

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

MEASURING THE UNIFORMITY OF MEASURABLE SUBSETS OF THE UNIT SQUARE

BHARATH KRISHNAN

Department of Mathematics, Indiana East University, bharathk98@gmail.com

ABSTRACT. Suppose set $R \subseteq [0, 1] \times [0, 1]$. We want to define a measure of uniformity of R in the unit square using dimension $d \in [0, 2]$ of the d -dimensional Hausdorff measure. In order to understand *uniformity*, we'll give examples in §0 where points of R are uniform in $[0, 1] \times [0, 1]$. Next in §1, we will define preliminary definitions (e.g. Hausdorff & Counting measure) to define a uniformity of measurable subsets of the unit square. Finally, in §2 we will define a measure of uniformity between 0 and 1 w.r.t a uniform R with Hausdorff-dimension d . (In this case, the larger the measure of uniformity, the smaller the non-uniformity w.r.t to uniform R)

Keywords: Spatial, Dimensions, Uniform Distribution, Measure Theory, Hausdorff Dimension, Sparse

0. INTRO

Suppose we define set $R \subseteq [0, 1] \times [0, 1]$. I want to find a measure of the "uniformity" of the distribution of points in R (on set $[0, 1] \times [0, 1]$) where the larger the measure (as it approaches value one or infinity), the smaller the non-uniformity. Note in order to understand *uniformity*, here are examples of R with points *uniform* in $[0, 1] \times [0, 1]$:

- (1) If R is finite, then for $n \in \mathbb{N}$, suppose we partition set $[0, 1] \times [0, 1]$ into n^2 squares with length $1/n$. If there exists $n \in \mathbb{N}$ such there is only one point at the center of each square, then we say the discrete points are *uniform* in $[0, 1] \times [0, 1]$.
- (2) If we have that set R has a Hausdorff dimension of 2 (§1, def. 3), then for all real x_1, x_2, y_1, y_2 , if $0 \leq x_1 < x_2 \leq 1$ and $0 \leq y_1 < y_2 \leq 1$ where the Lebesgue measure (on the Lebesgue sigma-algebra) of $([x_1, x_2] \times [y_1, y_2]) \cap R$ is $(x_2 - x_1)(y_2 - y_1)$, then set R is *uniform* in $[0, 1] \times [0, 1]$.

For R with a Hausdorff dimension between 0 and 2, it's unclear what a *uniform* distribution of points in R covering $[0, 1] \times [0, 1]$ looks like. (Despite this, we will try to define the following in §2)

Note I want to find a general theory of uniformity (for distributions of points in the unit square) such that, depending on the Hausdorff dimension of the set of points (§1, def. 3), these points have a measure of uniformity between zero and one or zero and infinity.

Further note there are already several measures of uniformity for *finite* points in the unit square (e.g. Wasserstein distance [2] or distance between empirical copula & independence copula [3]) but no measure for *infinite* points in the unit square.

1. PRELIMINARY DEFINITIONS

Definition 1 (Counting Measure). Suppose $|\cdot|$ is the cardinality of a set. If $A \subseteq \mathbb{R}$ and μ is some measure defined on the sigma-algebra $\mathcal{P}(\mathbb{R})$, then the counting measure is:

E-mail address: bharathk98@gmail.com.

Date: December 14, 2022.

$$\mu(A) = \begin{cases} |A| & A \text{ is finite} \\ +\infty & A \text{ is infinite} \end{cases} \quad (1.0.1)$$

Definition 2 (Hausdorff Measure). Let (X, φ) be a metric space, $d \in [0, \infty)$ and C, E be arbitrary sets where $C, E \subseteq X$ and the diameter of C is:

$$\text{diam}(C) := \sup\{\varphi(x, y) : x, y \in C\}, \quad \text{diam}(\emptyset) := 0$$

If $i \in \mathbb{N}$ and $\delta \in \mathbb{R}$ such that $\delta > 0$, where the Euler's Gamma function is Γ and constant \mathcal{N}_d is:

$$\mathcal{N}_d = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2} + 1\right)} \quad (1.0.2)$$

then we define:

$$\mathcal{H}_\delta^d(E) = \mathcal{N}_d \inf \left\{ \sum_{i=1}^{\infty} (\text{diam}(C_i))^d : \text{diam}(C_i) \leq \delta, E \subseteq \bigcup_{i=1}^{\infty} C_i \right\} \quad (1.0.3)$$

such if the infimum of the equation is taken over the countable covers of sets $C_i \subset X$ of E (satisfying $\text{diam}(C_i) \leq \delta$), then the Hausdorff Outer Measure is:

$$\mathcal{H}^d(E) = \sup_{\delta > 0} \mathcal{H}_\delta^d(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^d(E)$$

such for $d \in \mathbb{N}$, $\mathcal{H}^d(E)$ coincides with the d -dimensional Lebesgue Measure, where we convert the Outer measure to the Hausdorff measure from restricting E to the σ -field of Carathéodory measurable sets [5].

