

La integral de Riemman-Stieltjes

La integral de Riemman-Stieltjes es una generalización de la integral de Riemman. Stieltjes observó que cualquier función creciente $F : \mathbb{R} \rightarrow \mathbb{R}$ origina una noción de medida de intervalos,

$$m_F((a, b]) = F(b) - F(a)$$

Para las aplicaciones a la teoría de probabilidades, nos interesa el caso en que F es la función de distribución de una variable aleatoria.

Stieltjes definió la integral

$$\int_a^b \varphi(x) dF(x) \quad (3.2)$$

generalizando la definición de la integral de Riemman de la siguiente manera: sea

$$\pi : a = x_0 < x_1 < x_2 < \dots < x_n = b$$

una partición del intervalo $[a, b]$ (Dar una partición no es otra cosa que elegir finitos puntos del intervalo en orden creciente) y elijamos puntos intermedios $\xi_i \in (x_i, x_{i+1}]$ en cada intervalito de la partición (En realidad, estamos trabajando con particiones con puntos marcados, pero no lo haremos explícito en la notación). Consideramos entonces las sumas de Riemman-Stieltjes

$$S_\pi(\varphi, F) = \sum_{i=0}^{n-1} \varphi(\xi_i)(F(x_{i+1}) - F(x_i))$$

Definición 3.2.1 *Diremos que la integral (3.2) existe y toma el valor $I \in \mathbb{R}$ si las sumas $S_\pi(\varphi, F)$ tienden al valor I cuando la norma*

$$|\pi| = \max_{0 \leq i \leq n-1} |x_{i+1} - x_i|$$

de la partición π tiende a cero, es decir si dado $\varepsilon > 0$, existe $\delta > 0$ tal que $|I - S_\pi(\varphi, F)| < \varepsilon$ para toda partición π con $|\pi| < \delta$.

Observemos que si $F(x) = x$, la integral de Riemman-Stieltjes se reduce a la integral de Riemman usual.

Algunas propiedades de la integral que son consecuencias más o menos inmediatas de las definiciones:

Lema 3.2.1 (Linealidad) 1. Si $\int_a^b \varphi_1(x) dF(x)$ y $\int_a^b \varphi_2(x) dF(x)$ existen, y $\varphi = \lambda_1 \varphi_1 + \lambda_2 \varphi_2$ entonces, $\int_a^b \varphi(x) dF(x)$ también existe, y tenemos que:

$$\int \varphi(x) dF(x) = \lambda_1 \int_a^b \varphi_1(x) dF(x) + \lambda_2 \int_a^b \varphi_2(x) dF(x)$$

2. Si $\int_a^b \varphi(x) dF_1(x)$ y $\int_a^b \varphi(x) dF_2(x)$ existen, y $F = \lambda_1 F_1 + \lambda_2 F_2$ con $\lambda_1, \lambda_2 \geq 0$, entonces $\int_a^b \varphi(x) dF$ existe, y vale que:

$$\int_a^b \varphi(x) dF(x) = \lambda_1 \int_a^b \varphi(x) dF_1(x) + \lambda_2 \int_a^b \varphi(x) dF_2(x)$$

Teorema 3.2.1 Si $\varphi : [a, b] \rightarrow \mathbb{R}$ es continua, y si $F : [a, b] \rightarrow \mathbb{R}$ es creciente, entonces la integral de Riemman-Stieltjes

$$\int_a^b \varphi(x) dF(x)$$

existe

Variables aleatorias absolutamente continuas

Lema 3.3.2 Supongamos que $F : [a, b] \rightarrow \mathbb{R}$ es una función creciente con derivada continua $F'(x) = f(x)$, entonces

$$\int_a^b \varphi(x) dF(x) = \int_a^b \varphi(x) f(x) dx$$

para toda función $\varphi \in C[a, b]$.

Prueba: Por el teorema del valor medio, $F(x_{i+1}) - F(x_i) = f(\xi_i)(x_{i+1} - x_i)$ para cierto $\xi_i \in (x_i, x_{i+1})$. Entonces, con esta elección de los puntos intermedios, la suma S_π se puede escribir como

$$S_\pi = \sum_{i=0}^{n-1} \varphi(\xi_i) f(\xi_i) (x_{i+1} - x_i)$$

y vemos que cuando la norma de la partición π tiende a cero, tiende a la integral de Riemman

$$\int_a^b \varphi(x) f(x) dx$$

□

Veamos algunos ejemplos, para familiarizarnos con esta idea:

$$H_{x_0}(x) = \begin{cases} 0 & \text{si } x < x_0 \\ 1 & \text{si } x \geq x_0 \end{cases}$$

H_{x_0} es la función de distribución de una variable aleatoria que toma el valor x_0 con probabilidad 1. Entonces tenemos:

Lema 3.3.1 Si $x_0 \in [a, b]$ y $\varphi \in C[a, b]$, entonces:

$$\int_a^b \varphi(x) dH_{x_0} = \varphi(x_0)$$

Prueba: En $S_\pi(\varphi, F)$ el único término no nulo corresponde al intervalo $[x_i, x_{i+1}]$ que contiene a x_0 , en consecuencia:

$$S_\pi(\varphi, F) = \varphi(\xi_i)$$

y cuando $|\pi| \rightarrow 0$, $\varphi(\xi_i) \rightarrow \varphi(x_0)$, por la continuidad de φ . \square

En consecuencia si X es una función de distribución de una variable discreta que toma finitos valores x_1, x_2, \dots, x_n con probabilidad $p_i = P\{X = x_i\}$, tenemos que:

$$F(x) = \sum_{i=1}^n p_i H_{x_i}(x)$$

En consecuencia, por la linealidad de la integral de Riemman-Stieltjes respecto a F :

$$\int_a^b \varphi(x) dF(x) = \sum_{i=1}^n p_i \int_a^b \varphi(x) dH_{x_i} = \sum_{i=1}^n p_i \varphi(x_i)$$

(donde $a \leq x_i \leq b \forall i$).

Si F es la función de distribución empírica de x_1, \dots, x_n entonces $p_i = 1/n$ obteniendo de esta forma un promedio de $\varphi(x_i)$.

Teorema 4.4.1 (Helly) Supongamos que $F_n : [a, b] \rightarrow \mathbb{R}$ es una sucesión de funciones de distribución tales que $F_n(x) \rightarrow F(x)$ en cada punto de continuidad de $F(x)$, entonces:

$$\int_a^b \varphi(x) dF_n(x) \rightarrow \int_a^b \varphi(x) dF(x) \quad (4.1)$$

para toda función continua $\varphi \in C[a, b]$.

Sea ahora F_n la función de distribución empírica de x_1, \dots, x_n , con distribución subyacente F .

Teorema de Glivenko-Cantelli: $\sup_x |F_n(x) - F(x)| \xrightarrow{\text{pp}} 0$, siendo F la función de distribución de las X_i 's.

Reuniendo todos los resultados anteriores estaríamos probando que si $f(x)$ es la función de densidad de F (distribución acumulada de x_i) entonces:

$$\frac{1}{n} \sum_{i=1}^n \varphi(x_i) \rightarrow \int_a^b \varphi(x) f(x) dx$$