Valor Esperado e Integral de Riemann-Stieltjes

Prof. Rafael de Mattos Grisi

23 de outubro de 2019

Este material foi escrito para dar suporte às aulas de Cálculo de Probabilidade da UFABC.

Nele apresentaremos de modo informal uma nova maneira de integrar, conhecida como integral de Riemann-Stieltjes. Trata-se apenas de uma generalização da já conhecida integral de Riemann, estudada no curso de FUV, mas que permite dar um passo grande na unificação de alguns conceitos estudados em IPE.

A seguir vamos mostrar como relacionamos este tipo de integral com o valor esperado (ou esperança) de uma variável aleatória, generalizando este conceito para além das variáveis discretas e contínuas.

1 Conectando dois Mundos

Até este momento do curso, e durante todo o curso de IPE, fizemos todo o estudo de variáveis aleatórias separando-as em dois mundos: as variáveis discretas e as contínuas.

A definição de esperança, por exemplo, é distinta para cada uma delas. As semelhanças são visíveis, mas a conexão não está completamente clara.

O primeiro elemento introduzido no curso que começa a fazer a ponte entre estes dois mundos é a função distribuição acumulada. Lembrando, se X é uma v.a., definimos a função distribuição acumulada de X como a função $F_X : \mathbb{R} \to [0, 1]$ dada por

$$F_X(x) := \mathbb{P}(X \leq x).$$

Essa definição é válida para qualquer tipo de v.a., não precisando ser nem discreta e nem contínua, e por isso é um bom local para iniciarmos nosso estudo.

Comecemos com um exemplo, e tomemos uma v.a. contínua X com densidade

$$f_X(x) = 2x,\ 0 \leq x \leq 1$$

e função distribuição

$$F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)dt = x^2, \quad 0 \leq x \leq 1.$$

Vamos esquecer por um momento a definição de esperança de uma variável contínua, e tentar calcular a esperança usando a definição de v.a. discreta.

Uma ideia seria tentar aproximar a variável X por variáveis discretas. Em outras palavras, usar X para construir v.a.'s discretas que "representem" a variável original.

Para isso, note primeiro que X assume valores em $[0,1]$, e em seguida escolha uma partição P_n de $[0,1]$ dada por pontos $0 = x_0 < x_1 < \cdots < x_n = 1$.

Se os intervalos forem suficientemente pequenos, podemos aproximar X por uma variável discreta X_n que assume valores $\xi_k = (x_{k-1} + x_k)/2$ com probabilidade

$$\mathbb{P}(X_n = \xi_k) := \mathbb{P}(x_{k-1} < x \leq x_k) = F_X(x_k) - F_X(x_{k-1}) = x_k^2 - x_{k-1}^2$$

para cada $k = 1, \ldots, n$.

Calculando agora o valor esperado de X_n encontramos

$$\begin{aligned}
\mathbb{E}[X_n] &= \sum_{k=1}^{n} \xi_k \mathbb{P}(X_n = \xi_k) \\
&= \sum_{k=1}^{n} \xi_k (x_k^2 - x_{k-1}^2) \\
&= \sum_{k=1}^{n} \xi_k (x_k + x_{k-1})(x_k - x_{k-1}) \\
&= \sum_{k=1}^{n} 2\xi_k \frac{(x_k + x_{k-1})}{2}(x_k - x_{k-1}) \\
&= \sum_{k=1}^{n} 2\xi_k^2 (x_k - x_{k-1}).
\end{aligned}$$

Mas a última soma na igualdade acima é justamente uma soma de Riemann da função $g(x) = 2x^2$ relativa à partição P_n. Fazendo agora $n \to \infty$, de modo que o tamanho dos intervalos de P_n tendam à 0, temos que

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \int_0^1 2x^2 dx = \frac{2}{3}.$$

Mas mais que o valor numérico, é interessante observar que

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \int_0^1 2x^2 dx = \int_0^1 x \cdot 2x dx = \int_0^1 x f_X(x) dx = \mathbb{E}[X].$$

Isso nos indica uma conexão forte entre os mundos das variáveis discretas e contínuas. Mostra que as definições de valor esperada, aparentemente distintas, são no fundo bastante similares.

