Open Journal of Mathematical Sciences

PSR Press

Article

A sampling-induced statistical Riemann sum functional

Fatih Nuray

Department of Mathematics, Afyon Kocatepe University, Afyonkarahisar, Turkey; fnuray@aku.edu.tr

Received: 23 Mar 2026; Revised: 25 Jun 2026; Accepted: 30 Jun 2026; Published: 19 Jul 2026

Abstract: In this paper, we introduce a summability-based functional, called the *sampling-induced statistical Riemann sum functional*, generated by uniform partitions with a fixed sampling rule. For a function $f : [a, b] \to \mathbb{R}$, we consider the sequence of uniformly sampled right-endpoint Riemann sums and define the functional value as the statistical limit of this sequence, whenever it exists. We show that every classically Riemann integrable function yields the same value as the classical integral, while the converse fails. Unlike the classical Riemann integral, which is tag-independent and intrinsic, the functional studied here is sampling-dependent and may change under modifications of the function on countable sets. We illustrate these phenomena with examples, discuss the relationship with the statistical derivative, and highlight the structural compatibility issues that arise when attempting a statistical Fundamental Theorem of Calculus. The paper concludes with open problems concerning sampling-scheme independence and the characterization of functions for which the functional is tag-invariant.

Keywords: statistical convergence, statistical Riemann sum functional, uniform partitions, sampling schemes, statistical derivative

MSC: Primary 40G15; Secondary 40A35, 26A42.

---

1. Introduction

In mathematical analysis, the concepts of limit, continuity, derivative, and integral form the foundation of calculus. In recent decades, various generalizations of these classical notions have been introduced using summability methods. Statistical convergence, first introduced by Steinhaus [1] and Fast [2], and later studied further by Fridy [3] and Šalát [4], has proven to be a fruitful alternative to ordinary convergence. Deep structural characterizations of statistical convergence have been developed by Fridy [5] and Miller [6], leading to the notions of statistical limit, statistical continuity, and statistical derivative.

The concept of statistical derivative was introduced by Nuray [7]. As we shall see in §2, the definition given there, which requires the statistical limit of difference quotients along *every* sequence $x_n \to 0^+$, is actually equivalent to the classical right derivative. This observation motivates us to introduce a weaker notion in §7, tailored to the discrete sampling scale of our construction. This weaker derivative is designed to be structurally compatible with the sampling-induced integral functional.

In this paper, we address the question of defining a summability-based integral by introducing a functional based on Riemann sums generated by uniform partitions with a fixed sampling rule (specifically, right-endpoint sampling). We then take the statistical limit of this sequence. This construction mirrors the idea of using statistical limits in derivative definitions.

Comparison with existing literature. The use of summability methods to generalize the Riemann integral is not new. However, the present work differs from existing approaches in three essential respects:

1. We employ *statistical convergence* rather than Cesàro means or other matrix summability methods. Statistical convergence is a more flexible tool, allowing divergent sequences to be summable as long as the set of indices where they deviate has natural density zero.
2. We explicitly emphasize and study the *sampling-rule dependence* of the resulting functional. To the best of our knowledge, this dependence is not discussed in classical summability-based integration theories.
3. Our construction is tied to a *fixed partition sequence* (the uniform partitions) together with a fixed tagging rule. This contrasts with ideal convergence approaches that often consider all partitions simultaneously.

Thus, the object introduced here is not intended as a generalization of the classical integral but rather as a new sampling-dependent functional that captures information about functions on specific countable dense subsets determined by the sampling scheme.

Terminology. Throughout the paper, we avoid the potentially misleading term "statistical integral" in the sense of an intrinsic integral. Instead, we refer to the *sampling-induced statistical Riemann sum functional*, denoted by $\mathrm{st}\int_a^b f(x)\,\mathrm{d}x$ with the understanding that the uniform right-endpoint sampling scheme is fixed unless otherwise stated. When we need to emphasize the sampling rule σ, we write $\mathrm{st}\int_a^b f(x)\,\mathrm{d}x\,(\sigma)$.

The paper is organized as follows. In §2, we recall necessary background on statistical convergence and statistical derivative, carefully distinguishing between one-sided and symmetric versions, and clarifying the equivalence of the usual definition with the classical right derivative. In §3, we define the sampling-induced functional and establish its basic properties, including the failure of additivity with a genuine counterexample. §4 clarifies the relationship with the classical Riemann integral. §5 presents examples illustrating sensitivity to sampling rules and establishes conditions under which tag-independence holds. §6 discusses measure-zero sensitivity with a rigorous quantitative proposition. In §7, we address the structural compatibility issues related to a possible statistical analogue of the Fundamental Theorem of Calculus, presenting this as a future research direction rather than a completed theorem. Finally, Section 8 concludes with refined open problems.

