PHYSICAL REVIEW E **81**, 011122 (2010)

Diversity of Poissonian populations

Iddo I. Eliazar[1,*] and Igor M. Sokolov[2,†]
[1]*Department of Technology Management, Holon Institute of Technology, P.O. Box 305, Holon 58102, Israel*
[2]*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany*
(Received 17 September 2009; published 15 January 2010)

Populations represented by collections of points scattered randomly on the real line are ubiquitous in science and engineering. The statistical modeling of such populations leads naturally to Poissonian populations—Poisson processes on the real line with a distinguished maximal point. Poissonian populations are infinite objects underlying key issues in statistical physics, probability theory, and random fractals. Due to their infiniteness, measuring the diversity of Poissonian populations depends on the lower-bound cut-off applied. This research characterizes the classes of Poissonian populations whose diversities are invariant with respect to the cut-off level applied and establishes an elemental connection between these classes and extreme-value theory. The measures of diversity considered are variance and dispersion, Simpson's index and inverse participation ratio, Shannon's entropy and Rényi's entropy, and Gini's index.

DOI: 10.1103/PhysRevE.81.011122 PACS number(s): 02.50.−r, 89.65.Gh

I. INTRODUCTION

Populations represented by collections of points scattered randomly on the real line (or on a part of it) are prevalent across all fields of science. Examples of such random populations include: earthquakes taking place in a given geological region measured by their magnitudes—each point representing the magnitude of an earthquake; stars in a given sector of space measured by their masses—each point representing the mass of a star; citizens of a given state measured by their wealth—each point representing the wealth of a citizen; insurance claims in a given insurance portfolio measured by their costs—each point representing the cost of a claim; species in a given ecosystem measured by their fitness—each point representing the fitness of a species.

There are two main statistical methods to quantitatively model random populations: probabilistic and Poissonian. The probabilistic method is based on the notion of probability laws. A single member of a given population is sampled at random; the value of the sampled member is a random variable, and the probability law of this random variable represents the population's random scattering. The probabilistic model of the population is a sequence of independent and identically distributed (iid) random variables governed by population's probability law. The Poissonian method regards the random population as is—a collection of points scattered randomly on the real line—and models it as a Poisson process on the real line [1].

Probability laws are governed by their associated density functions—which are non-negative valued and normalized. Poisson processes, on the other hand, are governed by their associated intensity functions—which are non-negative valued but need not (necessarily) be normalized. Poisson processes with integrable intensities are essentially equivalent to probability laws. However, Poisson processes with nonintegrable intensities yield infinite random populations which cannot be modeled via the probabilistic method.

Poisson processes with nonintegrable intensities are of prime importance. These processes underlie: (a) the class of Lévy-stable probability laws—the only possible stochastic scaling limits of sums of iid random variables with infinite variance [2–4]; (b) the classes of extreme-value probability laws—Gumbel, Fréchet, and Weibull—the only possible stochastic scaling limits of maxima of iid random variables [5–7]; (c) nonlinear shot-noise systems [8,9]; (d) the definition and the exploration of fractality in the context of random populations [10,11]; (e) the definition and the exploration of the resilience of random populations to the action of random perturbations [12]; (f) mechanisms which universally generate fractal stochastic processes [13,14]. (This list of applications of Poisson processes with nonintegrable intensities is far from being exhaustive.)

Poisson processes with nonintegrable intensities arise naturally when considering random populations with a distinguished maximum—as is the case in all the aforementioned examples of random populations (earthquakes magnitudes, star masses, citizens wealth, insurance-claims costs, and species fitness). Indeed, the probability law of the maximal point of a given Poisson process is nondegenerate if and only if the intensity of the Poisson process is nonintegrable. Henceforth, we refer to Poisson processes with nondegenerate maxima as "Poissonian populations." A Poissonian population has the following topological structure: the population points can be ordered decreasingly—highest point, second-highest point, third-highest point, etc.; there are finitely many population points residing above any given level l ($l \in \mathcal{R}$, where $\mathcal{R}$ is the range of the population points), and infinitely many points residing below the level.

Science has developed a whole "toolbox" for quantifying the diversity of random populations represented by probability laws. Well-established measures of diversity include: (a) variance and dispersion—measuring the fluctuations of a population around its mean; (b) Simpson's index [15] and inverse participation ratio [16]—the former applied in Biology and Ecology to measure population heterogeneity, and the latter applied in Physics to measure population-localization; (c) Shannon's entropy [17] and Rényi's entropy

*eliazar@post.tau.ac.il
†igor.sokolov@physik.hu-berlin.de

[27]—applied in Physics and Information Theory to measure population-randomness; (d) Gini's index [18]—applied in Economics and Social Sciences to measure population-evenness ("societal egalitarianism"). How can these measures of diversity be applied to Poissonian populations? The answer to this question stems from the topological structure of Poissonian populations:

As noted above, a Poissonian population has finitely many population points residing above any given level $l(l \in \mathcal{R})$. Moreover, the points above a given level l turn out to be iid random variables governed by a common probability law. That is, to each level l corresponds an associated probability law—whose diversity can be measured by any of the aforementioned measures of diversity. Hence, applying a measure of diversity D to a Poissonian population $\mathcal{P}$ yields a level-dependent diversity function $D_{\mathcal{P}}(l)(l \in \mathcal{R}) - D_{\mathcal{P}}(l)$ being the diversity of the subpopulation of points residing above the level l. Our goal in this research is to: *characterize and explore Poissonian populations* $\mathcal{P}$ whose diversities are invariant with respect to the "cut-off" level l. Namely, Poissonian populations whose diversity functions are invariable: $D_{\mathcal{P}}(l) \equiv \text{const}$ $(l \in \mathcal{R})$. For such populations the measurement of diversity is independent of the cut-off level l applied—rendering diversity, in their case, a global quantitative gauge.

Henceforth, we refer to the cut-off level l applied as our "resolution level" (assuming values in the range $\mathcal{R}$), and term Poissonian populations with invariable diversity functions as "resolution invariant." Clearly, resolution invariance depends on the measure of diversity D considered. The analysis presented in this research yields three classes of resolution-invariant Poissonian populations, which are intimately related to the extreme-value probability laws noted above:

(1) The Gumbel class. Poissonian populations defined on the entire real line and governed by exponential intensities—which is resolution invariant with respect to the following measures of diversity: variance and dispersion; Simpson's index and inverse participation ratio; Shannon's entropy and Rényi's entropy. This class is termed "Gumbel" due to the fact that it constitutes of all Poissonian populations whose maximal point is governed by the Gumbel extreme-value probability law.

(2) The Fréchet class. Poissonian populations whose points are bounded from below and governed by power-law intensities—which is resolution invariant with respect to Gini's index. This class is termed "Fréchet" due to the fact that it constitutes of all Poissonian populations whose maximal point is governed by the Fréchet extreme-value probability law.

(3) The Weibull class. Poissonian populations whose points are bounded from above and governed by power-law intensities—which is resolution invariant with respect to Gini's index. This class is termed "Weibull" due to the fact that it constitutes of all Poissonian populations whose maximal point is governed by the Weibull extreme-value probability law.

