Journal of Physics Communications

PAPER • **OPEN ACCESS**

Mathematical diversity of parts for a continuous distribution

To cite this article: R Rajaram *et al* 2024 *J. Phys. Commun.* **8** 025008

IOP Publishing J. Phys. Commun. 8 (2024) 025008 https://doi.org/10.1088/2399-6528/ad2560

Journal of Physics Communications

RECEIVED
14 November 2023

REVISED
23 January 2024

ACCEPTED FOR PUBLICATION
2 February 2024

PUBLISHED
15 February 2024

PAPER

Mathematical diversity of parts for a continuous distribution

R Rajaram[1], N Ritchey[1] and B Castellani[2]

[1] Department of Math. Sci., Kent State University, United States of America
[2] Durham Research Methods Centre, Durham University, Durham, United Kingdom

E-mail: rrajaram@kent.edu

Keywords: Ecological diversity, Shannon entropy, Hill numbers

Abstract

The current paper is part of a series exploring how to link diversity measures (e.g., Gini-Simpson index, Shannon entropy, Hill numbers) to a distribution's original shape and to compare parts of a distribution, in terms of diversity, with the whole. This linkage is crucial to understanding the exact relationship between the density of an original probability distribution, denoted by $p(x)$, and the diversity D in non-uniform distributions, both within parts of a distribution and the whole. Empirically, our results are an important advance since we can compare various parts of a distribution, noting that systems found in contemporary data often have unequal distributions that possess multiple diversity types and have unknown and changing frequencies at different scales (e.g. income, economic complexity ratings, rankings, etc.). To date, we have proven our results for discrete distributions. Our focus here is continuous distributions. In both instances, we do so by linking case-based entropy, a diversity approach we developed, to a probability distribution's shape for continuous distributions. This allows us to demonstrate that the original probability distribution g_1, the case-based entropy curve g_2, and the slope of diversity g_3 ($c_{(a,x)}$ versus the $c_{(a,x)}{}^{*}\ln A_{(a,x)}$ curve) are one-to-one (or injective). Put simply, a change in the probability distribution, g_1, leads to variations in the curves for g_2 and g_3. Consequently, any alteration in the permutation of the initial probability distribution, which results in a different form, will distinctly define the graphs g_2 and g_3. By demonstrating the injective property of our method for continuous distributions, we introduce a unique technique to gauge the level of uniformity as indicated by D/c. Furthermore, we present a distinct method to calculate D/c for different forms of the original continuous distribution, enabling comparison of various distributions and their components.

1. Introduction

As we have explained elsewhere (Rajaram and Castellani 2020, Rajaram *et al* 2023), probability distributions are often the first quantitative characteristics of many systems and datasets, which, as Sornette and others have articulated (Newman 2010, Sornette 2009), makes them useful ways to explore diversity, as measurements on a wide range of systems and datasets are well approximated by their shape, particularly as the sample size increases. Given their value, we have developed a program of research exploring diversity within probability distributions. Specifically, we have sought new ways to link diversity measures (e.g., Gini-Simpson index, Shannon entropy, Hill numbers) to a distribution's original shape and to compare parts of a distribution, in terms of diversity, with the whole (Rajaram and Castellani 2020, Rajaram *et al* 2023). As we have shown across this research, this linkage is crucial to understanding the exact relationship between the density of an original probability distribution, denoted by $p(x)$, and the diversity D in non-uniform distributions, both within parts of a distribution and the whole—something the current field has yet to sufficiently address (Chao and Jost 2015, Hsieh *et al* 2016, Jost 2006, 2018, Leinster and Cobbold 2012, Pavoine *et al* 2016). This linkage is also empirically useful across the natural and social sciences, given that, in terms of probably distributions, most real-world systems have unequal distributions and consist of multiple diversity types with unknown and changing frequencies at different levels of scale (e.g., income diversity, economic complexity indices, rankings). As part of our program of research, we have proven our results for discrete distributions. Our focus for this paper is continuous distributions. In both instances, our strategy for establishing our diversity linkage is our engagement with the literature on Hill numbers (Jost 2018, Gaggiotti *et al* 2018, Jost 2006, macArthur 1965, Hill 1973, Peet 1974).

1.1. Research strategy

As we have explained in a series of papers, Hill numbers are defined by a parameter q that gives preference to types with either lower or higher frequencies (Rajaram and Castellani 2020, Rajaram *et al* 2023). This depends on whether $0 < q < 1$ or $q > 1$, respectively. Choosing $q = 1$ means that each type is assigned a weight proportion to its relative frequency by ${}^1D$. We also have that ${}^1D = e^H$, where H is the Shannon entropy of the distribution (Leinster 2021). In terms of advancing our understanding of diversity within distributions, Hill numbers hold a special place because they provide an all-encompassing structure to seize the various dimensions of diversity (MacArthur 1965, Hill 1973, Peet 1974), which they do by unifying the principles of richness, evenness, and dominance into a single numeric index. In doing so, Hill numbers enable the assessment and classification of diverse systems across the natural and social sciences, including diversity in ecosystems, where they are most widely used (Alberdi and Gilbert 2019, Gaggiotti *et al* 2018).

Still, the limitation of Hill numbers is that the precise relationship between the probability of each type within a distribution and the Hill number itself remains undeveloped. Moreover, the original concept of diversity, as proposed by Hill and Jost, is actually insensitive to permutations. This means a shuffling of the original probabilities in g_1 will not change the diversity of the entire distribution.

Hence the purpose of our program of research. In (Rajaram and Castellani 2016) we introduced our new measure, case-based entropy Cca modification of the ShannonWiener entropy measure H. As a next step, in (Rajaram and Castellani 2020) we proved a result relating the probability of each type p_i and the total diversity ${}^1D_K$ for a discrete probability distribution with K types. In a more recent paper (Rajaram *et al* 2023) we extended the results by explicitly proving a one-to-one relationship between the original probability distribution g_1, the case-based entropy curve g_2 and the slope of diversity curve g_3. We also showed that the ratio of diversity of a part to its cumulative probability denoted by $\frac{{}^1D_P}{c_P}$ is a measure of the degree of uniformity of the part P. Lastly, we also showed that the original probability distribution can be explicitly reconstructed by looking at the slopes of secants in the slope of diversity curve g_3.

In the current paper, we will show that analogous results hold true for continuous distributions with finite entropy (differential entropy to be more exact). We will show that the case-based entropy curve g_2 and the $c_{(a,x)}$ versus the $c_{(a,x)}{*}\ln A_{(a,x)}$ curve g_3, which we call the *slope of diversity* are one-to-one (or injective), i.e., a different probability distribution g_1 gives a different curve for g_2 and g_3. This means that the graphs g_2 and g_3 are determined uniquely by the original probability distribution. A proof of the injectivity will establish the uniqueness of the degree of uniformity of parts as measured by D_P/c_P. It will also create a unique way to compute D_P/c_P for arbitrary probability distributions. We also show that the original density $p(x)$ can be reconstructed by looking at the slope of tangents in the slope of diversity curve. We note once again, that analogous results have been proven for discrete distributions in (Rajaram *et al* 2023). Hence, this paper is an extension of those results for continuous distributions which have not been proven before.

We consider a general continuous probability distribution with finite entropy with a random variable X with support (a,b) (with $a = -\infty$ and $b = +\infty$ allowed) and probability density given by $p(x)$. We ask the following question: Is it possible to determine a connection (direct or indirect) between the probability density $p(x)$ and the case-based entropy curve (C_c versus c)? More to the point, does a connection exist between the shape of the case-based entropy curve (C_c versus c) and the probability density $p(x)$? How can we use the slope of diversity curve g_3 ($c_{(a,x)}$ versus $c_{(a,x)} \cdot \ln(A_{(a,x)})$ to compute the degree of uniformity of a given part P and furthermore, how can we reconstruct the original probability distribution g_1 from the slope of diversity curve g_3?

