PAPER • **OPEN ACCESS**

On the degree of uniformity measure for probability distributions

To cite this article: R Rajaram *et al* 2024 *J. Phys. Commun.* **8** 115003

Journal of Physics Communications

CrossMark

OPEN ACCESS

RECEIVED
20 September 2024

REVISED
23 October 2024

ACCEPTED FOR PUBLICATION
5 November 2024

PUBLISHED
14 November 2024

PAPER

On the degree of uniformity measure for probability distributions

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: shannon entropy, degree of uniformity, degree of inequality

Abstract

A key challenge in studying probability distributions is quantifying the inherent inequality within them. Certain parts of the distribution have higher probabilities than others, and our goal is to measure this inequality using the concept of mathematical diversity, a novel approach to examining inequality. We introduce a new measure $m_D(P)$, called the *degree of uniformity* measure on a given probability space that generalizes the idea of the slope of secant of the *slope of diversity* curve. This measure generalizes the idea of degree of uniformity of a contiguous part ($P = \{k_1, k_2\}$ in the discrete case or $P = (a, b)$ in the continuous case) in a probability space related to a random variable X, to an arbitrary measurable part P. We also demonstrate the truly scale free and self-contained nature of the concept of *degree of uniformity* by relating the measure of two parts P_1 and P_2 from completely unrelated distributions with random variables X_1 and X_2 that have completely different scales of variation.

1. Introduction

A fundamental characteristic of probability distributions is the inequality that is inherent in the distribution. Some parts of the distribution are more likely than others. This inequality is visually apparent in the shape of the distribution as the probabilities increase or decrease from left to right. While there are several ways to quantify this inherent inequality in a distribution through its shape, such as statistical measures of center and spread etc., there is a need to quantify this inequality for the parts (measurable subsets) or the whole of the distribution, which stems from the inherent uncertainty (information) that is present in the part or whole. Such a quantification can lead to (a) comparing the degree of uniformity (inequality) of the parts or the whole, not just within the same distribution but across different distributions; and (b) a truly scale-free and self-contained way of describing the inequality of a part or whole based solely on its inherent uncertainty.

Distributions that are uniform to begin with have a mathematical diversity that is equal to the support of the distribution itself. For discrete uniform distributions that have a support of $\{1, \ldots K\}$, the diversity ${}^1D$ evaluates to K and for continuous uniform distributions on the interval (a,b) the diversity ${}^1D$ evaluates to $(b - a)$. Distributions that deviate from uniformity will have a diversity that is necessarily less than a uniform distribution on the same support. Furthermore, the diversity ${}^1D$ in these cases have the intuition that we can redraw the original non-uniform distribution into a Shannon Equivalent Equi-probable (SEE) uniform distribution whose support has a size equal to ${}^1D$. We note that this equivalence is abstract yet useful in terms of visualization of the idea of diversity. The idea of diversity can also be visualized for parts of a distribution, as seen in Rajaram *et al* (2023, 2024). In short, a part P that has a cumulative probability of c_P and a diversity of D_P, can be visualized as a SEE part that is uniform, has a support length of D_P and an equal probability of $\frac{c_P}{D_P}$.

We use the term *Shannon Equivalent* to emphasize the fact that the original part or whole of the distribution has the same conditional Shannon Entropy as the equivalent abstract uniform part of the whole that is being visualized. We have chosen to use the exponential of Shannon Entropy (also called *mathematical diversity*) as a way to quantify the inequality in distributions. The main reason is because amongst all the Hill numbers ${}^qD$ with q being a parameter, which are measures of diversity, the one corresponding to $q = 1$ weights both the richness and evenness equally. Also, it is well known that Shannon Entropy is a measure of probabilistic (or information theoretic) uncertainty in a distribution.

Instead of quantifying the inequality of the original part or whole, which may not be uniform, the abstract SEE equivalent allows us to compute and compare the degree of uniformity of the part or whole on a level playing field. This is because the redrawn SEE equivalents are (a) uniform distributions and (b) have the same entropic uncertainty as the original part or whole that we started with. It is much easier to compare and quantify the degree of uniformity of SEE parts (albeit abstract) due to the uniformity in the distribution. This was the main content of our exposition in Rajaram *et al* (2023, 2024).

We have developed a quantification called *degree of uniformity* or *inequality* in our previous papers (Rajaram *et al* 2023, 2024) for parts that are contiguous in nature i.e. a consecutive set of indices $\{k_1, k_2\}$ or a single sub-interval (a,b). Among other things, we introduced the *slope of diversity* curve whose slope of secant between the end points of the contiguous part is a direct way of measuring the degree of uniformity of the given part. The idea of degree of uniformity of parts that are contiguous got us thinking along the following lines:

(1) Is there a way to generalize the idea of the slope of secant of the slope of diversity curve to parts that are not contiguous but are arbitrary measurable subsets of the given probability space?

(2) Can this generalization be used to compare arbitrary measurable parts or whole of entirely different distributions corresponding to random variables X_1 and X_2 that have no relationship between each other, and thereby show the truly scale-free and self contained nature of the idea of degree of uniformity.

The main focus of the current paper is to answer the above two questions for discrete distributions. We introduce a new measure called the *degree of uniformity* measure $m_D(P)$ that generalizes the concept of the slope of secant of the slope of diversity curve (all defined later in the paper) for contiguous parts $\{k_1, k_2\}$ with consecutive indices. We also show how this newly constructed measure can be used to compute and compare the degree of uniformity of arbitrary measurable parts (not just contiguous ones).

The paper is organized as follows. In section 2, we recall the relevant definitions and results from Rajaram *et al* (2023, 2024) that we are trying to generalize in this paper. In section 3, we introduce the newly discovered *degree of uniformity* measure $m_D(P)$ and state and prove some properties that elucidate why this measure is a generalization of results in the previous section. In sections 4 and 5, we state and prove some further properties of $m_D(P)$ that illustrate the truly scale-free and self-contained nature of the degree of uniformity measure. In section 6, we show some computational examples to demonstrate how $m_D(P)$ can be used to compute the degree of uniformity of parts across distributions. We conclude the paper in section 7 with some insights into the results and future work, and a short discussion on potential applications in physics and engineering.

2. A formal introduction to diversity: background material

Mathematical diversity (denoted by ${}^1D_K$ or ${}^1D$) is a quantification of the interplay between the richness or number of categories in a distribution, and its evenness which denotes the equi-probable occurrence of each type, as studied in Jost (2006), macArthur (1965), Hill (1973), Peet (1974). A majority of probability distributions which are not uniform, can be redrawn as a Shannon Equivalent Equiprobable distribution which is uniform but has the same probabilistic uncertainty as the original distribution. Mathematical diversity is based on the idea that when all categories in a discrete probability distribution have an equal likelihood of occurring, the diversity equals the number of categories. In a continuous distribution, the diversity of a uniform distribution is determined by the Lebesgue measure of its support. Any deviation from uniform probabilities reduces diversity. For background on mathematical diversity, we refer to Chao and Jost (2015), Hsieh *et al* (2016), Jost (2006, 2018), Leinster and Cobbold (2012), Pavoine *et al* (2016). We recall a few definitions and theorems related to mathematical diversity from Rajaram *et al* (2023, 2024) that are pertinent to this paper. In what follows, we will use the convention of a subscript of K to denote a discrete distribution, $\{k_1, k_2\}$ to denote a discrete part, a subscript of (a,b) to denote a continuous distribution or similar intervals for its parts. When something is true for both continuous and discrete distributions, we will say so and omit the subscripts. We will use I to denote the entire support of a distribution for both discrete and continuous case i.e. $I = \{1,\ldots,K\}$ or $I = (a, b)$ whenever it is pertinent.

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

Shannon entropy for discrete and continuous distributions is defined as below:

$$H_I = -\sum_{i=1}^{K} p_i \ln(p_i); \quad H_I = -\int_{(a,b)} p(x) \ln(p(x)) dx. \tag{1}$$

It was shown (Jost 2006, MacArthur 1965, Hill 1973, Peet 1974) that definition 2.1 implies that the total diversity ${}^1D$ (for both continuous and discrete distributions) is given by:

$${}^1D_I = e^{H_I}. \tag{2}$$

We will only consider the case $q = 1$ for the Hill numbers and hence, we will omit the left superscript of 1 while referring to the diversity as D. The reason for this choice of $q = 1$ is because for this choice, both richness and evenness are equally weighted. We recall the diversity of parts theorem for discrete distributions below.

Theorem 2.1. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\mathrm{p(x)}$ *be a probability density function (pdf) on* $I = (a, b)$, *with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P = \bigcup_i P_i$ *be a disjoint partition of a part* $P \subseteq I$. *Then the following is true for both discrete and continuous distributions:*

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \prod_{P_i \in P} \left(\frac{D_{P_i}}{c_{P_i}}\right)^{c_{P_i}}. \tag{3}$$

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

Definition 2.2. We define

$$A_P = \frac{D_P}{c_P \cdot D_I} \text{ and } A_{P_i} = \frac{D_{P_i}}{c_{P_i} \cdot D_I} \tag{4}$$

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

We next define the degree of uniformity of a part $P = \{k_1, k_2\}$ or $P = (x_1, x_2)$.