A chave para tal conexão está na soma

$$\sum_{k=1}^{n} \xi_k [F_X(x_k) - F_X(x_{k-1})],$$

que é similar à uma soma de Riemann, mas onde substituímos o comprimento $(x_k - x_{k-1})$ pela diferença $F_X(x_k) - F_X(x_{k-1})$.

Olhando geometricamente (ver Figura 1), se aproximarmos a região embaixo da curva $\varphi(x) = x$ pela união de retângulos de altura $\varphi(\xi_k)$, a área $\xi_k[x_k - x_{k-1}]$ do k-ésimo retângulo seria substituída pela "massa" deste mesmo retângulo, dada agora por por $\xi_k[F_X(x_k) - F_X(x_{k-1})]$.

O limite encontrado quando $n \to \infty$ é conhecido como integral de Riemann-Stieltjes, que estudaremos melhor a seguir.

2 A Integral de Riemann Stieltjes

Seja $F : \mathbb{R} \to \mathbb{R}$ uma função real com as seguintes propriedades

1. F é não decrescente ($x < y \Rightarrow F(x) \leq F(y)$);

2. F é contínua pela direita.

Deste ponto em diante nos referiremos à este tipo de função como *função peso*.

Dada uma partição P_n de um intervalo $[a, b] \subseteq \mathbb{R}$, definida por pontos $a = x_0 < x_1 < \cdots < x_n = b$, e uma função $\varphi : \mathbb{R} \to \mathbb{R}$, defina

$$S(P, \varphi, F) = \sum_{k=1}^{n} \varphi(\xi_k)[F((x_k) - F(x_{k-1})],$$

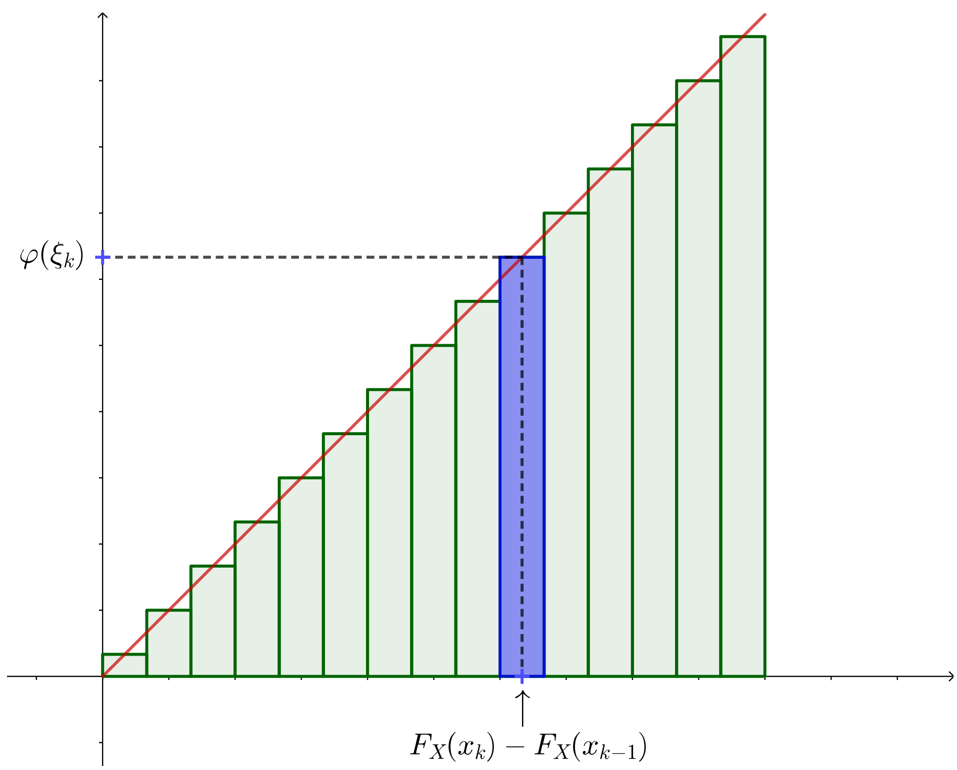

Figura 1: Cálculo do centro de massa por aproximação. Usamos aqui uma espécie de soma de Riemann, mas com a área do retângulo substituída pela "massa" $\xi_k[F_X(x_k) - F_X(x_{k-1})]$

onde $\xi_k \in [x_{k-1}, x_k],\ k = 1, \ldots, n$.

