Riemann–Stieltjes Integral, its use case in the Formulation of the Spectral Theorem - An Analysis

Prof. N. Shivaraju, Asst Professor of Mathematics, GFG College for Womens, Davanagere

Abstract

This paper looks at Riemann–Stieltjes integral as a generalization of the Riemann integral. It serves as an instructive and useful precursor of the Lebesgue integral. The Riemann-Riesz representation theorem is a remarkable result which describes the continuous linear functionals acting on the space of continuous functions defined on a set K. Basically, it says that any bounded linear functional \( T \) on the space of compactly supported continuous functions on \( X \) is the same as integration against a measure \( \mu \),

\[
Tf = \int f \, d\mu.
\]

Here, the integral is the Lebesgue integral.

Because linear functionals form a vector space, and are not "positive," the measure \( \mu \) may not be a positive measure. But if the functional \( T \) is positive, in the sense that \( f \geq 0 \) implies that \( Tf \geq 0 \), then the measure \( \mu \) is also positive. In the generality of complex linear functionals, the measure \( \mu \) is a complex measure. The measure \( \mu \) is uniquely determined by \( T \) and has the properties of a regular Borel measure. It must be a finite measure, which corresponds to the boundedness condition on the functional. In fact, the operator norm of \( T \), \( \|T\| \), is the total variation measure of \( X, |\mu|(X) \).

Naturally, there are some hypotheses necessary for this to make sense. The space \( X \) has to be locally compact and a T2-Space, which is not a strong restriction. In fact, for unbounded spaces \( X \), the theorem also applies to functionals on continuous functions which vanish at infinity, in the sense that for any \( \varepsilon > 0 \), there is a compact set \( K \) such that for any \( x \) not in \( K \), \( |f(x)| < \varepsilon \) (which is the notion from calculus of \( \lim_{x \to \infty} f(x) = 0 \)).

The Riesz representation theorem is useful in describing the dual vector space to any space which contains the compactly supported continuous functions as a dense subspace. Roughly speaking, a linear functional is modified, usually by convolving with a bump function, to a bounded linear functional on the compactly supported continuous functions. Then it can be realized as integration against a measure. Often the measure must be absolutely continuous, and so the dual is integration against a function.

Key words: Riemann-Riesz representation, theorem, hypotheses, absolutely continuous, bump function
Introduction

A measure $\lambda$ is absolutely continuous with respect to another measure $\mu$ if $\lambda(E) = 0$ for every set with $\mu(E) = 0$. This makes sense as long as $\mu$ is a positive measure, such as Lebesgue measure, but $\lambda$ can be any measure, possibly a complex measure.

By the Radon-Nikodym theorem, this is equivalent to saying that

$$\lambda(E) = \int_E f \, d\mu,$$

where the integral is the Lebesgue integral, for some integrable function $f$. The function $f$ is like a derivative, and is called the Radon-Nikodym derivative $d\lambda/d\mu$.

The measure supported at 0 ($\mu(E) = 1$ iff $0 \in E$) is not absolutely continuous with respect to Lebesgue measure, and is a singular measure.

A measure which takes values in the complex numbers. The set of complex measures on a measure space $X$ forms a vector space. Note that this is not the case for the more common positive measures. Also, the space of finite measures ($|\mu(X)| < \infty$) has a norm given by the total variation measure $||\mu|| = |\mu|(X)$, which makes it a Banach space.

Using the polar representation of $\mu$, it is possible to define the Lebesgue integral using a complex measure,

$$\int f \, d\mu = \int e^{i\theta} f \, d|\mu|.$$

Sometimes, the term "complex measure" is used to indicate an arbitrary measure. The definitions for measure can be extended to measures which take values in any vector space. For instance in spectral theory, measures on $\mathbb{C}$, which take values in the bounded linear maps from a Hilbert space to itself, represent the operator spectrum of an operator.

Objective:

This paper intends to explore The Riemann–Stieltjes integral of a real-valued function invaluable tool in unifying equivalent forms of statistical theorems that apply to discrete and continuous probability.

Hilbert space vector dot product

A Hilbert space is a vector space $H$ with an inner product $\langle f, g \rangle$ such that the norm defined by

$$|f| = \sqrt{\langle f, f \rangle}$$

turns $H$ into a complete metric space. If the metric defined by the norm is not complete, then $H$ is instead known as an inner product space.