Definition 2.3. Let P stand for a part of the form $\{k_1, k_2\}$ for a discrete probability distribution or (x_1, x_2) for a continuous distribution. The ratio $\frac{D_P}{c_P}$ is termed as degree of uniformity of the part P.

In Rajaram *et al* (2023), the validity of the ratio $\frac{D_P}{c_P}$ as a quantitative measure of the degree of uniformity of a part P of a discrete distribution was established. The slope of diversity curve was shown to be useful to compute and compare the degrees of uniformity of continguous parts of a distribution of the form $\{k_1, k_2\}$ or (x_1, x_2) in Rajaram *et al* (2023, 2024), by comparing the slopes of secants of the corresponding parts from this curve. We recall the version of that theorem for discrete distributions below. We first define the slope of secant of the slope of diversity curve, for a discrete distribution.

Definition 2.4. The graph of $c_{\{1,k\}}$ versus $c_{\{1,k\}} \cdot \ln(A_{\{1,k\}})$ in the discrete case or $c_{(a,x)}$ versus $c_{(a,x)} \cdot \ln(A_{(a,x)})$ is known as the slope of diversity curve. Given the slope of diversity curve, we define $S_{\{k_1,k_2\}}$ or $S_{(x_1,x_2)}$ as the slope of the secant of this curve between the two points given by

$$(c_{\{1,k_1\}}, c_{\{1,k_1\}} \ln(A_{\{1,k_1\}})) \quad \text{and} \quad (c_{\{1,k_2\}}, c_{\{1,k_2\}} \ln(A_{\{1,k_2\}}))$$

in the discrete case or

$$(c_{(a,x_1)}, c_{(a,x_1)} \ln(A_{a,x_1})) \quad \text{and} \quad (c_{(a,x_2)}, c_{(a,x_2)} \ln(A_{(a,x_2)}))$$

in the continuous case.

Theorem 2.2. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\{k_1, k_2\}$ *and* $\{k_3, k_4\}$ *be parts that are subsets of* I. *Then the following are true:*

$$\frac{D_{\{k_1+1,k_2\}}}{c_{\{k_1+1,k_2\}}} \begin{pmatrix} < \\ = \\ > \end{pmatrix} \frac{D_{\{k_3+1,k_4\}}}{c_{\{k_3+1,k_4\}}} \iff S_{\{k_1,k_2\}} \begin{pmatrix} < \\ = \\ > \end{pmatrix} S_{\{k_3,k_4\}} \tag{5}$$

$$\frac{D_{\{k_1+1,k_2\}}}{c_{\{k_1+1,k_2\}}} = De^{S_{\{k_1,k_2\}}}. \tag{6}$$

Let $\mathrm{p(x)}$ *be a probability density function (pdf) on* (a,b), *with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* (x_1, x_2) *and* (x_3, x_4) *be parts that are subsets of* (a,b). *Then the following are true:*

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}}\begin{pmatrix}<\\=\\>\end{pmatrix}\frac{D_{(x_3,x_4)}}{c_{(x_3,x_4)}} \Longleftrightarrow S_{(x_1,x_2)}\begin{pmatrix}<\\=\\>\end{pmatrix}S_{(x_3,x_4)}. \tag{7}$$

$$\frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}} = De^{S_{(x_1,x_2)}}. \tag{8}$$

Remark 2.1. Theorem 2.1 relates the degree of uniformity $\frac{D_P}{c_P}$ of a given part P of a discrete 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 2.2 means that when comparing the slopes of secants $S_{\{k_1,k_2\}}$ and $S_{\{k_3,k_4\}}$ of the slope of diversity curve, we are also comparing the degrees of uniformity in the parts $\{k_1+1, k_2\}$ and $\{k_3+1, k_4\}$. It also means that we can compute the degree of uniformity $\frac{D_{\{k_1+1,k_2\}}}{c_{\{k_1+1,k_2\}}}$ of an arbitrary part $P=\{k_1+1, k_2\}$ directly from the slope of secant $S_{\{k_1,k_2\}}$ of the slope of diversity curve. Similar statements are true for continuous distributions. This is the main importance of the two results in this section.

3. Degree of uniformity measure

We consider a discrete probability distribution on $I = \{1, 2, 3,\ldots,K\}$, where K can be infinite, and a continuous probability distribution on $I = (a, b)$ where $a = -\infty$ and $b = +\infty$ being permitted. We also assume that the entropy H and diversity D are finite. We recall that we showed the following for the discrete case in Rajaram *et al* (2023, 2024),

$$c_P \ln\left(\frac{D_P}{c_P}\right) = -\sum_{i\in P} p_i \ln(p_i).$$

We also recall from Rajaram *et al* (2023) that if $S_P = S_{\{k_1,k_2\}}$ denotes the slope of the secant line for the part denoted by $P = \{k_1, k_2\}$, then

$$\ln(A_P) = S_P,\ \ \text{or} A_P = e^{S_P} \text{ or } \frac{D_P}{c_P} = De^{S_P}.$$

We define the *degree of uniformity* measure as follows:

Definition 3.1. Let p_i with $i \in I = \{1,\ldots K\}$ be a discrete probability distribution, with $K = +\infty$ permitted. Alternatively, let $p(x)$ be a probability density function (pdf) on $I = (a, b)$, with $a = -\infty$ and $b = +\infty$ permitted. Let P be a general measurable subset (not necessarily contiguous) of I. We define a new signed point mass measure on I called the *degree of uniformity* or *degree of ineqality* measure for such a measurable subset $P \subseteq I$ (irrespective of whether it is from a discrete or continuous distribution) by the following:

$$m_D(P) = c_P \ln(A_P),\ \forall \text{ measurable } P \subseteq I. \tag{9}$$

The degree of uniformity measure is a signed measure as seen in definition 3.1. The sign of the measure indicates whether the given measurable subset P has a degree of uniformity that is less than, equal to or greater than D which is the degree of uniformity of the entire distribution. We state and prove a theorem below that shows this fact. We label equations with a (D) for discrete distributions and (C) for continuous distributions for results that are slightly different for the respective kinds.

Theorem 3.1. *Let* p_i *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $p(x)$ *be a probability density function (pdf) on* $I = (a, b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* P *be a general measurable subset (not necessarily contiguous) of* I*. Then the following are true:*

(1)

$$\text{(Disc.) } c_P \ln(A_P) = m_D(P) = -\sum_{i\in P} p_i \ln(p_i D) \text{ and hence } A_P = \exp\left\{\frac{m_D(P)}{c_P}\right\}, \tag{10}$$

$$\text{(Cont.)} c_P \ln(A_P) = m_D(P) = -\int_P p(x)\ln(p(x)D) \text{ and hence } A_P = \exp\left\{\frac{m_D(P)}{c_P}\right\}, \tag{11}$$

(2)

$$m_D(P)\begin{pmatrix} < \\ = \\ > \end{pmatrix}0 \Leftrightarrow A_P\begin{pmatrix} < \\ = \\ > \end{pmatrix} 1. \tag{12}$$

Proof.

(1) (Disc.) Applying theorem 2.1 to $P_i = \{i\}\ \forall\, i \in P$, we have

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \prod_{i\in P}\left(\frac{1}{p_i}\right)^{p_i}.$$

Taking logarithms, we have

$$c_P \ln\left(\frac{D_P}{c_P}\right) = -\sum_{i\in P} p_i \ln(p_i).$$

Hence,

$$\begin{aligned} c_P \ln\left(\underbrace{\frac{D_P}{c_P D}}_{A_P}\right) &= -\sum_{i\in P} p_i \ln(p_i) - c_P \ln(D) \\ &= -\sum_{i\in P} p_i \ln(p_i) - \sum_{i\in P} p_i \ln(D) \\ &= -\sum_{i\in P}(p_i \ln(p_i) + p_i \ln(D)) \\ \Rightarrow c_P \ln(A_P) &= \underbrace{-\sum_{i\in P} p_i \ln(p_i D)}_{\text{We are calling this } m_D(P)}. \end{aligned}$$

Hence,

$$c_P \ln(A_P) = m_D(P) = -\sum_{i\in P} p_i \ln(p_i D).$$

Finally, taking the exponential on both sides, we have

$$A_P = \exp\left\{\frac{m_D(P)}{c_P}\right\}.$$

This proves the discrete case of the first part. Next, we prove the continuous case. (Cont.) 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,

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

So,

$$c_P \ln\left(\underbrace{\frac{D_P}{c_P D}}_{A_P}\right) = -\int_P p(x)\ln(p(x))dx - c_P\ln(D)$$
$$= -\int_P p(x)\ln(p(x))dx - \left(\int_P p(x)dx\right)\ln(D)$$
$$= -\int_P (p(x)\ln(p(x)) + p(x)\ln(D))dx,$$
$$\Rightarrow c_P\ln(A_P) = \underbrace{-\int_P (p(x)\ln(p(x)) + p(x)\ln(D))dx.}_{\text{We are calling this } m_D(P)}$$

Then,

$$c_P\ln(A_P) = m_D(P) = -\int_P p(x)\ln(p(x)D)dx.$$

Finally, taking the exponential on both sides, we have

$$A_P = \exp\left\{\frac{m_D(P)}{c_P}\right\}.$$

This proves the continuous case of the first part.

(2) The proof of the second part is directly observed from

$$A_P = \exp\left\{\frac{m_D(P)}{c_P}\right\}.$$

This proves the Theorem.