2. Preliminaries

We begin by recalling the concept of statistical convergence for sequences of real numbers. For detailed treatments and structural properties, we refer to [1–3,5,6].

Definition 1. A sequence $x = (x_k)$ of real numbers is said to be *statistically convergent* to a number L if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \left|\{k \le n : |x_k - L| \ge \varepsilon\}\right| = 0.$$

In this case, we write $\text{st-}\lim_{k\to\infty} x_k = L$.

Definition 2. Let $E \subseteq \mathbb{N}$. If the limit

$$d(E) = \lim_{n\to\infty} \frac{|E \cap \{1,\ldots,n\}|}{n},$$

exists, it is called the *natural density* of E. The *upper natural density* is defined as

$$\overline{d}(E) = \limsup_{n\to\infty} \frac{|E \cap \{1,\ldots,n\}|}{n}.$$

Statistical convergence generalizes ordinary convergence: every convergent sequence is statistically convergent to the same limit, but the converse is not true. For example, the sequence defined by

$$x_k = \begin{cases} 1, & \text{if } k \text{ is a perfect square,} \\ 0, & \text{otherwise,} \end{cases}$$

is statistically convergent to 0 (since the set of perfect squares has density 0) but is not convergent in the ordinary sense.

Remark 1. A statistically convergent sequence need not be bounded in the ordinary sense. For example, the sequence

$$x_n = \begin{cases} n, & \text{if } n \text{ is a perfect square,} \\ 0, & \text{otherwise,} \end{cases}$$

is statistically convergent to 0 (perfect squares have density 0) but is unbounded. However, every statistically convergent sequence is *statistically bounded*: there exists $M > 0$ such that the set $\{n : |x_n| > M\}$ has density zero.

Definition 3 (Nuray [7])**.** A function $f : \mathbb{R} \to \mathbb{R}$ has a *statistical derivative* $w \in \mathbb{R}$ at a point $x_0 \in \mathbb{R}$ if for every sequence (x_n) with $x_n > 0$ and $\lim_{n\to\infty} x_n = 0$,

$$\text{st-}\lim_{n\to\infty} \frac{f(x_0 + x_n) - f(x_0)}{x_n} = w.$$

Remark 2 (One-sided vs. symmetric derivatives)**.** It is important to note that the one-sided statistical derivative (Definition 3) and a symmetric statistical derivative defined by

$$\text{st-}\lim_{n\to\infty} \frac{f(x_0 + x_n) - f(x_0 - x_n)}{2x_n} = w,$$

are *not equivalent* in general. For example, consider $f(x) = |x|$ at $x_0 = 0$. The symmetric difference quotient is identically zero, so the symmetric statistical derivative exists and equals 0. However, the right-hand difference quotient is identically 1, so the one-sided statistical derivative (Definition 3) exists and equals 1. Thus, the two concepts are distinct.

Moreover, Definition 3, as stated, is equivalent to the existence of the classical right derivative. Indeed, suppose that the right-hand difference quotient of f at x_0 does not converge to w as $h \to 0^+$. Then there exists $\varepsilon_0 > 0$ such that, for every $\delta > 0$, one can find $h \in (0, \delta)$ satisfying

$$\left| \frac{f(x_0 + h) - f(x_0)}{h} - w \right| \geq \varepsilon_0.$$

For each $n \in \mathbb{N}$, choose $x_n \in (0, 1/n)$ with this property. Then $x_n \to 0^+$, while

$$\left| \frac{f(x_0 + x_n) - f(x_0)}{x_n} - w \right| \geq \varepsilon_0,$$

for every n. Hence the exceptional set is all of $\mathbb{N}$, and therefore has natural density 1, contradicting Definition 3. Conversely, the existence of the classical right derivative clearly implies Definition 3, since ordinary convergence implies statistical convergence. Thus, Definition 3 is equivalent to the classical right derivative.

3. Sampling-induced statistical Riemann sum functional

In this section, we introduce the main concept of this paper.

3.1. Definition

The classical Riemann integral of a function f over an interval $[a, b]$ is defined as the limit of Riemann sums as the mesh of the partition goes to zero. For a partition $P : a = x_0 < x_1 < \cdots < x_n = b$ with tags $t_i \in [x_{i-1}, x_i]$, the Riemann sum is

$$S(f, P, t) = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}).$$

If $\lim_{\|P\|\to 0} S(f, P, t)$ exists and is independent of the choice of partitions and tags, this limit is the Riemann integral.

