Concept Paper

Not peer-reviewed version

Measuring the uniformity of Measurable Subsets of The Unit Square

Bharath Krishnan *

Posted Date: 12 July 2023

doi: 10.20944/preprints202306.0834.v2

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

Concept Paper

Measuring the Uniformity of Measurable Subsets of the Unit Square

Bharath Krishnan

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

Abstract: Suppose set $A \subseteq [0,1] \times [0,1]$. We want to define a measure of uniformity of A in the unit square. In order to understand *uniformity*, we'll give examples in §0 where A is uniform in $[0,1]^2$. Next in §1, we will define preliminary definitions (e.g. Hausdorff measure) to define uniformity of measurable subsets of the unit square. Finally, in §2 we'll define a measure of uniformity that measures the "distance" from uniformity between 0 and 1 (where the larger the "distance", the larger the non-uniformity).

Keywords: spatial; dimensions; uniform distribution; measure theory; hausdorff dimension; sparse

---

0. Intro

The aim of this paper is to measure the "uniformity" of measurable subsets of $[0,1]^2$. If set $A \subseteq [0,1] \times [0,1]$; we want to define a measure of uniformity for A.

Here are some example of a "uniform A":

(1) A_n is the uniform distribution on the uniform grid $G_n := \left\{\left(\frac{i}{n}, \frac{j}{n}\right) : i = 0, \ldots, n,\ j = 0, \ldots, n\right\}$ such that with large n, if $A_n \to A$, then A is *uniform* in $[0,1]^2$

(2) For all real x_1, x_2, y_1, y_2, if $0 \le x_1 < x_2 \le 1$ and $0 \le y_1 < y_2 \le 1$ where the Lebesgue measure (on the Lebesgue sigma-algebra) of $([x_1, x_2] \times [y_1, y_2]) \cap A$ is $(x_2 - x_1)(y_2 - y_1)$, then set A is *uniform* in $[0,1] \times [0,1]$.

(In general, we shall define a "uniform" subset of $[0,1]^2$ in def. 4, §2,).

Note we wish for a measure of uniformity to be between (and including) zero and one or zero and infinity such that the larger the measure, the larger the *non-uniformity*.

Further note, there are already several measures of uniformity for *finite* points in the unit square (e.g. wasserstein distance [1] or distance between empirical copula & independence copula [2]) but no measure for *infinite* points in the unit square.

1. Preliminary Definitions

Definition 1 (Hausdorff Measure). *Let (X, α) be a metric space, $\alpha \in [0, \infty)$. For every $C \in X$, define the diameter of C as:*

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

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

$$\mathcal{N}_\alpha = \frac{\pi^{\alpha/2}}{\Gamma\left(\frac{\alpha}{2} + 1\right)} \tag{1.0.1}$$

we define:

$$\mathcal{H}^\alpha_\delta(E) = \mathcal{N}_\alpha \inf\left\{\sum_{i=1}^{\infty} (diam(C_i))^\alpha : diam(C_i) \le \delta, E \subseteq \bigcup_{i=1}^{\infty} C_i\right\} \tag{1.0.2}$$

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202306.0834.v2

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

where for $\alpha \in \mathbb{N}$, $\mathcal{H}^{\alpha}(E)$ coincides with the α-dimensional Lebesgue Measure, where we convert the Outer measure to the Hausdorff measure from restricting E to the σ-field of Carathéodory measurable sets [3].

Definition 2 (Hausdorff Dimension). *The Hausdorff Dimension of E is defined by $\phi(E)$ where:*

$$\mathcal{H}^{d}(E) = \begin{cases} \infty & \text{if } 0 \le d < \phi(E) \\ 0 & \text{if } \phi(E) < d < \infty \end{cases} \tag{1.0.3}$$

1.1. Generalized Hausdorff Measure

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

Definition 3 (Generalized Hausdorff Measure). *Suppose (X, d) is a metric space. Let $h : [0, \infty) \to [0, \infty)$ be an (exact) dimension function (or gauge function) which is monotonically increasing, strictly positive, and right continuous [4].*