We note that in theorem 3.1 part 2, the comparison of A_P to be less than, equal to or greater than 1 is equivalent to the statement that the degree of uniformity of P is less than, equal to or greater than D, the degree of uniformity of the entire distribution.

Our focus next, is to determine how to generalize the idea of comparing slopes of secants on the slope of diversity curve to glean out the comparisons of degree of uniformity for parts P that don't look like $\{k_1, k_2\}$ or (x_1, x_2). In other words, we want to generalize theorem 2.2 for parts P that are not a set of consecutive indices such as $\{k_1, k_2\}$, but general measurable subsets of I.

From theorem 3.1, we have:

$$\ln(A_P) = \frac{m_D(P)}{c_P}.$$

In the above equation P is any general discrete or continuous part (or event) not necessarily consisting of consecutive indices. For such arbitrary measurable subsets, the slope of secant as in theorem 2.2 will not make sense as we cannot draw a secant for such subsets.

We have the following:

$$A_P = e^{\frac{m_D(P)}{c_P}} \Rightarrow \frac{D_P}{c_P} = De^{\frac{m_D(P)}{c_P}}.$$

Comparing $P_{\{k_1,k_2\}}$ or $P_{(x_1,x_2)}$ from theorem 2.2 with

$$\frac{D_{\{k_1,k_2\}}}{c_{\{k_1,k_2\}}} = De^{S_{\{k_1-1,k_2\}}}, \quad \frac{D_{(x_1,x_2)}}{c_{(x_1,x_2)}} = De^{S_{(x_1,x_2)}}.$$

it is clear that the ratio $\frac{m_D(P)}{c_P}$ takes the role of $S_{\{k_1-1,k_2\}}$ or $S_{(x_1,x_2)}$, if P is any general measurable subset that does not look like a contiguous part. In fact, this is the meaning and use for the signed measure $m_D(P)$.

In figure 1 above, the union of red portions indicate a general measurable set that is not just a consecutive set of indices or a contiguous interval. For such subsets, we compute $\frac{m_D(P)}{c_P}$ instead of the slope of secant, where P is the part in red.

We state and prove a theorem below that shows that for arbitrary measurable subsets P, the ratio $\frac{m_D(P)}{c_P}$ in fact generalizes the slope of secant of the *slope of diversity* curve for parts $P = \{k_1, k_2\}$ or $P = (x_1, x_2)$ that are contiguous.

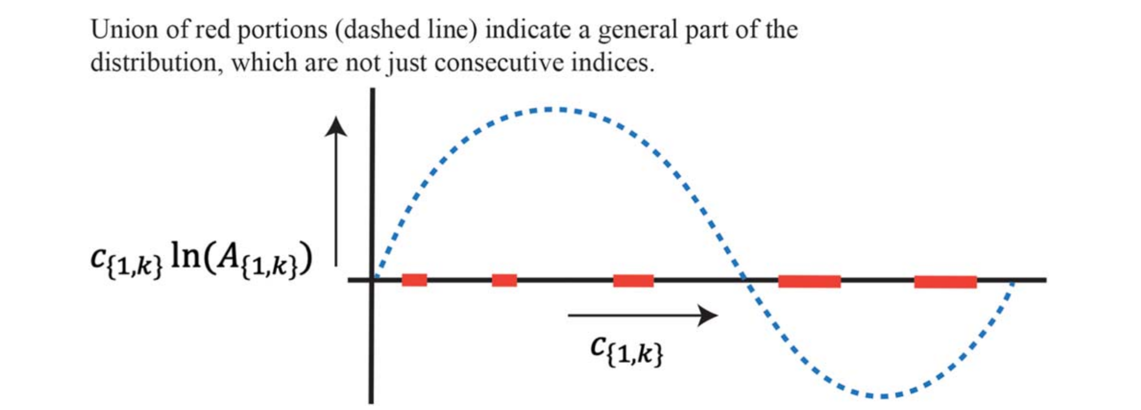

Figure 1. Picture showing an arbitrary measurable subset of the probability space.

Theorem 3.2. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $P = \{k_1,\ k_2\}$ *Then, we have the following:*

$$\frac{m_D(\{k_1,\ k_2\})}{c_{\{k_1,k_2\}}} = S_{\{k_1-1,k_2\}}. \tag{13}$$

Let $\mathrm{p}(\mathrm{x})$ *be a probability density function (pdf) on* $I = (a,\ b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P = (x_1,\ x_2)$*. Then, we have the following:*

$$\frac{m_D((x_1,\ x_2))}{c_{(x_1,x_2)}} = S_{(x_1,x_2)}. \tag{14}$$

Proof. (Disc.) Let $P = \{k_1,\ k_2\}$. Then from theorem 3.1, we have:

$$m_D(P) = c_P \ln(A_P) \Rightarrow \ln(A_P) = \frac{m_D(P)}{c_P}.$$

Also, by definition 3.1, $m_D(\{k_1,\ k_2\}) = -\sum_{i=k_1}^{k_2} p_i \ln(p_i D)$. Thus,

$$\begin{aligned} m_D(\{k_1,\ k_2\}) &= -\sum_{i=k_1}^{k_2} p_i \ln(p_i D) \\ &= \left(-\sum_{i=1}^{k_2} p_i \ln(p_i D)\right) - \left(-\sum_{i=1}^{k_1-1} p_i \ln(p_i D)\right) \\ &= m_D(\{1,\ k_2\}) - m_D(\{1,\ k_1-1\}) \\ &= c_{\{1,k_2\}} \ln(A_{\{1,k_2\}} - c_{\{1,k_1-1\}} \ln(A_{\{1,k_1-1\}}. \end{aligned}$$

Hence, we have:

$$\frac{m_D(\{k_1,\ k_2)\}}{c_{\{k_1,k_2\}}} = \frac{c_{\{1,k_2\}} \ln(A_{\{1,k_2\}} - c_{\{1,k_1-1\}} \ln(A_{\{1,k_1-1\}}}{\underbrace{c_{\{1,k_2\}} - c_{\{1,k_1-1\}}}_{c_{\{k_1,k_2\}}}}$$

$$\underbrace{=}_{\text{By Definition 2.4}} S_{\{k_1-1,k_2\}}.$$

This proves the discrete part of the Theorem. (Cont.) Let $P = (x_1,\ x_2)$. Then from theorem 3.1, we have:

$$m_D(P) = c_P \ln(A_P) \Rightarrow \ln(A_P) = \frac{m_D(P)}{c_P}.$$

Also, by definition 3.1, $m_D((x_1,\ x_2)) = -\int_{(x_1,x_2)} p(x)\ln(p(x)D)dx$. Thus,

$$\begin{aligned} m_D((x_1,\ x_2)) &= -\int_{(x_1,x_2)} p(x)\ln(p(x)D)dx \\ &= \left(-\int_{(a,x_2)} p(x)\ln(p(x)D)dx\right) - \left(-\int_{(a,x_1)} p(x)\ln(p(x)D)dx\right) \\ &= m_D((a,\ x_2)) - m_D((a,\ x_1)) \\ &= c_{(a,x_2)} \ln(A_{(a,x_2)} - c_{(a,x_1)} \ln(A_{(a,x_1)}. \end{aligned}$$

Hence, we have:

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

$$\underset{\text{By Definition 5}}{=} S_{(x_1,x_2)}.$$

This proves the continuous part of the Theorem.

Next, we state and prove a generalization of theorem 2.2 for general measurable sets P:

Theorem 3.3. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\mathrm{p(x)}$ *be a probability density function (pdf) on* $I = (a, b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P = (x_1, x_2)$*. Let* P_1 *and* P_2 *be arbitrary measurable subsets of* I*. Then, we have the following:*

$$\frac{D_{P_1}}{c_{P_1}}\begin{pmatrix} < \\ = \\ > \end{pmatrix}\frac{D_{P_2}}{c_{P_2}} \Leftrightarrow \frac{m_D(P_1)}{c_{P_1}}\begin{pmatrix} < \\ = \\ > \end{pmatrix}\frac{m_D(P_2)}{c_{P_2}}. \tag{15}$$

Proof. From theorem 3.1, we have:

$$\frac{D_P}{c_P} = De^{\frac{m_D(P)}{c_P}}.$$

Hence, we have the following:

$$\frac{D_{P_1}}{c_{P_1}}\begin{pmatrix} < \\ = \\ > \end{pmatrix}\frac{D_{P_2}}{c_{P_2}} \Leftrightarrow \frac{m_D(P_1)}{c_{P_1}}\begin{pmatrix} < \\ = \\ > \end{pmatrix}\frac{m_D(P_2)}{c_{P_2}}.$$

This proves the Theorem.

Hence, $m_D(P)$ allows us to compute and compare the degree of uniformity of parts P_1 and P_2 that don't look like a set of consecutive indices. This is a generalization of theorem 2.2.

Next, we prove a theorem that computes the degree of uniformity measure of a countable disjoint union of measurable sets.