The reminder of the paper is organized as follows. In Sec. II we concisely review: measures of diversity (Sec. II A); extreme-value probability laws (Sec. II B); Poisson processes (Sec. II C). In Sec. III we present the analysis of resolution-invariant Poissonian populations: methodology and general results (Sec. III A); the Gumbel class (Sec. III B); the Fréchet and Weibull classes (Sec. III C). In Sec. IV we introduce and analyze the notion of "global diversity" of Poissonian populations. The proofs of the main results are given in the Appendix.

II. PRELIMINARIES

In this section we concisely review the following preliminaries: measures of diversity (Sec. II A), extreme-value probability laws (Sec. II B), and Poisson processes (Sec. II C).

A. Measures of diversity

Different fields of Science use different gauges to quantify the diversity of random populations governed by arbitrary probability laws. Henceforth, consider a random population whose probability law is represented by the density function $f(x)$ (x real). We focus—to start with—on the following four "elemental" measures of diversity: variance, Simpson's index, Shannon's entropy, and Gini's index.

Variance:

$$V(f) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} x f(x)dx\right]^2. \quad (1)$$

The variance is the most basic measure of diversity and is applied across all fields of science. It is a positive-valued quantitative measure for the fluctuations of the population considered: the greater the variance—the more dispersed is the population around its mean $\mu(f) = \int_{-\infty}^{\infty} x f(x)dx$. The root of the variance is the population's standard deviation (σ).

Simpson's index [15]:

$$S(f) = \frac{1}{\int_{-\infty}^{\infty} f(x)^2 dx}. \quad (2)$$

Simpson's index $S(f)$ is applied in biology and ecology. It is a positive-valued quantitative measure for the heterogeneity of the population considered: the greater Simpson's index—the more heterogeneous and less uniform the population. For example, in the case of the Gaussian probability law Simpson's index is proportional to the population's standard deviation σ.

Remark 1. To attain intuition regarding the rational underlying Simpson's index, consider a random population governed by a discrete probability law $\{p_k\}_k$ (p_k being the occurrence probability of outcome k). In the discrete setting the reciprocal of Simpson's index is given by $S^{-1} = \Sigma_k p_k^2$. Namely, S^{-1} is the probability that two independent samples from the population will yield the same outcome. Hence, if there is only one possible outcome then $S^{-1} = 1$ and $S = 1$. On the other hand, if there are n outcomes—all occurring with equal probability $p_k = 1/n$—then $S^{-1} = 1/n$ and $S = n$. In general, if there are n possible outcomes then Simpson's index S assumes values in the range $1 \leq S \leq n$ where: the lower bound $S = 1$ corresponds to the deterministic scenario (where one single outcome has probability 1) and the upper bound

$S=n$ corresponds to the uniform scenario (where all outcomes occur with equal probability $1/n$). In physics the quantities S and S^{-1} are referred to, respectively, as the "participation ratio" and the "inverse participation ratio" [16]. These ratios are commonly applied (in physics) as measures of localization, and we shall elaborate on them in Sec. III B below.

Shannon's entropy [17]:

$$H(f) = -\int_{-\infty}^{\infty} \ln[f(x)]f(x)dx. \qquad (3)$$

Shannon's entropy $H(f)$ is applied in statistical physics and information theory [19]. It is a real-valued quantitative measure for the randomness of the population considered: the greater Shannon's entropy—the more random and haphazard the population. For example, in the case of the Gaussian probability law Shannon's entropy is proportional to the logarithm $\ln(\sigma)$ of the population's standard deviation σ.

Gini's index [18]:

$$G(f) = 1 - \frac{\int_0^{\infty} F(x)^2 dx}{\int_0^{\infty} F(x)dx}, \qquad (4)$$

where the function $F(x)=\int_x^{\infty} f(x')dx'(x\geq 0)$ is the tail probability corresponding to the density function $f(x)$ (see remark 2 below). Gini's index $G(f)$ is applied in economics and social sciences. It is a quantitative measure—taking values in the unit interval [0,1]—for the evenness of the population considered: the greater Gini's index—the more unequal and less egalitarian the population. In recent years the application of Gini's index has extended beyond economics and social sciences—where it is commonly applied as a measure of societal egalitarianism—and has attained popularity in other scientific fields as a general quantitative gauge of evenness. Examples include: astrophysics—the analysis of galaxy morphology [20]; medical chemistry—the analysis of kinase inhibitors [21]; ecology—the effect of biodiversity on ecosystem functioning [22].

Remark 2. Gini's index is defined only for random populations with non-negative values—i.e., density functions $f(x)$ supported on the non-negative half line $(x\geq 0)$. The definition of Gini's index can be extended to random populations with values bounded from either below or above—by considering the distance of the population members from the population's bound (rather than their sheer values). Hence, for a population bounded from below by the bound b—represented by the density function $f(x)(x\geq b)$—we set $F(x)=\int_{b+x}^{\infty} f(x')dx'(x\geq 0)$ in Eq. (4). And, for a population bounded from above by the bound b—represented by the density function $f(x)(x\leq b)$—we set $F(x)=\int_{-\infty}^{b-x} f(x')dx'(x\geq 0)$ in Eq. (4).

B. Extreme-value probability laws

Extreme-value theory studies the stochastic scaling limits of sequences of iid random variables [5–7]. Extreme-value theory is of major importance in the modeling and analysis of rare and catastrophic events such as floods in hydrology, large claims in insurance, crashes in finance, and material failure in corrosion analysis [23,24].

Given a sequence $\{X_n\}_{n=1}^{\infty}$ of iid random variables, consider the following affine scaling of the sequence's maxima:

$$M_n = \frac{\max\{X_1,\ldots,X_n\} - b_n}{a_n} \qquad (5)$$

$(n=1,2,\ldots)$, where $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are arbitrary scaling coefficients. Extreme-value theory seeks nontrivial stochastic limits $M=\lim_{n\to\infty} M_n$ (the limit being in law) of the scaled maxima. The "central limit theorem" of extreme-value theory asserts that there are three possible classes of nontrivial stochastic limits—referred to as the extreme-value probability laws [7]:

(i) *Gumbel*—admitting values on the entire real line, and governed by the cumulative distribution function

$$\mathrm{Prob}(M \leq x) = \exp[-c\exp(-\varepsilon x)] \qquad (6)$$

(x real; the coefficient c and the exponent ε are positive parameters).

(ii) *Fréchet*—its values bounded from below, and governed by the cumulative distribution function

$$\mathrm{Prob}(M \leq x) = \exp[-c(x-b)^{-\varepsilon}] \qquad (7)$$

($x\geq b$, where b is the lower bound; the coefficient c and the exponent ε are positive parameters).

(iii) *Weibull*—its values bounded from above, and governed by the cumulative distribution function

$$\mathrm{Prob}(M \leq x) = \exp[-c(b-x)^{\varepsilon}] \qquad (8)$$

($x\leq b$, where b is the upper bound; the coefficient c and the exponent ε are positive parameters).