2. Understanding diversity

As a measure, diversity is used to evaluate the *richness* and *evenness* of diversity in probability distributions (Jost 2006, MacArthur 1965, Hill 1973, Peet 1974). Richness refers to the quantity of types in a distribution; evenness refers to the equal likelihood of each type of diversity occurring, as highlighted in various studies. As we have explained elsewhere (Rajaram and Castellani 2020, Rajaram *et al* 2023), this concept of diversity is rooted in the understanding that if all the K types in a discrete probability distribution have the same probability of occurrence, then the diversity should be equivalent to the number of types K. On the other hand, any departure from uniformity in probabilities will invariably lead to a decrease in the value of diversity.

Definition 2.1. (Shannon Diversity corresponding to $q = 1$ for Hill numbers) Given a continuous random variable X with support (a,b) (with $a = -\infty$ and $b = +\infty$ allowed) and its probability density $p(x)$, the diversity of the entire distribution ${}^1D_{(a,b)}$ is defined as the length of the support of an equivalent uniform distribution that yields the same value of Shannon entropy H.

Differential Shannon entropy for continuous distributions with a density $p(x)$ is defined as below:

$$H_{(a,b)} = -\int_{(a,b)} p(x)\ln(p(x))dx. \quad (1)$$

Remark 2.1. To avoid mathematical pathologies, we will only consider probability densities $p(x)$ that have a finite value for the Shannon entropy $H_{(a,b)}$As previously demonstrated by others (Jost 2006, MacArthur 1965, Hill 1973, Peet 1974) that definition 2.1 suggests that the total diversity ${}^1D_{(a,b)}$ is given by:

$${}^1D_{(a,b)} = e^H. \quad (2)$$

2.1. An example of biodiversity

In (Jost 2006) a comparison of species of butterfly in two communities was carried out to illustrate the purpose of using diversity instead of entropy to study the similarities in the communities. A case was made that the Hill number ${}^1D$ is a better index of diversity than Shannon entropy. Data from the canopy and understory communities of fruit-feeding butterflies was used to illustrate the point of the multiplication principle. Instead of repeating the same example, let us consider two communities of birds. Let us assume that the first community has 8 species of birds and each species has 50 birds, and the second community has 10 species of birds each of which has 50 birds as well. Let us assume furthermore, that the species in the two communities are distinct. The diversity of the first community is intuitively 8 and that of the second community if 10. When we pool the two communities, the diversity of the pooled community should be 18 since we will then have 18 distinct species that are uniformly distributed. This is exactly what happens if we use the diversity ${}^1D$ instead of Shannon entropy if the original distributions are not uniform. Then ${}^1D$ will still be the right diversity index to use, where now each species will be counted in a manner proportional to the relative abundance in the pooled community. We extend this notion in this paper by proving results for general continuous distributions where different parts are being pooled with different relative abundances. We also definitively show that the notion of diversity ${}^1D$ for continuous distributions and its corresponding case-based diversity and slope of diversity curves are one-to-one, and the slope of diversity curve can be used to measure the degree of uniformity of a continuous distribution. This establishes for the first time, important results for continuous distributions that need a separate consideration due to the intricacies involved in proving results using the probability density.

In this paper, we have four objectives:

1. Just like we showed in (Rajaram *et al* 2023) for discrete distributions, we show a similar way to compute the ratio $\frac{D_P}{c_P}$ for arbitrary parts P from the graph of the slope of diversity curve ($c_{(a,x)}$ versus $c_{(a,x)}{}^*\ln A_{(a,x)}$ or g_3) for continuous distributions. This will be an important step towards calculating the extent of uniformity of parts of a continuous distribution.
2. We prove that the slope of the secant $S_{(x_1,x_2)}$ of the slope of diversity curve can be used to compute the degree of uniformity of an arbitrary part $P = (x_1, x_2)$ of the original continuous distribution denoted by $\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}$.
3. We show that the original continuous distribution g_1 can be reconstructed using the slope of the tangent of the slope of diversity curve g_3.
4. Finally, we show that the natural map between the original continuous distribution g_1, the case-based entropy curve g_2, and the slope of diversity curve g_3 is one-to-one or injective, thereby establishing that two different original distributions g_1 will always lead to different curves g_2 and g_3. This will bridge the gap in connecting the Hill numbers to the form of the original continuous distribution.

In essence, this paper is an extension of (Rajaram *et al* 2023) for continuous distributions.

The paper is organized as follows: In section 3 we prove the results in the first two objectives. In section 4, we prove the third objective above. In section 5, prove the fourth objective. In section 6 we demonstrate our results for the example of the continuous exponential distribution. In section 7, we will end the paper with some observations on our findings.

Table 1. General dataset with complexity types x_i each having a probability p_i and a frequency f_i.

X	P	F
x_1	p_1	f_1
x_2	p_2	f_2
x_3	p_3	f_3
$\vdots$	$\vdots$	$\vdots$
x_J	p_J	f_J
$\vdots$	$\vdots$	$\vdots$
x_K	p_K	f_K

3. Computing $\frac{D_P}{c_P}$ for parts P of a continuous distribution

We begin by recalling two 'parts-to-whole' formulae for discrete distributions, which we proved in (Rajaram and Castellani 2020).

Theorem 3.1. *Given a discrete probability distribution similar to table 1, the diversity of the distribution ${}^qD_K$ for a system or dataset (be it complex or otherwise), and the diversities of its disjoint parts ${}^qD_{P_i}$ and their respective cumulative probabilities c_{P_i} are associated as follows:*

$$ {}^1D_K = \prod_{P_i \in \mathcal{P}} \left(\frac{{}^1D_{P_i}}{c_{P_i}} \right)^{c_{P_i}}, \tag{3} $$

and

$$ {}^qD_K = \left(\sum_{P_i \in P} c_{P_i} \left(\frac{{}^qD_{P_i}}{c_{P_i}} \right)^{(1-q)} \right)^{\frac{1}{1-q}}. \tag{4} $$

We note that equations (3) and (4) are simply the weighted geometric and arithmetic means (of order $1-q$) respectively of the ratio $\left(\frac{{}^qD_{P_i}}{c_{P_i}}\right)$. We also note that ${}^1D_K = \lim_{q \to 1}^q D_K$. The following corollary can be easily proved using the same technique as in the proof of theorem 3.1 in (Rajaram and Castellani 2020).

Corollary 3.1. *Given a discrete probability distribution similar to table 1, let the part $P = \bigcup_i P_i$ be a disjoint union of sub-parts* $\mathrm{P_i}$. *Then, the diversity of the part ${}^qD_P$ and the diversities of disjoint sub-parts ${}^qD_{P_i}$ and their respective cumulative probabilities c_{P_i} are related as follows:*

$$ \left(\frac{{}^1D_P}{c_P} \right)^{c_P} = \prod_{P_i \in P} \left(\frac{{}^1D_{P_i}}{c_{P_i}} \right)^{c_{P_i}}, \tag{5} $$

and

$$ c_P \left(\frac{{}^qD_P}{c_P} \right)^{1-q} = \sum_{P_i \in P} c_{P_i} \left(\frac{{}^qD_{P_i}}{c_{P_i}} \right)^{(1-q)}. \tag{6} $$

Remark 3.1. We remark that in general, there is no monotonic relationship between the diversity of continuous and discrete distributions. For example, we could consider the uniform distribution in the discrete case wherein $p_i = \frac{1}{N}$ for $i=1,\ldots,N$ and its counterpart in the continuous case where $p(x) = \frac{1}{b-a}$ on the interval (a,b). The diversity of the discrete uniform distribution is N and that of the continuous one is simply $b-a$. One can adjust N or $(b - a)$ to make the diversity of the discrete uniform distribution to be equal to, less than or larger than the diversity of the continuous uniform distribution. In general, due to the wide variation in shapes of distributions, theres no universal comparison that can be made between all continuous and all discrete distributions. However, given that the development of continuous distributions requires a separate mathematical treatment due to the intricacies involved in using a probability density, the proofs of the results are different and need to be written separately. For example, to reconstruct the original probability density from the *slope of diversity* curve in theorem 4.1, we have to use the slope of the tangent instead of the secant. Hence, the material in this paper for continuous distributions requires a separate consideration from discrete distributions.

We now state and prove the main theorem for continuous distributions. This is the first time that it has been proven for continuous distributions. We note that we will only consider the case $q = 1$ and hence omit the left superscript in ${}^1D$ and simply denote the diversity by D from now on.

