Concept Paper

Defining the Most Generalized, Natural Extension of the Expected Value on Measurable Functions

Bharath Krishnan

ABSTRACT. In this paper, we will extend the expected value of the function w.r.t the uniform probability measure on sets measurable in the Carathèodory sense to be finite for a larger class of functions, since the set of measurable functions with infinite or undefined expected values might form a prevalent subset of the set of all measurable functions. Before we get to the specific problem (or main question) of the paper, we will outline some preliminary definitions. We will then define a precise main question that will attempt to offer a unique solution along with a partial solution to the question. Along the way, we will ask a series of questions that will clarify our understanding of the paper.

Keywords: Expected Value, Uniform Measure, Measure theory, Prevalence, Entropy, Sample, Linear, Superlinear, Choice Function, Bernard's Paradox, Pseudo-random

1. Preliminaries

Suppose A is a set measurable in the Carathèodory sense [7], such for $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, and function $f : A \to \mathbb{R}$.

1.1. **Motivation.** It seems the set of measurable functions with infinite or undefined expected values (def. 1), using the uniform measure [16, p.32-37], might be a prevalent subset [14, 11] of the set of all measurable functions, meaning "almost every" measurable function has an infinite or undefined expected value. Furthermore, when the Lebesgue measure of A, measurable in the Caratheodory sense, has zero or infinite volume (or undefined measure), there might be multiple, conflicting ways of defining a "natural" uniform measure on A.

Below I will attempt to define a question regarding an extension of the expected value (when it's undefined or infinite) which allows for a finite value instead.

Note the reason the question will be so long is there are plenty of "meaningless" extensions of the expected value (e.g. if the expected value is infinite or undefined we can just replace it with zero).

Therefore we must be more specific about what is meant by "meaningful" extension but there are some preliminary definitions we must clarify.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

1.2. **Preliminary Definitions.**

Definition 1 (**Expected value w.r.t the Uniform Probability Measure**). *From an answer to a question in cross validated (a website for statistical questions)* [10] *, let* $\mathbf{X} \sim \mathit{Uniform}(A)$ *denote a uniform random variable on set* $A \subseteq \mathbb{R}^n$ *and* $p_{\mathbf{X}}$ *denote the probability density function from the radon-nikodym derivative* [2, p.419-427] *of the uniform probability measure on* A *measurable in the Carathèodory sense. If* $\mathbb{I}(x \in A)$ *denotes the indicator function on* $x \in A$*:*

$$\mathbb{I}(x \in A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

then the radon-nikodym derivative of uniform probability measure must have the form $\mathbb{I}(x \in A)/U'(A)$*. (Note* U' *is not the derivative of* U *in the sense of calculus but is rather different from the uniform probability measure defined as* U*.)*

Therefore, using the law of the unconscious statistician, we should get

$$\begin{aligned} \mathbb{E}[f(\mathbf{X})] &= \int\limits_{\mathbb{R}^n} f(x) \cdot p_{\mathbf{X}}(x)\, d\mathbf{x} && (1.2.1) \\ &= \int\limits_{\mathbb{R}^n} f(\mathbf{x}) \cdot \frac{\mathbb{I}(x \in A)}{U'(A)}\, d\mathbf{x} \\ &= \frac{1}{U'(A)} \int\limits_{A} f(x)\, d\mathbf{x} && (P1) \\ &= \mathbb{E}_{U'}[f(\mathbf{X})] \end{aligned}$$

such the expected value is undefined when A *does not have a uniform probability distribution or* f *is not integrable w.r.t the measure* U'*.*

Definition 2 (**Defining the pre-structure**). *Since there's a chance that* $\mathbf{X} \sim \mathit{Uniform}(A)$ *does not exist or* f *is not integrable w.r.t to* U'*, using def.* 1 *we define a sequence of sets* $\{F_r\}_{r\in\mathbb{N}}$ *where if:*

(1) $\bigcup\limits_{r=1}^{\infty} F_r = A$

(2) *For all* $r \in \mathbb{N}$*,* $\mathbf{X}_r \sim \mathit{Uniform}(F_r)$ *exists (if* A *is countable infinite then for every* $r \in \mathbb{N}$*,* F_r *must be a finite set since* $\mathbf{X}_r$ *is a discrete uniform distribution of* F_r*; otherwise, if* A *is uncountable, then* $\mathbf{X}_r$ *is the normalized Lebesgue measure or another uniform measure on* F_r *(e.g.* [8]*) such for every* $r \in \mathbb{N}$ *the Lebesgue measure or some other uniform measure on* F_r *exists and is finite.* [16, p.32-37].

(3) *For all* $r \in \mathbb{N}$*,* $U'(F_r)$ *is positive and finite where* U' *is intrinsic. (For countably infinite* A*,* U' *is the counting measure where* $U'(F_r)$ *is positive and finite since* F_r *is finite. For uncountable* A*,* U' *is the Lebesgue or radon-nikodym derivative on some other uniform measure on* F_r *(e.g.* [8]*) where either of the measures on* F_r *are positive and finite.)*

$\{F_r\}_{r\in\mathbb{N}}$ *is then a* ***pre-structure*** *of* A*, since for every* $r \in \mathbb{N}$ *the sequence does not equal* A*, but "approaches"* A *as* r *increases.*

Definition 3 (**Expected value of Pre-Structure**). *If* $\{F_r\}_{r\in\mathbb{N}}$ *is a pre-structure of* A *(def.* 2*), then for* $r \in \mathbb{N}$*, if*

$$\mathbb{E}_{U'}[f(\mathbf{X}_r)] = \frac{1}{U'(F_r)} \int_{F_r} f\, d\mathbf{x} \qquad (1.2.2)$$

we then have that the expected value of the pre-structure could be described as $\mathbb{E}_{U'}\left[f(\mathbf{X}_r)\right] \to \mathbb{E}^{\star}_{U'}[f]$ *(def. 1) where:*

$$\forall(\epsilon > 0)\exists(N \in \mathbb{N})\forall(r \in \mathbb{N})\left(r \geq N \Rightarrow \left|\mathbb{E}_{U'}[f(\mathbf{X}_r)] - \mathbb{E}^{\star}_{U'}[f]\right| < \epsilon\right) \implies \tag{1.2.3}$$

$$\forall(\epsilon > 0)\exists(N \in \mathbb{N})\forall(r \in \mathbb{N})\left(r \geq N \Rightarrow \left|\frac{1}{U'(F_r)}\int_{F_r} f \, d\mathbf{x} - \mathbb{E}^{\star}_{U'}[f]\right| < \epsilon\right) \tag{1.2.4}$$