For $i \in \mathbb{N}$, where $\delta \in \mathbb{R}$ and $\delta > 0$, if the Hausdorff dimension is $\phi(E)$; we define:

$$\mathcal{H}^{h}_{\delta}(E) = \mathcal{N}_{\phi(E)} \inf\left\{ \sum_{i=1}^{\infty} h(diam(C_i)) : diam(C_i) \le \delta, E \subseteq \bigcup_{i=1}^{\infty} C_i \right\} \tag{1.1.1}$$

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

$$\mathcal{H}^{h}(E) = \sup_{\delta>0} \mathcal{H}^{h}_{\delta}(E) = \lim_{\delta\to 0} \mathcal{H}^{h}_{\delta}(E) \tag{1.1.2}$$

where when $\phi(E) \in \mathbb{N}$, $\mathcal{H}^{h}(E)$ should coincide with the $\phi(E)$-dimensional Lebesgue Measure such that we define the "outer h-Hausdorff measure" as h-Hausdorff measure by restricting the Outer Measure to E measurable in the sense of carathèodory, where $\mathcal{H}^{h}(E)$ is strictly positive and finite.

2. Measuring "Uniformity" of a Measurable Subset of $[0,1] \times [0,1]$

Using this answer [5], let $S := [0,1]^2$ be the unit square. "Partition" S naturally into four congruent squares $S_{1,j}$ (with side length $1/2$ each), where $j = 1, \ldots, 4$; the quotation marks are used here because the $S_{1,j}$'s will have some common boundary points. Next, "partition" each $S_{1,j}$ naturally into four congruent squares (with side length $1/2^2$ each), so that we get 4^2 squares $S_{2,j}$ for $j = 1, \ldots, 4^2$. Continue doing so, so that at the kth step we get 4^k squares $S_{k,j}$ for $j = 1, \ldots, 4^k$, for each $k = 1, 2, \ldots$.

Take any subset A of S. For each $k = 1, 2, \ldots$ and each $j = 1, \ldots, 4^k$, let

$$A_{k,j} := (A \cap S_{k,j}) - s_{k,j},$$

where $s_{k,j}$ is the southwest vertex of the square $S_{k,j}$, so that $A_{k,j} \subseteq S_k := 2^{-k}S$.

Suppose that for each k we have a "measure" D_k of dissimilarity for subsets of S_k, so that for any two subsets B and C of S_k we have a nonnegative real number $D_k(B, C)$, which is the greater the more "dissimilar" B and C are (and is 0 if $B = C$); here the term "measure" is used in the general sense, not necessarily in the sense of measure theory. For instance, $D_k(B, C)$ may depend on the Hausdorff distance [6] between B and C or on some "measure" of the symmetric difference of the sets B and C or on some combination thereof.

Then the distance of the set A from uniformity can be defined by the formula

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202306.0834.v2

$$D(A) := \sum_{k=1}^{\infty} \frac{1}{L^k} \sum_{j=1}^{4^k} \sum_{m=1}^{4^k} \frac{D_k(A_{k,j}, A_{k,m})}{1 + D_k(A_{k,j}, A_{k,m})},$$

where L is a real number > 16 (to ensure the convergence of the series). Then $D(A)$ will be small if, for "most" levels k of "zooming", "most" of the intersections of the set A with all the "k-level" small squares $S_{k,j}$ "look similar" to one another. In other terms:

Definition 4 (Definition of Uniform A in The Unit Square). *If $D(A) = 0$, A is uniform in $[0,1]^2$.*

(Of course, $D(A)$ will depend on the choices of L and the dissimilarity "measures" D_k.)

For instance, for any L and any D_k's we have $D(S) = 0$ – of course, the unit square S is at distance 0 from uniformity (in itself).