C. Poisson processes

This paper studies the diversity of populations represented by collections of points scattered randomly on the real line (or on a part of it). The common statistical method for the random scattering of points in general domains is that of Poisson processes [1]. Poisson processes have a wide spectrum of applications ranging from insurance and finance [23] to queueing systems [25].

Henceforth, we consider a real range $\mathcal{R}=(r_*,r^*)$ of values, where $r_*(r_*\geq -\infty)$ is the range's lower bound and $r^*(r^*\leq\infty)$ is the range's upper bound. Informally, a Poisson process $\mathcal{P}$ defined on the range $\mathcal{R}$, with intensity $\lambda(x)(x\in\mathcal{R})$, is a random collection of points scattered randomly as follows: the infinitesimal interval $(x,x+dx)$ contains a single point with probability $\lambda(x)dx$ and is empty with probability $1-\lambda(x)dx$ (independently of all other infinitesimal intervals). More precisely, the points of the Poisson process $\mathcal{P}$ are scattered as follows [1]: (i) the number of points residing in an interval $I \subset \mathcal{R}$ is a Poisson-distributed random variable with mean $\int_I \lambda(x)dx$;[1] (ii) the number of points residing in disjoint intervals are independent random variables.

The maximal point of the Poisson process $\mathcal{P}$—henceforth denoted $M_{\mathcal{P}}$—is governed by the cumulative distribution function

$$\text{Prob}(M_{\mathcal{P}} \leq x) = \exp[-\Lambda(x)] \tag{9}$$

($x \in \mathcal{R}$), where $\Lambda(x) = \int_x^{r^*} \lambda(x')dx'$ is the *tail intensity* of the Poisson process $\mathcal{P}$.[2] For the cumulative distribution function of Eq. (9) to be nondegenerate—i.e., to be monotone increasing from the level 0 (in the limit $x \to r_*$) to the level 1 (in the limit $x \to r^*$)—it is required that the intensity $\lambda(x)$ be nonintegrable at the lower bound r_* and integrable at the upper bound r^*. This, in turn, implies that: (i) the function $\Lambda(x)$ is monotone decreasing from the level $\infty = \lim_{x \to r_*} \Lambda(x)$ to the level $0 = \lim_{x \to r^*} \Lambda(x)$; (ii) the random population $\mathcal{P}$ is infinite—consisting of infinitely many points. As stated in the introduction, we term Poisson processes with nondegenerate maxima Poissonian populations.

Comparing the cumulative distribution function of Eq. (9) to the cumulative distribution functions of Eqs. (6)–(8) leads to an intimate and profound relation between extreme-value theory and Poissonian populations. Indeed, the extreme-value stochastic limits M turn out to be equal (in law) to the maximal points $M_{\mathcal{P}}$ of special Poissonian populations:

(i) *Gumbel*—the probability law of the maximal point of Poissonian populations, defined on the entire real line, with tail intensity

$$\Lambda(x) = c \exp(-\varepsilon x) \tag{10}$$

(x real; the coefficient c and the exponent ε are positive parameters).

(ii) *Fréchet*—the probability law of the maximal point of Poissonian populations, defined on the half-line (b,∞), with tail intensity

$$\Lambda(x) = c(x-b)^{-\varepsilon} \tag{11}$$

($x > b$; the coefficient c and the exponent ε are positive parameters).

(iii) *Weibull*—the probability law of the maximal point of Poissonian populations, defined on the half-line $(-\infty, b)$, with tail intensity

$$\Lambda(x) = c(b-x)^{\varepsilon} \tag{12}$$

($x \leq b$; the coefficient c and the exponent ε are positive parameters).

[1]Namely, the probability that the interval I will contain exactly k points is given by $\frac{1}{k!}(\mu_I)^k \exp(-\mu_I)(k=0,1,2,\ldots)$, where $\mu_I = \int_I \lambda(x)dx$.

[2]The derivation of Eq. (9) is as follows. The maximal point $M_{\mathcal{P}}$ satisfies $\{M_{\mathcal{P}} \leq x\}$ if and only if the random population $\mathcal{P}$ has no points exceeding the level x—i.e., has zero points residing in the interval $I=(x,r^*)$. However, the number of points residing in the interval I is Poisson-distributed with mean $\Lambda(x) = \int_x^{r^*} \lambda(x')dx'$—implying that the probability that the interval I contains zero points is $\exp(-\Lambda(x))$.

III. DIVERSITY OF POISSONIAN POPULATIONS

In this section we present the analysis of Poissonian populations with resolution-invariant diversity functions: methodology and general results (Sec. III A), the Gumbel class (Sec. III B), and the Fréchet and Weibull classes (Sec. III C).

A. Methodology and general results

Consider a Poissonian population $\mathcal{P}$ defined on the range $\mathcal{R}$ and governed by the intensity $\lambda(x)(x \in \mathcal{R})$. Given a level $l(l \in \mathcal{R})$, the Poissonian population $\mathcal{P}$ consists of finitely many points residing above the level l and of infinitely many points residing below the level l. The "existence theorem" of the theory of Poisson processes ([1], Sec. 2.5) implies that the points of the Poissonian population $\mathcal{P}$ which reside above the level l are iid random variables—their probability law governed by the common density function

$$f_l(x) = \begin{cases} 0 & (r_* < x < l) \\ \dfrac{\lambda(x)}{\Lambda(l)} & (l \leq x < r^*). \end{cases} \tag{13}$$

Thus, to each level $l(l \in \mathcal{R})$ corresponds a density function $f_l(x)$—given by Eq. (13)—which represents the probability law of the sub-population of points residing above this level. Applying a measure of diversity D to the density function $f_l(x)$ yields, in turn, the diversity $D(f_l)$ of the sub-population of points residing above the level l. Hence, when trying to quantify the diversity of a Poissonian population $\mathcal{P}$, we obtain a diversity function $D_{\mathcal{P}}(l) \coloneqq D(f_l)$ which is dependent on the resolution level l applied ($l \in \mathcal{R}$). We refer to the cut-off level l applied as our "resolution level" and focus in this research on Poissonian populations whose diversities are resolution invariant:

Definition 3. A Poissonian population $\mathcal{P}$, defined on the range $\mathcal{R}$, is resolution-invariant with respect to the measure of diversity D if its diversity function $D_{\mathcal{P}}(l)$ is invariable with respect to the resolution level l applied: $D_{\mathcal{P}}(l) \equiv \text{const}$ ($l \in \mathcal{R}$).

A statistical analysis yields the following characterizations of resolution-invariant Poissonian populations:

Proposition 4. The Gumbel class. A Poissonian population $\mathcal{P}$, defined on the real line, is resolution invariant with respect to the following measures of diversity—variance, Simpson's index, and Shannon's entropy—if and only if it is governed by the "Gumbel tail intensity" of Eq. (10).