Theorem 3.4. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\mathrm{p(x)}$ *be a probability density function (pdf) on* $I = (a, b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P = \bigcup_n P_n$*, where* $P_i \cap P_j = \phi$ *for* $i \neq j$ *be a countable disjoint union of measurable subsets of* I*. Then, we have the following:*

$$\frac{D_P}{c_P} = D\exp\left(\frac{\sum_i m_D(P_i)}{\sum_i c_{P_i}}\right). \tag{16}$$

Proof. Using the countable additivity of disjoint unions for measures, we can expand on $m_D(P)$. Let $P = \bigcup_n P_n$, where $P_i \cap P_j = \phi$ for $i \neq j$. Then,

$$m_D(\bigcup_n P_n) = \sum_n m_D(P_n)$$
$$c_P = \sum_n c_{P_n}$$
$$\ln(A_P) = \frac{m_D(P)}{c_P} = \frac{\sum_i m_D(P_i)}{\sum_i c_{P_i}}$$
$$A_P = \exp\left(\frac{\sum_i m_D(P_i)}{\sum_i c_{P_i}}\right).$$

The latter equation implies that

$$\frac{D_P}{c_P} = D\exp\left(\frac{\sum_i m_D(P_i)}{\sum_i c_{P_i}}\right).$$

This proves the Theorem.

Hence, we can compute the degree of uniformity of a part $P = \cup_n P_n$ which is a disjoint union of parts P_n using the above formula. All we need is the value of the signed measure $m_D(P_n)$ and the cumulative probability c_{P_n} for the parts P_n.

Corollary 3.1. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\mathrm{p}(\mathrm{x})$ *be a probability density function (pdf) on* $I = (a,\ b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P = \{k_1,\ k_2\} \cup \{k_3,\ k_4\}$ *with* $k_2 < k_3$ *or* $P = (x_1,\ x_2) \cup (x_3,\ x_4)$*. Then we have the following:*

$$\text{(Disc.) } m_D(P) = \frac{m_D(\{k_1,\ k_2\}) + m_D(\{k_3,\ k_4\})}{c_{\{k_1,k_2\}} + c_{\{k_3,k_4\}}} \tag{17}$$

$$= \frac{S_{\{k_1,k_2\}} + S_{\{k_3,k_4\}}}{c_{\{k_1,k_2\}} + c_{\{k_3,k_4\}}} \tag{18}$$

$$\text{(Cont.) } m_D(P) = \frac{m_D((x_1,\ x_2)) + m_D((x_3,\ x_4))}{c_{(x_1,x_2)} + c_{(x_3,x_4)}} \tag{19}$$

$$= \frac{S_{(x_1,x_2)} + S_{(x_3,x_4)}}{c_{(x_1,x_2)} + c_{(x_3,x_4)}}. \tag{20}$$

Proof. Follows directly from theorem 3.4.

4. Properties of $m_D(P)$

Let $I = \{1,\ldots,K\}$ in the discrete case with $K = +\infty$ allowed and $I = (a, b)$ with $a = -\infty$ and $b = +\infty$ allowed. Let P be an arbitrary measurable subset of I and let P_1 be an arbitrary measurable subset of P. We want to establish a relationship between the degree of uniformity measure of P_1 with respect to I and with respect to P, where $P_1 \subseteq P \subseteq I$. What is the relationship between $m_D(P_1)$ and $m_{\frac{D_P}{c_P}}(P_1)$?

The key thing to understand is that the degree of uniformity measure of P_1 with respect to I is in general different than the degree of uniformity measure of P_1 with respect to P. Hence, we want to quantify the relationship between $m_D(P_1)$ and $m_{\frac{D_P}{c_P}}(P_1)$.

Theorem 4.1. *Let* $\mathrm{p_i}$ *with* $i \in I = \{1,\ldots K\}$ *be a discrete probability distribution, with* $K = +\infty$ *permitted. Let* $\mathrm{p}(\mathrm{x})$ *be a probability density function (pdf) on* $I = (a,\ b)$*, with* $a = -\infty$ *and* $b = +\infty$ *permitted. Let* $P_1,\ P$ *be arbitrary measurable subsets with* $P_1 \subseteq P \subseteq I$*. Then we have the following:*

1.

$$m_{\frac{D_P}{c_P}}(P_1) = m_D(P_1) - \left(\frac{c_{P_1}}{c_P}\right) m_D(P) \tag{21}$$

2.

$$m_{\frac{D_P}{c_P}}(P_1) = c_{P_1} \ln\left(\frac{\left[\frac{D_{P_1}}{c_{P_1}}\right]}{\left[\frac{D_P}{c_P}\right]}\right). \tag{22}$$

Proof.

1. (Disc.) Starting from

$$\frac{D_P}{c_P} = D \exp\left\{\frac{m_D(P)}{c_P}\right\},$$

we have

$$p_i \frac{D_P}{c_P} = p_i D \exp\left\{\frac{m_D(P)}{c_P}\right\}$$
$$\ln\left(p_i \frac{D_P}{c_P}\right) = \ln(p_i D) + \frac{m_D(P)}{c_P}$$
$$p_i \ln\left(p_i \frac{D_P}{c_P}\right) = p_i \ln(p_i D) + p_i \frac{m_D(P)}{c_P}$$
$$-\sum_{i\in P_1} p_i \ln\left(p_i \frac{D_P}{c_P}\right) = -\sum_{i\in P_1} p_i \ln(p_i D) - \left(\sum_{i\in P_1} p_i\right)\frac{m_D(P)}{c_P}.$$

That is,

$$m_{\frac{D_P}{c_P}}(P_1) = m_D(P_1) - \left(\frac{c_{P_1}}{c_P}\right) m_D(P).$$

This proves the discrete case of part (1) of the Theorem. We next prove the continuous case next: (Cont.) Starting from

$$\frac{D_P}{c_P} = D \exp\left\{\frac{m_D(P)}{c_P}\right\},$$

we have

$$p(x)\frac{D_P}{c_P} = p(x) D \exp\left\{\frac{m_D(P)}{c_P}\right\}$$
$$\ln\left(p(x)\frac{D_P}{c_P}\right) = \ln(p(x)D) + \frac{m_D(P)}{c_P}$$
$$p(x)\ln\left(p(x)\frac{D_P}{c_P}\right) = p(x)\ln(p(x)D) + p(x)\frac{m_D(P)}{c_P}$$
$$-\int_{P_1} p(x)\ln\left(p(x)\frac{D_P}{c_P}\right)dx = -\int_{P_1} p(x)\ln(p(x)D) - \left(\int_{P_1} p(x)\right)\frac{m_D(P)}{c_P}.$$

That is,

$$m_{\frac{D_P}{c_P}}(P_1) = m_D(P_1) - \left(\frac{c_{P_1}}{c_P}\right) m_D(P).$$

This proves the continuous case of part (1).

2. We start by noting the following first:

$$\frac{D_P}{c_P} = D \exp\left(\frac{m_D(P)}{c_P}\right)$$

and

$$\left(\frac{D_P}{c_P}\right)^{c_P} \underset{\text{discrete}}{=} \prod_{i\in P}\left(\frac{1}{p_i}\right)^{p_i}, \quad \left(\frac{D_P}{c_P}\right)^{c_P} \underset{\text{continuous}}{=} \exp\left\{-\int_P p(x)\ln(p(x))dx\right\}$$

is true for P and P_1. We first prove the discrete case of part (2). (Disc.) Recall that

$$\left(\frac{D_P}{c_P}\right)^{c_P} = \prod_{i\in P}\left(\frac{1}{p_i}\right)^{p_i} \Rightarrow c_P \ln\left(\frac{D_P}{c_P}\right) = -\sum_{i\in P} p_i \ln(p_i)$$
$$m_D(P) = -\sum_{i\in P} p_i \ln(p_i D)$$
$$m_{\frac{D_P}{c_P}}(P_1) = -\sum_{i\in P_1} p_i \ln\left(p_i \frac{D_P}{c_P}\right).$$

From the right hand side of the last equation we have

$$
\begin{aligned}
m_{\frac{D_P}{c_P}}(P_1) &= -\sum_{i\in P_1} p_i \ln\left(p_i \frac{D_P}{c_P}\right)\\
&= -\sum_{i\in P_1} p_i \ln(p_i) - \left[\sum_{i\in P_1} p_i\right]\ln\left(\frac{D_P}{c_P}\right)\\
&= c_{P_1}\ln\left(\frac{D_{P_1}}{c_{P_1}}\right) - c_{P_1}\ln\left(\frac{D_P}{c_P}\right)\\
&= c_{P_1}\ln\left(\frac{\left[\frac{D_{P_1}}{c_{P_1}}\right]}{\left[\frac{D_P}{c_P}\right]}\right).
\end{aligned}
$$

This proves the discrete case of part (2) of the Theorem. We next prove the continuous case of part(2) of the Theorem. (Cont.) Recall that

$$
\begin{aligned}
&\left(\frac{D_P}{c_P}\right)^{c_P} = \exp\left\{-\int_P p(x)\ln(p(x))dx\right\} \Rightarrow c_P\ln\left(\frac{D_P}{c_P}\right) = -\int_P p(x)\ln(p(x))\\
&m_D(P) = -\int_P p(x)\ln(p(x)D)\\
&m_{\frac{D_P}{c_P}}(P_1) = -\int_{P_1} p(x)\ln\left(p(x)\frac{D_P}{c_P}\right).
\end{aligned}
$$

From the right hand side of the last equation we have

$$
\begin{aligned}
m_{\frac{D_P}{c_P}}(P_1) &= -\int_{P_1} p(x)\ln\left(p(x)\frac{D_P}{c_P}\right)\\
&= -\int_{P_1} p(x)\ln(p(x)) - \left[\int_{P_1} p(x)\right]\ln\left(\frac{D_P}{c_P}\right)\\
&= c_{P_1}\ln\left(\frac{D_{P_1}}{c_{P_1}}\right) - c_{P_1}\ln\left(\frac{D_P}{c_P}\right)\\
&= c_{P_1}\ln\left(\frac{\left[\frac{D_{P_1}}{c_{P_1}}\right]}{\left[\frac{D_P}{c_P}\right]}\right).
\end{aligned}
$$

This proves the continuous case of part (2).