Theorem 3.2. *Let $p(x)$ be a probability density function (pdf) on (a, b) with finite entropy and with $a = -\infty$ and $b = +\infty$ permitted. Let $P = \bigcup_i P_i$ be a disjoint partition of a part $P \subset (a, b)$. Then the following is true:*

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \prod_{P_i \in P} \left(\frac{D_{P_i}}{c_{P_i}}\right)^{c_{P_i}}. \tag{7}$$

PROOF. Recall the following definitions.

$$\text{ShannonEntropy:} \qquad H = -\int_{(a,b)} p(x)\ln(x)dx$$
$$\text{Diversity:} \qquad D = e^H$$

Let $c_P = \int_P p(x)dx$ be the probability mass function of P. Note that c_{P_i} is defined in a similar manner. Also, $p_{P_i}(x) = p_{P_i}(x)/c_{P_i}$ is the normalized probability density for the part P_i (same definition for P).Then we have the following:

$$\begin{aligned} D_P &= \exp\left\{-\int_P p_P(x)\ln(p_P(x))dx\right\} \\ &= \exp\left\{-\int_P \left(\frac{p(x)}{c_P}\right)[\ln(p(x)) - \ln(c_P)]dx\right\} \\ &= \exp\left\{-\frac{1}{c_P}\int_P p(x)\ln(p(x))dx + \ln(c_P)\overbrace{\int_P p_P(x)dx}^{=1}\right\} \\ &= c_P \exp\left\{-\frac{1}{c_P}\int_P p(x)\ln(p(x))dx\right\}. \end{aligned}$$

Hence, we have $\frac{D_P}{c_P} = \exp\left\{-\frac{1}{c_P}\int_P p(x)\ln(p(x))dx\right\}$. Thus,

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \exp\left\{-\int_P p(x)\ln(p(x))dx\right\}. \tag{8}$$

Note that following the same steps for the part P_i we have

$$\left(\frac{D_{P_i}}{c_{P_i}}\right)^{c_{P_i}} = \exp\left\{-\int_{P_i} p(x)\ln(p(x))dx\right\}. \tag{9}$$

Using information from the two equations above, and recalling that $\bigcup_i P_i$ is a disjoint partition

$$\begin{aligned} \left(\frac{D_P}{c_P}\right)^{c_P} &= \exp\left\{-\int_P p(x)\ln(p(x))dx\right\} \\ &= \exp\left\{-\int_{\bigcup_i P_i} p(x)\ln(p(x))dx\right\} \\ &= \exp\left\{-\sum_i \int_{P_i} p(x)\ln(p(x))dx\right\} \\ &= \prod_i \exp\left\{-\int_{P_i} p(x)\ln(p(x))dx\right\} \\ &= \prod_i \left(\frac{D_{P_i}}{c_{P_i}}\right)^{c_{P_i}}. \end{aligned}$$

This proves the Theorem. □

We make some definitions to establish some notation to prove our next theorem.

Definition 3.1. We define

$$A_P = \frac{D_P}{c_P \cdot D_{(a,b)}} \quad \text{and} \quad A_{P_i} = \frac{D_{P_i}}{c_{P_i} \cdot D_{(a,b)}} \tag{10}$$

to be the average case-based entropy per unit cumulative frequency for the part P and the sub-part P_i respectively.

Definition 3.2. Let $P = (a, x)$ be a part for a continuous probability distribution on (a, b), with $a = -\infty$ and $b = +\infty$ allowed. The graph of $c_{(a,x)}$ on the x-axis versus $c_{(a,x)}{*}\ln(A_{(a,x)})$ on the y-axis is defined as the slope of diversity curve. Also, the slope of the secant joining the points $(c_{(a,x_1)}, c_{(a,x_1)}{*}\ln(A_{(a,x_1)}))$ and $(c_{(a,x_2)}, c_{(a,x_2)}{*}\ln(A_{(a,x_2)}))$ on the slope of diversity curve is denoted by $S_{(x_1,x_2)}$.

We next define the degree of uniformity of a part $P = (x_1, x_2)$

Definition 3.3. Let $P = (x_1, x_2)$ be a part for a continuous probability distribution on (a, b), with $a = -\infty$ and $b = +\infty$ allowed. The ratio $\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}$ is termed as degree of uniformity of the part $P = (x_1, x_2)$.

Remark 3.2. In (Rajaram *et al* 2023), we have justified the usage of this terminology by showing the intuition behind how the ratio $\frac{D_P}{c_P}$ is an accurate measure of the degree of uniformity of the part P in a discrete distribution. The same intuition carries over for a continuous distribution with a part $P = (x_1, x_2)$.

Definition 3.4. Let $p(x)$ be a probability density function (pdf) on (a, b) with $a = -\infty$ and $b = +\infty$ permitted. Then the graph of $c_{(a,x)}$ on the x-axis versus $C_{(a,x)} = \frac{D_{(a,x)}}{D_{(a,b}}$ on the y-axis is called the case-based entropy curve. We denote these curves by g_2. The graph of $c_{(a,x)}$ versus $c_{(a,x)}\cdot \ln(A_{(a,x)})$ is called the slope of diversity curve. This is denoted by g_3.

We now state and prove a theorem that relates the slope of secant $S_{(x_1,x_2)}$ and the degree of uniformity.

Theorem 3.3. *Let $p(x)$ be a probability density function (pdf) on (a, b) with finite entropy and with $a = -\infty$ and $b = +\infty$ permitted. Let $P = \bigcup_i P_i$ be a disjoint partition of a part $P \subset (a, b)$. Then the following are true:*

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}\begin{pmatrix}<\\=\\>\end{pmatrix}\frac{D_{(x_3,x_4)}}{c_{(x_3,x_4)}}\Longleftrightarrow S_{(x_1,x_2)}\begin{pmatrix}<\\=\\>\end{pmatrix}S_{(x_3,x_4)}. \tag{11}$$

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}} = De^{S_{(x_1,x_2)}}. \tag{12}$$

Proof 3.1 Recall that

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \prod_i \left(\frac{D_{P_i}}{c_{P_i}}\right)^{c_{P_i}}, \quad \text{with} \quad c_P = \sum_i c_{P_i}. \tag{13}$$

Thus,

$$D^{c_P} = \prod_i D^{c_{P_i}}. \tag{14}$$

Dividing both sides of equation (13) by the corresponding sides of equation (14),

$$A_P^{c_P} = \prod_i (A_{P_i})^{c_{P_i}}.$$

Taking the natural logarithm of both sides, we get

$$c_P \ln(A_P) = \sum_i c_{P_i} \ln(A_{P_i}).$$

Let $P = (a, x_2)$; $P_1 = (a, x_1)$ and $P_2 = (x_1, x_2)$ with $a \leqslant x_1 \leqslant x_2$. Then we have,

$$c_{(a,x_2)}\ln(A_{(a,x_2)}) = c_{(a,x_1)}\ln(A_{(a,x_1)}) + c_{(x_1,x_2)}\ln(A_{(x_1,x_2)}).$$

Rearranging we get

$$\ln(A_{(x_1,x_2)}) = \frac{c_{(a,x_2)}\ln(A_{(a,x_2)}) - c_{(a,x_1)}\ln(A_{(a,x_1)})}{(c_{(a,x_2)} - c_{(a,x_1)})}.$$

Noticing that the right-hand-side of this equation is the slope of the secant line $S_{(x_1,x_2)}$ for the graph of $c_{(a,x)}$ versus $c_{(a,x)}\ln(A_{(a,x)})$ as defined in definition 3.2. By the same development as in the discrete case, let $S_{(x_1,x_2)}$ be the slope of the secant line joining the points $(c_{(a,x_1)}, c_{(a,x_1)}\ln(A_{(a,x_1)}))$ and $(c_{(a,x_2)}, c_{(a,x_2)}\ln(A_{(a,x_2)}))$. Then we have $\ln(A_{(x_1,x_2)}) = S_{(x_1,x_2)}$, or

$$A_{(x_1,x_2)} = e^{S_{(x_1,x_2)}},$$

or,

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}} = De^{S_{(x_1,x_2)}}.$$

Thus,

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}\begin{pmatrix} < \\ = \\ > \end{pmatrix}\frac{D_{(x_3,x_4)}}{c_{(x_3,x_4)}} \Longleftrightarrow S_{(x_1,x_2)}\begin{pmatrix} < \\ = \\ > \end{pmatrix}S_{(x_3,x_4)},$$

which is the continuous version of the discrete result. □

Remark 3.3. Theorem 3.1 relates the degree of diversity $\frac{D_P}{c_P}$ of a given part P of a continuous distribution as the weighted geometric mean of the degree of diversity of $\frac{D_{P_i}}{c_{P_i}}$ of its sub-parts P_i with the cumulative probabilities c_{P_i} as the weights. Theorem 3.3 means that when comparing the slopes of secants $S_{(x_1,x_2)}$ of the slope of diversity curve, we are also comparing the degrees of uniformity in the parts (x_1, x_2) and (x_3, x_4). It also means that we can compute the degree of uniformity $\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}$ of an arbitrary part $P = (x_1, x_2)$ directly from the slope of secant $S_{(x_1,x_2)}$ of the slope of diversity curve. This is the main importance of the two results in this section.