Proposition 5. The Fréchet class. A Poissonian population $\mathcal{P}$, defined on the half-line (b,∞), is resolution invariant with respect to Gini's index if and only if it is governed by the "Fréchet tail intensity' of Eq. (11).

Proposition 6. The Weibull class. A Poissonian population $\mathcal{P}$, defined on the half-line $(-\infty,b)$, is resolution invariant with respect to Gini's index if and only if it is governed by the "Weibull tail intensity" of Eq. (12).

The proofs of propositions 4–6 are given in the Appendix, Secs. 1–3. In what follows, we further explore the Gumbel, Fréchet, and Weibull classes of Poissonian populations.

B. Gumbel class

Poissonian populations, governed by the Gumbel tail intensity of Eq. (10), were shown to be resolution invariant with respect to the following measures of diversity: variance, Simpson's index, and Shannon's entropy. These measures of diversity, in turn, are special cases of the following—more general—measures of diversity:

Dispersion:

$$D_\alpha(f) = \int_{-\infty}^{\infty} |x-\mu(f)|^\alpha f(x)dx, \tag{14}$$

where $\mu(f)=\int_{-\infty}^{\infty} x f(x)dx$ is the mean of density function $f(x)$ and where α is a positive exponent. The dispersion measures the fluctuation around the mean using the L_α metric. The variance $V(f)$ is a special case of the dispersion, corresponding to the L_2 metric: $V(f)=D_2(f)$.

Inverse participation ratio [16]:

$$P_\alpha(f) = \int_{-\infty}^{\infty} f(x)^\alpha dx, \tag{15}$$

where α is a positive exponent. The inverse participation ratio $P_\alpha(f)$ is the distance—measured in the L_α metric—of the density function $f(x)$ from the zero function. The inverse participation ratio $P_\alpha(f)$ and the participation ratio $1/P_\alpha(f)$ are commonly used in physics as a quantitative measure of localization of wave functions in the disordered media [26]. Simpson's index $S(f)$ is a special case of the participation ratio, corresponding to the L_2 metric: $S(f)=1/P_2(f)$.

Rényi's entropy [27]:

$$R_\alpha(f) = \frac{-1}{\alpha-1}\ln\left[\int_{-\infty}^{\infty} f(x)^\alpha dx\right], \tag{16}$$

where $\alpha\neq 1$ is a positive exponent. Rényi's entropy $R_\alpha(f)$ is proportional to the logarithm of the inverse participation ratio $P_\alpha(f)$. Shannon's entropy $H(f)$ is a special case of the Rényi's entropy, corresponding to the L_1 metric: $H(f)=\lim_{\alpha\to 1} R_\alpha(f)$.

Statistical analysis asserts that proposition 4 can be extended to include also diversities with respect to dispersion $D_\alpha(f)$, inverse participation ratio $P_\alpha(f)$, and Rényi's entropy $R_\alpha(f)$:

Proposition 7. The Gumbel class. A Poissonian population $\mathcal{P}$, defined on the real line, is resolution-invariant with respect to the following measures of diversity—dispersion, inverse participation ratio, and Rényi's entropy—if and only if it is governed by the Gumbel tail intensity of Eq. (10).

The proof of proposition 7 is given in the Appendix, Sec. 1. The resolution invariance of Poissonian populations, governed by the Gumbel tail intensity of Eq. (10), can be approached also from a different perspective which we now describe. Substituting the resolution-dependent density function of Eq. (13) into Eqs. (14)–(16) yields, respectively:

Dispersion:

$$D_\alpha(f_l) = \int_0^{\infty}\left|x-\int_0^{\infty} x'\left(\frac{\lambda(l+x')}{\Lambda(l)}\right)dx'\right|^\alpha\left(\frac{\lambda(l+x)}{\Lambda(l)}\right)dx. \tag{17}$$

Inverse participation ratio:

$$P_\alpha(f_l) = \int_0^{\infty}\left(\frac{\lambda(l+x)}{\Lambda(l)}\right)^\alpha dx. \tag{18}$$

Rényi's entropy:

$$R_\alpha(f) = \begin{cases} \dfrac{-1}{\alpha-1}\ln\left[\displaystyle\int_0^{\infty}\left(\frac{\lambda(l+x)}{\Lambda(l)}\right)^\alpha dx\right] & \alpha\neq 1 \\ \displaystyle\int_0^{\infty}\ln\left(\frac{\lambda(l+x)}{\Lambda(l)}\right)\cdot\left(\frac{\lambda(l+x)}{\Lambda(l)}\right)dx & \alpha=1. \end{cases} \tag{19}$$

Note that the resolution-dependent measures of diversity appearing in Eqs. (17)–(19) are all functionals of the density function

$$\psi_l(x) = \frac{\lambda(l+x)}{\Lambda(l)} \tag{20}$$

$(x\geq 0)$. This observation leads to the following result:

Proposition 8. The density function $\psi_l(x)$ of Eq. (20) is independent of the resolution level l (l real) if and only if the tail intensity $\Lambda(x)$ (x real) admits the Gumbel form of Eq. (10)—in which case $\psi_l(x)=\psi(x)$, where

$$\psi(x) = \varepsilon\exp(-\varepsilon x) \quad (x\geq 0). \tag{21}$$

The proof of proposition 8 is given in the Appendix, Sec. 1. Note that the density function $\psi(x)$ appearing in Eq. (21) is that of the *Exponential* probability law with mean $\frac{1}{\varepsilon}$. The meaning of proposition 8 is the following:

Let $X_l(X_l\geq l)$ denote the random value of an arbitrary point of the Poissonian population $\mathcal{P}$ residing above the resolution level l (l real). The density function of Eq. (20), in effect, governs the probability law of the scaled random variable $\hat{X}_l=X_l-l(\hat{X}_l\geq 0)$. Proposition 8 asserts that the probability law of the scaled random variable $\hat{X}_l$ is independent of the resolution level l if and only if the tail intensity $\Lambda(x)$ (x real) admits the Gumbel form of Eq. (10)—in which case the scaled random variables $\{\hat{X}_l\}_{-\infty<l<\infty}$ are all governed by a common *Exponential* probability law with mean $\frac{1}{\varepsilon}$.

C. Fréchet and Weibull classes

Approaches analogous to the one leading to proposition 8 can be followed also in the cases of the Fréchet and Weibull classes of Poissonian populations.

