

Теорема на Радон Никодим. Условно Математическо Очакване

1. Увод

В този файл ще дискутираме накратко едно красиво доказателство на теоремата на Радон-Никодим и ще въведем условното математическо очакване (УМО). Този подход към УМО не е толкова интуитивен, но пък показва интересно преплитане на области на математиката.

2. ТЕОРЕМА НА РАДОН-НИКОДИМ

В тази част ще разглеждаме измеримо пространство $\{\Omega, \mathcal{F}\}$, което в частност може да бъде реалната права с Бореловата сигма алгебра $(\mathbb{R}, \mathcal{B})$. Припомням, че \mathcal{F} е сигма алгебра от подмножества на множеството Ω . Например когато $\Omega = \mathbb{R}$, тогава \mathcal{B} е най-малката сигма алгебра, която съдържа всички интервали на реалната права.

Върху $\{\Omega, \mathcal{F}\}$ ще предполагаме, че са дефинирани две крайни, неотрицателни мерки $\mu, \nu : \mathcal{F} \mapsto [0, \infty)$. Тогава е в сила:

Теорема 2.1. *За две крайни, неотрицателни мерки $\mu, \nu : \mathcal{F} \mapsto [0, \infty)$ съществува $f : \Omega \mapsto [0, \infty)$ такава, че f е измерима спрямо \mathcal{F} , и съществува $A \in \mathcal{F}$ такава, че $\mu(A^c) = 0$, така че за всяко $B \in \mathcal{F}$*

$$(2.1) \quad \nu(B) = \int_{A \cap B} f d\mu + \nu(B \cap A^c).$$

Ако $\nu \ll \mu$, т.е. ν е абсолютно непрекъснат спрямо μ или $\mu(A) = 0 \implies \nu(A) = 0$, тогава е в сила следствието:

Следствие 2.2. *За две крайни, неотрицателни мерки $\mu, \nu : \mathcal{F} \mapsto [0, \infty)$, такива че ν е абсолютно непрекъснат спрямо μ , то съществува $f : \Omega \mapsto [0, \infty)$ такава, че f е измерима спрямо \mathcal{F} , и такава че за всяко $B \in \mathcal{F}$*

$$(2.2) \quad \nu(B) = \int_B f d\mu.$$

f се нарича производна на Радон-Никодим на ν спрямо μ и се означава чрез $d\nu/d\mu := f$.

Доказателство Теорема 2.1:

Идея: Уравнение (2.1) свързва реално интеграл по ν с интеграл по μ . Интегралите сами по себе си са линейни функционали, чието представяне на Риц, особено върху Хилбертово пространство е добре известно. Припомняме, че ако \mathcal{H} е Хилбертово пространство и $\Lambda : \mathcal{H} \mapsto \mathbb{R}$ е линеен, непрекъснат функционал, то $\Lambda h = \langle h, f \rangle$ за всяко $h \in \mathcal{H}$ и фиксирано $f \in \mathcal{H}$. С прости думи, линейните функционали са скалярни произведения, а те във функционални пространства са интеграли! Как да намерим подходящи Хилбертови пространства? Линейните функционали ще бъдат интегралите.

Нека $\pi = \mu + \nu$ и е ясно, че π е крайна мярка на $\{\Omega, \mathcal{F}\}$. За произволна крайна мярка ξ , $L^2(\xi) = \{g : \int_{\Omega} g^2 d\xi < \infty\}$ е Хилбертово пространство със скалярно произведение

$$\langle g, h \rangle_{\xi} = \int_{\Omega} gh d\xi.$$

От дефиницията на $\pi = \mu + \nu$ следва, че $L^2(\pi) \subseteq L^2(\nu)$ понеже, ако една функция е квадратично интегрируема спрямо π , то трябва и да е такава спрямо всяко събираемо на π , т.е. μ, ν . Но тогава $\Lambda g := \int_{\Omega} g d\nu$ е линеен, непрекъснат функционал както върху $L^2(\nu)$, така и върху подпространството му $L^2(\pi)$, т.е. $\Lambda : L^2(\pi) \mapsto \mathbb{R}$. Тогава по представянето на Риц, приложено върху $L^2(\pi)$,

имаме $h \in L^2(\pi)$, така че за всяко $g \in L^2(\pi)$

$$\Lambda g = \int_{\Omega} g d\nu = \langle g, h \rangle_{\pi} = \int_{\Omega} g h d\pi = \int_{\Omega} g h d\nu + \int_{\Omega} g h d\mu.$$

Така получаваме, че

$$(2.3) \quad \int_{\Omega} g(1-h) d\nu = \int_{\Omega} g h d\mu.$$

За да получим (2.1), ще трябва да разберем свойствата на h и да опростим израза (2.3).

