Probabilidade e Estatística

Aula 2 – Integração e Expectativa

Luis A. F. Alvarez

20 de janeiro de 2025

Integral de Lebesgue

- Nesta aula, partiremos de um espaço de medida (Ω, Σ, μ) e definiremos, para uma função mensurável $f : \Omega \mapsto \mathbb{R}$, a integral de Lebesgue com respeito a μ.
  - Como veremos, essa noção de integral *estende* a noção de integral de Riemann para uma classe mais ampla de funções e espaços subjacentes.
  - Esse conceito de integral será fundamental para a definição formal de esperança condicional.

Funções simples

- Seja (Ω, Σ, μ) um espaço de medida.
- Uma função $f : \Omega \mapsto \mathbb{R} \cup \{-\infty, \infty\}$ é dita simples se existem $k \in \mathbb{N}$ e $E_1, E_2, \dots, E_k \in \Sigma$, $a_1, \dots, a_k \in \mathbb{R} \cup \{-\infty, \infty\}$ tais que:

$$f(\omega) = \sum_{i=1}^{k} a_i \mathbf{1}_{E_1}(\omega)\,, \quad \omega \in \Omega$$

onde $\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{se } \omega \notin A \end{cases}$ é a função indicadora do conjunto A.

- **Convenção:** se $\mathbf{1}_{E_i}(\omega) = 0$ e $a_i = \pm\infty$, então $\mathbf{1}_{E_i}(\omega) a_i = 0$
- Fácil ver que f é mensurável.
  - Também fácil ver que, sem perda de generalidade, podemos tomar os E_i como disjuntos.

Integral de Lebesgue de funções simples não negativas

- Seja f uma função simples **não negativa**. A integral de Lebesgue com respeito a μ, denotada por $\mu(f)$ ou $\int f(\omega)\mu(d\omega)$ ou $\int f d\mu$, é definida por:
$$\int f(\omega)\mu(d\omega) := \sum_{i=1}^{k} a_i \mu(E_i) \, .$$
- Fácil de ver que integral está bem-definida (para duas expressões distintas da função simples em termos de conjuntos finitos, expressão dará a mesma coisa).
- Integral será um elemento de $[0, \infty]$

Integral de Lebesgue de funções mensuráveis não negativas

- Seja $f : \Omega \mapsto [0, \infty]$ uma função mensurável não negativa. Definimos a integral de Lebesgue como:

$$\mu(f) \coloneqq \sup(\mu(g) : g \text{ simples e não negativa}, g \leq f) \,.$$

Lema

Seja $f \geq 0$ mensurável.

- *Se $\mu(f) < \infty$, então, $\mu(\{\omega : f(\omega) = \infty\}) = 0$.*
- *Se $\mu(f) = 0$, então $\mu(\{\omega : f(\omega) > 0\}) = 0$.*
- *Seja $g \geq 0$ mensurável. Se $\mu(\{\omega : g(\omega) \neq f(\omega)\}) = 0$, então $\mu(f) = \mu(g)$.*

- **Obs:** Quando uma afirmação é válida a não ser em um conjunto de pontos ω de medida zero, dizemos que ela vale em μ-quase todo ponto (μ-q.t.p.).
  - Se μ é medida de probabilidade, equivalente ao qualificador "quase certamente" visto em aula anterior.

Teorema da Covergência Monótona

- Seja $(f_n)_n$ uma sequência de funções mensuráveis. Dizemos que f_n converge a uma função f em μ-quase todo ponto se $\mu(\{\omega : f_n(\omega) \nrightarrow f(\omega)\})) = 0$.
  - Medida do evento em que não há convergência é zero.
- Como f é mensurável (por quê?), se $f_n \geq 0$ para todo n, podemos nos perguntar se:

$$\int f_n d\mu \to \int f d\mu \,,$$

  i.e. podemos passar o limite por "debaixo" da integral?
- Resposta é verdadeira se a convergência for **monótona** em μ-quase todo ponto, i.e. $\mu(\{\omega : f_n(\omega) \uparrow f(\omega)\}^{\complement}) = 0$.

Teorema

Seja $(f_n)_n$ uma sequência de funções não negativas mensuráveis tais que $f_n \uparrow f$ em μ-quase todo ponto. Então.

$$\mu(f_n) \uparrow \mu(f) \leq \infty$$

Construindo aproximação monotônicas

- Em alguns contextos, é interessante construir funções monotônicas que aproximam uma dada função não negativa f.
- Uma construção bastante comum é dada por $f_n = s_n \circ f$, onde s_n são funções escada da forma:

$$s_n(y) = \begin{cases} 0 & \text{se } y = 0 \\ \frac{(j-1)n}{2^n}, & \text{se } y \in \left(\frac{(j-1)n}{2^n}, \frac{jn}{2^n}\right], j \in \{1, \ldots, 2^n\} \\ n, , & \text{se } y > n \end{cases}$$

- Sequência é tal que $f_n \uparrow f$.