1. Fréchet class

Consider a Poissonian population $\mathcal{P}$ defined on the half-line (b,∞). Equation (A24) of the Appendix implies that the population's resolution-dependent Gini index is given by

$$G(f_l) = 1 - \frac{1 + \int_1^{\infty} \left\{ \frac{\Lambda[b + (l-b)x]}{\Lambda(l)} \right\}^2 dx}{1 + \int_1^{\infty} \left\{ \frac{\Lambda[b + (l-b)x]}{\Lambda(l)} \right\} dx} \tag{22}$$

$(l > b)$. Noting that the resolution-dependent Gini index of Eq. (22) is a functional of the tail probability

$$\Psi_l(x) = \frac{\Lambda(b + (l-b)x)}{\Lambda(l)} \tag{23}$$

$(x \geq 1)$ leads to the following result:

Proposition 9. The tail probability $\Psi_l(x)$ of Eq. (23) is independent of the resolution level $l(l > b)$ if and only if the tail intensity $\Lambda(x)(x > b)$ admits the 'Fréchet form' of Eq. (11)—in which case $\Psi_l(x) = \Psi(x)$, where

$$\Psi(x) = x^{-\varepsilon} \quad (x \geq 1). \tag{24}$$

The proof of proposition 9 is given in the Appendix, Sec. 2. Note that the tail probability $\Psi(x)$ appearing in Eq. (24) is that of the *Pareto* probability law with exponent ε. The meaning of proposition 9 is the following:

Let $X_l(X_l \geq l)$ denote the random value of an arbitrary point of the Poissonian population $\mathcal{P}$ residing above the resolution level l $(l > b)$. The tail probability of Eq. (23), in effect, governs the probability law of the scaled random variable $\hat{X}_l = (X_l - b)/(l - b)(\hat{X}_l \geq 1)$. Proposition 9 asserts that the probability law of the scaled random variable $\hat{X}_l$ is independent of the resolution level l if and only if the tail intensity $\Lambda(x)(x > b)$ admits the Fréchet form of Eq. (11)—in which case the scaled random variables $\{\hat{X}_l\}_{l>b}$ are all governed by a common *Pareto* probability law with exponent ε.

2. Weibull class

Consider a Poissonian population $\mathcal{P}$ defined on the half-line $(-\infty, b)$. Equation (A38) of the Appendix implies that the population's resolution-dependent Gini index is given by

$$G(f_l) = 1 - \frac{\int_0^1 \left(\frac{\Lambda(b - (b-l)x)}{\Lambda(l)} \right)^2 dx}{\int_0^1 \left(\frac{\Lambda(b - (b-l)x)}{\Lambda(l)} \right) dx} \tag{25}$$

$(l < b)$. Noting that the resolution-dependent Gini index of Eq. (25) is a functional of the cumulative distribution function

$$\Psi_l(x) = \frac{\Lambda[b - (b-l)x]}{\Lambda(l)} \tag{26}$$

$(0 \leq x \leq 1)$ leads to the following result:

Proposition 10. The cumulative distribution function $\Psi_l(x)$ of Eq. (26) is independent of the resolution level l $(l < b)$ if and only if the tail intensity $\Lambda(x)(x \leq b)$ admits the 'Weibull form' of Eq. (12)—in which case $\Psi_l(x) = \Psi(x)$, where

$$\Psi(x) = x^{\varepsilon} \quad (0 \leq x \leq 1). \tag{27}$$

The proof of proposition 10 is given in Appendix, Sec. 3. Note that the cumulative distribution $\Psi(x)$ appearing in Eq. (27) is that of the *inverse Pareto* probability law with exponent ε. The meaning of proposition 10 is the following:

Let $X_l(l \leq X_l \leq b)$ denote the random value of an arbitrary point of the Poissonian population $\mathcal{P}$ residing above the resolution level $l(l < b)$. The cumulative distribution function of Eq. (26), in effect, governs the probability law of the scaled random variable $\hat{X}_l = (b - X_l)/(b - l)(0 \leq \hat{X}_l \leq 1)$. Proposition 10 asserts that the probability law of the scaled random variable $\hat{X}_l$ is independent of the resolution level l if and only if the tail intensity $\Lambda(x)(x \leq b)$ admits the Weibull form of Eq. (12)—in which case the scaled random variables $\{\hat{X}_l\}_{l \leq b}$ are all governed by a common *inverse Pareto* probability law with exponent ε.

IV. GLOBAL DIVERSITY

So far we followed a resolution-invariance approach to explore the diversity of Poissonian populations—focusing on Poissonian populations whose diversity functions $D_{\mathcal{P}}(l)$ are invariable with respect to the resolution level $l(l \in \mathcal{R})$. In this section we follow an alternative approach based on the notion of global diversity:

Definition 11. A Poissonian population $\mathcal{P}$, defined on the range $\mathcal{R}$, has a global diversity $D_{\mathcal{P}}$—with respect to the measure of diversity D—if the limit $D_{\mathcal{P}} := \lim_{l \to r_*} D_{\mathcal{P}}(l)$ exists.

The statistical analysis of global diversity involves the notion of *regular variation* [28]. A real function $\phi(x)$ is said to be regularly varying at the limit point $x \to p$ if the limit $\lim_{x \to p} \phi(cx)/\phi(x)$ exists for all positive constants c. Theory shows that if the function $\phi(x)$ is regularly varying at the limit point $x \to p$ then $\lim_{x \to p} \phi(cx)/\phi(x) = c^{\nu}$, where the exponent ν is a real parameter called the "exponent of regular variation." Regularly varying functions are generalizations of power-law functions and play a key role in many fields of probability theory (see Chap. 8 in [28]). With the notion of regular variation at hand, we are in position to present the global diversity counterparts of the Gumbel, Fréchet, and Weibull classes explored so forth.

The Gumbel counterpart. Equations (17)–(19) imply that a Poissonian population $\mathcal{P}$, defined on the entire real line, has a global diversity with respect to the following measures of diversity—variance and dispersion, Simpson's index and inverse participation ratio, Shannon's entropy and Rényi's entropy—if and only if the limit

$$\psi(x) = \lim_{l \to -\infty} \frac{\lambda(x + l)}{\Lambda(l)} \tag{28}$$

exists for all $x \geq 0$. The limit of Eq. (28) exists if and only if the function $\tilde{\lambda}(\theta) = \lambda[\ln(\theta)](\theta > 0)$ is regularly varying at the limit $\theta \to 0$—in which case the function $\psi(x)(x \geq 0)$ admits the exponential form of Eq. (21). This case is the asymptotic counterpart of the Gumbel case of propositions 4 and 7.

The Fréchet counterpart. Equation (22) implies that a Poissonian population $\mathcal{P}$, defined on the half-line (b, ∞), has a global diversity with respect to Gini's index if and only if the limit

$$\Psi(x) = \lim_{l \to b} \frac{\Lambda[b + (l - b)x]}{\Lambda(l)} \tag{29}$$

exists for all $x \geq 1$. The limit of Eq. (29) exists if and only if the tail intensity $\Lambda(l)(l > b)$ is regularly varying at the limit $l \to b$—in which case the function $\Psi(x)(x \geq 1)$ admits the Pareto form of Eq. (24). This case is the asymptotic counterpart of the Fréchet case of proposition 5.

The Weibull counterpart. Equation (25) implies that a Poissonian population $\mathcal{P}$, defined on the half-line $(-\infty, b)$, has a global diversity with respect to Gini's index if and only if the limit

$$\Psi(x) = \lim_{l \to -\infty} \frac{\Lambda[b - (b - l)x]}{\Lambda(l)} \tag{30}$$

exists for all $0 \leq x \leq 1$. The limit of Eq. (30) exists if and only if the tail intensity $\Lambda(l)(l < b)$ is regularly varying at the limit $l \to -\infty$—in which case the function $\Psi(x)(0 \leq x \leq 1)$ admits the Beta form of Eq. (27). This case is the asymptotic counterpart of the Weibull case of proposition 6.