Examples of finite-dimensional Hilbert spaces include

1. The real numbers $\mathbb{R}^n$ with $\langle v, u \rangle$ the vector dot product of $v$ and $u$. 
2. The complex numbers $\mathbb{C}^d$ with $\langle v, u \rangle$ the vector dot product of $v$ and the complex conjugate of $u$.

An example of an infinite-dimensional Hilbert space is $L^2$, the set of all functions $f : \mathbb{R} \to \mathbb{R}$ such that the integral of $f^2$ over the whole real line is finite. In this case, the inner product is

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x) \, dx.$$  

A Hilbert space is always a Banach space, but the converse need not hold.

A set of functions $\{f_1(n,x), f_2(n,x)\}$ is termed a complete biorthogonal system in the closed interval $\mathbb{R}$ if, they are biorthogonal, i.e.,

$$\int_{\mathbb{R}} f_1(m,x)f_1(n,x) \, dx = c_m \delta_{mn}$$

$$\int_{\mathbb{R}} f_2(m,x)f_2(n,x) \, dx = d_m \delta_{mn}$$

$$\int_{\mathbb{R}} f_1(m,x)f_2(n,x) \, dx = 0$$

$$\int_{\mathbb{R}} f_1(m,x) \, dx = 0$$

$$\int_{\mathbb{R}} f_2(m,x) \, dx = 0$$

and complete.

A complete biorthogonal system has a very special type of generalized Fourier series. The prototypical example of a complete biorthogonal system is $\{\sin(nx), \cos(nx)\}_{n=0}^{\infty}$ over $R = [-\pi, \pi]$, which can be used as a basis for constructing "the" Fourier series of an arbitrary function.

A Banach space is a complete vector space $\mathcal{B}$ with a norm $\| \cdot \|$. Two norms $\| \cdot \|_1$ and $\| \cdot \|_2$ are called equivalent if they give the same topology, which is equivalent to the existence of constants $c$ and $C$ such that

$$\|v\|_1 \leq c \|v\|_2$$

and

$$\|v\|_2 \leq C \|v\|_1$$

hold for all $v$.

In the finite-dimensional case, all norms are equivalent. An infinite-dimensional space can have many different norms.

A basic example is $n$-dimensional Euclidean space with the Euclidean norm. Usually, the notion of Banach space is only used in the infinite dimensional setting, typically as a vector space of functions. For example, the set of continuous functions on closed interval of the real line with the norm of a function $f$ given by
is a Banach space, where \( \sup \) denotes the supremum.

On the other hand, the set of continuous functions on the unit interval \([0, 1]\) with the norm of a function \( f \) given by

\[
\|f\| = \int_0^1 |f(x)| \, dx
\]

is not a Banach space because it is not complete. For instance, the Cauchy sequence of functions

\[
f_n = \begin{cases} 
1 & \text{for } x \leq \frac{1}{2} \\
\frac{1}{2} n + 1 - n x & \text{for } \frac{1}{2} < x \leq \frac{1}{2} + \frac{1}{n} \\
0 & \text{for } x > \frac{1}{2} + \frac{1}{n}
\end{cases}
\]

does not converge to a continuous function.

Hilbert spaces with their norm given by the inner product are examples of Banach spaces. While a Hilbert space is always a Banach space, the converse need not hold. Therefore, it is possible for a Banach space not to have a norm given by an inner product. For instance, the supremum norm cannot be given by an inner product.

**complete orthogonal system**

expansion of a function based on the special properties of a complete orthogonal system of functions. The prototypical example of such a series is the Fourier series, which is based on the biorthogonality of the functions \( \cos (n \pi x) \) and \( \sin (n \pi x) \) (which form a complete biorthogonal system under integration over the range \([-\pi, \pi]\)). Another common example is the Laplace series, which is a double series expansion based on the orthogonality of the spherical harmonics \( Y_l^m (\theta, \phi) \) over \( \theta \in [0, \pi] \) and \( \phi \in [0, 2 \pi] \).

