathbf{Pr}(X > x)$, abbreviated as CCDF. When X denotes time T, the CCDF is called a *survival function*, or $S(t)$.[7] In a log-log linear graph this model has the form

$$\log\bigl[1 - \Phi(x)\bigr] = a' - (b-1)\log x, \tag{6.33}$$

with $a' = \log a$, which yields the c.d.f.

$$\Phi(x) = 1 - \frac{a}{x^{(b-1)}} = 1 - ax^{1-b} \tag{6.34}$$

and corresponding p.d.f. given by

$$p(x) = \frac{a(b-1)}{x^b} \quad \text{(Type IV power law)}. \tag{6.35}$$

Note that in this type of power law the elasticity in Eq. (6.33) is $\eta = (b-1)$, *not* just b as in previous models—a critical difference to remember! Table 6.1 provides a comparison of the defining probability functions of a Type IV power law model (top row) with respect to other distribution models of social phenomena. Note that the negative exponential p.d.f. also corresponds to a Poisson process, which is common in many social phenomena such as riots, onsets of warfare, and organizational

[6] Note that Type II (absolute frequency) and Type III (relative frequency) yield the same slope b, although the functions on the left side are not mathematically identical.

[7] Also, strictly speaking, the event "$X \geq x$" makes more sense than "$X > x$" when X is a discrete (count) variable. This is because $0.99999\ldots$ is not computable and 0 is mathematically impossible, so 1 is the base count for social processes such as events, riots, wars, and other social count processes.

Table 6.1 The Type IV power law model of social complexity compared to other common social processes and distributions

Model	p.d.f. $p(x)$	c.d.f. $\Phi(x)$	h.f.f. $H(x)$	Mean $E(x)$
Power law	$\frac{a(b-1)}{x^b}$	$1 - ax^{b-1}$	$\frac{b-1}{x}$	$\frac{a(b-1)}{2-b} x^{2-b} \vert_{x_{\min}}^{\infty}$
Exponential	$\lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	λ	$\frac{1}{\lambda}$
Weibull	$\lambda \gamma x^{\gamma-1} \exp(-\lambda x^\gamma)$	$1 - \exp(-\lambda x^\gamma)$	$\lambda \gamma x^{\gamma-1}$	$\lambda^{-1/\gamma} \Gamma(\frac{1}{\gamma} + 1)$
Lognormal	$\frac{1}{\sigma x \sqrt{2\pi}} \times \exp[-(\ln(x/m))^2/(2\sigma^2)]$	$1 - \frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty \frac{p(u)}{u} du$	$\frac{p(x)}{1-\Phi(x)}$	$\exp(0.5\sigma)$
Gaussian	$\frac{1}{\sigma x \sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$	$1 - \frac{1}{\sqrt{2\pi}} \times \int_x^\infty \exp[-\frac{1}{2}(\frac{u-\mu}{\sigma})^2] du$	$\frac{p(x)}{1-\Phi(x)}$	μ

turnover. The intensity or hazard force functions (h.f.f.) corresponding to power law, exponential, and Weibull models are of major interest in practical applications. The lognormal and Gaussian cases are also computed as $p(x)/[1 - \Phi(x)]$ but are omitted from the table due to space constraints and infrequent use. The graphs of probability density functions in Table 6.1 were shown earlier in Fig. 6.4.

Equation (6.35) looks deceptively similar to a Type III power law (compare with Eq. (6.31)), with the crucial difference that the proportionality constant is partially dependent on the exponent (b) or slope ($b - 1$). This fourth type of power law, based on the complementary c.d.f., has been reported for the size of firms in terms of revenue, for fatalities that occur in warfare (both civil wars and international wars), as well as for a variety of natural phenomena including the magnitude of earthquakes (Gutenberg-Richter Law).