V. CONCLUSIONS

This paper studied the diversity of Poissonian populations. Considering populations represented by collections of points scattered randomly on the real line and modeling such populations by Poisson processes with distinguished maxima led us to Poissonian populations. The infiniteness of these random populations renders them beyond the quantitative description of probability laws—thus making it impossible to study their statistical diversity via measures of diversity of probability laws.

However, exploiting the topological structure of Poissonian populations gave rise to a resolution-dependent measurement of their diversities. A Poissonian population $\mathcal{P}$, defined on the range $\mathcal{R}$, has finitely many points above any given level $l(l \in \mathcal{R})$ and: the subpopulation of points residing above the resolution level l are iid random variables governed by a common resolution-dependent probability law [with density function $f_l(x)$ given by Eq. (13)]. Hence—with respect to a given measure of diversity D—to each level l we obtain the diversity $D(f_l)$ of the subpopulation of points residing above the level l. The diversity of the entire Poissonian population $\mathcal{P}$ is given, in turn, by the level-dependent diversity function $D_{\mathcal{P}}(l) = D(f_l)(l \in \mathcal{R})$.

The goal of this research was to characterize Poissonian populations whose diversity functions are invariable with respect to the level: $D_{\mathcal{P}}(l) \equiv \text{const}$ $(l \in \mathcal{R})$. For such populations the diversity is independent of the cut-off level l applied and is thus a global quantitative gauge. We termed the cut-off level l our resolution level and referred to the aforementioned Poissonian populations as resolution-invariant. Statistical analysis led to the following three classes of resolution-invariant Poissonian populations:

(1) *The Gumbel class*—Poissonian populations defined on the entire real line, and governed by exponential intensities—which is resolution-invariant with respect to the following measures of diversity: variance and dispersion; Simpson's index and inverse participation ratio; Shannon's entropy and Rényi's entropy.

(2) *The Fréchet class*—Poissonian populations whose points are bounded from below and governed by power-law intensities—which is resolution-invariant with respect to Gini's index.

(3) *The Weibull class*—Poissonian populations whose points are bounded from above and governed by power-law intensities—which is resolution-invariant with respect to Gini's index.

The Gumbel, Fréchet and Weibull classes are termed such due to the fact that they constitute, respectively, all Poissonian populations whose maximal points are governed by the Gumbel, Fréchet, and Weibull extreme-value probability laws. Thus, this paper presents an explicit characterization of Poissonian populations with resolution-invariant diversities and establishes an elemental connection between these Poissonian populations and extreme-value theory.

APPENDIX

1. Gumbel class

a. Proposition 4: Variance

Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (1) yields the resolution-dependent variance

$$V(f_l) = \frac{1}{\Lambda(l)}\left[\int_l^\infty x^2\lambda(x)dx\right] - \frac{1}{\Lambda(l)^2}\left[\int_l^\infty x\lambda(x)dx\right]^2. \tag{A1}$$

Assuming that the variance $V(f_l)$ is resolution-independent—i.e., $V(f_l) \equiv v$ (for all resolution levels l)—implies that

$$\Lambda(l)\left[\int_l^\infty x^2\lambda(x)dx\right] - \left[\int_l^\infty x\lambda(x)dx\right]^2 = v1\Lambda(l)^2. \tag{A2}$$

Differentiating Eq. (A2) with respect to the resolution level l leads—after some basic algebra—to

$$\Lambda''(l) = \frac{1}{v}\Lambda(l). \tag{A3}$$

Since the function $\Lambda(l)$ (l real) is monotone decreasing, the solution of the differential Eq. (A3) is of the form

$$\Lambda(l) = c\exp(-\varepsilon l), \tag{A4}$$

where the coefficient c and the exponent ε are positive parameters.

b. Proposition 4: Simpson's index

Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (2) yields the resolution-dependent Simpson index

$$S(f_l) = \frac{\Lambda(l)^2}{\int_l^\infty \lambda(x)^2 dx}. \quad \text{(A5)}$$

Assuming that the Simpson index $S(f_l)$ is resolution independent—i.e., $S(f_l) \equiv s$ (for all resolution levels l)—implies that

$$\int_l^\infty \lambda(x)^2 dx = \frac{1}{s}\Lambda(l)^2. \quad \text{(A6)}$$

Differentiating Eq. (A6) with respect to the resolution level l leads—after some basic algebra—to

$$\Lambda'(l) = -\frac{2}{s}\Lambda(l). \quad \text{(A7)}$$

The solution of the differential Eq. (A7) is of the form

$$\Lambda(l) = c \exp(-\varepsilon l), \quad \text{(A8)}$$

where the coefficient c and the exponent ε are positive parameters.

c. Proposition 4: Shannon's entropy

Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (3) yields—after some basic algebra—the resolution-dependent Shannon entropy

$$H(f_l) = \frac{1}{\Lambda(l)}\int_l^\infty \lambda(x)\ln(\lambda(x))dx - \ln(\Lambda(l)). \quad \text{(A9)}$$

Assuming that the Shannon entropy $H(f_l)$ is resolution independent—i.e., $H(f_l) \equiv h$ (for all resolution levels l)—implies that

$$\int_l^\infty \lambda(x)\ln[\lambda(x)]dx - \Lambda(l)\ln[\Lambda(l)] = h\Lambda(l). \quad \text{(A10)}$$

Differentiating Eq. (A10) with respect to the resolution level l leads—after some basic algebra—to

$$\Lambda'(l) = -\exp(1+h)\Lambda(l). \quad \text{(A11)}$$

The solution of the differential Eq. (A11) is of the form

$$\Lambda(l) = c \exp(-\varepsilon l), \quad \text{(A12)}$$

where the coefficient c and the exponent ε are positive parameters.

d. Proposition 7: Inverse participation ratio and Rényi's entropy ($\alpha \neq 1$)

Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (15) yields the resolution-dependent inverse participation ratio

$$P_\alpha(f_l) = \Lambda(l)^{-\alpha}\int_l^\infty \lambda(x)^\alpha dx. \quad \text{(A13)}$$

Assuming that the inverse participation ratio $P_\alpha(f_l)$ is resolution-independent—i.e., $P_\alpha(f_l) \equiv p$ (for all resolution levels l)—implies that

$$\int_l^\infty \lambda(x)^\alpha dx = p\Lambda(l)^\alpha. \quad \text{(A14)}$$

Differentiating Eq. (A14) with respect to the resolution level l leads—after some basic algebra—to

$$\Lambda'(l) = -(p\alpha)^{1/(\alpha-1)}\Lambda(l). \quad \text{(A15)}$$

The solution of the differential Eq. (A15) is of the form

$$\Lambda(l) = c \exp(-\varepsilon l), \quad \text{(A16)}$$

where the coefficient c and the exponent ε are positive parameters.