Diremos então que a *integral de Riemann-Stieltjes de* φ *com respeito à* F é igual à L, se

$$\lim_{\|P\| \to 0} S(P, \varphi, F) = L,$$

onde $\|P\| = \max\{|x_k - x_{k-1}| : k = 1, \cdots, n\}$ é o comprimento do maior intervalo da partição P.

Em outras palavras, se para todo $\epsilon > 0$, existir $\delta > 0$ tal que para toda partição P com $\|P\| < \delta$, e toda escolha de pontos $\xi_k \in [x_{k-1}, x_k],\ k = 1, \ldots, n$, vale que

$$|S(P, \varphi, F) - L| < \epsilon.$$

Neste caso, escreveremos que

$$\int_a^b \varphi(x) dF(x) = L.$$

A existência das integrais de Riemann-Stieltjes são um pouco mais complicadas que as integrais de Riemann, dado que agora precisamos lidar também com o comportamento da função F, além da função φ. Mas para o que precisamos neste texto, basta sabermos que se φ tem finitos pontos de descontinuidade em $[a, b]$, e nenhum deles é também um ponto de descontinuidade de F, então a integral existe.

Nas aplicações que faremos aqui, lidaremos apenas com funções φ contínuas, de modo que não encontraremos problemas.

Exemplo 2.1. Seja $F : \mathbb{R} \to \mathbb{R}$ dada por

$$F(x) = \begin{cases} 0; & x < 0 \\ \frac{1}{3}; & 0 \leq x < 1 \\ 1; & 1 \leq x < 1 \end{cases},$$

e $\varphi : \mathbb{R} \to \mathbb{R}$ uma função contínua em 0 e 1.

Dado $a > 1$, para calcular $\int_{-a}^{a} \varphi(x) dF(x)$, tome primeiro uma partição P definida por pontos $-a = x_0 < x_1 < \cdots < x_n = a$, e considere $\|P\| < 1$, de modo que 0 e 1 pertençam a intervalos distintos da partição.

Tome i, j tais que $0 \in (x_{i-1}, x_i]$ e $1 \in (x_{j-1}, x_j]$ e observe que se $k \neq i$ e $k \neq j$, então $F(x_{k-1} - F(x_k) = 0$ enquanto $F(x_{i-1} - F(x_i) = \frac{1}{3}$ e $F(x_{j-1} - F(x_j) = \frac{2}{3}$, de modo que

$$S(P, \varphi, F) = \sum_{k=1}^{n} \varphi(\xi_k)[F((x_k) - F(x_{k-1})] = \varphi(\xi_i)\frac{1}{3} + \varphi(\xi_j)\frac{2}{3}.$$

Agora, quando $\|P\| \to 0$, temos $\xi_i \to 0$ e $\xi_j \to 1$, e como φ é contínua em 0 e 1, segue que $\varphi(\xi_i) \to \varphi(0)$ e $\varphi(\xi_j) \to \varphi(1)$,

Deste modo podemos concluir que

$$\int_{-a}^{a} \varphi(x) dF(x) = \varphi(0)\frac{1}{3} + \varphi(1)\frac{2}{3}.$$

A solução do exemplo acima pode ser facilmente generalizada para funções do tipo "escada". Basta para isso escolhermos uma partição com intervalos suficientemente finos, de modo que cada pontos de descontinuidade de F pertença a um intervalo distinto da partição.