An important result that links this type of power law model to other classical distributions models (e.g., Weibull) is given by the following theorem:

Theorem 6.1 (Intensity Function of a Power Law) *Given a Type IV power law with p.d.f. as in Eq.* (6.35) *and c.d.f. as in Eq.* (6.34), *then the associated intensity function or hazard force function* $H(x)$ *is given by*

$$H(x) = \frac{b-1}{x}, \tag{6.36}$$

where $H(x)$ *is defined as* $p(x)/[1 - \Phi(x)]$, *which is*:

1. *linear in* b
2. *hyperbolically decreasing in* x *with power law exponent* 1 (*scale-free*),
3. *independent of* a
4. *a special case of the Weibull distribution for* $\gamma(\text{shape}) = -1$ *and* $\lambda(\text{scale}) = b - 1$, *or slope of the CCDF in log-log space*
5. *has an associated stress or load function* $\Lambda(x)$ *given by*

$$\Lambda(x) = \int_0^x H(u) du = (b - 1) \ln x. \tag{6.37}$$

Proof By substituting Eqs. (6.34) and (6.35) into the definition of $H(x)$ and simplifying the resulting expression to obtain Eq. (6.36). □

Theorem 6.1 is interesting because it provides a simple and direct link between social complexity theory on the one hand, and risk analysis and uncertainty on the other. The principle says that *all complex social phenomena are generated by inverse intensity*. The Weibull model includes one such instance of an inverse function, as do other stochastic processes with hyperbolically decreasing intensity or hazard rate. Conversely, using Eq. (6.36), the intensity function theorem allows us to express a power law as a function of the many features associated with $H(x)$, such as moments and other characteristics.

Types III and IV power laws should never be referred to as "Zipf's Law for $b = 1$," because such terminology implies that these models contain ranked variables; they do not.

6.3.2.6 Type V: Algebraic Models

Finally, a fifth type of power law model found in the literature is based on the linear plot of two ordinary ratio-level variables, so

$$\log y(x) = a' - b \log x \tag{6.38}$$

and

$$y(x) = \frac{a}{x^b}. \tag{6.39}$$

Note that in this case there is no difference between the log-linear slope and the hyperbolic exponent—a property that differs from the previous cases. Although most social scientists do not think of ordinary algebraic expressions such as Eq. (6.39) as a power law, in the natural sciences (and in elementary mathematics) the study of power laws includes these models as well. For example, the relation between the number of routers y and the number of nodes x in the Internet is governed by Eq. (6.38) with $b \approx 1.9$ (Faloutsos's Law). If the class of power laws includes these algebraic relationships or hyperbolic models (type V), then all inverse empirical relationships that are linear in log-log space also qualify as power laws (e.g., Polachek's Law of international conflict and trade, and social gravity models in human geography and regional economics).

It should be reiterated that the preceding five types of power laws share a great deal in common—the right side of the equation is always a term inversely proportional to a given variable x—but the mappings are different because what is modeled on the left side of each equation varies across types. Such variations are sometimes relatively minor, as between Type II (absolute frequencies) and Type III (relative frequencies). Other times they are more significant, as between Type III (p.d.f.-based) and Type IV (c.d.f.-based), or between ratio variables, frequency-based, and variable-based models.[8] Beyond the formal differences highlighted by the preceding taxonomy, all power laws are susceptible to empirical analysis, as discussed in the next section.

6.4 Power Law Analysis

Power laws of social complexity are susceptible to various forms of empirical, data-oriented analysis, as well as theoretical, mathematically-oriented analysis. Both approaches are necessary and synergistic for understanding complexity in social phenomena.

6.4.1 Empirical Analysis: Estimation and Assessing Goodness of Fit

Suppose a given data sample or set of observations $\{x\}$ of a variable X yields a power law of some type (I–IV). From an empirical perspective a review of current practices in the extant literature shows that there are two common procedures for assessing the goodness of fit of a power law model in relation to empirical data: (1) visual inspection of the log-log plot to see if it approximates a straight line, and (2) judging goodness of fit on the basis of a high value for the R^2 statistic. These procedures deserve close scrutiny, because they can be misused, resulting in false inferences.

6.4.1.1 Visual Assessments

Visual assessments are useful, informal, and always subjective. A common problem that is often highlighted by data plotted on log-log scales is "bending" away from the log-linear model at lower and upper ranges of the distribution (see Fig. 6.6).