Definition 3 (Hausdorff Dimension). The Hausdorff Dimension of E is defined by $\phi(E)$ where:

$$\mathcal{H}^d(E) = \begin{cases} \infty & \text{if } 0 \leq d < \phi(E) \\ 0 & \text{if } \phi(E) \leq d < \infty \end{cases} \quad (1.0.4)$$

which is needed since constant $\mathcal{N}_{\phi(E)}$ (eq. 1.1.1) is required for the next definition.

1.1. Generalized Hausdorff Measure. If $\mathcal{H}^{\phi(E)}(E)$ is zero or infinity, consider the following:

Definition 4 (Generalized Hausdorff Measure). Suppose (X, d) is a metric space and $C, E \subseteq X$. Let $h : [0, \infty) \rightarrow [0, \infty)$ be an (exact) dimension function (or gauge function) which is monotonically increasing, strictly positive, and right continuous [6]. If the diameter of C is:

$$\text{diam}(C) := \sup\{\varphi(x, y) : x, y \in C\}, \quad \text{diam}(\emptyset) := 0$$

such that for $i \in \mathbb{N}$, where $\delta \in \mathbb{R}$ and $\delta > 0$, if the Hausdorff dimension is $\phi(E)$ where Euler's Gamma function is Γ , and $\mathcal{N}_{\phi(E)}$ is the constant where:

$$\mathcal{N}_{\phi(E)} = \frac{\pi^{\phi(E)/2}}{\Gamma\left(\frac{\phi(E)}{2} + 1\right)} \quad (1.1.1)$$

we then define:

$$\mathcal{H}_\delta^h(E) = \mathcal{N}_{\phi(E)} \inf \left\{ \sum_{i=1}^{\infty} h(\text{diam}(C_i)) : \text{diam}(C_i) \leq \delta, E \subseteq \bigcup_{i=1}^{\infty} C_i \right\} \quad (1.1.2)$$

such that if the infimum of the equation above is taken over the countable covers of sets $C_i \subset X$ of E (which satisfy $\text{diam}(C_i) \leq \delta$), then the h -Hausdorff Outer Measure follows:

$$\mathcal{H}^h(E) = \sup_{\delta > 0} \mathcal{H}_\delta^h(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^h(E) \quad (1.1.3)$$

such that for $\phi(E) \in \mathbb{N}$, $\mathcal{H}^h(E)$ coincides with the $\phi(E)$ -dimensional Lebesgue Measure where we define the outer h -Hausdorff measure as the h -Hausdorff measure from restricting the Outer Measure to E measurable in the sense of Carathéodory, and defining h so $\mathcal{H}^h(E)$ is strictly positive and finite.

2. MEASURING "UNIFORMITY" OF A MEASURABLE SUBSET OF $[0, 1] \times [0, 1]$

Here is the attempt to measure the *uniformity* of R in $[0, 1] \times [0, 1]$ w.r.t a *uniform* R with a Hausdorff-dimension of $d \in [0, 2]$ (def. 3).

Now suppose $\mu(A)$ is the counting measure of A (def. 1) where:

$$s = \begin{cases} |A| & \mu(A) < +\infty \\ 1 & \mu(A) = +\infty \end{cases}$$

For $n, j \in \{sk : k \in \mathbb{N}\}$ (where $j \ll n$), suppose we partition set $[0, 1] \times [0, 1]$ into n^2 squares with length n , such that we combine these squares to form larger squares with area $1/(j^2)$. Furthermore, we define $i, m, r \in \mathbb{N} \cup \{0\}$, where $x_{i+1} - x_i = 1/n$ and $0 \leq i + r \leq n - 1$ such that the side of each of the squares (parallel to the x-axis) is the interval $[x_i, x_{i+r}]$ with length $1/(n+r)$, and we define $y_{m+1} - y_m = 1/n$ where $0 \leq m + r \leq n - 1$ and the side of each square (parallel to the y-axis) is the interval $[y_m, y_{m+r}]$ with length $1/(n+r)$.

If $d \in [0, 2]$, where the d -dimensional Hausdorff measure defined on the σ -algebra of caratheodory-measurable sets is \mathcal{H}^d (def. 2), I would like to divide measure \mathcal{H}^d of each $([x_i, x_{i+r}] \times [y_m, y_{m+r}]) \cap R$ (such that the area of $[x_i, x_{i+r}] \times [y_m, y_{m+r}]$ is $1/(j^2)$) by the d -Hausdorff measure of all uniform R , with Hausdorff-dimension d , in all squares or $((\mathcal{H}^d(\text{dom}(R)) \cdot \mathcal{H}^d(\text{range}(R))) / (j^2))$. This results in a discrete probability distribution \mathbb{P} :

$$S(n, j) = \{(i, m, r) : i, m, r \in \mathbb{N} \cup \{0\}, x_{i+1} - x_i = y_{m+1} - y_m = 1/n, \text{Area}([x_i, x_{i+r}] \times [y_m, y_{m+r}]) = 1/j^2\} \quad (2.0.1)$$

$$\mathbb{P}(n, j) = \left\{ \frac{\mathcal{H}^d([x_i, x_{i+r}] \times [y_m, y_{m+r}] \cap R)}{((\mathcal{H}^d(\text{dom}(R)) \cdot \mathcal{H}^d(\text{range}(R))) / (j^2)) \cdot |S(n, j)|} : (i, m, r) \in S(n, j) \right\} \quad (2.0.2)$$