The object defined below is *not* a tag-free integration theory. Rather, it is a sampling-dependent functional built from a specific sequence of partitions and a specific tagging rule.

Definition 4. Let $f : [a, b] \to \mathbb{R}$ be a function. For each $n \in \mathbb{N}$, consider the uniform partition of $[a, b]$ into n subintervals of equal length:

$$x_i^{(n)} = a + i\frac{b-a}{n}, \quad i = 0, 1, \ldots, n.$$

Define the *right-endpoint Riemann sum* as

$$R_n(f) = \sum_{i=1}^{n} f\left(a + i\frac{b-a}{n}\right) \cdot \frac{b-a}{n}.$$

The sequence $(R_n(f))_{n\in\mathbb{N}}$ is called the *uniform right-endpoint Riemann sum sequence* associated with f.

Definition 5 (Sampling-induced functional)**.** A function $f : [a,b] \to \mathbb{R}$ is said to yield a *statistical Riemann sum value* (with respect to the right-endpoint sampling rule) on $[a,b]$ if the sequence $(R_n(f))$ is statistically convergent. In this case, we define

$$\mathrm{st}\int_a^b f(x)\,\mathrm{d}x = \text{st-}\lim_{n\to\infty} R_n(f).$$

When we need to emphasize the sampling rule σ, we write $\mathrm{st}\int_a^b f(x)\,\mathrm{d}x\,(\sigma)$.

Remark 3. The choice of uniform partitions is natural because it produces a canonical single sequence of Riemann sums, allowing statistical convergence to be applied directly. If one replaces the uniform partitions with an arbitrary sequence of partitions (P_n) with $\|P_n\| \to 0$, one obtains a different functional. Thus, the present theory is not based on a universal class of partitions but on one fixed partition sequence together with one fixed tagging rule. This is analogous to fixing a summability method before studying its properties: the definition is well-posed only after the underlying sampling scheme has been specified.

3.2. Basic properties

We now establish some elementary properties.

Proposition 1 (Linearity)**.** *If f and g yield statistical Riemann sum values on $[a,b]$ and $\alpha,\beta \in \mathbb{R}$, then $\alpha f + \beta g$ also yields a statistical Riemann sum value, and*

$$\mathrm{st}\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha\,\mathrm{st}\int_a^b f(x)\,\mathrm{d}x + \beta\,\mathrm{st}\int_a^b g(x)\,\mathrm{d}x.$$

Proof. For each n, $R_n(\alpha f + \beta g) = \alpha R_n(f) + \beta R_n(g)$. Let $I_f = \text{st-}\lim R_n(f)$ and $I_g = \text{st-}\lim R_n(g)$. For any $\varepsilon > 0$, we have

$$\left|(\alpha R_n(f) + \beta R_n(g)) - (\alpha I_f + \beta I_g)\right| \le |\alpha|\left|R_n(f) - I_f\right| + |\beta|\left|R_n(g) - I_g\right|.$$

Define

$$A_n = \left\{k \le n : \left|R_k(f) - I_f\right| \ge \frac{\varepsilon}{2(|\alpha|+1)}\right\}, \quad B_n = \left\{k \le n : \left|R_k(g) - I_g\right| \ge \frac{\varepsilon}{2(|\beta|+1)}\right\}.$$

Then

$$\left\{k \le n : \left|(\alpha R_k(f) + \beta R_k(g)) - (\alpha I_f + \beta I_g)\right| \ge \varepsilon\right\} \subseteq A_n \cup B_n.$$

Since $\lim_{n\to\infty} |A_n|/n = 0$ and $\lim_{n\to\infty} |B_n|/n = 0$, we have $\lim_{n\to\infty} |A_n \cup B_n|/n = 0$. Therefore, $\text{st-}\lim(\alpha R_n(f) + \beta R_n(g)) = \alpha I_f + \beta I_g$. □

Remark 4. The linearity proof does not require boundedness of the sequences; it uses only the definition of statistical convergence and the fact that the union of two sets of density zero has density zero.

Remark 5 (Failure of additivity)**.** Unlike the classical Riemann integral, the sampling-induced functional does *not* automatically satisfy additivity over subintervals. For $a < c < b$, it is not generally true that

$$\mathrm{st}\int_a^b f = \mathrm{st}\int_a^c f + \mathrm{st}\int_c^b f.$$

The reason is that the uniform partitions used to define the functional on $[a,c]$ and $[c,b]$ may not align with the uniform partition on $[a,b]$, and the statistical limit does not commute with such restrictions. When $c-a$ is a rational multiple of $b-a$, the relevant uniform partitions align along a suitable arithmetic subsequence, which may permit an additivity result under additional hypotheses. In general, however, additivity fails, as shown by the following example.