Bending of an empirical distribution at *lower* quantiles can occur because there might be missing observations for small values that are lost or hard to measure. For example, in a dataset of war magnitudes the smallest wars may not be recorded. This is a form of measurement error that can arise for many reasons. Bending at the lower quantiles can be acceptable if the claim that the smallest observations are incomplete can be supported; otherwise, lower quantile bending presents a serious problem with accepting the research hypothesis that the observed data conforms to a power law.

Bending can be found in empirical data that approximate a power law, but can also be diagnostic of an exponential or lognormal tail. Also, a *uniform distribution* (which is far from being a power law!) plotted on log-log space yields a curved pattern with both lower and upper quantile bending, so the problem in such cases may not be due to missing observations or finite size—it may be because the distribution is close to uniform, not at all a power law or even exponential.

[8]The basic point is that care must be taken to specify which type of power law model is being discussed or presented; this should not have to be deciphered from poorly labeled plots or misnamed equations.

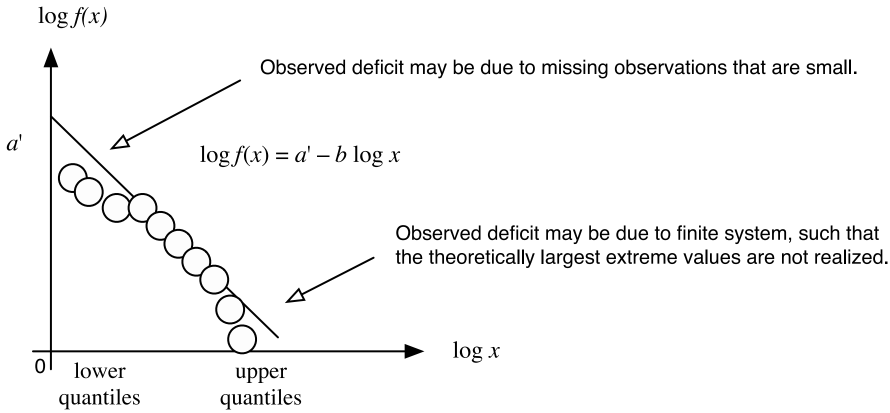

Fig. 6.6 "Bending" is frequently observed in visual assessment of empirical power law distributions

6.4.1.2 R-Squared

In much of the extant literature, goodness of fit is often assessed using the coefficient of determination, R^2. However, R^2 is best avoided as a measure of goodness of fit and the most recent specialized statistical works on size distributions do not discuss it. Other statistics and methods, such as the standard error of the coefficients or the Anderson-Darling test, are preferable when necessary. Still, a good use of the R^2 statistic is for comparing different empirical models that have the same functional form but are estimated using different data samples.

6.4.1.3 Good Practices: Multiple Lines of Evidence

As is normally the case for various estimators, goodness of fit also should be assessed on the basis of multiple methods that provide diverse lines of evidence: small standard errors, large t-ratios, the Kolmogorov-Smirnov test, the Anderson-Darling test, among other methods. The estimation of power law models using *maximum likelihood* methods is recommended, such as based on the **Hill estimator**. Table 6.2 compares various statistical assessments for power laws.

By way of summary, some good practices in the empirical analysis of power laws with statistical data include the following:

1. Use disaggregated data values $\{x\}$ of the observed variable X to construct the relevant frequency distribution plots ensuring that all axes and units of measurement are properly labeled. Report the standard errors of all coefficients when conducting an estimation. Specifically:
   (a) For the Type I power law (Eq. (6.27)), data values are ordered from largest to smallest and the resulting plot should resemble a simple harmonic function with a long upper tail. In log-log space the same data should approximate a straight line with slope value of 1.
   (b) For Type II (Eq. (6.29)), the data values should be used directly to construct a histogram of value frequencies and the results plotted in log-log space. The plot should approximate a straight line, as in Eq. (6.30). Note that in this case

Table 6.2 Goodness of fit statistics used for assessment of an empirical power law