As another example, for the uniform grid G_n (defined in §0, crit. (0)) with $n = 2^K$ for a natural K, any real $L > 16$, and any D_k's we have

$$D(G_n) \leq \sum_{k=K+1}^{\infty} \frac{1}{L^k} 16^k = Cn^{-p} \to 0$$

as $n = 2^K \to \infty$, where $C := \frac{16}{L-16}$ and $p := \log_2 \frac{L}{16}$.

2.1. *Specific Example of Measure of Uniformity for Measurable Subsets of* $[0,1]^2$

For example, if D_k is $\mathcal{H}^h$ (def. 3) where h is *the dimension function of* A and $L = 32$, one measure of uniformity is:

$$D(A) := \sum_{k=1}^{\infty} \frac{1}{32^k} \sum_{j=1}^{4^k} \sum_{m=1}^{4^k} \frac{\mathcal{H}^h(A_{k,j} \,\Delta\, A_{k,m})}{1 + \mathcal{H}^h(A_{k,j} \,\Delta\, A_{k,m})},$$

where if $\mathcal{H}^h(A_{k,j} \,\Delta\, A_{k,m}) = +\infty$ for some $k, j, m \in \mathbb{N}$ or if $\mathcal{H}^h(A_{k,j} \,\Delta\, A_{k,m}) = 0$ for all $k, j, m \in \mathbb{N}$, take §3 of this paper [7] (if version 3 of the paper exists, consider *that* version instead).

In [7], for (§3, eq. 4.1.9) if we replace k with z, such that $(F_z^{\star\star\star})$ is a chosen sequence from a set of equivalent $\star$-sequence of sets (§2, def. 4), this should give us:

$$\forall(\epsilon > 0)\exists(N \in \mathbb{N})\forall(z \in \mathbb{N}) \left(z \geq N \Rightarrow \left| D^\star(A) - \sum_{k=1}^{\infty} \frac{1}{32^k} \sum_{j=1}^{4^k} \sum_{m=1}^{4^k} \frac{\mathcal{H}^h\left(\left(A_{k,j} \,\Delta\, A_{k,m}\right) \cap F_z^{\star\star\star}\right)}{1 + \mathcal{H}^h\left(\left(A_{k,j} \,\Delta\, A_{k,m}\right) \cap F_z^{\star\star\star}\right)} \right| \leq \epsilon \right)$$

where $D^\star(A)$ is the **final measure of uniformity** for subset of the unit square (if it exists).

References

1. Bonis, T. Improved rates of convergence for the multivariate Central Limit Theorem in Wasserstein distance, 2023, [arXiv:math.PR/2305.14248]. https://arxiv.org/pdf/2305.14248.pdf.
2. (https://stats.stackexchange.com/users/346978/jaewon lee), J.L. Is there a way to measure uniformness of points in a 2D square? Cross Validated, [https://stats.stackexchange.com/q/599023]. https://stats.stackexchange.com/q/599023 (version: 2022-12-14).
3. M., T. The Caratheodory Construction of Measures. https://mtaylor.web.unc.edu/wp-content/uploads/sites/16915/2018/04/measch5.pdf.
4. Wikipedia. Dimension Function. https://en.wikipedia.org/wiki/Dimension_function.
5. (https://mathoverflow.net/users/36721/iosif pinelis), I.P. Defining a measure of uniformity for measurable subsets of $[0,1]^2$ w.r.t dimension $\alpha \in [0,2]$. MathOverflow, [https://mathoverflow.net/q/449772]. https://mathoverflow.net/q/449772 (version: 2023-06-29).

 doi:10.20944/preprints202306.0834.v2

6. Rockafellar, R.; Wets, R. *Variational Analysis*; Vol. 317, 2004; p. p.117. https://sites.math.washington.edu/~rtr/papers/rtr169-VarAnalysis-RockWets.pdf, doi:10.1007/978-3-642-02431-3.
7. Krishnan, B. Finding an Extension of the Expected Value That Is Unique, Finite, and Natural for All Functions in Prevalent Subset of the Set of All Functions. *Preprints.org* **2023**. https://www.preprints.org/manuscript/202307.0560/v2, doi:10.20944/preprints202307.0560.v2.