Para deixar mais claro, tomando-se F com um total finito de pontos de descontinuidade em $[a, b]$, denotados por $\{c_1, \cdots, c_m\}$, e φ um função contínua em cada c_i, escolha uma partição P dada por $a = x_0 < x_1 < \cdots < x_n = b$, com $\|P\|$ suficientemente pequeno, de modo que $\{c_1, \cdots, c_m\}$ pertençam a m intervalos distintos da partição.

Assim, tomando $\{i_k\}_{k=1}^m$ de modo que $c_k \in (x_{i_k-1}, x_{i_k}]$, e observando que $[F((x_k) - F(x_{k-1})] = 0$ para $k \notin \{i_k\}_{k=1}^m$, temos

$$S(P, \varphi, F) = \sum_{k=1}^{m} \varphi(\xi_k)[F((x_{i_k}) - F(x_{i_k-1})].$$

Como $x_{k-1} < c_k \leq x_k$, quando $\|P\| \to 0$, temos $F((x_{i_k}) - F(x_{i_k-1}) = F(c_k) - F(c_k-)$, onde $F(c_k-) = \lim_{x \to c_k^-} F(x)$ é o limite pela esquerda de F em c_k. Além disso, como $\xi_k \to c_k$, e φ é contínua em c_k, para cada $k = 1, \cdots, m$, concluímos o seguinte resultado.

Proposição 2.1. *Sejam $F : [a, b] \to \mathbb{R}$ é uma função peso constante por partes, com descontinuidades em $\{c_1, \cdots, c_m\} \subset [a, b]$, e $\phi : [a, b] \to \mathbb{R}$ contínua por partes, e contínua em $\{c_1, \cdots, c_m\}$. Nestas condições*

$$\int_{a}^{b} \varphi(x) dF(x) = \sum_{k=1}^{m} \varphi(c_k)[F(c_k) - F(c_k-)]. \tag{1}$$

Outro caso interessante acontece no caso em que F é diferenciável em (a, b), com derivada $F' = f$, e contínua em $[a, b]$.

Para começar, tome novamente uma partição P de $[a, b] \subseteq \mathbb{R}$, definida por pontos $a = x_0 < x_1 < \cdots < x_n = b$. Note agora, como F é diferenciável em $[a, b]$, pelo Teorema do Valor Médio, para cada $k = 1, \ldots, n$, existe $c_k \in [x_{k-1}, x_k]$ tal que $f(c_k)(x_k - x_{k-1}) = F(x_k) - F(x_{k-1})$, e portanto, tomando $\xi_k = c_k$, temos

$$S(P, \varphi, F) = \sum_{k=1}^{n} \varphi(c_k)[F((x_k) - F(x_{k-1})] = \sum_{k=1}^{n} \varphi(c_k) f(c_k)(x_k - x_{k-1}).$$

Agora, fazendo $\|P\| \to 0$, encontramos que:

Proposição 2.2. *Se $F : [a, b] \to \mathbb{R}$ é uma função peso contínua, diferenciável em todos, menos um total finito de pontos de $[a, b]$, com derivada $F' = f$,então*

$$\int_a^b \varphi(x) dF(x) = \int_a^b \varphi(x) f(x) dx. \tag{2}$$

Exemplo 2.2. Seja F a função distribuição de uma v.a. Normal Padrão. Ou seja

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Segue assim que $F'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ e

$$\int_{-a}^{a} \varphi(x) dF(x) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \varphi(x) e^{-\frac{x^2}{2}} dx.$$

Em particular

$$\int_{-a}^{a} x dF(x) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} x e^{-\frac{x^2}{2}} dx = 0.$$

Para finalizar precisamos apenas definir a integral em intervalos não limitados.

Assim como nas integrais de Riemann, definimos

$$\int_a^{\infty} \varphi(x) dF(x) = \lim_{b \to \infty} \int_a^b \varphi(x) dF(x),$$

$$\int_{-\infty}^{b} \varphi(x)dF(x) = \lim_{a\to\infty} \int_{a}^{b} \varphi(x)dF(x),$$

e

$$\int_{-\infty}^{\infty} \varphi(x)dF(x) = \lim_{a\to-\infty} \lim_{b\to\infty} \int_{a}^{b} \varphi(x)dF(x),$$

desde que os limites sejam finitos.