Statistic	Pros	Cons	References
Hill estimator	MLE	Can be unstable for small sample size	Alfarano et al. (2008); Hill (1975)
Anderson-Darling	Sensitive to upper tail values	Rarely used; not well known; Type I error risk	Anderson and Darling (1954)
Kolmogorov-Smirnov	Widely known	Insensitive to upper tail values; Type II error risk	Chakravarti et al. (1967, pp. 392–394)
R^2	Commonly used; good for comparing samples	Not a proper goodness of fit statistic	King (1986)

the estimated slope $\hat{b}$ in Eq. (6.30) is exactly the value of the exponent b in Eq. (6.29)—i.e., without the $(+1)$ transformation that is necessary with the Type IV law.

(c) For Type III (Eq. (6.31)), the procedure is the same as for the Type II power law, except that it is necessary to compute relative as opposed to absolute frequencies.

(d) For Type IV (Eq. (6.35)), which is arguably the most important case, the data values are again used directly, this time to construct the normalized complementary cumulative frequencies—i.e., the values of the function $[1 - \Phi(x)]$, without binning.[9] The log-log plot should then approximate a straight line with slope $(b + 1)$. Accordingly, a *slope* of $(b + 1)$ for the distribution of the complementary c.d.f $[1 - \Phi(x)]$ in log-log space yields an *exponent* of b in the Type IV power law (Eq. (6.35)).[10] That is: *slope* $(b + 1) \rightleftarrows$ *exponent* b.

2. Inspect the upper and lower quantiles for excessive bending. Significant bending should be accounted for (e.g., are there missing observations? is finite size somehow involved?). Otherwise, the power law model simply may not fit the data and other models should therefore be considered (e.g., lognormal?).
3. Inspect the number of orders of magnitude (sometimes called "decades") covered by the domain of values. In general, the larger the number of orders of magnitude the more interesting the model because the scale-free property (discussed in the next section) will extend over several orders. Ensure that the range of orders of magnitude is not an artifact of the units of measurement.

[9]"Binning" refers to the procedure of classifying values into equal and finite intervals, which creates problems when the distribution of the underlying population is unknown. It is unnecessary in power law analysis that uses raw data. The direct construction of the histogram of normalized cumulative frequencies is often feasible and always preferable because no binning is necessary. However, sometimes binning is unavoidable when using official statistics such as provided by government agencies.

[10]The exact value of the exponent b is of great theoretical relevance, as explained below in Sect. 6.4.2.1, so reporting the standard error of b is another good practice.

4. Rely on the most valid and reliable data available, especially when N is not very large, because other issues such as bending and goodness of fit can be greatly affected by data quality.
5. Use the standard errors to assess the coefficient estimates, as well as other methods for assessing goodness of fit, such as the Hill estimator. (Ignore significance tests for the slope estimates of Type IV models, since, by definition, cumulative data will always yield slopes greater than zero.)
6. Avoid the R^2 for purposes of assessing goodness of fit, but use it to compare models that have the same functional form—as a comparative measure.[11]
7. Develop familiarization with standards and methods in various fields where power laws are used to gain a better perspective and improve the quality of empirical analysis in social power law modeling.

These good practices—based on multiple lines of evidence and complementary approaches demonstrated over the past century—are susceptible to improvement as social scientists and other modelers gain experience with empirical applications of power law models. Important scientific goals will be achieved as good practices emerge.

6.4.2 Theoretical Analysis: Deriving Implications

A power law is important, *inter alia*, because of the set of intriguing theoretical implications it can generate, not just because it establishes an empirical regularity based on empirical evidence. This is increasingly relevant as social scientists gain experience in the exploitation of synergies between formal models and empirical data. Among the theoretical implications that can be drawn from finding a power law in a given set of data, the following are especially significant in terms of understanding social complexity.