4. Reconstruction of the original probability distribution using the slope of tangent s_x of the slope of diversity curve g_3

So far, all of the results so far from the discrete case have carried over. In this section, we show that the slope of the tangent in the slope of diversity curve allow us to reconstruct the original density $p(x)$. We note that every point on the slope of the diversity curve is of the form $(c_{(a,x)}, c_{(a,x)}\ln(A_{(a,x)}))$.

Definition 4.1. Given the slope of diversity curve, we define s_x as the slope of the tangent of this curve at a general point given by $(c_{(a,x)}, c_{(a,x)}\ln(A_{(a,x)}))$.

Theorem 4.1. *Let $p(x)$ be a probability density function (pdf) on (a, b) with finite entropy and with $a = -\infty$ and $b = +\infty$ permitted. Let s_x represent the slope of the tangent at a general point on the slope of diversity curve denoted by $(c_{(a,x)}, c_{(a,x)}\ln(A_{(a,x)}))$. Then the following is true:*

$$p(x) = \frac{e^{-s_x}}{D_{(a,b)}}. \tag{15}$$

Proof 4.1 By the definition of $A_{(a,x)}$, and by taking the natural log of equation (21), we have the following:

$$\begin{aligned} A_{(a,x)} &= \frac{D_{(a,x)}}{c_{(a,x)}D_{(a,b)}}; \\ \Rightarrow c_{(a,x)}\ln(A_{(a,x)}) &= c_{(a,x)}\ln\Big(\frac{D_{(a,x)}}{c_{(a,x)}}\Big) - c_{(a,x)}\ln(D_{(a,b)}) \\ &= -\int_{(a,x)} p(t)\ln(p(t))dt - c_{(a,x)}\ln(D_{(a,b)}). \end{aligned}$$

Differentiate with respect to $c_{(a,x)}$. Recall that $c_{(a,x)}$ is the x-axis on the slope of diversity curve, using the Chain Rule, dividing, and employing the Fundamental theorem of Calculus, we have

$$\begin{aligned} s_x &= \frac{d}{d[c_{(a,x)}]}\{c_{(a,x)}\ln(A_{(a,x)})\} = \frac{\frac{-d}{dx}\left\{\int_{(a,x)} p(t)\ln(p(t))dt\right\}}{\frac{d}{dx}(c_{(a,x)})} - \ln(D) \\ &= -\frac{p(x)\ln(p(x))}{p(x)} - \ln(D) \\ &= -\ln(D \cdot p(x)). \end{aligned}$$

Thus,

$$D \cdot p(x) = e^{-s_x} \Rightarrow p(x) = \frac{e^{-s_x}}{D}.$$

□

This means that the slope of the tangent of the slope of diversity curve at $c_{(a,x)}$ explicitly determines the value of $p(x)$ at x. Figure 1 illustrates the last two theorems.

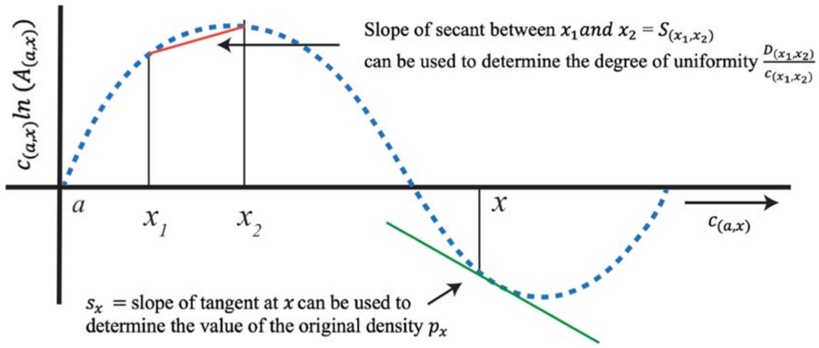

Figure 1. Slope of tangent of the slope of diversity curve.

Remark 4.1. We note that the result in theorem 4.1 is the continuous analog of a similar result that was proven in (Rajaram *et al* 2023) which states that the original discrete probability can be reconstructed using the slope of the secant of the slope of diversity curve. The secant in the discrete version became the tangent in the continuous version in theorem 4.1. Furthermore, theorem 4.1 explicitly relates the slope of tangent s_x of the slope of diversity curve g_3 back to the original continuous probability distribution g_1. This is the main importance of theorem 4.1.

Remark 4.2. In (Rajaram *et al* 2017) a direct comparison of the Boltzmann, Fermi and Bose–Einstein distributions was made using the case-based entropy idea. The celebrated Boltzmann distribution in one dimension is given as follows:

$$p_{B,1D}(E) = \left(\frac{1}{k_B T}\right) e^{-\frac{E}{k_B T}} = \frac{\beta}{e^{\beta E}} \tag{16}$$

where k_B is the Boltzmann constant and $\beta = \left(\frac{1}{k_B T}\right)$. We notice a striking similarity between equation (16) and equation (15). More specifically, $\lambda = \beta$, $x = E$ and hence, $s_x = \beta E - 1$. This shows an interesting relationship between the slope of tangent s_x and the energy E. Also, equation (15) resembles the general relationship between the Hamiltonian of the canonical ensemble and the probability of states, In general, for various choices of ensembles in statistical mechanics, it would be interesting to see if the slope of tangent s_x for the distribution of states can be related to the Hamiltonian. We will try and explore this in future papers.

5. Injectiveness of the graphs g_1, g_2, and g_3

In this section, we prove that there is a unique injective correspondence between the original density g_1, the case-based-entropy curve g_2 and the slope of diversity curve g_3. This means that the shape of original continuous distribution uniquely determines the shapes of both case-based entropy and slope of diversity curves.

Theorem 5.1. *Let $p(x)$ be a probability density function (pdf) on (a, b) with finite entropy and with $a = -\infty$ and $b = +\infty$ permitted. Also let $\mathcal{G}_1$ be the set of graphs of the original probability density $p(x)$, $\mathcal{G}_2$ be the set of graphs of the corresponding case-based entropy curves, and $\mathcal{G}_3$ be the set of graphs of the corresponding slope of diversity curves, with g_1, g_2 and g_3 denoting elements (graphs) in $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ respectively. In addition, let $T_{j\to k}$ be the map from the graph $\mathcal{G}_j$ to the graph $\mathcal{G}_k$ where $j, k = 1, 2, 3$. Then we have the following:*

$$T_{j\to k}\colon \mathcal{G}_j \xrightarrow{\sim} \mathcal{G}_k \tag{17}$$

is injective (or one-to-one).

Remark 5.1. We note that the map $T_{j\to k}\colon \mathcal{G}_j \xrightarrow{\sim} \mathcal{G}_k$ is taken to be the map between $g_j \in \mathcal{G}_j$ and $g_k \in \mathcal{G}_k$ with points taken as they appear from left to right.

PROOF.