Exemplo 2.3. Seja X uma v.a. Normal Padrão. Ou seja, X possui função distribuição

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Segue do exemplo anterior que

$$\int_{-\infty}^{\infty} xdF(x) = \lim_{a\to\infty} \lim_{b\to\infty} \int_{-a}^{b} xdF(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-\frac{x^2}{2}} dx = \mathbb{E}[X].$$

---

3 Valor Esperado e a Integral de Riemann-Stieltjes

A esta altura já deve estar ficando clara a relação entre uma v.a. e a integral de Riemann-Stieltjes. As Proposições 2.1 e 2.2 fazem uma ponte direta entre a distribuição de uma v.a. contínua ou discreta e seu valor esperado.

De fato, se X uma v.a. discreta, assumindo valores em $\{x_1, x_2, \ldots\}$ e função distribuição acumulada F_X, então a Proposição 2.1 nos dá que

$$\int_{-\infty}^{\infty} xdF_X(x) = \sum_{k=1}^{\infty} x_k[F_X(x_k) - F_X(x_k-)] = \sum_{k=1}^{\infty} x_k\mathbb{P}(X = x_k) = \mathbb{E}[X].$$

Do mesmo modo, se X é uma v.a. contínua de densidade f_X e função distribuição F_X então, pela Proposição 2.2

$$\int_{-\infty}^{\infty} xdF_X(x) = \int_{-\infty}^{\infty} xf_X(x)dx = \mathbb{E}[X].$$

Outra relação interessante está no cálculo de algumas probabilidades relacionadas à distribuição de X. Tome, por exemplo, $a < b$ e note que para qualquer partição P, tomando $\phi(x) \equiv 1$, temos

$$S(P, \phi, F) = \sum_{k=1}^{n} [F_X(x_k) - F_X(x_{k-1})] = F_X(b) - F_X(a) = \mathbb{P}(a < X \le b).$$

E portanto

$$\int_a^b dF_X(x) = \mathbb{P}(a < X \leq b).$$

Esta conexão, apesar de ainda incompleta (não conseguimos calcular $\mathbb{P}(X \in A)$ para qualquer evento A deste modo), está na direção do que estamos buscando e permite ao menos um passo muito grande: uma expressão única para esperança de uma v.a., seja ela discreta ou contínua.

Podemos agora, sem medo, usar esta identidade para definir o valor esperado de uma variável X qualquer.

Definição 3.1. Dada uma v.a. X com função distribuição acumulada $F_X(x) = \mathbb{P}(X \leq x)$ defina o valor esperado (ou esperança) de X por

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x dF_X(x).$$

Ótimo! Temos uma definição única para o valor esperado de uma v.a.! Mas como usá-la para efetivamente calcular uma esperança? Afinal, só temos expressões fechadas para variáveis contínuas e variáveis discretas!

Para entender como fazer isso, precisaremos de uma propriedade importante das integrais de Riemann-Stieltjes.

Suponha então que $F = G + H$, onde G e H são duas funções peso. Segue naturalmente que F é uma função peso e

$$\begin{aligned}
S(P, \varphi, F) &= \sum_{k=1}^{n} \varphi(\xi_k)[F((x_k) - F(x_{k-1})] \\
&= \sum_{k=1}^{n} \varphi(\xi_k)[G((x_k) + H(x_k) - G(x_{k-1}) - H(x_{k-1})] \\
&= \sum_{k=1}^{n} \varphi(\xi_k)[G((x_k) - G(x_{k-1})] + \sum_{k=1}^{n} \varphi(\xi_k)[H((x_k) - H(x_{k-1})] \\
&= S(P, \varphi, G) + S(P, \varphi, H)
\end{aligned}$$

Isso nos mostra que se $\int_a^b \varphi(x)dG(x)$ e $\int_a^b \varphi(x)dH(x)$ existem, então $\int_a^b \varphi(x)dF(x)$ existe e

$$\int_a^b \varphi(x)dF(x) = \int_a^b \varphi(x)dG(x) + \int_a^b \varphi(x)dH(x).$$

Isso nos abre as portas para calcular facilmente uma grande quantidade de integrais de Riemann-Stieltjes e, consequentemente, o valor esperado de uma gama enorme de variáveis aleatórias.