Definition 4 (**Uniform ε coverings of each term of the pre-structure**). *We define the uniform ε coverings of each term of the pre-structure $\{F_r\}_{r\in\mathbb{N}}$ (def. 2) or F_r as a group of pair-wise disjoint sets that cover F_r for every $r \in \mathbb{N}$, such the measure U' of each of the sets that cover F_r have the same value of $\varepsilon \in range(U')$, where $\varepsilon > \inf\left(range(U')\right)$ and the total sum for U' of the coverings is minimized. As a shortcut, if*

- *The element $t \in \mathbb{N}$*
- *The set $T \supset \mathbb{N}$ is arbitrary and uncountable.*

and set Ω is defined as:

$$\Omega = \begin{cases} \{1, \cdots, t\} & \text{if there are } t \text{ ways of writing uniform } \varepsilon \text{ coverings of } F_r \\ \mathbb{N} & \text{if there are countably infinite ways of writing uniform } \varepsilon \text{ coverings of } F_r \\ T & \text{if there are uncountable ways of writing uniform } \varepsilon \text{ coverings of } F_r \end{cases} \tag{1.2.5}$$

then for every $\omega \in \Omega$, the set of uniform ε coverings is defined using the notation $\mathcal{U}(\epsilon, F_r, \omega)$ where ω "enumerates" all possible uniform ε coverings of F_r for every $r \in \mathbb{N}$.

Definition 5 (**Sample of the uniform ε coverings of each term of the pre-structure**). *The sample of uniform ε coverings of each term of the pre-structure $\{F_r\}_{r\in\mathbb{N}}$ or F_r is the set of points, such for every $\varepsilon \in range(U')$ and $r \in \mathbb{N}$, we take a point from each of pair-wise disjoint set in the uniform ε coverings of F_r (def. 4). As a shortcut, if*

- *The element $k \in \mathbb{N}$*
- *The set $\mathcal{K} \supset \mathbb{N}$ is arbitrary and uncountable.*

and set Ψ_ω is defined as:

$$\Psi_\omega = \begin{cases} \{1, \cdots, k\} & \text{if there are } k \text{ ways of writing the sample of uniform } \varepsilon \text{ coverings of } F_r \\ \mathbb{N} & \text{if there are countably infinite ways of writing the sample of uniform } \varepsilon \text{ coverings of } F_r \\ \mathcal{K} & \text{if there are uncountable ways of writing the sample of uniform } \varepsilon \text{ coverings of } F_r \end{cases} \tag{1.2.6}$$

then for every $\psi \in \Psi_\omega$, the set of all samples of the set of uniform ε coverings is defined using the notation $\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)$, where ψ "enumerates" all possible samples of $\mathcal{U}(\epsilon, F_r, \omega)$.

Definition 6 (**Entropy on the sample of uniform coverings of each term of the pre-structure**). *Since there are finitely many points in the sample of the uniform ε coverings of each term of pre-structure $\{F_r\}_{r\in\mathbb{N}}$ (def. 5), we:*

(1) *Arrange the x-value of the points in the sample of uniform ε coverings from least to greatest. This is defined as:*

$$Ord(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))$$

(2) *Take the multi-set of the absolute differences between all consecutive pairs of elements in (1). This is defined as:*

$$\boldsymbol{\nabla} Ord(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))$$

(3) *Normalize (2) into a probability distribution, where for multi-set X, we have $|X|$ as the cardinality of all elements in the multi-set, including repeated ones. This is defined as:*

$$\mathbb{P}(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)) = \left\{y / \left|\boldsymbol{\nabla} Ord(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))\right| : y \in \boldsymbol{\nabla} Ord(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))\right\}$$

(4) *Take the entropy of (3), (for further reading, see* [12, p.61-95]*). This is defined as:*

$$E(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)) = -\sum_{x \in \mathbb{P}(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))} x \log_2 x$$

where (4) is the entropy on the sample of uniform coverings of F_r.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

Definition 7 (**Pre-Structure Converging Uniformly to** A). *For every $r \in \mathbb{N}$ (using def. 4, 5, and 6) if set A is finite:*

$$\lim_{\varepsilon \to \inf(range(U'))} \sup_{r\in\mathbb{N}} \sup_{\omega\in\Omega} \sup_{\psi\in\Psi_\omega} E(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)) \geq E(F_r)$$

and if set A is non-finite:

$$\lim_{\varepsilon \to \inf(range(U'))} \sup_{r\in\mathbb{N}} \sup_{\omega\in\Omega} \sup_{\psi\in\Psi_\omega} E(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)) = +\infty$$

we then say pre-structure $\{F_r\}_{r\in\mathbb{N}}$ ***converges uniformly*** *to A (or in shorter notation):*

$$F_r \overset{r\in\mathbb{N}}{\rightrightarrows} A \tag{1.2.7}$$

(Note we wish to define a uniform convergence of a sequence of sets to A since the definition is analogous to a uniform measure.)