6.4.2.1 Average Size

The first moment (average or mean value) of a power law distribution exhibits some unusually interesting behavior. This is given by

$$E(x) = \int_{\min\{x\}}^{\infty} x p(x) dx = a(b-1) \int_{\min\{x\}}^{\infty} x^{1-b} dx \tag{6.40}$$

$$= \frac{a(b-1)}{2-b} x^{2-b} \bigg|_{\min\{x\}}^{\infty} = \frac{x_{\min}(b-1)}{b-2}, \tag{6.41}$$

which goes to infinity when $b \leq 2$. In other words, there is no mean size (no expected value $E(x)$ exists) for social phenomena that are governed by a power law with exponent in the range $0 < b < 2$, or $(b-1) < 1$ (below unit elasticity). This is an insightful theoretical result for social patterns such as organizational sizes,

[11] However, recall that the standard error of estimates contains essentially the same information.

fatalities in warfare, and terrorist attacks. The threshold $b = 2$ is therefore theoretically critical, as it marks the boundary between social phenomena that have a finite average and computable size ($b > 2$) and those phenomena that lack an expected value or mean size ($b \leq 2$). This is a theoretical insight derived directly from the empirically estimated value of the power law exponent b.

6.4.2.2 Inequality

By definition, a power law is a model of *inequality* (the "many-some-rare" pattern discussed earlier in this chapter), so every power law model has an associated **Lorenz curve** given by:

$$L(\Phi) = 1 - \left[1 - \Phi(x)\right]^{1-1/(b-1)} \tag{6.42}$$

and a corresponding **Gini index** given by

$$G(b) = 1 - 2\int_0^1 L(\Phi)d\Phi = \frac{1}{2b-3}, \tag{6.43}$$

which can be estimated by the empirical equation (Kleiber and Kotz 2003: 35):

$$\hat{G} = \frac{1}{n^2 E(x)} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|. \tag{6.44}$$

These interesting and insightful theoretical links between the exponent b of a power law and its corresponding Gini index G of inequality can be summarized by the following two relations in reference to the tail of a distribution:

$$\begin{array}{c} \text{heavy tail} \\ (b \to 0) \end{array} \Longleftrightarrow \left\{ \begin{array}{l} \text{more inequality} \\ \text{less equality} \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} \text{smaller } b \\ \text{larger } G \end{array} \right.$$

$$\begin{array}{c} \text{thin tail} \\ (b \to \infty) \end{array} \Longleftrightarrow \left\{ \begin{array}{l} \text{more equality} \\ \text{less inequality} \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} \text{larger } b \\ \text{smaller } G \end{array} \right.$$

6.4.2.3 Entropy

By extension, the greater inequality of a heavy tail also implies greater **Shannon entropy** in the distribution of values, or

$$U(b) = \ln\left(\frac{b-1}{\min\{x\}}\right) - \frac{1}{b-1} - 1, \tag{6.45}$$

where $\min\{x\}$ is the smallest value in the distribution of X. This last expression establishes a direct connection between complexity theory and information theory by linking Shannon's entropy U to the power law exponent b. Equation (6.45) guarantees the existence of as yet unknown information-theoretic properties of social power laws.

6.4.2.4 Self-Similarity

When a given variable X obeys a power law, a recurring pattern of constant proportion occurs across the entire range of values of X, as highlighted earlier by the linear graph in Fig. 6.3(b). The graph of the transformed function $f^*(x) = \log f(x)$ is as linear in the low range of values as it is in the high range and everywhere in between. This type of global symmetry is called **self-similarity** in complexity theory. Self-similarity is also said to be an "emergent" property, because it applies to a whole set of values, not to individual values or elements.

Self-similarity is also a property of structural laws of social complexity. For example, a system of first-order conjunctions (or disjunctions) embedded by higher-order conjunctions (or disjunctions) is self-similar. A policy process is a classical example of self-similar structural social complexity in terms of overall policy response (first-order), programs (second-order), activities (third-order), down to the smallest required events (nth-order) that produce policy results.

6.4.2.5 Scaling

The property of self-similarity is also known as **scaling**, which has prompted the term "scale-free phenomena." Vilfredo Pareto discovered that wealth and income scale. Lewis F. Richardson discovered in the late 1940s (possibly earlier) that warfare ("deadly quarrels") scales with respect to magnitude μ. Since then, it has been shown that not just international wars but civil wars also scale, as do certain features of terrorism. "Artificial" wars generated by agent-based models also scale. Do other dimensions besides war fatalities, such as time of onset and conflict duration, scale? The answer is: generally, no. Time durations are more often exponentially or Weibull-distributed, as we will discuss in Chap. 9.