- Нека $C = \{\omega \in \Omega : h(\omega) < 0\}$. Ако $\mu(C) > 0$, то от (2.3) получаваме с индикаторната функция $g = 1_C$ и $1 - h > 1$

$$\nu(C) = \int_{\Omega} 1_C d\nu \leq \int_{\Omega} 1_C(1-h) d\nu \stackrel{(2.3)}{=} \int_{\Omega} 1_C h d\mu = \int_C h d\mu < 0$$

или $\nu(C) < 0$, което не е възможно. Ако $\nu(C) > 0$, то по аналогичен начин получаваме противоречие от

$$\nu(C) = \int_{\Omega} 1_C d\nu < \int_{\Omega} 1_C(1-h) d\nu \stackrel{(2.3)}{=} \int_{\Omega} 1_C h d\mu = \int_C h d\mu \leq 0.$$

- Нека $C = \{\omega \in \Omega : h(\omega) > 1\}$. Ако $\mu(C) > 0$, то от (2.3) получаваме с индикаторната функция $g = 1_C$ и $1 - h < 0$ противоречието

$$0 \geq \int_{\Omega} 1_C(1-h) d\nu \stackrel{(2.3)}{=} \int_{\Omega} 1_C h d\mu = \int_C h d\mu > \mu(C),$$

защото $0 > \mu(C)$ не е възможно. Ако $\nu(C) > 0$, то противоречието идва от

$$0 > \int_{\Omega} 1_C(1-h) d\nu \stackrel{(2.3)}{=} \int_{\Omega} 1_C h d\mu = \int_C h d\mu \geq \mu(C).$$

Дотук имаме, че множеството $\{h > 1\} \cup \{h < 0\}$ има нулева мярка спрямо μ, ν и можем да го пренебрегнем и да допуснем, че $\{0 \leq h \leq 1\}$.

Да означим, $A^c = \{\omega \in \Omega : h(\omega) = 1\}$. От (2.3) получаваме с $g = 1_{A^c}$, че

$$0 = \int_{\Omega} 1_{A^c}(1-h) d\nu = \int_{\Omega} 1_{A^c} h d\mu = \mu(A^c).$$

Представяме $\Omega = A \cup A^c$ и отбелязваме, че $A = \{0 \leq h < 1\}$. Тогава, понеже $\mu(A^c) = 0$ и $A^c = \{h = 1\}$, (2.3) става

$$(2.4) \quad \int_{\Omega} g(1-h) d\nu = \int_A g(1-h) d\nu = \int_A g h d\mu + \int_{A^c} g h d\mu = \int_A g h d\mu.$$

Нека $B \in \mathcal{F}$, тогава *формално* с $g = 1_{A \cap B}/(1-h)$ и $f = h/(1-h)$, получаваме, че

$$(2.5) \quad \int_A 1_{A \cap B} d\nu = \nu(A \cap B) = \int_A 1_{A \cap B} \frac{h}{1-h} d\mu = \int_{A \cap B} f d\mu.$$

В последната стъпка има дупка, която е хубаво да се запълни. Не знаем априори, че $g = 1_{A \cap B}/(1-h) \in L^2(\pi)$, но с приближение става вярно равенството (2.5).

Така получаваме, че за всяко $B \in \mathcal{F}$

$$\nu(B) = \nu(A \cap B) + \nu(A^c \cap B) = \int_{A \cap B} f d\mu + \nu(A^c \cap B)$$

или (2.1) е в сила.

Доказателство Следствие 2.2: Понеже в предходното доказателство $\mu(A^c) = 0$, то и $\nu(A^c) = 0$ от допускането за абсолютна непрекъснатост. Така $\Omega = A$ и директно

$$\nu(B) = \int_B f d\mu.$$

3. УСЛОВНО МАТЕМАТИЧЕСКО ОЧАКВАНЕ (УМО)

Нека $\{\Omega, \mathcal{F}, \mathbb{P}\}$ е вероятностно пространство и нека $\mathcal{G} = \sigma(\eta) \subseteq \mathcal{F}$ е сигма алгебра (информация) генерирана от наблюдавана случайна величина η . Ако ξ е случайна величина в предходен файл видяхме, че $\mathbb{E}[\xi|\mathcal{G}]$ е случайната величина изцяло препределима от η , която минимизира $\mathbb{E}[(\xi - \tau)^2]$ по всички случайни величини τ предопределими от η .