Integral de Lebesgue no caso geral

- Seja $f : \Omega \mapsto \mathbb{R} \cup \{\infty, -\infty\}$ uma função mensurável. Defina as partes positiva e negativa de f como:

$$f^+ = \max\{f, 0\}\,, \quad f^- = -\min\{f, 0\}\,,$$

- Dizemos que f é **integrável** se $\mu(|f|) = \mu(f^+) + \mu(f^-) < \infty$. Nesse caso, a integral de f é definida como:

$$\mu(f) = \mu(f^+) - \mu(f^-)$$

- Vamos denotar por $L^1(\Omega, \Sigma, \mu)$ o espaço de funções Lebesgue-integráveis.

Integral de Lebesgue: propriedades

Proposição

Sejam $f, g \in L^1(\Omega, \Sigma, \mu)$ *e* $\alpha \in \mathbb{R}$. *Temos que:*

1. *(Monotonicidade)* $f \leq g \implies \mu(f) \leq \mu(g)$
2. *(Linearidade)* $f + \lambda g \in L^1(\Omega, \Sigma, \mu)$ *e* $\mu(f + \lambda g) = \mu(f) + \lambda\mu(g)$.

Proposição

Seja $(f_n)_n$ *uma sequência de funções em* $L^1(\Omega, \Sigma, \mu)$, *tais que* $f_n \to f$ *em* μ-*q.t.p. Se existe* $g \geq$ *mensurável tal que* $\mu(g) < \infty$ *e:*

$$|f_n(\omega)| \leq g(\omega)\, \forall n \in \mathbb{N}, \omega \in \Omega\,,$$

então $f \in L^1(\Omega, \Sigma, \mu)$ *e:*

$$\mu(|f_n - f|) \to 0\,,$$

$$\mu(f_n) \to \mu(f)$$

Integral de Lebesgue e integral de Riemann

- Considere o espaço $([0,1], \mathcal{B}[0,1], \lambda)$, com λ a medida de Lebesgue.
- Nesse espaço, podemos tanto calcular a integral de Lebesgue $\int f(x)\lambda(dx)$ como a integral de Riemann:

$$\int_0^1 f(x)dx$$

- Qual a relação entre as duas integrais?

Proposição

Seja $f \geq$ uma função real com domínio em $[0,1]$ Riemann integrável. Então f é mensurável e $\lambda(f) = \int_0^1 f(x)dx$.

- A recíproca do resultado acima nem sempre verdadeira. Por exemplo, a função

$$\mathbf{1}_{[0,1]\setminus\mathbb{Q}},$$

não é Riemann-integrável, embora $\lambda(\mathbf{1}_{[0,1]\setminus\mathbb{Q}}) = 0$.

Densidade com respeito a uma medida

- Seja (Ω, Σ, μ) um espaço de medida, e $f \geq 0$ uma função mensurável.
- O conjunto de integrais:

$$\int_A f d\mu := \mu(f \mathbf{1}_A)\,, A \in \Sigma$$

  define uma medida sobre (Ω, Σ) (verifique).
- Reciprocamente, se Φ é uma medida sobre (Ω, Σ, μ), dizemos que Φ admite uma densidade com respeito a μ se existe $g \geq 0$ mensurável tal que, para todo $A \in \Sigma$:

$$\Phi(A) = \int_A g d\mu$$

- Condição necessária e suficiente para existência de densidade é dada pelo teorema de Radon-Nikodyn.

Teorema

Seja (Ω, Σ) um espaço mensurável, e μ e Φ duas medidas σ-finitas. Φ admite uma densidade com respeito a μ se, e somente se, para todo $A \in \Sigma$, $\mu(A) \implies \Phi(A)$.

Densidade e uma fórmula padrão

Lema

Seja (Ω, Σ) *um espaço mensurável, e* μ *e* Φ *duas medidas. Se* Φ *admite densidade* g *com respeito a a* μ*, então para qualquer* $h \in L^1(\Omega, \Sigma, \Phi)$*, temos que:*

$$\Phi(f) = \int h(\omega) g(\omega) \mu(d\omega)$$

- Fórmula acima nos permite calcular esperança diretamente da integral com respeito a μ.
- **Demonstração:** primeiro verificamos a expressão para funções simples, depois para funções não negativas usando uma aproximação por funções-escada, e por fim estendemos para funções gerais.

Esperança

- Seja $(\Omega, \Sigma\, \mathbb{P})$ espaço de probabilidade.
- Neste caso, damos à integral de Lebesgue o nome de expectativa ou esperança, denotando-a, para $X \in L^1(\Omega, \Sigma, \mathbb{P})$ por:

$$\mathbb{E}[X] := \mathbb{P}(X)$$

.

- Como uma medida de probabilidade é finita, um corolário imediato do teorema da convergência dominada é:

Corolário (Teorema da convergência limitada)

Seja $(X_n)_n \in L^1(\Omega, \Sigma, \mathbb{P})^{\mathbb{N}}$ tais que $X_n \to X$ quase certamente. Se existe $K > 0$ tal que:

$$\mathbb{P}[\{\omega : |X_n(\omega)| \leq K,\ \forall n\}] = 1\,,$$

então $X \in L^1(\Omega, \Sigma, \mathbb{P})$ e: $\mathbb{E}[X_n] \to \mathbb{E}[X]$

Esperança: desigualdades fundamentais

- No que segue, considere um espaço de probabilidade $(\Omega, \Sigma, \mathbb{P})$.