Remark 4.1. If we choose $P = P_1$ in equation (22), we get

$$m_{\frac{D_P}{c_P}}(P) = 0 \text{ Special Case is} m_D(I) = 0. \tag{23}$$

This tells us that the part P has a degree of uniformity equal to $\frac{D_P}{c_P}$. However, equation (21) tells us how the degree of uniformity measure of P_1 with respect to P is related to degree of uniformity measure of P_1 with respect to I. Figure 2 shows the relationship between P_1, P and I where $P_1 \subseteq P \subseteq I$. It also shows the equation relating the degree of uniformity of the respective parts with the respective wholes, that are appropriately labeled.

We note that $P_1 \subseteq P \subseteq I$. Equation (21) says that the degree of uniformity measure of P_1 with respect to P given by $m_{\frac{D_P}{c_P}}(P_1)$ equals the degree of uniformity of P_1 with respect to I given by $m_D(P_1)$ minus a proportional amount of degree of uniformity of P with respect to I given by $m_D(P)$, with $\frac{c_{P_1}}{c_P}$ being the proportion.

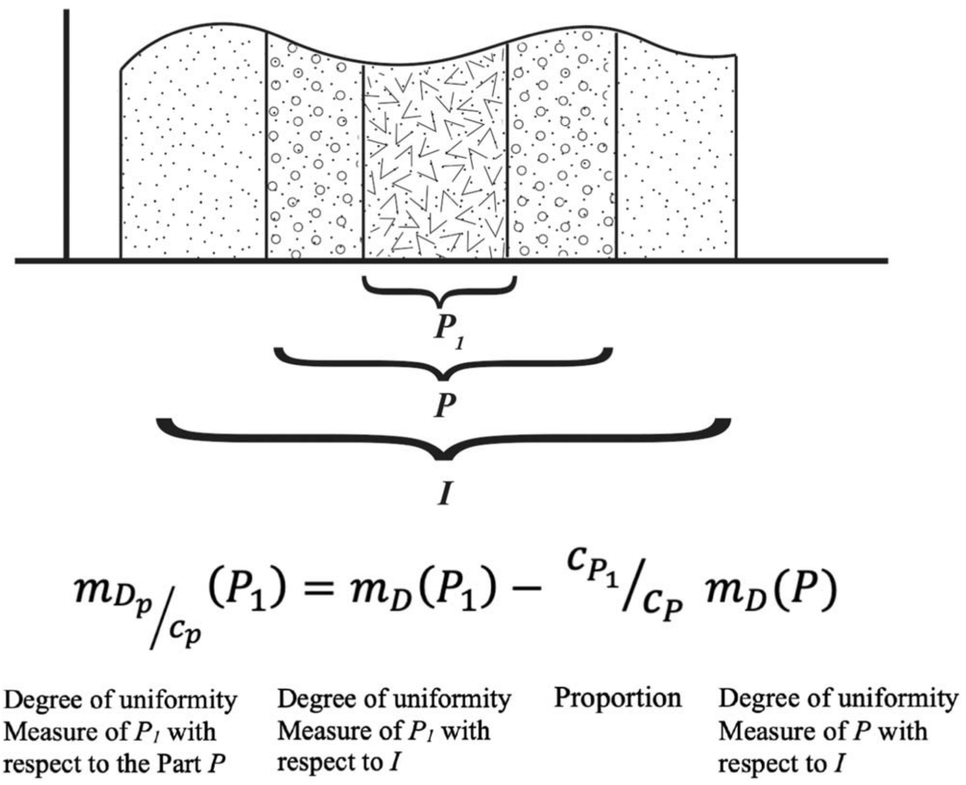

Figure 2. Picture showing an arbitrary measurable subparts $P_1 \subseteq P \subseteq I$.

Remark 4.2. We also have an alternate proof for equation (22) starting from equation (21),

$$\frac{m_{\frac{D_P}{c_P}}(P_1)}{c_{P_1}} = \frac{m_D(P_1)}{c_{P_1}} - \frac{m_D(P)}{c_P}$$

$$= \ln\left(\frac{D_{P_1}}{c_{P_1}D}\right) - \ln\left(\frac{D_P}{c_P D}\right)$$

$$= \ln\left(\frac{\left[\frac{D_{P_1}}{c_{P_1}}\right]}{\left[\frac{D_P}{c_P}\right]}\right).$$

Next, we prove a very important theorem that relates parts P_1 and P_2 from two different distributions as shown in figure 3. The main point in this theorem is that the parts P_1 and P_2 can come from two different random variables X_1 and X_2 that are completely unrelated to each other.

Theorem 4.2. *Let* $\mathrm{p_i}$ *with* $i \in I_1 = \{1,\ldots K_1\}$ *and* $\mathrm{q_i}$ *with* $i \in I_2 = \{1,\ldots,K_2\}$ *be two different discrete probability distributions corresponding to random variables* X_1 *and* X_2 *respectively, with* $K_1 = +\infty$ *and* $K_2 = +\infty$ *permitted. Similarly let* $\mathrm{p(x)}$ *and* $\mathrm{q(x)}$ *be two different continuous probability distributions on* $I_1 = (a,\ b)$ *and* $I_2 = (c,\ d)$ *corresponding to random variables* X_1 *and* X_2 *respectively, with* $a,\ c = -\infty$ *and* $b,\ d = +\infty$ *allowed. Let* P_1 *and* P_2 *be measurable subsets from the probability spaces of* X_1 *and* X_2 *respectively. Then the following is true:*

(1)

$$m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) = c_{P_2} \ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right) \tag{24}$$

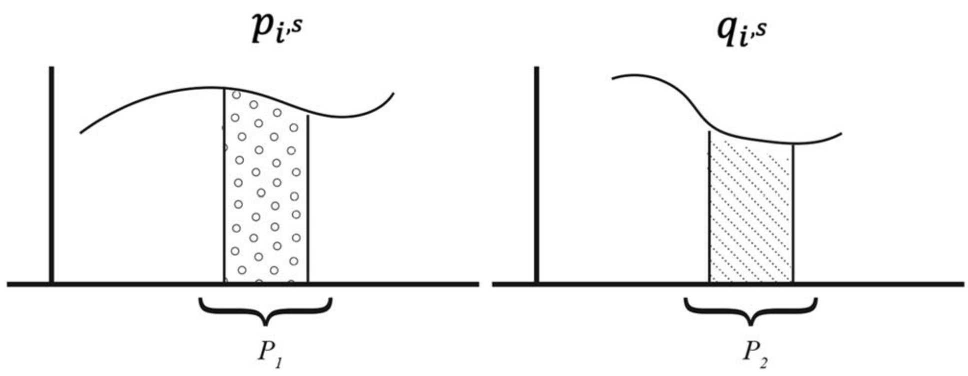

Figure 3. Picture showing an arbitrary measurable subparts P_1 and P_2 from two different distributions.

(2)

$$\frac{D_{P_2}}{c_{P_2}} = \left(\frac{D_{P_1}}{c_{P_1}}\right)\exp\left(\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}}\right) \tag{25}$$

(3)

$$m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)\begin{pmatrix} > \\ = \\ < \end{pmatrix}0 \Leftrightarrow \frac{D_{P_2}}{c_{P_2}}\begin{pmatrix} > \\ = \\ < \end{pmatrix}\frac{D_{P_1}}{c_{P_1}} \tag{26}$$

(4)

$$\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}} + \frac{m_{\frac{D_{P_2}}{c_{P_2}}}(P_1)}{c_{P_1}} = 0. \tag{27}$$

Proof.

(1) Remembering that P_2 is a measurable subset of the probability space of X_2, we have the following for the discrete case:

$$\begin{aligned}
\text{(Disc.)}\ \ m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) &= -\sum_{i\in P_2} q_i \ln\left(q_i\frac{D_{P_1}}{c_{P_1}}\right)\\
&= -\sum_{i\in P_2} q_i \ln(q_i) - \left(\sum_{i\in P_2} q_i\right)\ln\left(\frac{D_{P_1}}{c_{P_1}}\right)\\
&= c_{P_2}\ln\left(\frac{D_{P_2}}{c_{P_2}}\right) - c_{P_2}\ln\left(\frac{D_{P_1}}{c_{P_1}}\right)\\
&= c_{P_2}\ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right).
\end{aligned}$$

And the following for the continuous case:

$$
\begin{aligned}
\text{(Cont.) } m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) &= -\int_{P_2} q(x)\ln\left(q(x)\frac{D_{P_1}}{c_{P_1}}\right)dx \\
&= -\int_{P_2} q(x)\ln(q(x)) - \left(\int_{P_2} q(x)\right)\ln\left(\frac{D_{P_1}}{c_{P_1}}\right) \\
&= c_{P_2}\ln\left(\frac{D_{P_2}}{c_{P_2}}\right) - c_{P_2}\ln\left(\frac{D_{P_1}}{c_{P_1}}\right) \\
&= c_{P_2}\ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right)
\end{aligned}
$$

That is,

$$
m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) = c_{P_2}\ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right).
$$

This proves part (1) of the Theorem.