Definition 8 (**Equivalent Pre-Structures**). *The pre-structures $\{F_r\}_{r\in\mathbb{N}}$ and $\{F'_j\}_{j\in\mathbb{N}}$ of A are* ***equivalent*** *if, from def. 3, ($\mathbb{E}_{U'}[f(\mathbf{X}_r)] \to \mathbb{E}^{\star}_{U'}[f]$ and $\mathbb{E}_{U'}[f(\mathbf{X}'_j)] \to \mathbb{E}^{\star\star}_{U'}[f]$) :*

$$\forall \left(f \in \mathbb{R}^A\right) \left(\mathbb{E}^{\star}_{U'}[f] = \mathbb{E}^{\star\star}_{U'}[f]\right)$$

Definition 9 (**Non-Equivalent Pre-Structures**). *The pre-structures $\{F_r\}_{r\in\mathbb{N}}$ and $\{F'_j\}_{j\in\mathbb{N}}$ of A are* ***non-equivalent*** *if, from def. 3, ($\mathbb{E}_{U'}[f(\mathbf{X}_r)] \to \mathbb{E}^{\star}_{U'}[f]$ and $\mathbb{E}_{U'}[f(\mathbf{X}'_j)] \to \mathbb{E}^{\star\star}_{U'}[f]$)*

$$\exists \left(f \in \mathbb{R}^A\right) \left(\mathbb{E}^{\star}_{U'}[f] \neq \mathbb{E}^{\star\star}_{U'}[f]\right)$$

Definition 10 (**Pre-Structures converging Sublinearly, Linearly, or Superlinearly to** A **compared to that of another Sequence**). *Suppose pre-structures $\{F_r\}_{r\in\mathbb{N}}$ and $\{F'_j\}_{j\in\mathbb{N}}$ are non-equivalent and converge uniformly to A; and suppose for every $\varepsilon \in range(U')$, where $\varepsilon > \inf(range(U'))$ and $r \in \mathbb{N}$:*

(a) *We take the cardinality of the sample of the uniform ε coverings of F_r (def. 5) divided by the smallest cardinality of the sample of the uniform ε coverings of F'_j (def. 5), where the entropy on the sample of uniform coverings on F'_j (def. 6) is larger than the entropy on the sample of uniform coverings on F_r (def. 6). In other words, if:*

$$\overline{|\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)|} = \tag{1.2.8}$$
$$\inf\left\{|\mathcal{S}(\mathcal{U}(\epsilon, F'_j, \omega'), \psi')| \,:\, j \in \mathbb{N},\ \omega' \in \Omega,\ \psi' \in \Psi_\omega,\ E(\mathcal{S}(\mathcal{U}(\epsilon, F'_j, \omega'), \psi')) \geq E(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))\right\}$$

then the ratio at the beginning of the paragraph is defined (using 1.2.8) as

$$\overline{\alpha}\,(\epsilon, r, \omega, \psi) = |\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))| \,/\, \overline{|\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)|} \tag{1.2.9}$$

(b) *We take the cardinality of the sample of uniform ε covering of F_r (def. 5) divided by the largest cardinality of the sample of the uniform ε covering of F'_j (def. 5), where the entropy on the sample of uniform coverings on F'_j (def. 6) is smaller then the entropy on the sample of uniform coverings on F_r (def. 6). In other words if:*

$$\underline{|\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)|} = \tag{1.2.10}$$
$$\sup\left\{|\mathcal{S}(\mathcal{U}(\epsilon, F'_j, \omega'), \psi')| \,:\, j \in \mathbb{N},\ \omega' \in \Omega,\ \psi' \in \Psi_\omega,\ E(\mathcal{S}(\mathcal{U}(\epsilon, F'_j, \omega'), \psi')) \leq E(\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))\right\}$$

then the ratio at the beginning of the paragraph is defined (using 1.2.10) as

$$\underline{\alpha}\,(\epsilon, r, \omega, \psi) = |\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi))| \,/\, \underline{|\mathcal{S}(\mathcal{U}(\epsilon, F_r, \omega), \psi)|} \tag{1.2.11}$$

(1) *If using equations* 1.2.9 *and* 1.2.11 *we have that:*

$$\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\overline{\alpha}(\epsilon,r,\omega,\psi)=\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\underline{\alpha}(\epsilon,r,\omega,\psi)=0$$

we say $\{F_r\}_{r\in\mathbb{N}}$ *converges uniformly to* A *at a* ***superlinear rate*** *to that of* $\{F'_j\}_{j\in\mathbb{N}}$.

(2) *If using equations* 1.2.9 *and* 1.2.11 *we have that:*

$$0<\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\overline{\alpha}(\epsilon,r,\omega,\psi)=\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\underline{\alpha}(\epsilon,r,\omega,\psi)<+\infty$$

we say $\{F_r\}_{r\in\mathbb{N}}$ *converges uniformly to* A *at a* ***linear rate*** *to that of* $\{F'_j\}_{j\in\mathbb{N}}$.

(3) *If using equations* 1.2.9 *and* 1.2.11 *we have that:*

$$\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\overline{\alpha}(\epsilon,r,\omega,\psi)=\lim_{\varepsilon\to\inf(\mathit{range}(U'))}\sup_{r\in\mathbb{N}}\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\underline{\alpha}(\epsilon,r,\omega,\psi)=+\infty$$

we say $\{F_r\}_{r\in\mathbb{N}}$ *converges uniformly to* A *at a* ***sublinear rate*** *to that of* $\{F'_j\}_{j\in\mathbb{N}}$.

I assume $\overline{\alpha}$ *and* $\underline{\alpha}$ *are always equal but I'm not sure how to prove this.*

1.3. Question on Preliminary Definitions.

(1) Are there "simpler" alternatives to either of the preliminary definitions? (Keep this in mind as we continue reading).

2. Main Question

Does there exist a unique extension (or a method that constructively defines a unique extension) of the expected value of f when the value's finite, using the uniform probability measure [16, p.32-37] on sets measurable in the Carathèodory sense, such we replace f with infinite or undefined expected values with f defined on a *chosen* pre-structure which depends on A where:

(1) The expected value of f on each term of the pre-structure is finite

(2) The pre-structure *converges uniformly* to A

(3) The pre-structure *converges uniformly* to A at a *linear or superlinear* rate to that of other non-equivalent pre-structures of A which satisfies (1) and (2).

(4) The *generalized expected value of* f on the pre-structure (i.e. an extension of def. 3 to answer the full question) satisfies (1), (2), and (3) while having an unique & finite value.

(5) A choice function is defined that chooses a pre-structure from A such the following satisfies (1), (2), (3), and (4) for the largest *possible* subset of $\mathbb{R}^A$.