Тук допускаме само че $\mathbb{E}[|\xi|] < \infty$ и разглеждаме събития $B \in \mathcal{G}$, т.е. събъждането на B се предопределя от наблюдаваната η . Интересува ни очакването на ξ върху B , т.е.

$$(3.1) \quad \nu(B) := \mathbb{E}[\xi 1_B] = \int_B \xi d\mathbb{P}.$$

Как можем да пресметнем тези средни? η по принцип не дава пълна информация за ξ . За целта ще получим алтернативна формула.

Първо забелязваме, че $\xi = \xi 1_{\xi \geq 0} - |\xi| 1_{\xi < 0} = \xi^+ - \xi^-$ или ξ е разлика на две неотрицателни случайни величини. Затова имаме

$$(3.2) \quad \nu(B) := \mathbb{E}[\xi 1_B] = \int_B \xi^+ d\mathbb{P} - \int_B \xi^- d\mathbb{P} = \nu^+(A) - \nu^-(A).$$

Можем да работим отделно с ν^\pm и затова допускаме изначално, че $\xi = \xi^+ \geq 0$. Това не променя разсъжденията.

Отбелязваме следните факти, които можете да проверите,

- ν е неотрицателна, мярка върху \mathcal{F} , защото $\nu(A) = \mathbb{E}[\xi 1_A] \leq \mathbb{E}[\xi] < \infty$;
- \mathbb{P} е неотрицателна мярка върху \mathcal{F} и $\mathbb{P}(B) = 0 \implies \nu(B) = 0$ или $\nu \ll \mathbb{P}$, защото $\mathbb{E}[\xi 1_B] = 0$;
- ν, \mathbb{P} са мерки и върху $\mathcal{G} \subseteq \mathcal{F}$.

От трите наблюдения следва, че Следствие 2.2 е в сила за (Ω, \mathcal{G}) с две мерки ν, \mathbb{P} . Тогава

$$(3.3) \quad \nu(B) = \mathbb{E}[\xi 1_B] = \int_B f d\mathbb{P} = \int_\Omega f 1_B d\mathbb{P}, B \in \mathcal{G},$$

където $f : \Omega \mapsto [0, \infty)$ е случайна величина измерима спрямо \mathcal{G} или спрямо информацията от η . С f означаваме $\mathbb{E}[\xi|\mathcal{G}]$ или $\mathbb{E}[\xi|\eta]$. Формула (3.3) е добре известното равенство

$$(3.4) \quad \mathbb{E}[\xi 1_B] = \mathbb{E}[1_B \mathbb{E}[\xi|\mathcal{G}]],$$

от което за $B = \Omega \in \mathcal{G}$ следва забележителната формула

$$\mathbb{E}[\xi] = \mathbb{E}[\mathbb{E}[\xi|\mathcal{G}]].$$

4. ЕЛЕМЕНТАРНО НАБЛЮДЕНИЕ ЗА КЛАСИЧЕСКАТА УСЛОВНА ВЕРОЯТНОСТ

Нека $\xi = 1_C$ и $\eta = B$, т.е. ξ, η са с разпределение на Бернули. Тогава $\mathcal{G} = \{B, B^c, \Omega, \emptyset\}$ и всички измерими спрямо η случайни величини са от вида

$$\tau = a 1_B + b 1_{B^c},$$

където $a, b \in \mathbb{R}$. Тогава понеже $\mathbb{E}[\xi|\mathcal{G}]$ е измерима спрямо η , имаме, че

$$\mathbb{E}[\xi|\mathcal{G}] = a1_B + b1_{B^c}$$

за някои a, b . Как да ги намерим? От (3.4) имаме, че

$$\mathbb{P}(B \cap C) = \mathbb{E}[1_C 1_B] = \mathbb{E}[1_B (a1_B + b1_{B^c})] = a\mathbb{P}(B)$$

или

$$a = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B)} := \mathbb{P}(C|B).$$

Аналогично $b = \mathbb{P}(C|B^c)$ откъдето следва, че

$$\mathbb{E}[\xi|\mathcal{G}] = \mathbb{E}[1_C|\mathcal{G}] = \mathbb{P}(C|B)1_B + \mathbb{P}(C|B^c)1_{B^c}.$$