1. $T_{1\to 2}$: Let $g_1^a, g_1^b \in \mathcal{G}_1$. It will be shown below that $T_{1\to 2}(g_1^a) = T_{1\to 2}(g_1^b)$ implies that $g_1^a = g_1^b$.

$$T_{1\to 2}(g_1^a) = T_{1\to 2}(g_1^b) \Rightarrow (c^a_{(a,x)}, C^a_{(a,x)}) = (c^b_{(a,x)}, C^b_{(a,x)})\ \forall x \in\ (a, b) \Rightarrow$$
$$c^a_{(a,x)} = c^b_{(a,x)}\ \forall x \in\ (a, b) \Rightarrow p^a(x) = p^b(x)\ \forall x \in (a, b).$$

$$\text{Hence, } T_{1\to 2}(g_1^a) = T_{1\to 2}(g_1^b) \Rightarrow g_1^a = g_1^b.$$

This shows that the map $T_{1\to 2}$ is injective.

2. Let $g_2^a, g_2^b \in \mathcal{G}_2$. It will be shown below that $T_{2\to 3}(g_2^a) = T_{2\to 3}(g_2^b)$ implies that $g_2^a = g_2^b$.

$$T_{2\to 3}(g_2^a) = T_{2\to 3}(g_2^b) \Rightarrow (c_{(a,x)}^a, c_{(a,x)}^a{}^*\ln(A_{(a,x)}^a)) = (c_{(a,x)}^b, c_{(a,x)}^b{}^*\ln(A_{(a,x)}^b))\ \forall x \in (a, b).$$

$$\Rightarrow c_{(a,x)}^a = c_{(a,x)}^b \text{ and } c_{(a,x)}^a{}^*\ln(A_{(a,x)}^a) = c_{(a,x)}^b{}^*\ln(A_{(a,x)}^b)\ \forall x \in (a, b)$$

$$\Rightarrow A_{(a,x)}^a = A_{(a,x)}^b \Rightarrow \frac{D_{(a,x)}^a}{D_{(a,b)}c_{(a,x)}^a} = \frac{D_{(a,x)}^b}{D_{(a,b)}c_{(a,x)}^b} \Rightarrow \frac{D_{(a,x)}^a}{D_{(a,b)}} = \frac{D_{(a,x)}^b}{D_{(a,b)}}\ \forall x \in (a, b)$$

$$C_{(a,x)}^a = C_{(a,x)}^b\ \forall x \in (a, b)$$

$$\text{Hence, } T_{2\to 3}(g_2^a) = T_{2\to 3}(g_2^b) \Rightarrow g_2^a = g_2^b.$$

This shows that the map $T_{2\to 3}$ is injective.

3. Let $g_3^a, g_3^b \in \mathcal{G}_3$. We will show below that $T_{3\to 1}(g_3^a) = T_{3\to 1}(g_3^b)$ implies that $g_3^a = g_3^b$.

$$T_{3\to 1}(g_3^a) = T_{3\to 1}(g_3^b) \Rightarrow p^a(x) = p^b(x)\ \forall \Rightarrow c_{(a,x)}^a = c_{(a,x)}^b\ \ \forall x \in (a, b).$$

$$\text{Also, } D_{(a,x)}^a = c_{(a,x)}^a \cdot \exp\left\{-\frac{1}{c_{(a,x)}^a}\int_{(a,x)} p^a(x)\ln(p^a(x))dx\right\}$$

$$= c_{(a,x)}^b \cdot exp\left\{-\frac{1}{c_{(a,x)}^b}\int_{(a,x)} p^b(x)\ln(p^b(x))dx\right\} = D_{(a,x)}^b\ \ \forall x \in (a, b).$$

$$\text{Hence, } A_{(a,x)}^a = \frac{D_{(a,x)}^a}{D_K c_{(a,x)}^a} = \frac{D_{(a,x)}^b}{D_K c_{(a,x)}^b} = A_{(a,x)}^b\ \ \forall x \in (a, b),$$

$$\text{and } (c_{(a,x)}^a, c_{(a,x)}^a{}^*\ln(A_{(a,x)}^a)) = (c_{(a,x)}^b, c_{(a,x)}^b{}^*\ln(A_{\{1,k_2\}}^b))\ \forall x \in (a, b).$$

$$\text{Thus, } T_{3\to 1}(g_3^a) = T_{3\to 1}(g_3^b) \Rightarrow g_3^a = g_3^b.$$

This shows that the map $T_{3\to 1}$ is injective.

Remark 5.2. Theorem 5.1 says that, just like in the discrete case, there is a one-to-one correspondence between the original density g_1, the case-based-entropy curve g_2 and the slope of diversity curve g_3. This means that the shape of g_1 uniquely determines the shapes of both g_2 and g_3 curves.

6. Examples

6.1. Exponential distribution

In this section, we compute the slope of diversity curve for the general exponential distribution and show that we can reconstruct the original distribution from the slope of diversity curve (and hence equivalently from the case-based entropy curve).

Suppose that $p(x) = \lambda e^{-\lambda x}; x \in (0, \infty)$.

We calculate entropy as follows:

Entropy:

$$\begin{aligned}
H_{(0,\infty)} &= -\int_{-\infty}^{\infty} p(x)\ln(p(x))dx \\
&= -\int_0^{\infty} \lambda e^{-\lambda x}(\ln(\lambda) - \lambda x)dx \\
&= -\lambda\ln(\lambda)\int_0^{\infty} e^{-\lambda x}dx + \lambda^2\int_0^{\infty} xe^-\lambda xdx \\
&= -\lambda\ln(\lambda)\left(\frac{e^{-\lambda x}}{\lambda}\right)|_{\infty}^{0} + \lambda^2\left\{\frac{-xe^{-\lambda x}}{\lambda}|_0^{\infty} + \frac{1}{\lambda}\int_0^{\infty} e^{-\lambda x}dx\right\} \\
&= -\ln(\lambda) + \lambda^2 \cdot \frac{1}{\lambda} \cdot \frac{e^{-\lambda x}}{\lambda}|_{\infty}^{0} \\
&= 1 - \ln(\lambda).
\end{aligned}$$

Since $H_{(0,\infty)} = 1 - \ln(\lambda)$, we know that

$$D_{(0,\infty)} = e^{H_{(0,\infty)}} = e^{(1-\ln(\lambda))} = \frac{e}{\lambda}.$$

For $x \in (0, \infty)$, let's consider $(0, x)$ as an interval part. Then

$$p_{(0,x)} = \frac{p(t)}{c_{(0,x)}}, \qquad \text{for } t \in (0, x) \qquad \text{and} \qquad c_{(0,x)} = \int_0^x \lambda e^{-\lambda t)dt} = 1 - \mathrm{e}^{-\lambda x}.$$

In other words,

$$p_{(0,x)} = \frac{\lambda e^{-\lambda t}}{(1 - \mathrm{e}^{-\lambda x)}}, \text{ for } t \in (0, x);$$

$$H_{(0,x)} = -\int_0^x p_{(0,t)} \ln(p_{(0,t)})dt = -\int_0^x \frac{\lambda e^{-\lambda t}}{(1 - e^{-\lambda t})} \ln\left(\frac{\lambda e^{-\lambda t}}{(1 - e^{-\lambda t})}\right)dt.$$

Now,

$$D_{(0,x)} = \exp\left\{-\frac{1}{(1 - e^{-\lambda x})} \int_{(0,x)} \lambda e^{-\lambda t} \ln(\lambda e^{-\lambda t)}dt\right\} \cdot (1 - e^{-\lambda x}).$$

Evaluating the inside integral by parts, we have

$$-\int_0^x \lambda e^{-\lambda t} \ln(\lambda e^{-\lambda t})dt = -\ln(\lambda)(1 - \mathrm{e}^{-\lambda x}) - \lambda x e^{-\lambda x} + (1 - e^{-\lambda x})$$
$$=(1 - e^{-\lambda x)}(1 - \ln(\lambda)) - \lambda x e^{-\lambda x}.$$

So,

$$D_{(0,x)} = \exp\left\{\frac{(1 - e^{-\lambda x})(1 - \ln \lambda) - \lambda x e^{-\lambda x}}{(1 - e^{-\lambda x})}\right\} \cdot \underbrace{(1 - e^{-\lambda x})}_{c_{(0,x)}}.$$

Dividing,

$$\frac{D_{(0,x)}}{c_{(0,x)}} = \overbrace{\exp\{1 - \ln(\lambda)\}}^{D_{(0,\infty)}} \cdot \exp\left\{-\frac{\lambda x e^{-\lambda x}}{\underbrace{(1 - e^{-\lambda x})}_{c_{(0,x)}}}\right\}.$$

Thus,

$$\left(\underbrace{\frac{D_{(0,x)}}{c_{(0,x)}} \frac{1}{D_{(0,\infty)}}}_{A_{(0,x)}}\right)^{c_{(0,x)}} = \exp\{-\lambda x e^{-\lambda x}\}.$$

Taking the logarithm of both sides

$$c_{(0,x)} \ln(A_{(0,x)}) = -\lambda x e^{-\lambda x},$$

where $c_{(0,x)}=(1 - e^{-\lambda x})$ for $x \in (0, \infty)$.