(6) If there is more than one choice function that satisfies (1), (2), (3), (4) and (5), we choose the choice function with the "simplest form", meaning for a general pre-structure of A, when each choice function is fully expanded, we take the choice function with the fewest variables/numbers (excluding those with quantifiers).

How do we answer this question? (See §3.1 & §3.3 for a partial answer.)

3. Informal Attempt to Answer Main Question

(I advise using computer programmings such as Mathematica, Python, JavaScript, or Matlab to understand the definitions of the answer below.)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

3.1. **Choice Function.** Suppose $\mathbb{S}'(A)$ is the set of all pre-structures of A which satisfy criteria (1) and (2) of the main question where the expected value of the pre-structures as they converge uniformly to A is finite, such the pre-structure $\{F''_r\}_{r\in\mathbb{N}} \in \mathbb{S}'(A)$ should be a sequence of sets that satisfy criteria (1), (2), (3), and (4) of the main question and pre-structure $\{F'_j\}_{j\in\mathbb{N}}$ is an element of $S'(A)$ but not an element of the set of equivelant pre-structures of $\{F_r\}_{r\in\mathbb{N}}$. (Note I'm unsure if the choice function I'll define satisfies criteria (5) and (6)).

Further note from (a), with equation 1.2.8 in def. 10, if we take:

$$\overline{|\mathcal{S}(\mathcal{U}(\epsilon,F''_r,\omega),\psi)|} = \inf\left\{|\mathcal{S}(\mathcal{U}(\epsilon,F'_j,\omega'),\psi')| \,:\, j\in\mathbb{N},\ \omega'\in\Omega,\ \psi'\in\Psi_\omega,\ E(\mathcal{S}(\mathcal{U}(\epsilon,F'_j,\omega'),\psi'))\geq E(\mathcal{S}(\mathcal{U}(\epsilon,F_r,\omega),\psi))\right\} \tag{3.1.1}$$

and from (b), with equation 1.2.10 in def. 10, we take:

$$\underline{|\mathcal{S}(\mathcal{U}(\epsilon,F''_r,\omega),\psi)|} = \sup\left\{|\mathcal{S}(\mathcal{U}(\epsilon,F'_j,\omega'),\psi')| \,:\, j\in\mathbb{N},\ \omega'\in\Omega,\ \psi'\in\Psi_\omega,\ E(\mathcal{S}(\mathcal{U}(\epsilon,F'_j,\omega'),\psi'))\leq E(\mathcal{S}(\mathcal{U}(\epsilon,F_r,\omega),\psi))\right\} \tag{3.1.2}$$

Then, using def. 5 with equations 3.1.1 and 3.1.2, if:

$$\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\mathcal{S}(\mathcal{U}(\epsilon,F''_r,\omega),\psi) = \mathcal{S}'(\varepsilon,F''_r) = \mathcal{S}' \tag{3.1.3}$$

$$\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\overline{\left|\mathcal{S}(\mathcal{U}(\epsilon,F''_r,\omega),\psi)\right|} = \overline{\left|\mathcal{S}'(\varepsilon,F''_r)\right|} = \overline{|\mathcal{S}'|} \tag{3.1.4}$$

$$\sup_{\omega\in\Omega}\sup_{\psi\in\Psi_\omega}\underline{\left|\mathcal{S}(\mathcal{U}(\epsilon,F''_r,\omega),\psi)\right|} = \underline{\left|\mathcal{S}'(\varepsilon,F''_r)\right|} = \underline{\left|\mathcal{S}'\right|} \tag{3.1.5}$$

where, using absolute value function $||\cdot||$, we take:

$$S(r) = \left(\sup(F''_{r+1}) - \sup(F''_r)\right)\left(\inf(F''_r) - \inf(F''_{r+1})\right)\left|\left|\left(\inf(F''_r) - \inf(F''_{r+1})\right)\left(\sup(F''_{r+1}) - \sup(F''_r) - 1\right)\right|\right| \tag{3.1.6}$$

such that

$$T(r) = \left(\sup(F''_{r+1})\inf(F''_r) - \sup(F''_r)\inf(F''_{r+1})\right)\Big(\left(\inf(F''_r) - \inf(F''_{r+1})\right) - \left(\sup(F''_{r+1}) - \sup(F''_r)\right) - 1\Big) \left(\inf(F''_r) - \inf(F''_{r+1})\right)\left(\sup(F''_{r+1}) - \sup(F''_r)\right) \tag{3.1.7}$$

then, using equations 3.1.3, 3.1.4, 3.1.5, 3.1.6 and 3.1.7, and the nearest integer function $[\cdot]$ we want:

$$K(\varepsilon,F''_r) = ||1-S(r)||\left(\left|\left|\frac{\left|\mathcal{S}'\right|\left(1+\left[\frac{|\mathcal{S}'|\left(\underline{|\mathcal{S}'|}+2|\mathcal{S}'|\right)}{\left(\underline{|\mathcal{S}'|}+|\mathcal{S}'|\right)\left(\underline{|\mathcal{S}'|}+|\mathcal{S}'|+\overline{|\mathcal{S}'|}\right)}\right]\right)\left(1+\left[\underline{|\mathcal{S}'|}/|\mathcal{S}'|\right]\right)}{\left(1+\left[|\mathcal{S}'|/\overline{|\mathcal{S}'|}\right]\right)\left(1+\left[\underline{|\mathcal{S}'|}/\overline{|\mathcal{S}'|}\right]\right)} - \left|\mathcal{S}'\right|\right|\right| + \left|\mathcal{S}'\right|\right) - T(r) \tag{3.1.8}$$

such, using equations 3.1.5 and 3.1.8, if set $S''(A)\subseteq S'(A)$ and $\mathcal{P}(\cdot)$ is the power-set, then set $C(A)$ is the largest element of:

$$\Big\{S''(A)\subseteq S'(A) \,:\, \forall(\epsilon_1>0)\exists(M\in\mathbb{N})\forall(\varepsilon\in\text{range}(U'))\exists(j\in\mathbb{N})\forall(r\in\mathbb{N})\exists(\{F''_r\}\in\mathbb{S}''(A)) \left(\inf(\text{range}(U'))<\varepsilon\leq M,\ r\geq j \Rightarrow \mathcal{S}'(\varepsilon,F''_r) - K(\varepsilon,F''_r) - \inf_{\{F_g\}\in\mathbb{S}'(A)}\left(\mathcal{S}'(\varepsilon,F_g) - K(\varepsilon,F_g)\right) < \epsilon_1\right)\Big\} \subseteq \mathcal{P}(S'(A)) \tag{3.1.9}$$

w.r.t to inclusion, such the **choice function** is $C(A)$ when the following contains just one element.