Given a complete orthogonal system of univariate functions \( \{\phi_n (x)\} \) over the interval \( R \), the functions \( \phi_n (x) \) satisfy an orthogonality relationship of the form

\[
\int_R \phi_m (x) \phi_n (x) w (x) \, dx = c_m \delta_{mn}
\]

over a range \( R \), where \( w (x) \) is a weighting function, \( c_m \) are given constants and \( \delta_{mn} \) is the Kronecker delta. Now consider an arbitrary function \( f (x) \). Write it as a series

\[
f (x) = \sum_{n=0}^{\infty} a_n \phi_n (x)
\]

and plug this into the orthogonality relationships to obtain
\[
\int_{R} f(x) \phi_n(x) w(x) \, dx = \int_{R} \sum_{n=0}^{\infty} a_n \phi_n(x) \phi_n(x) w(x) \, dx = \sum_{n=0}^{\infty} a_n \int_{R} \phi_n(x) \phi_n(x) w(x) \, dx = \sum_{n=0}^{\infty} a_n c_m \delta_{mn} = a_n c_n.
\]

Note that the order of integration and summation has been reversed in deriving the above equations. As a result of these relations, if a series for \( f(x) \) of the assumed form exists, its coefficients will satisfy
\[
a_n = \frac{1}{c_n} \int_{R} f(x) \phi_n(x) w(x) \, dx.
\]

Given a complete biorthogonal system of univariate functions, the generalized Fourier series takes on a slightly more special form. In particular, for such a system, the functions \( f_1(n, x) \) and \( f_2(n, x) \) satisfy orthogonality relationships of the form
\[
\int_{R} f_1(m, x) f_1(n, x) w(x) \, dx = c_m \delta_{mn}
\]
\[
\int_{R} f_2(m, x) f_2(n, x) w(x) \, dx = d_m \delta_{mn}
\]
\[
\int_{R} f_1(m, x) f_2(n, x) w(x) \, dx = 0
\]
\[
\int_{R} f_1(m, x) w(x) \, dx = 0
\]
\[
\int_{R} f_2(m, x) w(x) \, dx = 0
\]
for \( m, n \neq 0 \) over a range \( R \), where \( c_m \) and \( d_m \) are given constants and \( \delta_{mn} \) is the Kronecker delta. Now consider an arbitrary function \( f(x) \) and write it as a series
\[
f(x) = \sum_{n=0}^{\infty} a_n f_1(n, x) + \sum_{n=0}^{\infty} b_n f_2(n, x)
\]
\[
= f_1(0) a_0 + \sum_{n=1}^{\infty} a_n f_1(n, x) + f_2(0) b_0 + \sum_{n=1}^{\infty} b_n f_2(n, x)
\]
\[
= [f_1(0) a_0 + f_2(0) b_0] + \sum_{n=1}^{\infty} a_n f_1(n, x) + \sum_{n=1}^{\infty} b_n f_2(n, x)
\]
\[
= e + \sum_{n=1}^{\infty} a_n f_1(n, x) + \sum_{n=1}^{\infty} b_n f_2(n, x)
\]
and plug this into the orthogonality relationships to obtain

\[
\int_R f(x) f_1(n, x) w(x) \, dx = e \int_R f_1(n, x) \, dx + \sum_{m=1}^{\infty} a_m f_1(m, x) f_1(n, x) w(x) \, dx + \sum_{m=1}^{\infty} b_m f_1(m, x) f_2(n, x) w(x) \, dx
\]

\[
= e \cdot 0 + \sum_{m=1}^{\infty} a_m \int_R f_1(m, x) f_1(n, x) w(x) \, dx + \sum_{m=1}^{\infty} b_m \int_R f_1(m, x) f_2(n, x) w(x) \, dx
\]

\[
= \sum_{m=1}^{\infty} a_m c_m \delta_{mn} + \sum_{m=1}^{\infty} b_m \cdot 0
\]

\[
= a_n c_n
\]

\[
\int_R f_2(n, x) w(x) \, dx = e \int_R f_2(n, x) \, dx + \sum_{m=1}^{\infty} a_m f_2(m, x) f_1(n, x) w(x) \, dx + \sum_{m=1}^{\infty} b_m f_2(m, x) f_2(n, x) w(x) \, dx
\]