Example 1. Let $f = \mathbf{1}_{\mathbb{Q}}$ be the Dirichlet function on $[0,1]$. Choose an irrational split point $c \in (0,1) \setminus \mathbb{Q}$, say $c = \sqrt{2} - 1 \approx 0.4142$. For the right-endpoint sampling rule on $[0,1]$, every sample point i/n is rational, so $R_n(f) = 1$ for all n, and thus $\mathrm{st}\int_0^1 f = 1$.

Now consider the subinterval $[0,c]$. The uniform right-endpoint Riemann sums on $[0,c]$ are

$$R_n^{(c)}(f) = \sum_{i=1}^{n} f\left(\frac{ic}{n}\right) \cdot \frac{c}{n}.$$

Since c is irrational, for any $i = 1, \ldots, n$, the number ic/n is irrational (otherwise $c = (n/i) \cdot (ic/n)$ would be rational). Hence $f(ic/n) = 0$ for all i and n, so every uniform right-endpoint Riemann sum on $[0,c]$ is zero. Therefore $\mathrm{st}\int_0^c f = 0$.

Similarly, on $[c,1]$, the sample points are

$$t_i = c + \frac{i(1-c)}{n} = \frac{i}{n} + \left(1 - \frac{i}{n}\right) c.$$

For $i < n$, the coefficient $1 - i/n$ is nonzero, so t_i is irrational; for $i = n$, $t_n = 1$ is rational. Thus

$$R_n^{(c,1)}(f) = f(1) \cdot \frac{1-c}{n} = \frac{1-c}{n} \to 0.$$

Hence $\mathrm{st}\int_c^1 f = 0$. Therefore,

$$1 = \mathrm{st}\int_0^1 f \neq \mathrm{st}\int_0^c f + \mathrm{st}\int_c^1 f = 0.$$

This provides a genuine counterexample to additivity.

4. Relationship with the classical Riemann integral

Statistical convergence of the Riemann sum sequence is determined by the sampled values $\{f(a + i(b-a)/n)\}$. In particular, if these sampled values generate a statistically convergent sequence of Riemann sums, then the functional yields a value even when f fails to be Riemann integrable.

Theorem 1. *If f is Riemann integrable on $[a,b]$ in the classical sense, then f yields a statistical Riemann sum value with respect to the right-endpoint sampling rule, and*

$$\mathrm{st}\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Proof. Let $I = \int_a^b f(x)\,\mathrm{d}x$ be the classical Riemann integral. For a Riemann integrable function, the Riemann sums converge to I as the mesh of the partition goes to zero, regardless of the choice of partitions and tags. In particular, for the uniform right-endpoint partitions, we have $\lim_{n\to\infty} R_n(f) = I$. Since ordinary convergence implies statistical convergence, we obtain $\mathrm{st}\text{-}\lim R_n(f) = I$. $\square$

The same result holds for any sampling rule that yields Riemann sums converging to the classical integral. The converse is false: a function can yield a statistical Riemann sum value without being Riemann integrable, as shown below.

Example 2. Consider the Dirichlet function $f : [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q}, \\ 0, & \text{otherwise.} \end{cases}$$

This function is not Riemann integrable because every upper sum equals 1 and every lower sum equals 0. However, for the right-endpoint uniform partitions, each sample point i/n is rational, so $f(i/n) = 1$ for all i and n. Consequently, $R_n(f) = 1$ for all n, and therefore $\mathrm{st}\text{-}\lim R_n(f) = 1$. Thus, the functional yields the value

1 on $[0,1]$. This example shows that the value is governed by the arithmetic nature of the sampling set rather than by measure-theoretic size.

5. Sampling rule dependence

We emphasize that left-, right-, and midpoint rules do not necessarily produce three different values for a given function. For the Dirichlet function on $[0,1]$, both the left-endpoint and midpoint rules still sample rational points and therefore yield the same value as the right-endpoint rule. The point is not that every standard rule gives a different value, but that changing the sampling scheme *can* change the value. The following examples make this explicit.

Example 3. Consider again the Dirichlet function f from Example 2. If we use left-endpoint uniform sums,

$$L_n(f) = \sum_{i=0}^{n-1} f\left(\frac{i}{n}\right) \cdot \frac{1}{n},$$

then for all n, $f(i/n) = 1$, so $L_n(f) = 1$. Thus, $\operatorname{st-lim} L_n(f) = 1$, the same as for right-endpoint sums.

To obtain a different value, we need a sampling scheme that consistently picks irrational points.