Otherwise, suppose for $k\in\mathbb{N}$, $C^k(A)$ represents the k-th iteration of the choice function of A, e.g. $C^3(A) = C(C(C(A)))$, where the infinite iteration of $C(A)$ (if it exists) is $\lim_{k\to\infty}C^k(A) = C^\infty(A)$. Therefore, when taking following:

$$C'(A) = \begin{cases} C(A) & \text{if } C(A) \text{ contains one element} \\ C^j(A) & \text{if } j\in\mathbb{N}, \text{ such for all } k\geq j,\ C^k(A) \text{ contains one element} \\ C^\infty(A) & \text{if it exists, and } C^\infty(A) \text{ contains one element} \end{cases} \tag{3.1.10}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

$C'(A)$ is the choice function.

3.2. **Questions on Choice Function.**

(1) What unique pre-structure would $C'(A)$ contain (if it exists) for:

(a) $\mathbb{Z}$ where if $\{F''_r\}_{r\in\mathbb{N}} \in C'(\mathbb{Z})$, we want $\{F''_r\}_{r\in\mathbb{N}} = \{\{m \in \mathbb{Z} : -r \le m \le r''\}\}_{r\in\mathbb{N}}$

(b) $\mathbb{Q}$ where if $\{F''_r\}_{r\in\mathbb{N}} \in C'(\mathbb{Q})$, we want $\{F''_r\}_{r\in\mathbb{N}} = \{\{s/r! : s \in \mathbb{Z},\ -r\cdot r! \le s \le r \cdot r!\}\}_{r\in\mathbb{N}}$

(c) $\mathbb{R}$ where we're not sure what $\{F''_r\}_{r\in\mathbb{N}} \in C'(\mathbb{R})$ would be in this case. What would $\{F''_r\}_{r\in\mathbb{N}}$ be if it's unique?

3.3. **Generalized Expected Values.** In case $C'(A)$ from equation 3.1.10 does not exist, using $C(A)$ from equation 3.1.9 (i.e. in section 3.1), if the image of f under A is $f[A] := \{f(x) : x \in A\}$, we then take $\{F''_r\} \in C(f[A])$ and also take the pre-image under f of F_r (defined as $f^{-1}\left[F''_r\right] := \left\{x \in A : f(x) \in F''_r\right\}$) such from def. 7:

$$f^{-1}\left[F''_r\right] \overset{r\in\mathbb{N}}{\rightrightarrows} A$$

However, the expected value of $f^{-1}\left[F''_r\right]$ (def. 3) may be infinite (e.g. unbounded f). Hence, for every $r \in \mathbb{N}$, we take $\left\{\left\{F'''_{t_r}\right\}_{t_r\in\mathbb{N}}\right\}_{r\in\mathbb{N}} \in C(f^{-1}\left[F''_r\right])$ such (using def. 7):

$$\forall(r\in\mathbb{N})\left(F'''_{t_r} \overset{t_r\in\mathbb{N}}{\rightrightarrows} F''_r\right)$$

Thus, if there exists a unique and finite $\ddot{\mathbb{E}}[f]$ where:

$$\forall(\epsilon>0)\exists(N\in\mathbb{N})\forall(r\in\mathbb{N})\forall\left(t_r\in\mathbb{N}\right)\forall\left(\left\{\left\{F'''_{t_r}\right\}_{t_r\in\mathbb{N}}\right\}_{r\in\mathbb{N}} \in C\left(f^{-1}\left[F''_r\right]\right)\right) \tag{3.3.1}$$
$$\left(r \ge N, t_r \ge N \Rightarrow \frac{1}{U'\left(F'''_{t_r}\right)}\int_{F'''_{t_r}} f\, d\mathbf{x} - \ddot{\mathbb{E}}[f] < \epsilon\right)$$

Then $\ddot{\mathbb{E}}[f]$ is the **generalized expected value w.r.t choice function** C, which answers criteria (1), (2), (3), (4), (perhaps (5)) of the question in §2; however, there is still a chance that the equation 3.3.1 fails to give an unique $\ddot{\mathbb{E}}[f]$. Hence; if $k \in \mathbb{N}$, we take the k-th iteration of the choice function C in 3.1.9, such there exists a $j \in \mathbb{N}$, where for all $k \ge j$, the new expected value $\mathbb{E}^{\dagger}[f]$ (or the **generalized expected value w.r.t finitely iterated** C) is unique and finite.

Hence, if the k-th iteration of C is represent as $C^{[k]}$ (where e.g. $C^3(f^{-1}[F''_r]) = C(C(C(f^{-1}[F''_r]))))$, we want a unique $\mathbb{E}^{\dagger}[f]$ where:

$$\forall(\epsilon>0)\exists(N\in\mathbb{N})\forall(r\in\mathbb{N})\forall\left(t_r\in\mathbb{N}\right)\exists\left(j\in\mathbb{N}\right)\forall(k\in\mathbb{N})\left(k\ge j \Rightarrow\right. \tag{3.3.2}$$
$$\left.\forall\left(\left\{\left\{F'''_{t_r}\right\}_{t_r\in\mathbb{N}}\right\}_{r\in\mathbb{N}} \in C^{[j]}\left(f^{-1}\left[F''_r\right]\right)\right)\left(r\ge N, t_r \ge N \Rightarrow \frac{1}{U'\left(F'''_{t_r}\right)}\int_{F'''_{t_r}} f\, d\mathbf{x} - \mathbb{E}^{\dagger}[f] < \epsilon\right)\right)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