Equation (16) implies that

$$R_\alpha(f_l) = \frac{-1}{\alpha - 1}\ln[P_\alpha(f_l)]. \quad \text{(A17)}$$

Hence, the Rényi entropy $R_\alpha(f_l)$ is resolution independent if and only if the inverse participation ratio $P_\alpha(f_l)$ is resolution independent—leading, once again, to Eq. (A16).

e. Proposition 8

Assume that the density function $\psi_l(x)(x \geq 0)$ of Eq. (20) is independent of the resolution level l. This means that

$$\frac{\Lambda'(l+x)}{\Lambda(l)} = -\psi(x) \quad \text{(A18)}$$

holds for all l real and $x \geq 0$, where $\psi(x)$ is a positive-valued function of the variable x. In particular, Eq. (A18) implies that

$$\Lambda'(l) = -\varepsilon\Lambda(l) \quad \text{(A19)}$$

(l real), where $\varepsilon = \psi(0)$. The solution of the differential Eq. (A19) is

$$\Lambda(l) = c \exp(-\varepsilon l) \quad \text{(A20)}$$

(l real). Substituting the tail intensity of Eq. (A20) back into Eq. (A18) further yields

$$\psi(x) = \varepsilon \exp(-\varepsilon x) \quad (x \geq 0). \quad \text{(A21)}$$

2. Fréchet class

a. Proposition 5

The Poissonian population considered is defined on the half-line (b,∞). Hence—due to remark 2—the tail probability corresponding to the density function $f_l(x)$ $(x > b)$ is given by

$$F_l(x) = \int_{b+x}^\infty f_l(x')dx' \quad \text{(A22)}$$

$(x \geq 0)$. Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (A22) yields the resolution-dependent tail probability

$$F_l(x) = \begin{cases} 1 & (0 \le x < l-b) \\ \dfrac{\Lambda(b+x)}{\Lambda(l)} & (l-b \le x < \infty). \end{cases} \tag{A23}$$

Substituting the resolution-dependent tail probability of Eq. (A23) into Eq. (4) further yields the resolution-dependent Gini index

$$G(f_l) = 1 - \frac{(l-b) + \int_{l-b}^{\infty} \left(\frac{\Lambda(b+x)}{\Lambda(l)}\right)^2 dx}{(l-b) + \int_{l-b}^{\infty} \left(\frac{\Lambda(b+x)}{\Lambda(l)}\right) dx}. \tag{A24}$$

Setting $\theta = l-b (l>b)$ and $\Phi(\theta) = \Lambda(b+\theta)(\theta>0)$, Eq. (A24) becomes

$$G(f_l) = 1 - \frac{\theta + \int_{\theta}^{\infty} \left(\frac{\Phi(x)}{\Phi(\theta)}\right)^2 dx}{\theta + \int_{\theta}^{\infty} \left(\frac{\Phi(x)}{\Phi(\theta)}\right) dx}. \tag{A25}$$

Assuming that the Gini index $G(f_l)$ is resolution independent—i.e., $G(f_l) \equiv g$ (for all resolution levels $l>b$)—implies that

$$\theta\Phi(\theta)^2 + \int_{\theta}^{\infty} \Phi(x)^2 dx = (1-g)\left[\theta\Phi(\theta)^2 + \Phi(\theta)\int_{\theta}^{\infty} \Phi(x) dx\right] \tag{A26}$$

holds for all $\theta>0$. Differentiating Eq. (A26) with respect to the variable θ leads—after some basic algebra—to

$$2\theta\Phi(\theta) = (1-g)\left[2\theta\Phi(\theta) + \int_{\theta}^{\infty} \Phi(x) dx\right]. \tag{A27}$$

Differentiating Eq. (A27) with respect to the variable θ leads—after some basic algebra—to

$$\frac{\Phi'(\theta)}{\Phi(\theta)} = -\left(\frac{1+g}{2g}\right)\frac{1}{\theta}. \tag{A28}$$

The solution of the differential Eq. (A28) is of the form

$$\Phi(\theta) = c\theta^{-\varepsilon} \tag{A29}$$

$(\theta>0)$, where the coefficient c and the exponent ε are positive parameters. Equation (A29), in turn, implies that

$$\Lambda(l) = c(l-b)^{-\varepsilon} \tag{A30}$$

$(l>b)$.

b. Proposition 9

Assume that the tail probability $\Psi_l(x)$ of Eq. (23) is independent of the resolution level $l(l>b)$. This means that

$$\frac{\Lambda(b+\theta x)}{\Lambda(b+\theta)} = \Psi(x) \tag{A31}$$

holds for all $\theta>0$ real and $x \ge 1$, where $\Psi(x)$ is a positive-valued function of the variable x. Equation (A31) can be extended also to $0<x\le 1$ via

$$\Psi(x) = \frac{\Lambda\left(b+\frac{\theta}{x}x\right)}{\Lambda\left(b+\frac{\theta}{x}\right)} = \frac{1}{\dfrac{\Lambda\left(b+\theta\frac{1}{x}\right)}{\Lambda(b+\theta)}} = \frac{1}{\Psi\left(\frac{1}{x}\right)} \tag{A32}$$

and Eq. (A31) implies that

$$\Psi(xy) = \frac{\Lambda(b+\theta xy)}{\Lambda(b+\theta)} = \frac{\Lambda(b+(\theta y)x)}{\Lambda(b+(\theta y))}\frac{\Lambda(b+(\theta y))}{\Lambda(b+\theta)} = \Psi(x)\Psi(y) \tag{A33}$$

holds for all $x, y \ge 1$.

Equations (A32) and (A33) imply that the function $\Psi(x)$ is a power law: $\Psi(x) = x^p (x>0)$. Since the tail intensity $\Lambda(l)(l>b)$ is monotone decreasing, we conclude that

$$\Psi(x) = x^{-\varepsilon} \tag{A34}$$

$(x>0)$, where ε is a positive exponent. Substituting Eq. (A34) back into Eq. (A31)—while setting $\theta=1$ and $x=l-b$—yields

$$\Lambda(l) = c(l-b)^{-\varepsilon} \tag{A35}$$

$(l>b)$, where $c = \Lambda(b+1)$.