Example 4. Fix an irrational number $\alpha \in (0,1)$, for instance $\alpha = \sqrt{2} - 1 \approx 0.4142$. For each $n \in \mathbb{N}$ and each subinterval $\left[\frac{i-1}{n}, \frac{i}{n}\right]$ (where $i = 1, \ldots, n$), define the tag

$$t_i^{(n)} = \frac{i-1+\alpha}{n}.$$

Observe that:

- $t_i^{(n)} \in \left[\frac{i-1}{n}, \frac{i}{n}\right]$ because $0 < \alpha < 1$.
- $t_i^{(n)}$ is irrational for all i and n. Indeed, if $t_i^{(n)}$ were rational, then $nt_i^{(n)} - (i-1) = \alpha$ would be rational, contradicting the irrationality of α.

For the Dirichlet function, $f(t_i^{(n)}) = 0$ for all i and n, so the corresponding Riemann sums are all 0. Hence the functional yields the value 0 for every irrational $\alpha \in (0,1)$, while it yields 1 for rational α. Therefore, viewed as a function of α, the resulting value is discontinuous at *every* $\alpha \in (0,1)$: at rational α the value is 1; at irrational α the value is 0. Thus, the set of discontinuity points is the entire interval $(0,1)$.

This example demonstrates that the same function can yield different statistical Riemann sum values depending on the chosen sampling rule. Therefore, this functional is not an intrinsic property of the function alone but depends on the specific sampling scheme employed.

Remark 6. A similar phenomenon occurs if one changes not only the tagging rule but also the underlying partition sequence. Once a different sequence of partitions (P_n) with $\|P_n\| \to 0$ is chosen, the sampled sums generally form a different sequence and may produce a different statistical limit. Therefore, the present notion is tied to two pieces of data: the partition sequence and the sampling rule. In this paper we fix the uniform partitions to keep the construction canonical on the partition side and isolate the effect of the tags.

Remark 7. The sampling-rule dependence is an inherent characteristic of the proposed construction: the functional captures information about the function on the specific countable set of sample points used in the construction. Different sampling schemes probe different countable subsets of the domain, potentially yielding different values.

In the remainder of this section, by "sampling-rule independence" we mean independence with respect to arbitrary choices of tags along the fixed uniform partitions (P_n).

Theorem 2. *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$. For each $n \in \mathbb{N}$, consider the uniform partition*

$$P_n : \ a = x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} = b, \qquad x_i^{(n)} = a + i\frac{b-a}{n}.$$

Let $\{t_i^{(n)}\}_{i=1}^n$ be any choice of tags with $t_i^{(n)} \in [x_{i-1}^{(n)}, x_i^{(n)}]$, and define

$$S_n(f;t) := \sum_{i=1}^{n} f\big(t_i^{(n)}\big)\,\big(x_i^{(n)} - x_{i-1}^{(n)}\big).$$

Then

$$\lim_{n\to\infty} S_n(f;t) = \int_a^b f(x)\,\mathrm{d}x,$$

in the ordinary sense, independently of the sampling rule $\{t_i^{(n)}\}$. Consequently,

$$\text{st-}\lim_{n\to\infty} S_n(f;t) = \int_a^b f(x)\,\mathrm{d}x,$$

so the statistical Riemann sum value along (P_n) is sampling-rule independent and coincides with the classical Riemann integral.

Proof. Since f is Riemann integrable, for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every tagged partition (P, t) with mesh $\|P\| < \delta$,

$$\left| S(f, P, t) - \int_a^b f(x)\,\mathrm{d}x \right| < \varepsilon.$$

For the uniform partitions P_n, $\|P_n\| = (b-a)/n \to 0$, so there exists N such that for all $n \geq N$, $\|P_n\| < \delta$. Hence, for all $n \geq N$ and any choice of tags,

$$\left| S_n(f;t) - \int_a^b f(x)\,\mathrm{d}x \right| < \varepsilon.$$

Thus $S_n(f;t)$ converges in the ordinary sense to the integral, which implies statistical convergence. □

Corollary 1. *If f is Riemann integrable on $[a,b]$, then for the uniform partitions (P_n) every statistical Riemann sum value defined using any fixed sampling rule (right-endpoint, left-endpoint, midpoint, irrational tags, etc.) exists and yields the same value $\int_a^b f(x)\,\mathrm{d}x$.*

Thus, Riemann integrability is a sufficient condition for sampling-rule independence (along (P_n)). Whether there are broader classes of functions for which this holds is an open question (see Section 8).

6. Measure-zero sensitivity

Another striking difference between this functional and classical integrals is its sensitivity to modifications of the function on sets of measure zero.