If this still does not give a unique and finite expected value, we then take the **most generalized expected value w.r.t an infinitely iterated** C i.e. $\mathbb{E}^{\ddagger}[f]$ where if the *infinite iteration* of C is stated as $\lim_{k\to\infty} C^{[k]}(f[A]) = C^{\infty}(f[A])$, we then take:

$$\forall(\epsilon > 0)\exists(N \in \mathbb{N})\forall(r \in \mathbb{N})\forall(t_r \in \mathbb{N}) \tag{3.3.3}$$
$$\forall\left(\left\{\left\{F_{t_r}'''\right\}_{t_r\in\mathbb{N}}\right\}_{r\in\mathbb{N}} \in C^{\infty}\left(f^{-1}\left[F_r''\right]\right)\right)\left(r \geq N, t_r \geq N \Rightarrow \frac{1}{U'\left(F_{t_r}'''\right)}\int_{F_{t_r}'''} f\, d\mathbf{x} - \mathbb{E}^{\ddagger}[f] < \epsilon\right)$$

However, the expected values $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, and $\mathbb{E}^{\ddagger}[f]$ in equations 3.3.1, 3.3.2, and 3.3.3 (respectively) should only be attempted for functions where *the expected value is infinite or undefined* or for **worst-case functions**—poorly behaved $f : A \to \mathbb{R}$ (where for $n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, and f is a function) defined on infinite points covering an infinite expanse of space. For example:

(1) For a worst-case f defined on countably infinite A (e.g. countably infinite "pseudo-random points" non-uniformly scattered across the real plane), one might typically use $\ddot{\mathbb{E}}[f]$ from equation 3.3.1 (since countable sets might need just one iteration of C for the generalized expected value to be unique); otherwise, one may use $\mathbb{E}^{\dagger}[f]$ from equation 3.3.2 for finite iterations of C.

(2) For a worst-case f defined on uncountable A, we might have to use $\mathbb{E}^{\ddagger}[f]$ from equation 3.3.3 as the function is so difficult to analyze. We can imagine this function as an uncountable number of "pseudo-random" points non-uniformly generated on a subset of the real plane (see §4.1 for a visualization.)

Note no matter how generalized and "meaningful" the extension of an expected value is, there will always be an f where the expected value does not exist.

3.4. Questions Regarding The Answer.

(1) Using prevalence and shyness [14, 11], can we say the set of f where $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, or $\mathbb{E}^{\ddagger}[f]$, from equations 3.3.1, 3.3.2 and 3.3.3 respectively, have unique and finite values that form either a **prevalent** or *neither prevalent nor shy* subset of $\mathbb{R}^A$? (If the subset is *prevalent*, this implies that either of the generalized expected values can be unique and finite for a "large" subset of $\mathbb{R}^A$; however, if the subset is *neither prevalent nor shy* we need a more precise definition of "size" which takes "an exact probability that the expected values are unique & finite"—some examples (which are shown in this answer [9]) being:

   (a) Fractal Dimension notions
   (b) Kolmogorov Entropy
   (c) Baire Category and Porosity

(2) There might be a total of 120 variables in the choice function C (excluding quantifiers). Is there a choice function with fewer variables (ignoring quantifiers) which answers criteria (1), (2), (3) & (4) of the main question in §2 for a "larger" subset of $\mathbb{R}^A$? (This might be impossible to answer since such a solution cannot be shown with prevalence or shyness [14, 11])—therefore, we need a more precise version of "size" with some examples, again, shown in [9].

(3) If question (2) is correct, what is the choice function C using the most generalized expected value $E^{\ddagger}[f]$ that fully answers the question in §2?

(4) Can either $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, or $\mathbb{E}^{\ddagger}[f]$ from equations 3.3.1, 3.3.2 and 3.3.3 respectively (when A is the set of all Liouville numbers [6] and $f = \mathrm{id}_A$) give a finite value? What would the value be?

(5) Similar to how definition 11 in §4 approximates the expected value in definition 1, how do approximate $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, and $\mathbb{E}^{\ddagger}[f]$ from equations 3.3.1, 3.3.2 and 3.3.3 respectively?

(6) Can programming be used to estimate $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, and $\mathbb{E}^{\ddagger}[f]$ from equations 3.3.1, 3.3.2 and 3.3.3 respectively (if an unique/finite result of either of the expected values exist)?

3.5. **Applications.**

(1) In Quanta magazine [3], Wood writes on Feynman Path Integrals: "No known mathematical procedure can *meaningfully average*[1] an infinite number of objects covering an infinite expanse of space in general. The path integral is more of a physics philosophy than an exact mathematical recipe."—despite Wood's statement, mathematicians Bottazzi E. and Eskew M. [5] found a constructive solution to the statement using integrals defined on filters over families of finite sets; however, the solution was not unique as one has to choose a value in a partially ordered ring of infinite and infinitesimal elements. In addition, although there were ways of preventing the use of the axiom of choice (within their integral), the axiom was still required for certain cases.

   (a) Perhaps, if Botazzi's and Eskew's Filter integral [5] is not enough to solve Wood's statement, could we replace the path integral with expected values $\ddot{\mathbb{E}}[f]$, $\mathbb{E}^{\dagger}[f]$, and $\mathbb{E}^{\ddagger}[f]$ from equations 3.3.1, 3.3.2 and 3.3.3 respectively? (See, again, §4.1 for a visualization of Wood's statement.)

(2) As stated in §1.1, "when the Lebesgue measure of A, measurable in the Caratheodory sense, has zero or infinite volume (or undefined measure), there may be multiple, conflicting ways of defining a "natural" uniform measure on A." This is an example of Bertand's Paradox which shows, "the principle of indifference (that allows equal probability among all possible outcomes when no other information is given) may not produce definite, well-defined results for probabilities if applied uncritically, when the domain of possibilities (i.e. the sample space) is infinite [15].