(2) proof of part (2) follows by exponentiation of both sides of equation (24).

$$
\begin{aligned}
m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) = c_{P_2}\ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right) &\Leftrightarrow \exp\left(\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}}\right) = \exp\left(\ln\left(\frac{\left[\frac{D_{P_2}}{c_{P_2}}\right]}{\left[\frac{D_{P_1}}{c_{P_1}}\right]}\right)\right) \\
&\Leftrightarrow \frac{D_{P_2}}{c_{P_2}} = \left(\frac{D_{P_1}}{c_{P_1}}\right)\exp\left(\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}}\right)
\end{aligned}
$$

(3) proof of part (3) follows by direct observation of equation (25).

(4) To prove part (4), we note the following:

$$
\begin{aligned}
m_{\frac{D_{P_1}}{c_{P_1}}}(P_2) &= c_{P_2}\ln\left(\frac{D_{P_2}/c_{P_2}}{D_{P_1}/c_{P_1}}\right) \\
m_{\frac{D_{P_2}}{c_{P_2}}}(P_1) &= c_{P_1}\ln\left(\frac{D_{P_1}/c_{P_1}}{D_{P_2}/c_{P_2}}\right) \\
\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}} + \frac{m_{\frac{D_{P_2}}{c_{P_2}}}(P_1)}{c_{P_1}} &= \ln\left(\frac{D_{P_2}/c_{P_2}}{D_{P_1}/c_{P_1}}\right) + \ln\left(\frac{D_{P_1}/c_{P_1}}{D_{P_2}/c_{P_2}}\right) \\
&= 0.
\end{aligned}
$$

That is,

$$
\frac{m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)}{c_{P_2}} + \frac{m_{\frac{D_{P_2}}{c_{P_2}}}(P_1)}{c_{P_1}} = 0.
$$

This ends the proof of all parts of the Theorem.

Again to emphasize, $P_1 \in X_1$ and $P_2 \in X_2$ where X_1 and X_2 are two completely different distributions with totally different scales. This shows that the degree of uniformity measure allows us to compare the degrees of uniformity of two completely unrelated parts in in a self contained manner.

We note that equation (22) in theorem 4.2 is a generalization of equation (22) in theorem 4.1. Indeed if we choose the part $P_1 = I_2$ (say), then $\frac{D_{P_1}}{c_{P_1}}$ is simply equal to D_2 (since the total probability is 1) and we recover equation (22) with the identification that $D_2 = D$.

We also note that equation (27) is indicative of the fact the ratio $\frac{m_D(P)}{c_P}$ is in the exponent of the formula for $\frac{D_P}{c_P}$ and hence $\frac{\frac{m_{D_{P_1}}(P_2)}{c_{P_1}}}{c_{P_2}}$ is of the opposite sign as $\frac{\frac{m_{D_{P_2}}(P_1)}{c_{P_2}}}{c_{P_1}}$. This is because if the degree of uniformity of P_1 is (say) x times the degree of uniformity of P_2, then that means P_2 has a degree of uniformity that is x^{-1} times that of P_1 and that is reflected in the anti-symmetry in the exponents given by $\frac{m_D(P)}{c_P}$ for the respective parts given by P_1 and P_2.

5. Scale free and self-contained nature of $\frac{D_P}{c_P}$

Equation (25) is a very interesting identity. It actually demonstrates the scale free nature of the ratio D/c for parts P_1 and P_2 from two entirely different distributions. By scale free here, we mean that there is no explicit dependence of $m_D(P)$ on the original random variable but only on the probabilities themselves. In other words, we can compute and compare the degree of uniformity of $P_1 \in X_1$ and $P_2 \in X_2$ using equation (25) without paying any heed to the scale of the original random variables X_1 or X_2. Also, by self-contained, we mean that only the probabilities from the part P are needed to compute $m_D(P)$.

We show below that the computation of the degree of uniformity is totally self-contained within the part P i.e. only requires the probabilities within the part P and also is scale-free i.e. does not require the knowledge of the random variable X.

$$\left(\frac{D_P}{c_P}\right)^{c_P} \underset{\text{discrete}}{=} \prod_{i \in P} \left(\frac{1}{p_i}\right)^{p_i}$$
$$\Rightarrow c_P \ln\left(\frac{D_P}{c_P}\right) = -\sum_{i \in P} p_i \ln(p_i)$$
$$\Rightarrow \frac{D_P}{c_P} = \exp\left[-\frac{1}{\left(\sum_{i \in P} p_i\right)} \sum_{i \in P} p_i \ln(p_i)\right] \tag{28}$$

$$\left(\frac{D_P}{c_P}\right)^{c_P} \underset{\text{continuous}}{=} \exp\left\{-\int_P p(x)\ln(p(x))\,dx\right\}$$
$$\Rightarrow c_P \ln\left(\frac{D_P}{c_P}\right) = -\int_P p(x)\ln(p(x))\,dx$$
$$\Rightarrow \frac{D_P}{c_P} = \exp\left[-\frac{1}{\left(\int_P p(x)\,dx\right)} \int_P p(x)\ln(p(x))\,dx\right] \tag{29}$$

In equations (28) and (29) above, we have derived explicit expressions for the ratio $\frac{D_P}{c_P}$ for both the continuous and discrete cases. We note that the final expressions do not have any reference to the value of the random variable directly (only probabilities), nor do they have any reference to probabilities outside of the part P. Hence, this definitively shows that the degree of uniformity is a scale free and self contained quantity.

This shows that we can compare the degree of uniformity (or inequality) of parts of different distributions in a self-normalized way. It also shows that the ratio D/c for a part P is an inherent characteristic of the part P's shape and does not need any information other than the part P itself i.e. the probabilities that are not in P are not required.

6. Examples

6.1. Example 1

In this section, we have chosen four examples of parts of distributions with two discrete and two continuous ones. Furthermore, the parts are all chosen to be union of two parts i.e. so that the slope of secant of the slope of diversity curve cannot be computed. Here are the examples listed below and shown in figure 4:

(1) **Binomial**distribution $B(50, 0.6)$ with the part being $k = \{20,..,25\} \cup \{30,..,35\}$.

(2) **Standard Normal** distribution $N(0, 1)$ with the part being $(-3, -1) \cup (2, 4)$.

(3) **Poisson** distribution with $\lambda = 20$ with the part being $k = \{5,..15\} \cup \{25,..35\}$.

(4) **Exponential** distribution with mean $\mu = 5$ with the part being $(0, 5) \cup (10, 15)$.

Table 1. The numbers are the values of the degree of uniformity measure for the column part with respect to the row part.

$m_{\text{row}}(\text{col})$	Binomial	Normal	Poisson	Exponential
Binomial	0	−0.0509	0.3225	−3.2447
Normal	0.1707	0	0.4109	−2.9619
Poisson	−0.6229	−0.2367	0	−4.2768
Exponential	1.9582	0.5332	1.3364	0

Table 2. The numbers are the values of the degree of uniformity measure for the column part with respect to the row part.

$\frac{\left(\frac{D_{\text{col}}}{c_{\text{col}}}\right)}{\left(\frac{D_{\text{row}}}{c_{\text{row}}}\right)}$	Binomial	Normal	Poisson	Exponential
Binomial	1	0.7537	2.8070	0.0390
Normal	1.3268	1	3.7224	0.0517
Poisson	0.3563	0.2685	1	0.0139
Exponential	25.6547	19.3351	72.0017	1

We used MATLAB to write functions that can compute the degree of uniformity measure for parts of a discrete or continuous distribution. The functions basically compute the degree of uniformity measure as stated in definition 3.1 and theorem 3.1.

We show a table of results below:

We make some general observations on the four examples shown above:

(1) In table 1, the diagonal entries are zero because the ratio of degree of uniformity of the row part corresponding to the column part is 1, since these two are the same parts. Correspondingly, in table 2 the diagonal entries are 1 reaffirming the same idea. This is a demonstration of theorem 4.2, equations (24) and (25).

(2) The off-diagonal entries in table 1 are positive or negative depending on whether the row part is more or less uniformly distributed compared to the column part. The corresponding entries in table 2 are either larger or smaller than 1. Also the (i,j)-th entry in table 2 is the reciprocal of the (j,i)-th entry as expected. This demonstrates equations (26) and (27) in theorem 4.2.

(3) It is interesting to note that the Poisson part is about 4 times more uniformly distributed than the Normal part, and the Binomial part is around 1.3 times more uniformly distributed than the Exponential part. Stated differently, the Poisson or the Binomial part has a localization or concentration of inequality that is 4 and 1.3 times more than the Normal or the Exponential parts respectively.

(4) We also note that the Exponential part has a lesser degree of uniformity (or localization of inequality) compared to all the other parts in the table. This is evident from the fact that all the numbers in the last row of table 2 are at least 19 signifying that the Exponential part is at least 19 times less uniformly distributed compared to all the other parts.