Scaling is empirically demonstrated for numerous other dimensions of social phenomena, but remains a deep theoretical notion. Scaling means that dichotomies of small versus large wars are false, because of the scale invariance given by the global power law. Scaling also means that it is a misconception to think that small and large wars share little or nothing in common; they are all—small and large—part of the same overall pattern, just different ranges of a power law governed by an identical set of parameter values. Note that scaling occurs if and *only* if a variable obeys a power law. (Most biological organisms do *not* scale.)

6.4.2.6 Fractal Dimension

If the exponent b of a power law equation were allowed to assume only integer values $(1, 2, 3, 4, \dots)$ then the frequencies associated with each value would decrease inversely by the power of such integer proportions. However, when b assumes fractional values (as many exponents reported in the empirical literature) the range of proportions is itself continuous and no longer discrete as in Euclidean space. This is why the b-value in a power law is often called Mandelbrot's **fractal dimension**. Note that scaling vanishes as $b \to 0$, because all values of X assume the same frequency when $b = 0$, so from a scaling perspective a uniform random variable exists in a 0-dimensional space. A Zipfian power law ($b = 1$) yields a 1-dimensional space. A quadratic power law ($b = 2$ or critical value) yields a 2-dimensional space.

In general, a b-power law yields a b-dimensional space and fractional values of b yield fractal dimensions embedded within Euclidean space. Thus, for $0 < b < 1$ the fractal dimensionality is between a point and a line; for $1 < b < 2$ it is between a line and a plane; for $2 < b < 3$ it is between a plane and a solid; and so on. Thus, the fractal dimension also offers another new classification scheme for social phenomena, an idea that physics has begun to exploit with intriguing insights (e.g., Sornette 2003).

6.4.2.7 Criticality and Driven Threshold Systems

Scaling phenomena can be produced by an underlying process that is driven to a phase of **criticality** by slowly evolving input forces that stress the system. Although the input driving the system can behave continuously, the state variables can change abruptly inside a critical region known as a **bifurcation set**, producing scaled phenomena. A precursor to this important insight was contributed over three decades ago by **Catastrophe Theory**, pioneered by mathematician René Thom [1923–2002]. Complexity theory supports and extends Catastrophe Theory by providing a new interpretation of bifurcation dynamics and metastability. For instance, when a power law is reported for a given social phenomenon, such a finding should prompt a set of catastrophe-theoretic questions that would otherwise not arise:

- Is the phenomenon governed by a **driven threshold** system in the sense of Complexity Theory?
- How is the bifurcation set of critical, metastable states to be interpreted?
- What is the form of the associated **potential function** $P(x)$ defined over the state-space?

The demonstration of extensive scaling in warfare, demography, and economics provides significant support for the idea of criticality and related insights on social complexity, such as metastability, long-range interactions, and universality.

6.4.2.8 Metastability

Social events never "come out of the blue"—they must develop potential before they can occur. Another important theoretical inference that can be drawn from the empirical demonstration of a power law in a given social domain is the complexity-theoretic condition known as "metastability." A system (or, more precisely, a given state $x \in X$ of a system) is said to be **Lyapunov-stable** if it is able to maintain its equilibrium under a range of perturbations. For instance, a positive social relation (e.g., a marriage, a friendship, an alliance) is stable in this sense if it is able to endure in spite of stresses that commonly affect social relations. By contrast, a social system is unstable it if falls apart when stressed, such as a polity or an alliance that ends under the pressure of conflict or unresolved issues. A broad range of social system theories—such as in the work of Pareto, Parsons, Samuelson, Deutsch, Easton, Flannery, Dahl and other social systems theorists—employ this Lyapunov concept of stability.