Using §3.2, perhaps if we take (from def. 3.1.9):

$$C'(A)=\begin{cases} C(A) & \text{if } C(A) \text{ contains one element} \\ C^{j}(A) & \text{if } j\in\mathbb{N}, \text{ such for all } k\geq j,\, C^{k}(A) \text{ contains one element} \\ C^{\infty}(A) & \text{if it exists, and } C^{\infty}(A) \text{ contains one element} \end{cases}$$

then for $\{F_r\}_{r\in\mathbb{N}}\in C'(A)$, if we want $S\subseteq A$ and we get the following:

$$\forall(\epsilon>0)\exists(N\in\mathbb{N})\forall(r\in\mathbb{N})\left(r\geq N \implies \frac{U'(S\cap F_r)}{U'(F_r)}-\mathscr{U}(S)<\epsilon\right) \tag{3.5.1}$$

Then $\mathscr{U}(S)$ might serve as a solution to Bertand's Paradox (unless there's a simpler $C'(A)$ and $\{F_r\}_{r\in\mathbb{N}}\in C'(A)$ which helps solve the main question in §2).

Now consider the following:

   (a) How do we apply $\mathscr{U}(S)$ (or a simpler solution) to the usual example which demonstrates the Bertand's Paradox as follows: for an equilateral triangle (inscribed in a circle), suppose a chord of the circle is chosen at random—what is the probability that the chord is longer than a side of the triangle? [4] (According to Bertand's Paradox there are three arguments which correctly use the principle of

[1] Meaningful Average—The measure inside the average is canonical when the measure is derived from the radon-nikodym derivative of the uniform probability measure [16, p. 32-37]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 March 2023 doi:10.20944/preprints202302.0367.v5

indifference yet give different solutions to this problem [4]:

(i) The “random endpoints" method: Choose two random points on the circumference of the circle and draw the chord joining them. To calculate the probability in question imagine the triangle rotated so its vertex coincides with one of the chord endpoints. Observe that if the other chord endpoint lies on the arc between the endpoints of the triangle side opposite the first point, the chord is longer than a side of the triangle. The length of the arc is one-third of the circumference of the circle, therefore the probability that a random chord is longer than a side of the inscribed triangle is $1/3$.

(ii) The "random radial point" method: Choose a radius of the circle, choose a point on the radius, and construct the chord through this point and perpendicular to the radius. To calculate the probability in question imagine the triangle rotated so a side is perpendicular to the radius. The chord is longer than a side of the triangle if the chosen point is nearer the center of the circle than the point where the side of the triangle intersects the radius. The side of the triangle bisects the radius, therefore the probability a random chord is longer than a side of the inscribed triangle is $1/2$.

(iii) The "random midpoint" method: Choose a point anywhere within the circle and construct a chord with the chosen point as its midpoint. The chord is longer than a side of the inscribed triangle if the chosen point falls within a concentric circle of radius $1/2$ the radius of the larger circle. The area of the smaller circle is one-fourth the area of the larger circle, therefore the probability a random chord is longer than a side of the inscribed triangle is $1/4$.

4. Glossary

4.1. **Example of Case (2) of Worst Case Functions.** (If the explanation below is difficult to understand, see this visualization to accompany the explanation [1], then when changing the sliders each time, wait a couple of seconds for the graph to load.)

We wish to create a function that appears to be a “pseudo-randomly" distributed but has infinite points that are non-uniform (i.e. does not have complete spatial randomness [13]) in the sub-space of $\mathbb{R}^2$, where the expected value or integral of the function w.r.t uniform probability measure [16][p.32-37] is non-obvious (i.e. not the center of the space the function covers nor the area of that space).

Suppose for real numbers x_1, x_2, y_1 and y_2, we generate an uncountable number of "nearly pseudo-random" points that are non-uniform in the subspace $[x_1, x_2] \times [y_1, y_2] \subseteq \mathbb{R}^2$.

We define the function as $f : [x_1, x_2] \to [y_1, y_2]$.

Now suppose $b \in \{2, 3, \cdots, 10\}$ where the base-b expansion of real numbers, in interval $[x_1, x_2]$, have infinite decimals that approach x from the right side so when $x_1 = x_2$ we get $f(x_1) = f(x_2)$.

Furthermore, for $\mathbb{N} \cup \{0\} = \mathbb{N}_0$, if $r \in \mathbb{N}_0$ and $\text{digit}_b : \mathbb{R} \times \mathbb{Z} \to \{0, 1, \cdots, b-1\}$ is a function where $\text{digit}_b(x, r)$ takes the digit in the b^r-th decimal fraction of the base-b expansion of x (e.g. $\text{digit}_{10}(1.789, 2) = 8$), then $\left\{{g_r}'\right\}_{r\in\mathbb{N}_0}$ is a sequence of functions such that ${g_r}' : \mathbb{N}_0 \to \mathbb{N}_0$ is defined to be:

$$g_r'(x) = \left\lceil \frac{10}{b} \sin(rx) + \frac{10}{b} \right\rceil \tag{4.1.1}$$

then for some large $k \in \mathbb{N}$ and $x_1, x_2 \in \mathbb{R}$, the intermediate function (before f) or $f_1 : [x_1, x_2] \to \mathbb{R}$ is defined to be

$$f_1(x) = \left|\left(\sum_{r=0}^{\infty} g'_{r+1}\left(\sum_{p=r}^{r+k} \text{digit}_b(x,p)\right)\Big/ b^r\right) - 10\right| = \left|\left(\left(\sum_{r=0}^{\infty}\left[\frac{10}{b}\sin\left((r+1)\left(\sum_{p=r}^{r+k}\text{digit}_b(x,p)\right)\right)+\frac{10}{b}\right]\right)\Big/ b^r\right) - 10\right| \tag{4.1.2}$$

where the points in f_1 are "almost pseudo-randomly" and non-uniformly distributed on $[x_1, x_2] \times [0, 10]$. What we did was convert every digit of the base-b expansion of x to a pseudo-random number that is non-equally likely to be an integer, including and also in-between, 0 and $(10 \cdot 10^s)/b$. Furthermore, we make the function truly "appear pseudo-random", by adding the b^r-th decimal fraction with the next k decimal fractions; however, we also want to control the end-points of $[0, 10^{s+1}]$ such if $y_1, y_2 \in \mathbb{R}$, we convert $[x_1, x_2] \times [0, 10]$ to $[x_1, x_2] \times [y_1, y_2]$ by manipulating equation 4.1.2 to get:

$$f(x) = y_2 - \frac{y_2 - y_1}{10} f_1(x) \quad y_2 - \left(\frac{y_2-y_1}{10}\right)\left|\left(\left(\sum_{r=0}^{\infty}\left[\frac{10}{b}\sin\left((r+1)\left(\sum_{p=r}^{r+k}\text{digit}_b(x,p)\right)\right)+\frac{10}{b}\right]\right)\Big/ b^r\right) - 10\right| \tag{4.1.3}$$

such the larger k is, the more pseudo-random the distribution of points in f in the space $[x_1, x_2] \times [y_1, y_2]$ but unlike most distributions of such points, f is uncountable.