Note that

$$s_x = \frac{d}{dx}\{c_{(0,x)} \ln(A_{(0,x)})\} \Big/ \frac{d\{c_{(0,x)}\}}{dx},$$

and

$$\frac{d}{dx}\{c_{(0,x)} \ln(A_{(0,x)})\} = -\lambda e^{-\lambda x} + \lambda^2 x e^{-\lambda x}$$
$$=\lambda e^{-\lambda x}(\lambda x - 1),$$

and

$$\frac{d}{dx}(c_{(0,x)}) = \lambda e^{-\lambda x}.$$

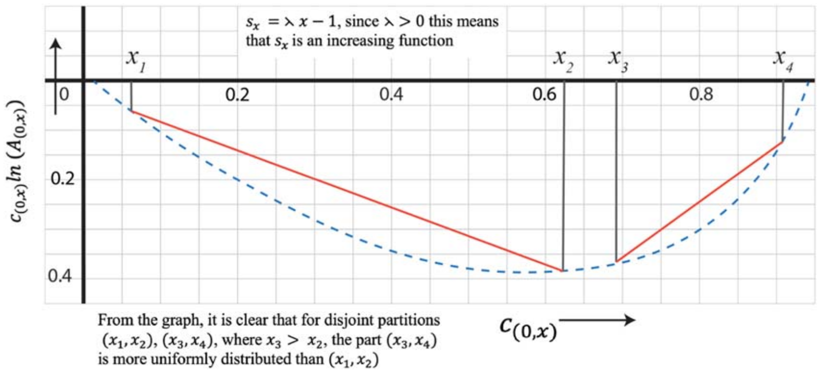

Figure 2. Slope of diversity curve for the exponential distribution.

So, $s_x = \lambda x - 1$, which implies that

$$p(x) = \frac{e^{-s_x}}{D_{(0,\infty)}} = \frac{e^{-\lambda x+1}}{\left(\frac{e}{\lambda}\right)} = \lambda e^{-\lambda x}.$$

In other words this example illustrates our theoretical work.

Also from figure 2, and since $s_x = \lambda x - 1$ is an increasing function of x since $\lambda > 0$, it is clear that for disjoint partitions (x_1, x_2) and (x_3, x_4) where $x_3 > x_2$, the part (x_3, x_4) is more uniformly distributed than (x_1, x_2).

6.2. Power law

In (Castellani and Rajaram 2016), an empirical comparison of diversity of power law distributions obtained from real data for various systems was done using case-based entropy. Here, we consider the power law from a theoretical standpoint. We consider the power law distribution as below:

Definition 6.1. A continuous random variable X is said to follow a power law distribution if its probability density function denoted by $p_z(x)$ satisfies the the following:

$$\begin{aligned} p_z(x) &= \frac{\alpha-1}{x_{\min}}\left(\frac{x}{x_{\min}}\right)^{-\alpha};\ \text{where } \alpha > 2;\ x \in (x_{\min}, \infty) \\ &= Cx^{-\alpha};\ \text{where } C = (\alpha-1)x_{\min}^{\alpha-1}. \end{aligned} \tag{18}$$

Theorem 6.1. *Given a power law distribution as in definition 6.1, its entropy is given by*

$$H = \ln\left(\frac{x_{\min}}{(\alpha-1)}\right) + \frac{\alpha}{(\alpha-1)}. \tag{19}$$

PROOF.

$$\ln(p_z(x)) = \ln(C) - \alpha\ln(x)$$

$$\begin{aligned} H &= -\int_{x_{\min}}^{\infty}(Cx^{-\alpha})\{\ln(C) - \alpha\ln(x)\}dx \\ &= -\ln(C)\int_{x_{\min}}^{\infty}\overbrace{Cx^{-\alpha}dx}^{p_z(x)} + \alpha C\int_{x_{\min}}^{\infty}x^{-\alpha}\ln(x)dx \\ &= -\ln(C) + \alpha C\left\{\frac{x^{-\alpha+1}}{\alpha+1}\cdot\ln(x) - \frac{x^{-\alpha+1}}{(-\alpha+1)^2}\right\}_{\min}^{\infty} \\ &= -\ln(C) + \alpha C\left\{\frac{x_{\min}^{-\alpha+1}}{(-\alpha+1)^2} - \frac{x_{\min}^{-\alpha+1}}{-\alpha+1}\cdot\ln(x_{\min})\right\} \end{aligned}$$

According to how we have defined C we have

$$\ln(C) = \ln(\alpha-1) + (\alpha-1)\ln(x_{\min})$$

Using this fact, we can rewrite H as

$$\begin{aligned}H &= -\ln(C) + \alpha C\cdot\frac{x_{\min}^{-\alpha+1}}{-\alpha+1}\left\{\frac{1}{(-\alpha+1)} - \ln(x_{\min})\right\}\\ &= \ln(\alpha-1) - (\alpha-1)\ln(x_{\min}) + \frac{\alpha}{(\alpha-1)} + \alpha\ln(x_{\min})\\ &= \ln(\alpha-1) + \ln(x_{\min}) + \frac{\alpha}{(\alpha-1)}\\ &= \ln\left(\frac{x_{\min}}{(\alpha-1)}\right) + \frac{\alpha}{(\alpha-1)}.\end{aligned}$$

This proves the Theorem. □

We denote the total diversity of the power law distribution $D_{(x_{\min},\infty)}$ by the symbol D just for simplicity. We can easily show that

$$\begin{aligned}c_{(x_{\min},x)} &= C\int_{x_{\min}}^{x} t^{-\alpha}dt\\ &=1-\left(\frac{x}{x_{\min}}\right)^{(1-\alpha)}.\end{aligned}$$

6.3. Slope of diversity curve for a power law distribution

In this section, we calculate an explicit formula for the slope of diversity curve of the power law distribution.

Theorem 6.2. *Given a power law distribution as in definition 6.1, the slope of diversity curve which plots $c_{(x_{\min},x)}$ on the x-axis and $c_{(x_{\min},x)}{}^{*}\ln(A_{(x_{\min},x)})$ on the y-axis has the following explicit formula:*

$$c_{(x_{\min},x)}\ln(A_{(x_{\min},x)}) = \frac{-\alpha}{(-\alpha+1)}(1-c_{(x_{\min},x)})\ln(1-c_{(x_{\min},x)}). \tag{20}$$

Also, the slope of the tangent s_x of the slope of diversity curve at $c_{(x_{\min},x)}$ is given by:

$$\begin{aligned}s_x &= \frac{\alpha}{\alpha-1)}\{-\ln(1-c_{(x_{\min},x)})-1\}\\ &=\alpha\ln\left(\frac{x}{x_{\min}}\right) - \frac{\alpha}{(\alpha-1)},\end{aligned}$$

Proof 6.2 From our previous paper we have

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \exp\left\{-\int_P p_z(x)\ln(p_z(x))dx\right\}. \tag{21}$$