Example 5. Let $g : [0,1] \to \mathbb{R}$ be the zero function: $g(x) = 0$ for all x. Clearly, g is Riemann integrable with integral 0, and by Theorem 1, its statistical Riemann sum value (with right-endpoint sampling) is also 0.

Now modify g on the set of rational numbers to obtain the Dirichlet function f from Example 2. The set of rationals has Lebesgue measure zero, so in the Lebesgue theory, f is equivalent to g and has Lebesgue integral 0. In the Riemann theory, f is not integrable.

For the right-endpoint sampling rule, however, the modification dramatically changes the value: $\mathrm{st}\!\int_0^1 f = 1 \neq 0$. This shows that modifying a function on a countable set (hence measure zero) can completely alter its statistical Riemann sum value.

Remark 8. The right-endpoint sampling set $S = \{a + i(b-a)/n : n \in \mathbb{N},\, 1 \le i \le n\}$ is countable and dense. The functional probes f only on this fixed countable set determined by the sampling scheme.

The following proposition provides a quantitative measure of this sensitivity.

Proposition 2. *Let $f, g : [a,b] \to \mathbb{R}$ be functions for which the statistical Riemann sum values $I_f = \mathrm{st}\int_a^b f$ and $I_g = \mathrm{st}\int_a^b g$ exist (with respect to the right-endpoint sampling rule).*

(a) *If f and g agree on the set*
$$S = \{a + i(b-a)/n : n \in \mathbb{N},\, 1 \le i \le n\},$$
of right-endpoint sample points, then $I_f = I_g$.

(b) *Suppose there exists $\gamma > 0$ such that the set*
$$A_\gamma = \{n \in \mathbb{N} : |R_n(f) - R_n(g)| \ge \gamma\},$$
has positive upper natural density. Then
$$|I_f - I_g| \ge \gamma.$$

Proof. Part (a) follows immediately from the definition. If f and g agree on S, then $R_n(f) = R_n(g)$ for every $n \in \mathbb{N}$, and hence their statistical limits coincide whenever they exist.

For part (b), assume that
$$\text{st-}\lim_{n\to\infty} R_n(f) = I_f \qquad \text{and} \qquad \text{st-}\lim_{n\to\infty} R_n(g) = I_g.$$

Suppose, for contradiction, that $|I_f - I_g| < \gamma$.

Choose $\varepsilon > 0$ such that
$$2\varepsilon + |I_f - I_g| < \gamma.$$

Define
$$F_\varepsilon = \{n \in \mathbb{N} : |R_n(f) - I_f| < \varepsilon\}, \qquad G_\varepsilon = \{n \in \mathbb{N} : |R_n(g) - I_g| < \varepsilon\}.$$

Since $R_n(f)$ and $R_n(g)$ are statistically convergent to I_f and I_g, respectively, both sets have natural density 1. Therefore,
$$H_\varepsilon = F_\varepsilon \cap G_\varepsilon,$$
also has natural density 1.

For any $n \in H_\varepsilon$,
$$\begin{aligned} |R_n(f) - R_n(g)| &\le |R_n(f) - I_f| + |I_f - I_g| + |I_g - R_n(g)| \\ &< 2\varepsilon + |I_f - I_g| \\ &< \gamma. \end{aligned}$$

Thus,
$$H_\varepsilon \subseteq \mathbb{N} \setminus A_\gamma, \quad \text{equivalently,} \quad A_\gamma \subseteq \mathbb{N} \setminus H_\varepsilon.$$

Since H_ε has natural density 1, its complement $\mathbb{N} \setminus H_\varepsilon$ has natural density 0. Hence A_γ has natural density 0, and therefore upper natural density 0. This contradicts the assumption that A_γ has positive upper natural density.

Therefore, $|I_f - I_g| \ge \gamma$. $\square$

Although the operator is highly sensitive to the arithmetic structure of the sampling set, it defines a bounded linear functional on the vector space of bounded functions whose right-endpoint Riemann sum sequences are statistically convergent, equipped with the supremum norm:
$$\left| \mathrm{st}\int_a^b f(x)\,\mathrm{d}x \right| \le (b-a)\|f\|_\infty.$$

However, this functional does not descend to the space $L^1([a,b])$, since it is not invariant under modifications on sets of measure zero. For instance, the zero function and the Dirichlet function represent the same element of $L^1([0,1])$, but their sampling-induced values are 0 and 1, respectively. Consequently, the functional cannot be regarded as an L^1-continuous functional.

7. Toward a structural statistical fundamental theorem of Calculus

Motivation. As noted in Remark 2, the statistical derivative in the sense of Nuray (Definition 3) is equivalent to the classical right derivative, and thus not a true extension. In order to obtain a genuine extension that is compatible with our sampling-induced integral, we must weaken the derivative concept to match the discrete sampling scale of the integral.