Lema

Desigualdade de Markov Seja Z uma variável aleatória, e $g : \mathbb{R} \mapsto [0, \infty]$ mensurável e não decrescente. Então, para todo $c \in \mathbb{R}$.

$$\mathbb{P}[Z \geq c]g(c) \leq \mathbb{E}[g(Z)]$$

Lema

Desigualdade de Jensen Seja X uma variável aleatória, e $c : C \mapsto \mathbb{R}$ uma função convexa onde $C \subseteq \mathbb{R}$ é um conjunto aberto. Suponha que:

$$\mathbb{E}[|X|] < \infty, \quad \mathbb{P}[X \in C] = 1, \quad \mathbb{E}[|c(X)|] < \infty\,,$$

então $\mathbb{E}[X] \in C$ e:

$$c(\mathbb{E}[X]) \leq \mathbb{E}[c(X)]\,.$$

A NORMA L^p

- Fixe $p \in [1, \infty]$. Para uma variável aleatória X, nós definimos:

$$\|X\|_p = (\mathbb{E}[|X|^p])^{1/p}$$

- Denotamos por $L^p(\Omega, \Sigma, \mathbb{P})$ o espaço de variáveis aleatórias $X \in m(\Sigma)$ tais que $\|X\|_p < \infty$.
- É possível mostrar que esse espaço é linear normado, com $\|\cdot\|_p$ definindo uma (semi)norma em $L^p(\Omega, \Sigma, \mathbb{P})$.
    - "Semi" vem do fato de que $\|X\|_p = 0 \implies \mathbb{P}[X = 0] = 1$, de modo que há múltiplas variáveis aleatórias com norma zero, embora todas-quase certamente iguais a zero.
        - Essa multiplicidade não é problemática.

Normas L_p: propriedades úteis

- Abaixo, elencamos algumas propriedades úteis das normas L_p.

Lema

1. *(Monotonicidade) Sejam* $1 \leq p \leq q$. *Se* $X \in L^q(\Omega, \Sigma, \mathbb{P})$ *, então* $X \in L^p(\Omega, \Sigma, \mathbb{P})$ *e* $\|X\|_p \leq \|X\|_q$.
2. *(Cauchy-Schwarz) sejam* $X, Y \in L^2(\Omega, \Sigma, \mathbb{P})$. *Então* $X \cdot Y \in L^1(\Omega, \Sigma, \mathbb{P})$ *e :*

$$|\mathbb{E}[XY]| \leq \mathbb{E}[|XY|] \leq \|X\|_2 \|Y\|_2 .$$

3. *(Hölder) Seja* $1 \leq p \leq q \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. *Sejam* $X \in L^p(\Omega, \Sigma, \mathbb{P})$ *e* $Y \in L^q(\Omega, \Sigma, \mathbb{P})$. *Então* $X \cdot Y \in L^1(\Omega, \Sigma, \mathbb{P})$ *e:*

$$|\mathbb{E}[XY]| \leq \mathbb{E}[|XY|] \leq \|X\|_p \|Y\|_q .$$

4. *(Minkowski) Seja* $p \geq 1$, *e* $X, Y \in L^p(\Omega, \Sigma, \mathbb{P})$, *então:*

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$$

- Cauchy-Schwarz é caso particular de Hölder.
- Minkowski garante desigualdade triangular (e que $\|\cdot\|_p$ é norma)

Esperança condicional

- Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade, Y uma variável aleatória, e $\mathcal{G} \subseteq \Sigma$ uma sub-σ-álgebra de Σ.
- Tome $Y \in L^1(\Omega, \Sigma, \mathbb{P})$.
- Definimos a esperança condicional de Y com respeito a $\mathcal{G}$, denotada por $\mathbb{E}[Y|\mathcal{G}]$ como a variável aleatória $\mathbb{E}[Y|\mathcal{G}] \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ que satisfaz, para todo $A \in \mathcal{G}$:

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\mathbf{1}_A] = \mathbb{E}[Y|\mathbf{1}_A]$$

- Fácil mostrar que esperança condicional está unicamente definida, a não por eventos de probabilidade zero.
  - Uma variável aleatória que satisfaz as condições acima é conhecida como *versão* da esperança condicional.
  - Sejam Z_1 e Z_2 duas versões de $\mathbb{E}[Y|\mathcal{G}]$, então $\{\omega : Z_1(\omega) > Z_2(\omega)\} \in \mathcal{G}$ e pela definição da esperança condicional $Z_1 \geq Z_2$ q.c.

Existência da esperança condicional

- No slide anterior, definimos a esperança condicional e verificamos que, caso exista, ela é única.
  - No entanto, cabe a pergunta: será que existe uma variável aleatória que satisfaz as condições requeridas?
  - Resposta é afirmativa, e dada por um teorema devido a Komogorov.

Esperança condicional: propriedades básicas

conteúdo...

Esperança condicional: propriedades básicas

conteúdo...