By contrast, a system is said to develop **metastability** when there exist one or more *potential states* $x' \in X$ or potential operating regimes (with $x \neq x'$), *other than the extant state*, to which the system could transition, given the realization of certain conditions. Metastability is common in many social systems, given their capacity for change. For example, a domestic political system or polity becomes metastable during an election or, even more dramatically, during a constitutional convention. *State failure* occurs when a polity that has first become metastable then loses governance capacity relative to accumulated or unresolved stresses. Similarly, an international system becomes metastable—sometimes increasingly so—in a time of crisis, because an alternate state of overt hostility or actual violence grows as the potential for war increases. In economics, financial markets become metastable when they develop a "bubble" capable of bringing about a market crash. Similarly, from a more positive viewpoint, a state of warfare becomes metastable when the potential for a return to peace increases; domestic turmoil and civil unrest also become metastable—as in state-building operations—as the state potential for governance (capacity) increases relative to stresses. Power laws are diagnostic of metastability because they model social situations where a broad range of states—not just the extant equilibrium or observed status quo—has the potential of being realized. Theories of social change should leverage the concept of metastability inherent in power laws.

6.4.2.9 Long-Range Interactions

Scaling phenomena are produced by systems that evolve into a critical phase where **long-range interactions** become possible and sometimes occur. A system governed by only nearest-neighbor interactions will tend to produce mostly normal or Gaussian-distributed phenomena, or other non-power law phenomena with significantly shorter or thinner tails in the upper (and lower) quantiles.

By contrast, a "globalized" system governed by long-range *spatio-temporal* interactions is subject to non-equilibrium dynamics and processes that produce power laws. In such systems the occurrence of extreme events is orders of magnitude higher (not just greater) than in "normal" (Gaussian) equilibrium systems. The spatial dimension of long-range interactions is fairly straightforward in terms of social or physical distance among social actors. Temporal long-range interactions refer to persistent memory of the past as well as future expectations, as already seen for the Hurst parameter in Sect. 5.5.2.2, Fig. 5.2.

The main purpose of these theoretical observations has been to alert readers to several significant potential implications that go beyond the demonstration of an empirical power law. This is not to suggest that each one of these theoretical implications is valid in every instance of an empirical power law, so these potential implications should be seen as a theoretical heuristic for discovering properties of social phenomena, not as proven properties.

6.5 Universality in Laws of Social Complexity

The social sciences have evolved from an initially unified tradition seeking to uncover universal scientific principles of human and social dynamics—which was the original spirit of the Age of Enlightenment and the rise of modern positive science in recent centuries—to today's condition of significant fragmentation along multiple dimensions: differences in empirical domains, disciplinary cultures, methodologies, even epistemologies. For those intrigued or motivated by the prospect of a unified science of the social universe, structural laws and power laws examined in this chapter offer robust and encouraging grounds for uncovering further universal principles to better understand human dynamics and social complexity based on a common set of empirical and theoretical features, such as those discussed in this chapters.

Self-similarity, scaling, fractal dimensionality, self-organized criticality, metastability, long-range interactions, and universality are all new perspectives surrounding power laws of social phenomena, based on Complexity Theory. These properties and insights were unknown at the time when the first power laws were discovered by Pareto, Zipf, Richardson, and other pioneers. Complexity Theory contains other properties of power laws that may prove insightful for the social sciences. In turn, discovery of power laws in the social sciences may contribute new insights for Complexity Theory and non-equilibrium dynamics.[12]

Recommended Readings

S. Alfarano, T. Lux, F. Wagner, Time variation of higher moments in a financial market with heterogeneous agents: an analytical approach. J. Econ. Dyn. Control **32**(1), 101–136 (2008)

G.A. Almond, S.J. Genco, Clouds, clocks, and the study of politics. World Polit. **29**, 489–522 (1977)

O.M. Ashford, H. Charnock, P.G. Drazin, J.C.R. Hunt, P. Smoker, I. Sutherland (eds.), *The Collected Papers of Lewis Fry Richardson. Volume 2: Quantitative Psychology and Studies of Conflict* (Cambridge University Press, Cambridge, 1993)