Choosing $P = (x_{\min}, x)$ above, and remembering that $A_{(x_{\min}}, x) = \frac{D_{(x_{\min},x)}}{c_{(x_{\min},x)} * D}$ we get:

$$c_{(x_{\min},x)}\ln(A_{(x_{\min}}, x)) = -\int_{(x_{\min,x})} p(t)\ln(p(t))dt - c_{(x_{\min},x)}\cdot\underbrace{\ln(D)}_{H},$$

with

$$\begin{aligned}&-\int_{x_{\min}}^{x} p(t)\ln(p(t))dt\\ &= -\int_{x_{\min}}^{x}(Ct^{-\alpha})(\ln(C)-\alpha\ln(t)dt\\ &= -\ln(C)\underbrace{\int_{x_{\min}}^{x}(Ct^{-\alpha})dt}_{c_{(x_{\min}}} + \alpha C\int_{x_{\min}}^{x} t^{-\alpha}\ln(t)dt\\ &= -\ln(C)\cdot c_{(x_{\min}} + \alpha C\left\{\frac{t^{(-\alpha+1}\ln(t)}{(-\alpha+1)} - \frac{t^{-\alpha+1}}{(-\alpha+1)^2}\right\}_{x_{\min}}^{x}.\end{aligned}$$

Therefore,

$$
\begin{aligned}
&c_{(x_{\min},x)}\ln(A_{(\min},x))\\
&= -c_{(x_{\min},x)}(\ln(C)+H)+\alpha C\left\{\frac{t^{(-\alpha+1}\ln(t)}{(-\alpha+1)}-\frac{t^{-\alpha+1}}{(-\alpha+1)^2}\right\}_{x_{\min}}^{x},
\end{aligned}
$$

where

$$
\begin{aligned}
&\ln(C)+H=\ln(\alpha-1)+(\alpha-1)\ln(x_{\min})+\ln(x_{\min})-\ln(\alpha-1)+\frac{\alpha}{(\alpha-1)}\\
&=\alpha\ln(x_{\min})+\frac{\alpha}{(\alpha-1)}.
\end{aligned}
$$

Now,

$$
\begin{aligned}
&\alpha C\left\{\frac{t^{(-\alpha+1)}\ln(t)}{(-\alpha+1)}-\frac{t^{-\alpha+1}}{(-\alpha+1)^2}\right\}_{x_{\min}}^{x}\\
&=\frac{\alpha C}{(-\alpha+1)}\left\{t^{(-\alpha+1)}\ln(t)-\frac{t^{(-\alpha+1)}}{(-\alpha+1)}\right\}_{x_{\min}}^{x}\\
&=\frac{\alpha}{(-\alpha+1)}\cdot(\alpha-1)x_{\min}^{(-\alpha+1)}\cdot\left\{\left(x^{(-\alpha+1)}\ln(x)-\frac{x^{(-\alpha+1)}}{(-\alpha+1)}\right)\right.\\
&\left.-\left((-\alpha+1)\ln(x)-\frac{x^{(-\alpha+1}}{(-\alpha+1)}\right)\right\}\\
&=-\alpha\left\{\left[\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\ln(x)-\frac{1}{(-\alpha+1)}\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\right]\right\}\\
&-\left[\ln(x_{\min})-\frac{1}{(-\alpha+1)}\right]\Bigg\}\\
&=\left(\frac{\alpha}{\alpha-1}\right)+\alpha\ln(x_{\min})-\alpha\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\ln(x)-\frac{\alpha}{(\alpha-1)}\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\\
&=\left(\frac{\alpha}{(\alpha-1)}\right)\left[\underbrace{1-\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}}_{c_{(x_{\min},x)}}\right]+\alpha\ln(x_{\min}-\alpha\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\ln(x)
\end{aligned}
$$

So,

$$
\begin{aligned}
&c_{(x_{\min},x)}\ln(A_{(x_{\min},x)})=-c_{(x_{\min},x)}\left[\alpha\ln(x_{\min})+\frac{\alpha}{(\alpha-1)}\right]\\
&+\frac{\alpha}{(\alpha-1}c_{(x_{\min},x)}+\alpha\ln(x_{\min})-\alpha\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\ln(x)\\
&=\alpha\left(\frac{x_{\min}}{x}\right)^{(-\alpha+1)}\ln(\frac{x_{\min}}{x})\\
&=-\alpha\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\ln(\frac{x}{x_{\min}})
\end{aligned}
$$

Recall that

$$
c_{(x_{\min},x)}=1-\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}.
$$

So,

$$s_x = -\alpha \frac{\frac{d}{dx}\left[\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} \ln\left(\frac{x}{x_{\min}}\right)\right]}{\underbrace{\frac{d}{dx}\left[1 - \left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}\right]}_{=p_z(x)}}$$

$$= -\alpha \frac{\frac{d}{dx}\left[\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} \ln\left(\frac{x}{x_{\min}}\right)\right]}{(\alpha - 1)x_{\min}^{(\alpha-1)}x^{-\alpha}}$$

We now exploit the derivative.

$$\frac{d}{dx}\left[\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} \ln\left(\frac{x}{x_{\min}}\right)\right]$$

$$= \ln\left(\frac{x}{x_{\min}}\right)\left((-\alpha + 1)\left(\frac{x}{x_{\min}}\right)^{(-\alpha)} \cdot \frac{1}{x_{\min}}\right) + \left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} \cdot \frac{1}{x}$$

$$= (-\alpha + 1)\left(\frac{x^{-\alpha}}{x_{\min}^{(\alpha-1)}}\right)\ln\left(\frac{x}{x_{\min}}\right) + \frac{x^{-\alpha}}{x_{\min}^{(-\alpha+1)}}$$

Thus,

$$s_x = \alpha \ln\left(\frac{x}{x_{\min}}\right) - \left(\frac{\alpha}{\alpha - 1)}\right)$$

We can also examine this by recalling that

$$c_{(x_{\min},x)} = 1 - \left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)}.$$

$$\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} = 1 - c_{(x_{\min},x)}$$

or

$$\ln\left(\frac{x}{x_{\min}}\right) = \frac{1}{-\alpha + 1)} \ln(1 - c_{(x_{\min},x)}.$$

Substituting this into our expression we have

$$c_{(x_{\min},x)}\ln(A_{(\min}, x)) = -\alpha\left(\frac{x}{x_{\min}}\right)^{(-\alpha+1)} \ln\left(\frac{x}{x_{\min}}\right)$$

$$= \frac{-\alpha}{(-\alpha + 1)}(1 - c_{(x_{\min},x)})\ln(1 - c_{(x_{\min},x)})$$

Differentiating $c_{(x_{\min},x)}\ln(A_{(\min}, x))$ with respect to $c_{(x_{\min},x)})$, we have

$$s_x = \frac{\alpha}{\alpha - 1)}\{-\ln(1 - c_{(x_{\min},x)}) - 1\}$$

$$= \frac{\alpha}{\alpha - 1)}\left\{(\alpha - 1)\ln\left(\frac{x}{x_{\min}}\right) - 1\right\}$$

$$= \alpha \ln\left(\frac{x}{x_{\min}}\right) - \frac{\alpha}{(\alpha - 1)},$$

which proves the Theorem. □

Remark 6.1. Recall that

$$p_z(x) = Cx^{-\alpha}; \text{ where } C = (\alpha - 1)x_{\min}^{(\alpha-1)}$$

We know from our previous paper that s_x is given by:

$$s_x = -\ln(p_z(x) \cdot D) = -\ln(p_z(x) - \ln(D).$$

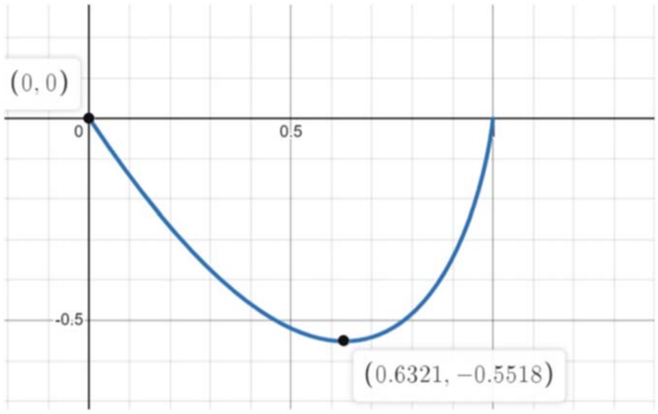

Figure 3. Slope of diversity for a power law distribution showing a minimum at $c_{(x_{\min},x)} = 0.632$.