(5) On the other hand, the Poisson part is at least 2.8 times more uniformly distributed than the other parts as seen in the third column of table 2, where the smallest ratio is 2.8 in the third column.

(6) The Normal part is 3.7 times less uniformly distributed compared to the Poisson part but it is almost 19 times more uniformly distributed than the Exponential part.

6.2. Example 2

In this section we consider the example of a power law ($\alpha = 3$ and $x_{\min} = 100$) with four communities with income ranges as below.

(1) P_1: $(105K, 110K) \cup (120K, 125K)$

(2) P_2: $(140K, 145K) \cup (160K, 165K)$

Table 3. The numbers are the values of the degree of uniformity measure for the column part with respect to the row part.

$m_{\text{row}}(\text{col})$	P1	P2	P3	P4
P1	0	0.0490	0.2282	0.2284
P2	−0.1144	0	0.1402	0.1530
P3	−0.2965	−0.0781	0	0.0329
P4	−0.3464	−0.0995	−0.0383	0

Table 4. The numbers are the values of the ratio of degree of uniformity the column part with respect to the row part.

$\frac{\left(\frac{D_{\text{col}}}{c_{\text{col}}}\right)}{\left(\frac{D_{\text{row}}}{c_{\text{row}}}\right)}$	P1	P2	P3	P4
P1	1	2.3326	8.9909	13.0047
P2	0.4287	1	3.8545	5.5753
P3	0.1112	0.2594	1	1.4464
P4	0.0769	0.1794	0.6914	1

(3) P_3: $(105K, 110K) \cup (160K, 165K)$

(4) P_4: $(120K, 125K) \cup (140K, 145K)$

Intuitively we expect the richer community P_2 to have a higher degree of inequality since there is more concentration of wealth there. Hence we expect $m_{\frac{D_{P_1}}{c_{P_1}}}(P_2)$ to be larger than zero and $m_{\frac{D_{P_2}}{c_{P_2}}}(P_1)$ is less than zero. Equivalently, we expect $\frac{D_{P_2}}{c_{P_2}}$ to be larger than $\frac{D_{P_1}}{c_{P_1}}$.

We show the results in tabular form below:

The second richer community P_2 has twice the degree of inequality as P_1. This means that there is twice as much concentration of wealth in P_2 corresponding to P_1. As expected the *degree of uniformity* measure of P_2 with respect to P_1 is positive and vice versa for P_1 with respect to P_2 which is negative as expected. However, due to the mixture of poor and rich sections in communities P_3 and P_4, intuition breaks down since it is not clear how to compare these communities with respect to the other three. However, mathematics comes to our rescue, and it is interesting to note that P_2, P_3 and P_4 all have a higher concentration of wealth than P_1, and P_3 and P_4 are both more concentrated in wealth than P_2, and lastly P_4 comes closest in concentration of wealth to P_1 (about 1.45 times more degree of uniformity than P_1). Also P_4 is 13 times more uniformly distributed compared to P_1 as well.

Looking at table 5, we see that while P_1 and P_2 both have similar diversities, P_2 has fewer people than P_1 as seen by the lower c value of 0.0579 for P_2 compared to P_1. This explains why P_2 has a larger degree of uniformity (more than twice) than P_1 in table 4. Similarly, P_4 has a larger diversity and the fewest people ($c = 0.0890$) and this explains why P_4 has the largest degree of uniformity and has 13 times more wealth concentration than P_1 for example. In a similar way, the numbers in table 5 can be used to explain the numbers in tables 3 and 4 respectively. The key point here is that a visual inspection of the communities from figure 5 is not going to help with a comparison of wealth concentration in the communities, especially given the mixture of rich and poor sections. We really need to rely on a precise mathematical quantification of inequality as measured by the degree of uniformity measure $m_D(P)$ and the degree of uniformity $\frac{D}{c}$.

We use P_1 and P_3 from the power law example to show how the ideas developed in the theorems in this paper work. We note that the same observations can be made for any two parts in this as well as the previous example.

(1) First, we note that we can compute the degree of uniformity for both parts P_1 and P_3 using theorem 3.4 and corollary 3.1 by computing $m_D(P)$ for both the binomial and normal examples. This can be used to match the answers obtained by directly computing the degree of uniformity of the parts from equations (28) and (29) respectively. The diversities computed by both methods match and are equal to 9.8060 and 67.8395 respectively.

(2) Next, we can verify theorem 4.2 by (a) directly computing the ratio of degree of uniformity of the Binomial to the Normal parts by computing $\frac{D_P}{c_P}$ for the Binomial and Normal parts and taking the ratio explicitly. This answer matches the answer by using equation (25) where 1 stands for Binomial and 2 stands for Normal.

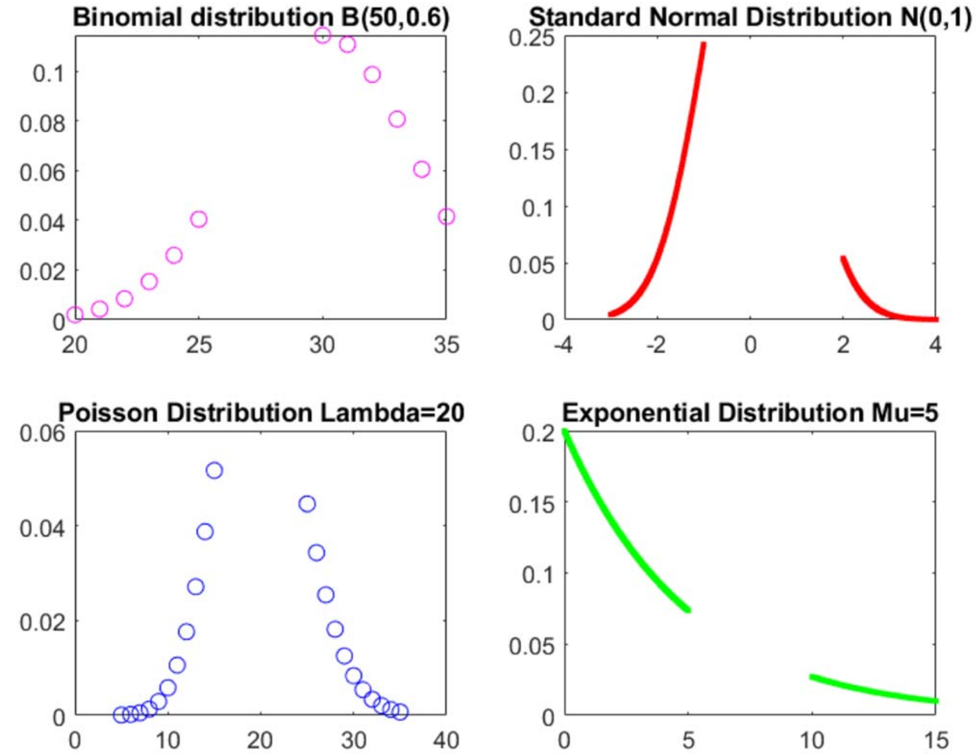

Figure 4. Two discrete and two continuous distributions chosen as examples.

The degrees of uniformity computed by both methods match and are equal to 72.6221 and 652.9385 respectively.

(3) We also note that from tables 3 and 4, that the degree of uniformity of the P_3 is larger than the Normal part (i.e. 8.9909 times more) and hence the corresponding degree of uniformity measure of the P_1 with respect to P_3 is positive (i.e. 0.2282). This verifies equations (25) and (26) in theorem 4.2.

(4) Lastly, equation (27) can be directly verified by seeing that $\ln(8.9909\ldots) + \ln(0.1112\ldots) = 0$. The '...' means that we need to used the unrounded versions.

Thus, we have used the parts P_1 and P_3 from the power law example to demonstrate the veracity of the theorems proved in this paper.

This power law example shows that while it is sometimes intuitively clear which parts of a distribution are more uniformly distributed compared to others visually for contiguous parts, such a luxury does not exist for non-contiguous parts. That is where our newly created degree of uniformity measure comes to the rescue by precisely quantifying the degree of uniformity and also allowing us to compute the ratio of degree of uniformity of parts.

The examples above demonstrate that the degree of uniformity measure is a generalization of the slope of secant of the slope of diversity curve that we introduced in Rajaram *et al* (2023, 2024) that allows us to compute the degree of uniformity (or degree of localization of inequality) of arbitrary measurable parts (or subsets) of the support of a given distribution that do not look like contiguous indices or intervals. Furthermore, we also see that we can compare the degree of uniformity of parts between discrete and continuous distributions as seen in theorem 3.1. This means that this measure is much more versatile in it usage as a quantification of degree of uniformity (or inequality) no matter the type of distribution that the part under consideration is. At the outset, we want reiterate the the degree of uniformity $\frac{D_P}{c_P}$ as discovered in Rajaram *et al* (2023, 2024) depends on both the SEE extent D_P or 'evenness' as well as the cumulative probability c_P or 'richness'. Hence, for arbitrary measurable parts such as in table 4 which is in general not a contiguous interval or set of indices, it is not intuitively clear visually whether a certain part has more or less degree of uniformity. The degree of uniformity measure serves as a way to quantify and compare this extent of uniformity in a mathematically sound manner, and hence its importance.