3. Weibull class

a. Proposition 6

The Poissonian population considered is defined on the half-line $(-\infty, b)$. Hence—due to remark 2—the tail probability corresponding to the density function $f_l(x)$ $(x<b)$ is given by

$$F_l(x) = \int_{-\infty}^{b-x} f_l(x') dx' \tag{A36}$$

$(x\ge 0)$. Substituting the resolution-dependent density function $f_l(x)$ [given by Eq. (13)] into Eq. (A36) yields the resolution-dependent tail probability

$$F_l(x) = \begin{cases} 1 - \dfrac{\Lambda(b-x)}{\Lambda(l)} & (0 \le x < b-l) \\ 0 & (b-l \le x < \infty). \end{cases} \tag{A37}$$

Substituting the resolution-dependent tail probability of Eq. (A37) into Eq. (4) further yields the resolution-dependent Gini index

$$G(f_l) = 1 - \frac{\int_0^{b-l} \left(1 - \frac{\Lambda(b-x)}{\Lambda(l)}\right)^2 dx}{\int_0^{b-l} \left(1 - \frac{\Lambda(b-x)}{\Lambda(l)}\right) dx}. \tag{A38}$$

Setting $\theta = b-l (l\le b)$ and $\Phi(\theta) = \Lambda(b-\theta)(\theta \ge 0)$, Eq. (A38) becomes

$$G(f_l) = 1 - \frac{\int_0^\theta \left(\frac{\Phi(x)}{\Phi(\theta)}\right)^2 dx}{\int_0^\theta \left(\frac{\Phi(x)}{\Phi(\theta)}\right) dx}. \quad (A39)$$

Assuming that the Gini index $G(f_l)$ is resolution-independent—i.e., $G(f_l)\equiv g$ (for all resolution levels $l\leq b$)—implies that

$$\theta\Phi(\theta)^2 - 2\Phi(\theta)\int_0^\theta \Phi(x)dx + \int_0^\theta \Phi(x)^2 dx = (1-g)\left[\theta\Phi(\theta)^2 - \Phi(\theta)\int_0^\theta \Phi(x)dx\right] \quad (A40)$$

holds for all $\theta\geq 0$. Differentiating Eq. (A40) with respect to the variable θ leads—after some basic algebra—to

$$2\theta\Phi(\theta) - 2\int_0^\theta \Phi(x)dx = (1-g)\left[2\theta\Phi(\theta) - \int_0^\theta \Phi(x)dx\right]. \quad (A41)$$

Differentiating Eq. (A41) with respect to the variable θ leads—after some basic algebra—to

$$\frac{\Phi'(\theta)}{\Phi(\theta)} = \left(\frac{1-g}{2g}\right)\frac{1}{\theta}. \quad (A42)$$

The solution of the differential Eq. (A42) is of the form

$$\Phi(\theta) = c\theta^\varepsilon \quad (A43)$$

($\theta\geq 0$), where the coefficient c and the exponent ε are positive parameters. Equation (A43), in turn, implies that

$$\Lambda(l) = c(b-l)^\varepsilon \quad (A44)$$

($l\leq b$).

b. Proposition 10

Assume that the cumulative distribution function $\Psi_l(x)$ of Eq. (26) is independent of the resolution level l ($l<b$). This means that

$$\frac{\Lambda(b-\theta x)}{\Lambda(b-\theta)} = \Psi(x) \quad (A45)$$

holds for all $\theta>0$ real and $0\leq x\leq 1$, where $\Psi(x)$ is a positive-valued function of the variable x. Equation (A45) can be extended also to $x\geq 1$ via

$$\Psi(x) = \frac{\Lambda\left(b - \frac{\theta}{x}x\right)}{\Lambda\left(b-\frac{\theta}{x}\right)} = \frac{1}{\frac{\Lambda\left(b-\theta\frac{1}{x}\right)}{\Lambda(b-\theta)}} = \frac{1}{\Psi\left(\frac{1}{x}\right)}. \quad (A46)$$

And, Eq. (A45) implies that

$$\Psi(xy) = \frac{\Lambda(b-\theta xy)}{\Lambda(b-\theta)} = \frac{\Lambda[b-(\theta y)x]}{\Lambda b-(\theta y)}\frac{\Lambda[b-(\theta y)]}{\Lambda(b-\theta)} = \Psi(x)\Psi(y) \quad (A47)$$

holds for all $0\leq x,y\leq 1$.

Equations (A46) and (A47) imply that the function $\Psi(x)$ is a power law: $\Psi(x)=x^p(x>0)$. Since the tail intensity $\Lambda(l)(l<b)$ is monotone decreasing, we conclude that

$$\Psi(x) = x^\varepsilon \quad (A48)$$

($x>0$), where ε is a positive exponent. Substituting Eq. (A48) back into Eq. (A45)—while setting $\theta=1$ and $x=b-l$—yields

$$\Lambda(l) = c(b-l)^\varepsilon \quad (A49)$$

($l<b$), where $c=\Lambda(b-1)$.

[1] J. F. C. Kingman, *Poisson Processes* (Oxford University Press, Oxford, 1993).
[2] P. Lévy, *Théorie de l'Addition des Variables Aleatoires* (Gauthier-Villars, Paris, 1954).
[3] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, London, 1954).
[4] I. A. Ibragimov and Yu. V. Linnik, *Independent and Stationary Sequences of Random Variables* (Walterss-Noordhoff, Groningen, The Netherlands, 1971).
[5] B. Gnedenko, Ann. Math. **44**, 423 (1943); Translated and reprinted in *Breakthroughs in Statistics I*, edited by S. Kotz and N. L. Johnson (Springer, New York, 1992), pp. 195–225.
[6] E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
[7] J. Galambos, Asymptotic Theory of Extreme Order Statistics, 2nd ed. (Krieger, New York, 1987).
[8] I. Eliazar and J. Klafter, Proc. Natl. Acad. Sci. U.S.A. **102**, 13779 (2005).
[9] I. Eliazar and J. Klafter, Phys. Rev. E **75**, 031108 (2007).
[10] I. Eliazar and J. Klafter, Physica A **383**, 171 (2007).
[11] I. Eliazar, Physica A **386**, 318 (2007).
[12] I. Eliazar and J. Klafter, Phys. Rev. E **79**, 011103 (2009).
[13] I. Eliazar and J. Klafter, Proc. Natl. Acad. Sci. U.S.A. **106**, 12251 (2009).
[14] I. Eliazar and J. Klafter, Phys. Rev. Lett. **103**, 040602 (2009).
[15] E. H. Simpson, Nature (London) **163**, 688 (1949).
[16] R. G. Bell, P. Dean, and D. C. Hibbins-Butler, J. Phys. C **3**, 2111 (1970); R. J. Bell and P. Dean, Philos. Mag. **25**, 1381 (1972).
[17] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
[18] C. Gini, Econ. J. **31**, 124 (1921).
[19] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[20] R. G. Abraham, S. van den Bergh, and P. Nair, Astrophys. J. **588**, 218 (2003).

[21] P. P. Graczyk, J. Med. Chem. **50**, 5773 (2007).
[22] L. Wittebolle *et al.*, Nature (London) **458**, 623 (2009); S. Naeem, *ibid.* **458**, 579 (2009).
[23] P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modeling Extremal Events for Insurance and Finance* (Springer, New York, 1997).
[24] M. Thomas and R. D. Reiss, *Statistical Analysis of Extreme Values* (Birkhauser, Boston, 2001).
[25] R. W. Wolff, *Stochastic Modeling and the Theory of Queues* (Prentice-Hall, London, 1989).
[26] B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).
[27] A. Rényi, *On measures of entropy and information*, in Proceeding of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1960 (University of California Press, Berkeley, 1961), Vol. 1, p. 547.
[28] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation* (Cambridge University Press, Cambridge, 1987).