4.2. **Question Regarding Section 4.1.** Let's be more specific, suppose for the function in equation 4.1.3 of §4.1, we have:

- $b = 3$
- $[x_1, x_2] \times [y_1, y_2] = [0, 1] \times [0, 1]$
- $k = 100$

(one can try simpler parameters); what is the expected value using either $\mathbb{E}^{\dagger}[f]$ or $\mathbb{E}^{\ddagger}[f]$ from equations 3.3.2 and 3.3.3 respectively if the answer is finite and unique?

What about for f in general (i.e. in terms of b, x_1, x_2, y_1, y_2 and k)?

(Note if $x_1, y_1 \to -\infty$ and $x_2, y_2 \to \infty$, then the function is an explicit example of the function that Wood [2] describes in Quanta Magazine)

4.3. **Approximating the Expected Value.**

Definition 11 (**Approximating the Expected Value**). *In practice, the computation of this expected value may be complicated if the set A is complicated. If analytic integration does not give a closed-form solution then a general and relatively simple way to compute the expected value (up to high accuracy) is with importance sampling. To do this, we produce values $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_M \sim$ IID g for some density function g with support $A \subseteq$ support$(g) \subseteq \mathbb{R}^n$ (hopefully with support fairly close to A) and we use the estimator:*

$$\hat{\mu}_M \equiv \frac{\sum_{i=1}^{M} \mathbb{I}(\mathbf{X}_i \in A) \cdot f(\mathbf{X}_i)/g(\mathbf{X}_i)}{\sum_{i=1}^{M} \mathbb{I}(\mathbf{X}_i \in A)/g(\mathbf{X}_i)} \tag{4.3.1}$$

From the law of large numbers, we can establish that $\mathbb{E}[f(\mathbf{X})] = \lim_{M\to\infty} \hat{\mu}_M$ so if we take M to be large then we should get a reasonably good computation of the expected value of interest.

[2]Wood wrote on Feynman Path Integrals: "No known mathematical procedure can *meaningfully average* [1] an infinite number of objects covering an infinite expanse of space in general."

Note importance sampling requires three things:

(1) *We need to know when point x is in set A or not*
(2) *We need to be able to generate points from a density g that is on a support that covers A but is not too much bigger than A*
(3) *We have to be able to compute $f(x)$ and $g(x)$ for each point $x \in A$*

References

[1] Krishnan B. Visualization of uncountable number of psuedo-random points generate on subset of the real plane, 2023. https://www.wolframcloud.com/obj/4e78f594-1578-402a-a163-ebb16319ada2.
[2] Patrick B. John Wiley & Sons, New York, 3 edition, 1995. https://www.colorado.edu/amath/sites/default/files/attached-files/billingsley.pdf.
[3] Wood C. Mathematicians prove 2d version of quantum gravity really works. *Quanta Magazine*. https://www.quantamagazine.org/mathematicians-prove-2d-version-of-quantum-gravity-really-works-20210617.
[4] Alon Drory. Failure and uses of jaynes' principle of transformation groups. *Foundations of Physics*, 45(4):439–460, feb 2015. https://arxiv.org/pdf/1503.09072.pdf.
[5] Bottazi E. and Eskew M. Integration with filters. https://arxiv.org/pdf/2004.09103.pdf.
[6] Adam Grabowski and Artur Kornilowicz. Introduction to liouville numbers. *Formalized Mathematics*, 25, 01 2017.
[7] Michael Greinecker (https://mathoverflow.net/users/35357/michael greinecker). Demystifying the caratheodory approach to measurability. MathOverflow. URL:https://mathoverflow.net/q/34007 (version: 2010-07-31).
[8] Mark McClure (https://mathoverflow.net/users/46214/mark mcclure). Integral over the cantor set hausdorff dimension. MathOverflow. https://mathoverflow.net/q/235609 (version: 2016-04-07).
[9] Dave L. Renfro (https://math.stackexchange.com/users/13130/dave-l renfro). Proof that neither "almost none" nor "almost all" functions which are lebesgue measurable are non-integrable. Mathematics Stack Exchange. https://math.stackexchange.com/q/4623168 (version: 2023-01-21).
[10] Ben (https://stats.stackexchange.com/users/173082/ben). In statistics how does one find the mean of a function w.r.t the uniform probability measure? Cross Validated. https://stats.stackexchange.com/q/602939 (version: 2023-01-24).
[11] Brian R. Hunt. Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. 1992. https://arxiv.org/abs/math/9210220.
[12] Gray M. Springer New York, New York [America];, 2 edition, 2011. https://ee.stanford.edu/~gray/it.pdf.
[13] Rokach L. Maimon O. Springer New York, New York [America];, 2 edition, 2010. https://doi.org/10.1007/978-0-387-09823-4.
[14] William Ott and James A. Yorke. Prevelance. *Bulletin of the American Mathematical Society*, 42(3):263–290, 2005. https://www.ams.org/journals/bull/2005-42-03/S0273-0979-05-01060-8/S0273-0979-05-01060-8.pdf.
[15] Nicholas Shackel. Bertrand's paradox and the principle of indifference. *Philosophy of Science*, 74(2):150–175, 2007. https://orca.cardiff.ac.uk/id/eprint/3803/1/Shackel%20Bertrand's%20paradox%205.pdf.
[16] Leinster T. and Roff E. The maximum entropy of a metric space. https://arxiv.org/pdf/1908.11184.pdf.

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