B.J.L. Berry, *Geography of Market Centers and Retail Distributions* (Prentice-Hall, Englewood Cliffs, 1967)

L.M.A. Bettencourt, The origins of scaling in cities. Science **340**, 1438–1441 (2013)

C. Cioffi-Revilla, *Politics and Uncertainty: Theory, Models, and Applications* (Cambridge University Press, Cambridge, 1998)

C. Cioffi-Revilla, I. Midlarsky Manus, Power laws, scaling and fractals in the most lethal international and civil wars, in *The Scourge of War: New Extensions on an Old Problem*, ed. by P.F. Diehl (University of Michigan Press, Ann Arbor, 2004), pp. 3–27

R.A. Dahl, *Polyarchy: Participation and Opposition* (Yale University Press, New Haven, 1972)

C. Gini, Sulla misura della concentrazione e della variabilità dei caratteri [On the measure of concentration and variability of traits]. Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti [Annals of the Royal Venetian Institute of Sciences, Letters, and Arts] **53**(2) (1913)

J.W. Kingdon, *Agendas, Alternatives and Public Policies*, 2nd edn. (Harper Collins, New York, 1995)

C. Kleiber, S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences* (Wiley, New York, 2003)

[12]Other important and insightful theoretical extensions of the power law functions discussed in this chapter consist of the gradient ∇f associated with several of the functions. For instance, the field $\mathbf{E}_b = -\nabla f(x; a, b)$ associated with the power law exponent has intrinsic interest for its social interpretation and relation to criticality, metastability, and other complexity concepts. Theoretical and empirical implications of these and other advanced extensions lie beyond the scope of this introductory textbook.

M. Landau, Redundancy, rationality, and the problem of duplication and overlap. Public Adm. Rev. **29**(4), 346–358 (1969)

M. Landau, *Political Theory and Political Science: Studies in the Methodology of Political Inquiry* (Humanities Press, New Jersey, 1979)

M.O. Lorenz, Methods of measuring the concentration of wealth. J. Am. Stat. Assoc. **9**, 209–219 (1905)

A.J. Lotka, The frequency distribution of scientific productivity. J. Wash. Acad. Sci. **16**, 317 (1926)

V. Ostrom, C. Tiebout, R. Warner, The organization of government in metropolitan areas. Am. Polit. Sci. Rev. **55**(4), 831–842 (1961)

J.F. Padgett, Bounded rationality in budgetary research. Am. Polit. Sci. Rev. **74**(2), 354–372 (1980)

V. Pareto, *Cours D'économie Politique* (Editions Rouge, Lausanne, 1897)

J.L. Pressman, A. Wildavsky, *Implementation: How Great Expectations in Washington are Dashed in Oakland* (University of California Press, Berkeley, 1973)

L.F. Richardson, Frequency and occurrence of wars and other fatal quarrels. Nature **148**, 598 (1941)

L.F. Richardson, The distribution of wars in time. J. R. Stat. Soc. A **107**(3–4), 242–250 (1945)

L.F. Richardson, Variation of the frequency of fatal quarrels with magnitude. J. Am. Stat. Assoc. **43**(244), 523–546 (1948)

L.F. Richardson, Is it possible to prove any general statements about historical fact? Br. J. Sociol. **3**(1), 77–84 (1952)

L.F. Richardson, *Statistics of Deadly Quarrels* (Boxwood Press, Pacific Grove, 1960)

W.H. Riker, *The Theory of Political Coalitions* (Yale University Press, New Haven, 1962)

H.A. Simon, On a class of skew distribution functions. Biometrika **42**(3/4), 425–440 (1955)

H.A. Simon, C.P. Bonini, The size distribution of business firms. Am. Econ. Rev. **48**(4), 607–617 (1958)

N.D. Singpurwalla, *Reliability and risk: a Bayesian perspective* (Wiley, New York, 2006)

A. Wohlstetter, The delicate balance of terror. Foreign Aff. J. **37**(1), 211–234 (1959)

G.K. Zipf, The unity of nature, least-action, and natural social science. Sociometry **5**(1), 48–62 (1942)

G.K. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, 1949)