We then want to take the entropy of the discrete distribution (2.0.3).

$$E(\mathbb{P}(n, j)) = \sum_{x \in \mathbb{P}(n, j)} -x \log_2 x \quad (2.0.3)$$

Next, we wish to define the *uniformity* of R in $[0, 1] \times [0, 1]$ using $\mu(A)$ and the dimension d ; however, it's difficult (when $n, j \rightarrow \infty$ and $j \leq \mu(A)$ such that $j \ll n$) to compare $E(\mathbb{P}(n, j))$ with the entropy [4, p.61-95] of a discrete uniform distribution with the same number of elements or $\log_2(|S(n, j)|)$. (Note the smaller the absolute difference of $E(\mathbb{P}(n, j))$ and $\log_2(|S(n, j)|)$, the *closer* R is to having the same uniformity in $[0, 1] \times [0, 1]$ as a uniform R , with Hausdorff-dimension d , in $[0, 1]^2$).

Therefore, the following might be more useful:

$$\overline{G}(n', j') = \inf_{j \in \mathbb{N}} \inf_{n \in \mathbb{N}} \{|\mathbb{P}(n, j)| : n', j' \in \mathbb{N}, E(\mathbb{P}(n, j)) \geq \log_2(|S(n', j')|)\}$$

$$\underline{G}(n', j') = \sup_{j \in \mathbb{N}} \inf_{n \in \mathbb{N}} \{|\mathbb{P}(n, j)| : n', j' \in \mathbb{N}, E(\mathbb{P}(n, j)) \leq \log_2(|S(n', j')|)\}$$

This gives us a measure of uniformity of R w.r.t a uniform d -dimensional space in $[0, 1] \times [0, 1]$ i.e. $\mathcal{U}^*(f, d, [0, 1] \times [0, 1])$ (when it exists):

$$\mathcal{U}^*(f, d, [0, 1] \times [0, 1]) = \lim_{j' \leq \mu(A), j' \rightarrow \infty} \lim_{n' \rightarrow \infty} \frac{\overline{G}(n', j')}{|S(n', j')|} = \lim_{j' \leq \mu(A), j' \rightarrow \infty, n' \rightarrow \infty} \frac{\underline{G}(n', j')}{|S(n', j')|}$$

Note the larger $\mathcal{U}'(f, d, [0, 1] \times [0, 1])$ is, the *closer* R is to have the same uniformity in $[0, 1] \times [0, 1]$ as a uniform R (in $[0, 1]^2$) with Hausdorff-dimension d .

2.1. **Summary:** If the Hausdorff Dimension (def. 3) of set $R \subseteq [0, 1] \times [0, 1]$ is $\phi(R)$ and:

- (1) If $d < \phi(R)$ then $\mathcal{U}^*(f, d, [0, 1] \times [0, 1]) = 1$
- (2) If $d = \phi(R)$ and $\mathcal{U}^*(f, d, [0, 1] \times [0, 1]) \in (0, 1)$ then $\mathcal{U}^*(f, d, [0, 1] \times [0, 1])$ is the *final measure of uniformity*; otherwise, we set dimension d to gauge function h (def. 4) or when neither of them work and R is fractal, we apply [1]
- (3) If $d > \phi(R)$ then $\mathcal{U}^*(f, d, [0, 1] \times [0, 1]) = 0$

For example, if R is a continuous function with domain $[0, 1]$ (and $\phi(R) = 1$), then if $d = 2$ the measure of uniformity should be 0. (This means w.r.t a uniform R in $[0, 1] \times [0, 1]$ with Hausdorff dimension 2, the previous set R is completely *non-uniform*.)

3. QUESTION:

Is there a simpler measure that measures the "non-uniformity" of set R w.r.t a uniform R that has a dimension $d \in [0, 2]$?

REFERENCES

- [1] Bedford B. and Fisher A. Ratio geometry, rigidity and the scenery process for hyperbolic cantor sets. <https://arxiv.org/pdf/math/9405217.pdf>.
- [2] Thomas Bonis. Improved rates of convergence for the multivariate central limit theorem in wasserstein distance, 2023. <https://arxiv.org/abs/2305.14248>.
- [3] JAEWON LEE ([https://stats.stackexchange.com/users/346978/jaewon lee](https://stats.stackexchange.com/users/346978/jaewon%20lee)). Is there a way to measure uniformness of points in a 2d square? Cross Validated. <https://stats.stackexchange.com/q/599023> (version: 2022-12-14).
- [4] Gray M. Springer New York, New York [America];, 2 edition, 2011. <https://ee.stanford.edu/~gray/it.pdf>.
- [5] Taylor M. The caratheodory construction of measures. <https://mtaylor.web.unc.edu/wp-content/uploads/sites/16915/2018/04/measch5.pdf>.
- [6] Wikipedia. Dimension function. https://en.wikipedia.org/wiki/Dimension_function.