But,

$$D = e^H,$$

so

$$-\ln(D) = -H.$$

Therefore,

$$\begin{aligned} s_x &= -\ln(p_z(x) - H \\ &= -\ln(\alpha - 1) - (\alpha - 1)\ln(x_{\min}) + \alpha\ln(x) - \ln(x_{\min}) + \ln(\alpha - 1) - \frac{\alpha}{\alpha - 1} \\ &= \alpha\ln(\frac{x}{x_{\min}}) - \frac{\alpha}{\alpha - 1}. \end{aligned}$$

Hence, the slope of tangent s_x can also be back-calculated from the probability density $p_z(x)$. We note that the slope of tangent of the slope of diversity curve at $c_{(x_{\min},x)}$ s_x increases in a logarithmic fashion with respect to $c_{(x_{\min},x)}$.

Remark 6.2. Setting s_x=0 we can obtain a minimum. In fact, no matter what α is, -$\ln(1 - c_{(x_{\min},x)}) - 1 = 0$, when s_x=0. And this occurs when $c_{(x_{\min},x)} = 1 - e^{-1} \approx 0.632$, as shown in figure 3. Let $x = c_{(x_{\min},x)}$ and $y = c_{(x_{\min},x)}\ln(A_{(x_{\min},x)})$. Then

$$y = \frac{\alpha}{\alpha - 1}(1 - x)\ln(1 - x))$$

Differentiating and setting $y' = 0$ we again find that a minimum occurs at $x = 1 - e^{-1}$. But, at $x = 1 - e^{-1}$ we have

$$\begin{aligned} y &= \frac{\alpha}{(\alpha - 1)}e^{-1}\ln(e^{-1}) \\ &= -\frac{\alpha}{(\alpha - 1)}e^{-1} \end{aligned}$$

Thus, the slope of the secant line joining the points (0, 0) and the minimum point $(1 - e^{-1}, -\frac{\alpha}{(\alpha - 1)}e^{-1})$ on the slope of diversity curve for the power law distribution is $-\frac{\alpha e^{-1}}{(\alpha - 1)1 - e^{-1}}$.

7. Conclusion

Accurately quantifying the degree of uniformity of probability distributions or its parts is a fundamental idea that is important due to its potential applications in the realm of studying inequality of resources. While the the Hill numbers ${}^qD$ provide a good starting point of such a quantification, there exist several limitations. First, the Hill numbers are insensitive to permutations and hence give the same value for a rearrangement of the original distribution. This is problematic since the shape of the distribution provides a very important characteristic of the distribution, namely the regions of abundance or scarcity. Second, the Hill numbers in their traditional sense, do not lend easily to comparison of degree of uniformity (or inequality) of parts of a distribution.

In this paper, we have shown that mathematical diversity of a probability distribution is a tool that allows us to quantify the degree of uniformity of a distribution or its parts for continuous distributions. In theorem 3.1 we established an explicit relationship between the degree of uniformity of a partition $P = \bigcup_i P_i$ and its sub-parts P_i for continuous probability distributions. We also established an explicit way to compute the degree of uniformity of a given arbitrary part $P = (x_1, x_2)$ of a continuous distribution using the slope of secant $S_{(x_1,x_2)}$ of the slope of diversity curve g_3 in theorem 3.3. We were able to completely reconstruct the original probability distribution using the slope of tangent s_x of the slope of diversity curve g_3 in theorem 4.1. Finally, we were also able to show that there exists a one-to-one correspondence between the original continuous probability distribution g_1, the case-based entropy curve g_2 and the slope of diversity curve g_3 in theorem 5.1. These results are the continuous counterparts of the results proved in (Rajaram *et al* 2023).

The main application of our work is in identifying regions of a given probability distribution that have the same degree of uniformity (we call this Shannon Equivalent Equiprobable or SEE parts) in a large dataset, based on our idea of mathematical diversity derived from information theory (or Shannon entropy to be more specific). Once such regions are identified, this gives researchers a starting point to further investigate such sub-sections of the original data to identify internal mechanisms or principles that led to such an equal degree of uniformity. One could start by looking at a single variable (which perhaps is an important characteristic of the dataset), and after identifying the SEE parts, can delve into the distribution of other variables of such parts to meaningfully explain the SEE behavior. Conversely, given two or more parts, we can compute and compare the degrees of uniformity of the given parts and say which part is more or less uniformly distributed compared to the others.

Another application could be to derive a much better measure of equality (or inequality) or uniformity (or non-uniformity) of a part or an entire distribution. For example, in the case of an income distribution, the slope of diversity curve g_3 can be used (by simply drawing secants of equal slope) to identify SEE parts of the distribution. We can compare the slopes of secant of parts to identify and also quantify the degree of uniformity of distribution of wealth. This is much more information than the GINI coefficient which (a) is an overall number and (b) is insensitive to the shape of the distribution. So in a sense, our technique will potentially prove to be more useful to analyze and quantify inequality in probability distributions by not only characterizing such an inequality for entire distributions, but also systematically dividing the distribution into SEE parts that have the same degree of uniformity of income.

In terms of where our program of research goes next, our goal is to advance these results to investigate distributions such as the power law, which has been well known to model the tails of several distributions in reality. For that matter, creating a quick toolbox that will quickly draw the three curves g_1, g_2 and g_3, along with the ability to draw the secants for g_3 and automatically filter out the SEE parts of the original distributional data for further investigation will prove very useful. We will endeavor to create such a computational toolbox. Lastly, we intend to apply our work to income distributions specifically to show that we can identify SEE parts of the distribution that will systematically divide the original data into parts that are SEE equivalent (and not just study the rich and poor parts). We strongly believe that this will lead to better policy formulation for the betterment of society towards equity in distribution of resources by using our information theoretic approach of diversity.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

R Rajaram https://orcid.org/0000-0003-1680-6706

References

Alberdi A and Gilbert M T P 2019 A guide to the application of hill numbers to dna-based diversity analyses *Mol. Ecol. Resour.* **19** 804–17
Castellani B and Rajaram R 2016 Complex systems and the limiting law of restricted diversity *Complexity* **21** 99–112
Chao A and Jost L 2015 Estimating diversity and entropy profiles via discovery rates of new species *Methods in Ecology and Evolution* **6** 873–82
Gaggiotti O E, Chao A, Peres-Neto P, Chiu C H, Edwards C, Fortin M J and Selkoe K A 2018 Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales *Evolutionary Applications* **11** 1176–93
Hill M 1973 Diversity and evenness: a unifying notation and its consequences *Ecology* **54** 427–32
Hsieh T, Ma K and Chao A 2016 inext: an r package for rarefaction and extrapolation of species diversity (h ill numbers) *Methods in Ecology and Evolution* **7** 1451–6
Jost L 2006 Entropy and diversity *Oikos* **113** 363–75
Jost L 2018 What do we mean by diversity? the path towards quantification *Mètode Science Studies Journal-Annual Review* 9
Leinster T 2021 *Entropy and diversity: The axiomatic approach* (Oxford University Press.)
Leinster T and Cobbold C A 2012 Measuring diversity: the importance of species similarity *Ecology* **93** 477–89

MacArthur R 1965 Patterns of species diversity *Biological Review* **40** 510–33

Newman M 2010 *Networks: An introduction* (Oxford University Press)

Pavoine S, Marcon E and Ricotta C 2016 equivalent numbers for species, phylogenetic or functional diversity in a nested hierarchy of multiple scales *Methods in Ecology and Evolution* **7** 1152–63

Peet R 1974 The measurement of species diversity *Annual Review of Ecological Systems* **5** 285–307

Rajaram R and Castellani B 2016 An entropy based measure for comparing distributions of complexity *Physica A.* **453** 35–43

Rajaram R and Castellani B 2020 Diversity in complex systems: measuring parts of the distribution to the whole *J. Phys. Commun.* 4

Rajaram R, Castellani B and Wilson A 2017 Advancing shannon entropy for measuring diversity in systems *Complexity* 2017

Rajaram R, Ritchey N and Castellani B 2023 On the comparison of diversity of parts of a distribution *Journal of Physics Communications* **7** 075006

Sornette D 2009 Probability distributions in complex systems *Encyclopedia of Complexity and Systems Science* ed R Meyers (Springer) 7009–24