Exemplo 3.1. Considere uma v.a. X com função de distribuição acumulada dada por

$$F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{4}, & -1 \leq x < 0 \\ \frac{1}{2} - \frac{1}{4}e^{-x}, & 0 \leq x < 1 \\ 1 - \frac{1}{4}e^{-x}, & x \geq 1. \end{cases}$$

É fácil ver que F_X tem exatamente dois pontos de descontinuidade, em -1 e 1, com

$$\mathbb{P}(X = -1) = F_X(-1) - F_X(-1^-) = \frac{1}{4},$$

e

$$\mathbb{P}(X = 1) = F_X(1) - F_X(1^-) = (1 - e^{-1}) - (\frac{1}{2} - e^{-x}) = \frac{1}{2}.$$

Defina então

$$G(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{4}, & -1 \leq x < 1 \\ \frac{3}{4}, & x \geq 1. \end{cases}$$

A função G assim definida não é mais uma função de distribuição acumulada ($\lim_{x\to\infty} G(x) \neq 1$), mas é ainda uma função peso.

Além disso, G é uma função escada com descontinuidades em -1 e 1 e saltos de mesmo tamanho que os saltos de F_X.

Em particular

$$\int_{-\infty}^{\infty} \varphi(x)dG(x) = \varphi(-1)[G(-1) - G(-1^-)] + \varphi(1)[G(1) - G(1^-)] = \varphi(-1)\frac{1}{4} + \varphi(1)\frac{1}{2}.$$

Defina então $H = F_X - G$, e note que

$$H(x) = \begin{cases} 0, & x < -1 \\ 0, & -1 \leq x < 0 \\ \frac{1}{4} - \frac{1}{4}e^{-x}, & 0 \leq x < 1 \\ \frac{1}{4} - \frac{1}{4}e^{-x}, & x \geq 1. \end{cases}$$

Ou seja

$$H(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}(1 - e^{-x}), & x \geq 0. \end{cases}$$

Segue que H é diferenciável em todo $x \neq 0$, e

$$H'(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}e^{-x}, & x \geq 0. \end{cases}$$

Assim, temos que

$$\int_{-\infty}^{\infty} \varphi(x) dH(x) = \frac{1}{4} \int_{0}^{\infty} \varphi(x) e^{-x} dx.$$

Juntando todas estas informações, como $F_X = G + H$, encontramos que

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x dF_X(x) \\ &= \int_{-\infty}^{\infty} x dG(x) + \int_{-\infty}^{\infty} x dH(x) \\ &= -1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + \frac{1}{4} \int_{0}^{\infty} x e^{-x} dx \\ &= \frac{1}{2}. \end{aligned}$$

---

4 Exercícios

1. Considere X uma variável aleatória com função distribuição acumulada dada por

$$F_X(x) = \begin{cases} 0 & , \ x < 0 \\ \dfrac{x}{4} & , \ 0 \leq x < 1 \\ \dfrac{x}{4} + \frac{1}{3} & , \ 1 \leq x < 2 \\ 1 & , \ x \geq 2 \end{cases},$$

   e calcule o que se pede.

   (a) $\mathbb{P}(X \leq 2)$;

   (b) $\mathbb{P}(X < 2)$;

   (c) $\mathbb{P}(0 < X \leq 1)$;

   (d) $\mathbb{P}(1 < X \leq 2)$;

   (e) $\mathbb{P}(0 \leq X \leq 1)$;

   (f) $\mathbb{P}(1 \leq X \leq 2)$;

   (g) $\mathbb{P}(X \geq 1)$;

   (h) $\mathbb{E}[X]$;