The sampling-induced functional uses the specific step size $(b-a)/n$. Therefore, it is natural to define a derivative using difference quotients taken over the same step size. This leads to the following definition.

Definition 6 (Sampling-Structured Statistical Derivative)**.** Let $f : [a,b] \to \mathbb{R}$. We say that f has a *sampling-structured statistical derivative* at $x \in (a,b)$ if there exists $w \in \mathbb{R}$ such that

$$\text{st-}\lim_{n\to\infty} \frac{f\left(x+\frac{b-a}{n}\right)-f(x)}{\frac{b-a}{n}} = w.$$

(For points near the right endpoint, we interpret $x+(b-a)/n$ as being taken only for those n sufficiently large so that the point remains within $[a,b]$; the statistical limit is unaffected by finitely many exceptions.)

This definition restricts the difference quotients to the same mesh sizes that generate the uniform partitions in the sampling-induced functional. Thus, both constructions are driven by the same discrete scale. Unlike Definition 3, this derivative is genuinely weaker than the classical derivative, because it ignores deviations that occur on subsets of indices of density zero.

Lemma 1. *Let $f : [a,b] \to \mathbb{R}$ be classically differentiable at $x \in (a,b)$. Then f admits a sampling-structured statistical derivative at x (with respect to the uniform step size $(b-a)/n$), and*

$$f'_{ss}(x) = f'(x).$$

Proof. Since f is differentiable at x, we have

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = f'(x).$$

In particular, for the sequence $h_n = (b-a)/n \to 0$,

$$\lim_{n\to\infty} \frac{f(x+h_n)-f(x)}{h_n} = f'(x).$$

Hence the difference quotients converge in the ordinary sense, and therefore also in the statistical sense. □

Example 6. The converse of Lemma 1 does not hold. To see this, fix $x_0 \in (a,b)$ and choose $N_0 \in \mathbb{N}$ large enough so that

$$\frac{b-a}{N_0} < b - x_0.$$

For $n \geq N_0$, let

$$h_n = \frac{b-a}{n},$$

so that $x_0 + h_n \in (a, b)$. Define $f : [a, b] \to \mathbb{R}$ by setting $f(x_0) = 0$ and

$$f(x_0 + h_n) = \begin{cases} h_n, & \text{if } n \text{ is a perfect square,} \\ 0, & \text{otherwise,} \end{cases} \qquad n \geq N_0,$$

and set $f(x) = 0$ at all remaining points of $[a, b]$.

Then, for every $n \geq N_0$,

$$\frac{f(x_0 + h_n) - f(x_0)}{h_n} = \begin{cases} 1, & \text{if } n \text{ is a perfect square,} \\ 0, & \text{otherwise.} \end{cases}$$

Since the set of perfect squares has natural density zero, the sequence of difference quotients is statistically convergent to 0. The omission of the finitely many indices $n < N_0$ does not affect the statistical limit. Hence $f'_{ss}(x_0) = 0$.

On the other hand, the ordinary right derivative of f at x_0 does not exist, because the difference quotients equal 1 along the subsequence $n = m^2$, for which $h_{m^2} \to 0$, and equal 0 along the subsequence of non-square indices. Thus, the sampling-structured statistical derivative is strictly weaker than the classical right derivative.

The natural question is the following.

Problem 1 (Structural Statistical FTC)**.** Suppose that f admits a sampling-structured statistical derivative f'_{ss} on (a, b) and that f'_{ss} yields a statistical Riemann sum value with respect to the uniform right-endpoint sampling scheme on $[a, b]$. Under what additional conditions does

$$\mathrm{st}\int_a^b f'_{ss}(t)\,\mathrm{d}t = f(b) - f(a),$$

hold?

For the purpose of forming the right-endpoint Riemann sums of f'_{ss} on $[a, b]$, we extend f'_{ss} to the endpoint b by assigning it an arbitrary finite value. This choice does not affect the statistical Riemann sum value, since the contribution of the endpoint is

$$f'_{ss}(b)\frac{b-a}{n},$$

which converges to zero as $n \to \infty$. Hence the problem is well-defined.

We restrict to the full interval $[a, b]$ here because both the derivative and the integral are generated by the same discrete scale $(b - a)/n$. Considering subintervals $[a, x]$ would introduce a different scale $(x - a)/n$, breaking the structural alignment.

Unlike the classical situation, statistical convergence allows fluctuations on sets of density zero. The essential difficulty is to determine whether such fluctuations accumulate or cancel when passing from discrete difference quotients to discrete Riemann sums.