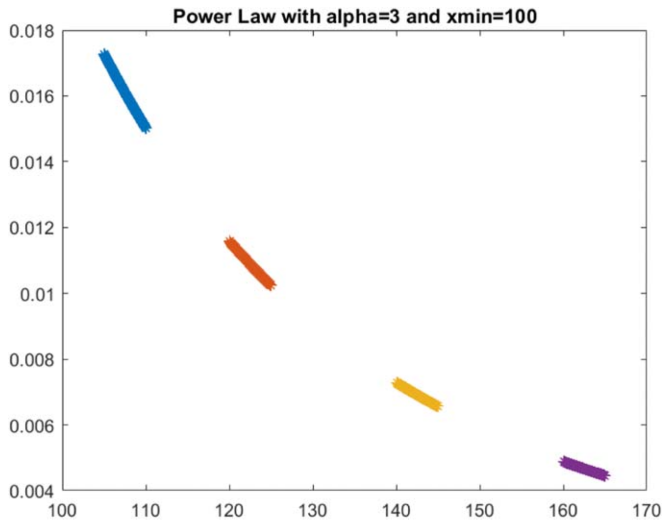

Figure 5. Union of disjoint intervals for the Power Law with alpha=3 and xmin = 100.

Table 5. The table shows the D, c and $\frac{D}{c}$ for the four communities in figure 5.

	P_1	P_2	P_3	P_4
D	9.8060	9.8072	67.8395	84.0769
c	0.1350	0.0579	0.1039	0.0890
$\frac{D}{c}$	72.6221	169.3954	652.9385	944.4270

7. Conclusion

We started the paper with two questions: (a) the question of generalizing the idea of degree of uniformity from contiguous parts (that required the slope of secant of the slope of diversity curve) to arbitrary measurable parts and (b) the question of comparison of degree of uniformity of two arbitrary measurable parts P_1 and P_2 corresponding to two random variables X_1 and X_2 that have no relationship between each other.

The answer to both questions came from the construction of a signed measure called the *degree of uniformity* measure in definition 3.1. The *degree of uniformity* measure was shown to generalize the slope of secant of the slope of diversity curve in theorem 3.2 that was used to measure the degree of uniformity of contiguous parts. More specifically the ratio $\frac{m_D(P)}{c_P}$ for an arbitrary part P is exactly equal to the slope of secant if the part P were contiguous. Hence, the newly constructed measure is a generalization of the slope of diversity curve for arbitrary measurable parts that are not contiguous.

In fact, as shown in theorem 3.3, comparing the ratio $\frac{m_D(P)}{c_P}$ for two different arbitrary parts leads to the same order of comparison for the degree of uniformity of the two parts, the key word here being arbitrary. The same result was proved to be true for contiguous parts with the slope of secant of the slope of diversity curve in Rajaram *et al* (2023, 2024). We also showed in 3.4 that the newly constructed measure can be used to compute the degree of uniformity of a part that is a countable union of indices or intervals.

Given parts P_1, P of the whole I that satisfy $P_1 \subseteq P \subseteq I$, in theorem 4.1, we established a relationship between the value of degree of uniformity measure for P_1 with respect to P, and with respect to I in equation (21). Finally in theorem 4.2, we have the most general relationship between the degree of uniformity measure of a part P_1 from a random variable X_1 to another part P_2 from a random variable X_2.

We have also shown the application of our newly discovered measure to four examples (two discrete and two continuous) of parts of distributions and compared and contrasted the degree of uniformity (or inequality) in said parts. We have also applied the degree of uniformity measure to a set of four communities in the distribution of wealth. A visual representation of the parts gives no intuitive way to compare the inequality in the parts, and hence, the importance of the degree of uniformity measure. The degree of uniformity measure allows us to precisely compute the degree of uniformity of these non-contiguous parts. The comparisons in the follow the intuition guided by the theoretical results that have been proved in this paper.

We also took the time to definitively showcase the scale-free and self-contained nature of the degree of uniformity measure in section 5. This is an important property of the measure as it illustrates the dependence of the measure only on the probabilities (or the density values) of the part P under study. The measure depends neither on the values of the random variable or its scale, nor on the value of probabilities that are outside of the part P. This is a very unique property of degree of uniformity, as it means that we can compare parts or whole of totally unrelated distributions. In a sense, the measure only depends on the 'y-axis' values (or probabilities) and hence is a quantification of inequality based solely on the inherent uncertainty (probability) that is present in the part, something that we have set out to accomplish in our series of papers.

Given the above developments, the *degree of uniformity measure* of a given part P with respect to a distribution whose degree of uniformity is $\frac{D}{c}$ given by $m_{\frac{D}{c}}(P)$ is a fundamental part of answering the two questions that we started the paper with. It not only allows us to generalize the idea of degree of uniformity to non-contiguous parts, but also allows us to compare parts P_1 and P_2 two entire different distributions governed by random variables X_1 and X_2 that have completely different scales.

In our future work, we aim to further develop the idea of the *degree of uniformity* measure $m_D(P)$ by looking at ways in which we could use the measure to systematically decompose the original distribution into parts that have an increasing or decreasing order of degree of uniformity. Such a decomposition will pave the way towards quantifying the inequality that is inherent in distributions in a systematic way. A systematic quantification of inequality of a distribution and a decomposition of the distribution based on said quantification will be an important step forward towards the study of inequality in distributions. We aim to lay down the foundation of such a decomposition in our future paper. We also plan to apply the decomposition theory to real data in a separate paper (or papers) to demonstrate the usage of our discovery.

We end the paper by discussing the value of our insights for engineering and physics, including two concrete examples where the study of inequality in distributions is important. Probability distributions are used widely in engineering and physics. The Normal, Binomial and Poisson distributions are used in quality control; Exponential, Weibull and Lognormal distributions are used in reliability engineering and failure analysis; the Poisson and Exponential are used in queuing theory and network analysis, the Gaussian, Rayleigh and Rice distributions are used in signal processing and communication systems; and the Normal, Lognormal and Extreme Value distributions are used in structural reliability and risk assessment. The celebrated Boltzmann, Bose–Einstein and Fermi distributions are used in Physics to describe the statistical mechanics to describe the kinetic energy of particles. Recently, probability models were used to describe the randomness associated with radioactive decay in Sanchez-Sanchez *et al* (2024).

As further evidence, consider these two examples. First, in kinetic theory of gases, the Maxwell-Boltzmann distribution describes the distribution of speeds among gas particles in thermal equilibrium. This distribution is highly non-uniform, with many particles having speeds around the most probable value, but fewer particles having very high or very low speeds. The probability distribution function of particle speeds v is given by:

$$f(v) = 4\pi\left(\frac{m}{2\pi k_B T}\right)^{\frac{3}{2}} v^2 e^{-\frac{v^2}{2k_B T}},$$

where m is the mass of the particle, T is the temperature and k_B is the Boltzmann constant. The distribution peaks at a certain speed but has a long tail, meaning there are always some particles with much higher or lower speeds than average. Studying the inequality in distribution of speeds is crucial for understanding phenomena like diffusion, viscosity, and thermal conductivity in gases. For example, even though the majority of particles move at moderate speeds, the fastest-moving particles can dominate energy transfer processes, making the variation in the speed distribution an important aspect of gas behavior.

Next, in structural engineering, the design of buildings, bridges, and other structures often relies on understanding the unequal distribution of loads (forces acting on a structure). Engineers need to account for the fact that while most of the time, loads like wind, traffic, or weight are moderate, there will be rare instances of extreme loads (e.g. during hurricanes or earthquakes) that could pose significant risks to the structure. Here, Extreme Value Theory (EVT) is used to model the probability distribution of maximum or minimum values of a dataset. Instead of assuming a uniform distribution of load magnitudes, EVT specifically focuses on the tail behavior of distributions—how likely extreme, rare events are. The Generalized Extreme Value (GEV) distribution, for example, models these extremes:

$$P(X \leqslant \xi) = e^{-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}},$$

where μ is the location parameter, σ is the scale parameter and ξ is the shape parameter controlling how 'fat' the tail is. Studying the inequality in the load distribution helps engineers design structures that can withstand not just average loads, but also rare extreme events. The non-uniformity and the rare, extreme values in load distributions are crucial for safety in engineering design.

In both examples, the variation of inequality in probability distributions—whether in gas particle speeds or in structural loads—plays a critical role in the physical behavior of systems and the design considerations that engineers must take into account. Our theory (specifically the decomposition that we mentioned for future work) can be used to precisely identify the ranges of speeds that have a larger localization of inequality thereby pinning down the exact variation of inequality in speeds of particles or the variation of inequality in the distribution of load.

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

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
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 2019 What do we mean by diversity? the path towards quantification *Mètode Science Studies Journal-Annual Review.* **9** 55–61
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
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, Ritchey N and Castellani B 2023 Mathematical diversity of parts for a continuous distribution *Journal of Physics Communications* **7** 075006
Rajaram R, Ritchey N and Castellani B 2024 On the comparison of diversity of parts of a distribution *Journal of Physics Communications* **8** 025008
Sanchez-Sanchez S, Cortez-Perez E and I M O V 2024 Binomial vs. poisson statistics: from a toy model to a stochastic model for radioactive decay *Physica A—Statistical Mechanics and Its Applications* **643** 129827