\[
= e \cdot 0 + \sum_{m=1}^{\infty} a_m \int_R f_1(m, x) f_2(n, x) w(x) \, dx + \sum_{m=1}^{\infty} b_m \int_R f_2(m, x) f_2(n, x) w(x) \, dx
\]

\[
= \sum_{m=1}^{\infty} a_m \cdot 0 + \sum_{m=1}^{\infty} b_m d_m \delta_{mn}
\]

\[
= b_n d_n
\]

\[
\int_R f(x) w(x) \, dx = e \int_R d x + \sum_{m=1}^{\infty} a_m f_1(m, x) w(x) \, dx + \sum_{m=1}^{\infty} b_m f_2(m, x) w(x) \, dx
\]

\[
= e \int_R d x + \sum_{m=1}^{\infty} a_m \int_R f_1(m, x) w(x) \, dx + \sum_{n=1}^{\infty} b_n \int_R f_2(n, x) w(x) \, dx
\]

\[
= e \int_R d x + \sum_{m=1}^{\infty} a_m \cdot 0 + \sum_{m=1}^{\infty} b_m \cdot 0
\]

\[
= e \int_R d x.
\]

As a result of these relations, if a series for \( f(x) \) of the assumed form exists, its coefficients will satisfy

\[
a_n = \frac{1}{c_n} \int_R f(x) f_1(n, x) w(x) \, dx
\]

\[
b_n = \frac{1}{d_n} \int_R f(x) f_2(n, x) w(x) \, dx
\]

\[
e = \frac{\int_R f(x) w(x) \, dx}{\int_R w(x) \, dx}.
\]

The usual Fourier series is recovered by taking \( f_1(n, x) = \cos(nx) \) and \( f_2(n, x) = \sin(nx) \) which form a complete orthogonal system over \([-\pi, \pi] \) with weighting function \( w(x) = 1 \) and noting that, for this choice of functions,
Therefore, the Fourier series of a function \( f(x) \) is given by

\[
f(x) = e + \sum_{n=1}^{\infty} a_n \cos(n x) + \sum_{n=1}^{\infty} b_n \sin(n x),
\]

where the coefficients are

\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(n x) \, dx
\]

\[
b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(n x) \, dx
\]

\[
e = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx.
\]

**Conclusion**

Riemann-Stieltjes integral on \( \mathbb{R} \) and \( \mathbb{R}^n \). Some natural and important applications in probability theory are discussed. The reason for discussing the Riemann-Stieltjes integral instead of the more general Lebesgue and Lebesgue-Stieltjes integrals are that most applications in elementary probability theory are satisfactorily covered by the Riemann-Stieltjes integral.

In particular there is no need for invoking the standard machinery of monotone convergence and dominated convergence that hold for the Lebesgue integrals but typically do not for the Riemann integrals. The reason for introducing Stieltjes integrals is to get a more unified approach to the theory of random variables, in particular for the expectation operator, as opposed to treating discrete and continuous random variables separately. Also it makes it possible to treat mixtures of discrete and continuous random variables: It is for instance not possible to show that the expectation of the sum of a discrete and a continuous r.v. is the sum of the expectations, without using Stieltjes integrals. There are also many advantages in inference theory, for instance in the discussion of plug-in estimators.
References

1. The Life of Bertrand Russell. Knopf. 1976. p. 119. ISBN 9780394490595. He became a relentless political activist during World War I, and throughout his life was an ardent advocate of parliamentary democracy through his support first of the Liberal Party and then of Labour.


5. Russell and G. E. Moore broke themselves free from British Idealism which, for nearly 90 years, had dominated British philosophy. Russell would later recall in "My Mental Development" that "with a sense of escaping from prison, we allowed ourselves to think that grass is green, that the sun and stars would exist if no one was aware of them ..."—Russell B, (1944) "My Mental Development", in Schilpp, Paul Arthur: The Philosophy of Bertrand Russell, New York: Tudor, 1951, pp. 3–20.


22. The Nobel Prize in Literature 1950 — Bertrand Russell: The Nobel Prize in Literature 1950 was awarded to Bertrand Russell "in recognition of his varied and significant writings in which he champions humanitarian ideals and freedom of thought". Retrieved on 22 March 2013.

23. "British Nobel Prize Winners (1950)". 13 April 2014 –