We emphasize that this compatibility condition is not merely technical: without synchronizing the discrete scales of differentiation and integration, a statistical Fundamental Theorem of Calculus cannot be expected to hold in general. Even with synchronized scales, uncontrolled oscillations of the difference quotients along subsets of indices of positive density may accumulate in the associated Riemann sums. Therefore, additional structural constraints—such as uniform statistical boundedness of the difference quotients or a suitable density-controlled variation condition—may be required to prevent error accumulation across the statistical limit.

A systematic investigation of this structural compatibility remains an open direction for future research. For instance, if $f(x) = x$ on $[0, 1]$, then $f'(x) = 1$. The sampling-induced value of f' over $[0, 1]$ (with respect to the right-endpoint scheme) equals

$$\mathrm{st}\int_0^1 f'(x)\,\mathrm{d}x = 1,$$

which coincides with $f(1) - f(0) = 1$. Thus, in this simple case, the classical Fundamental Theorem of Calculus holds even under the sampling-induced functional, showing that the obstruction is not universal but scheme-dependent.

8. Conclusion and open problems

In this paper, we have introduced a sampling-induced statistical Riemann sum functional based on uniform partitions with right-endpoint sampling. We have established its basic properties, explored its relationship with classical Riemann integrability, and highlighted two key features that distinguish it from classical integration theory:

- *Sampling-rule dependence*: The value of the functional depends on the choice of tags used in the Riemann sums. Different sampling rules may yield different values for the same function.
- *Measure-zero sensitivity*: Modifying a function on a countable set can drastically change its value, contrary to the behavior of Lebesgue and Riemann integrals.

These observations suggest that this functional is not a generalization of the classical integral but rather a new summability-based object that captures information about functions on specific countable dense subsets. This perspective opens several avenues for future research.

1. *Characterization of sampling-rule independence*: Characterize the class of functions for which all reasonable sampling rules yield the same statistical Riemann sum value along the uniform partitions (P_n). Theorem 2 shows that every Riemann integrable function has this property. Are there larger function classes (e.g., functions with only discontinuities of the first kind, or functions that are Riemann integrable on a dense set of subintervals) for which tag-independence holds? This characterization problem is the most fundamental open question in the present framework, as it directly addresses the distinction between intrinsic and sampling-dependent behavior.
2. *Statistical Lebesgue-type functional*: Develop a summability-based analog of the Lebesgue integral using simple functions and statistical limits. Such a construction might overcome some of the sampling-rule dependence issues by averaging over all possible sampling schemes or by considering all countable dense subsets simultaneously.
3. *Statistical Fundamental Theorem of Calculus*: Establish a rigorous version of the Fundamental Theorem of Calculus connecting the sampling-structured statistical derivative (Definition 6) with the sampling-induced functional. This would require a careful analysis of the conditions under which the functional of the derivative recovers the original function, including necessary and sufficient conditions on the oscillations of the difference quotients.
4. *Applications*: Explore potential applications in signal processing (where functions are often sampled at discrete points with fixed sampling rates) and in number theory (where statistical behavior on sets of density zero is relevant). The sensitivity of the functional to values on countable sets may be useful for studying functions with number-theoretic properties. However, these remain speculative at this stage and are presented as possible motivations rather than established contributions.

The sampling-induced statistical Riemann sum functional, despite its limitations, offers a new tool for analyzing functions through the lens of summability theory and serves as a sampling-based counterpart to derivative constructions arising from statistical convergence. We hope that this paper will stimulate further research in this direction.

References

[1] Steinhaus, H. (1951). Sur la convergence ordinaire et la convergence asymptotique. In *Colloquium Mathematicum,* (Vol. 2, No. 1, pp. 73-74).

[2] Fast, H. (1951). Sur la convergence statistique. In *Colloquium Mathematicum* (Vol. 2, pp. 241-244). Instytut Matematyczny Polskiej Akademii Nauk.

[3] Fridy, J. A. (1985). On statistical convergence. *Analysis, 5*(4), 301-314.

[4] Šalát, T. (1980). On statistically convergent sequences of real numbers. *Mathematica Slovaca, 30*(2), 139-150.

[5] Fridy, J. A. (1993). Statistical limit points. *Proceedings of the American Mathematical Society, 118*(4), 1187-1192.

[6] Miller, H. I. (1995). A measure theoretical subsequence characterization of statistical convergence. *Transactions of the American Mathematical Society, 347*(5), 1811-1819.

[7] Nuray, F. (2021). Cesàro and statistical derivative. *Facta Universitatis, Series: Mathematics and Informatics*, 1393-1398.