Теория вероятностей. Лекция десятая Измеримые множества и отображения

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

13.11.2018

Условное матожидание отн-но σ-алгебры

Пусть имеются полная группа событий $H_1, \ldots, H_k, \ldots$ $(\Omega = \bigsqcup_{i=1}^{\infty} H_i)$ и порожденная ей σ-алгебра $\mathcal{H} = \sigma(H_1, \ldots, H_k, \ldots)$. Пусть $q_i \triangleq \mathbb{P}(H_i) > 0$.

Условным математическим ожиданием $\mathbb{E}(\xi|\mathcal{H})$ дискретной случайной величины ξ относительно σ-алгебры $\mathcal{H} = \sigma(H_1, \ldots, H_k, \ldots)$ называют отображение $\mathbb{E}(\xi|\mathcal{H}) : \Omega \to \mathbb{R}$, действующее по правилу:

$$\Omega \ni \omega \mapsto \mathbb{E}(\xi|\mathcal{H})(\omega) \triangleq \begin{cases} \sum_{i \in \mathbb{N}} x_i \mathbb{P}(\xi = x_i | H_1), & \text{если } \omega \in H_1; \\ \ldots \\ \sum_{i \in \mathbb{N}} x_i \mathbb{P}(\xi = x_i | H_k), & \text{если } \omega \in H_k; \\ \ldots \end{cases}$$

будем говорить, что условное матожидание существует, если, конечно, все выражения выше определены и конечны.

Условное матожидание отн-но случайной величины

Рассмотрим некоторые дискретные случайные величины ξ, η; случайная величина η, принимающая на каждом H_i некоторое свое значение, порождает σ-алгебру

$$\sigma(\eta) = \sigma\{\eta^{-1}(x) \,|\, x \in \mathbb{R}\} = \sigma\{H_1, \dots, H_k, \dots\}.$$

Условное матожидание ξ относительно случайной величины η зададим формулой $\mathbb{E}(\xi|\eta) \equiv \mathbb{E}(\xi|\sigma(\eta))$, то есть как функцию из Ω в $\mathbb{R}$. Иногда его вводят как функцию из области значений случайной величины η в $\mathbb{R}$, т.е. как регрессию.

Характеристическое свойство условного матожидания

18^0 Для случайной величины ξ $(\mathbb{E}\xi^2 < +\infty)$ и случайной величины η среди всех таких $h : \mathbb{R} \to \mathbb{R}$, что $\mathbb{E}(h^2(\eta)) < +\infty$, невязка $\mathbb{E}(\xi - h(\eta))^2$ достигает минимума при $h_0(\eta) = \mathbb{E}(\xi|\eta)$.

Снова условное матожидание в табличке

По совместному распределению пишется регрессия

(Y,X)	x_1	x_2	x_3	$\dots$	$\mathbb{P}(Y=y_j)$	$\mathbb{E}(X\mid Y)$
y_1	p_{11}	p_{12}	p_{13}	$\dots$	q_1	$h_0(y_1)\triangleq\frac{1}{q_1}\sum_{i=1}^{\infty}x_ip_{1i}$
y_2	p_{21}	p_{22}	p_{23}	$\dots$	q_2	$h_0(y_2)\triangleq\frac{1}{q_2}\sum_{i=1}^{\infty}x_ip_{2i}$
y_3	p_{31}	p_{32}	p_{33}	$\dots$	q_3	$h_0(y_3)\triangleq\frac{1}{q_3}\sum_{i=1}^{\infty}x_ip_{3i}$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

Подумать: можно ли было минимизировать что-то другое? Чем лучше, или хуже, использовать невязки $\mathbb{E}|\xi-h(\eta)|$, $\mathbb{E}|\xi-h(\eta)|^r\,(r>1)$?
Подумать: сумеете получить схожие формулы для минимизации $\mathbb{E}|\xi-h(\eta)|$, $\mathbb{E}|\xi-h(\eta)|^4$ или $\mathbb{E}|\xi-h(\eta)|^{3/2}$?

Стандартные свойства условного матожидания: [с-но]

Если не касаться вопросов существования (свойств 8^0-13^0)

0^0 $\mathbb{E}(1_A|\mathcal{H}) = \mathbb{P}(A|\mathcal{H})$ для всех $A \in \mathcal{F}$.

1^0 $\mathbb{E}(\xi|\mathcal{H}) \geq 0$, если $\xi \geq 0$ для всех $\omega \in \Omega$.

2^0 $\mathbb{E}(\xi_1|\mathcal{H}) \geq \mathbb{E}(\xi_2|\mathcal{H})$, если $\xi_1(\omega) \geq \xi_2(\omega)$ для всех $\omega \in \Omega$.

3^0 $\mathbb{E}(c|\mathcal{H}) = c$ для вырожденной случайной величины $c \in \mathbb{R}$;
более того, $\mathbb{E}(\xi|\mathcal{H}) = \xi$ для $\mathcal{H}$-измеримой ξ,
в частности, всегда $\mathbb{E}(f(\xi)|\xi) = f(\xi)$, $\mathbb{E}(\xi|\xi) = \xi$.

4^0 $\mathbb{E}(c\xi|\mathcal{H}) = c\mathbb{E}(\xi|\mathcal{H})$ для $c \in \mathbb{R}$,
более того, $\mathbb{E}(\eta\xi|\mathcal{H}) = \eta\mathbb{E}(\xi|\mathcal{H})$, если η $\mathcal{H}$-измерима.

5^0 $\mathbb{E}(\xi_1|\mathcal{H}) + \mathbb{E}(\xi_2|\mathcal{H}) = \mathbb{E}(\xi_1 + \xi_2|\mathcal{H})$.

6^0 $\eta_1\mathbb{E}(\xi_1|\mathcal{H}) + \eta_2\mathbb{E}(\xi_2|\mathcal{H}) = \mathbb{E}(\eta_1\xi_1 + \eta_2\xi_2|\mathcal{H})$ для $\mathcal{H}$-измеримых η_1, η_2.

7^0 $\mathbb{E}(\xi\eta|\mathcal{H}) = \mathbb{E}(\xi|\mathcal{H})\mathbb{E}(\eta|\mathcal{H})$, если $\xi 1_H, \eta 1_H$ независимы для всех $H \in \mathcal{H}$.

“Условные” свойства условного матожидания [с-но]

Предполагая, что все нужные условные матожидания существуют...

14^0 $\mathbb{E}(\xi|\mathcal{G}) = \xi$ тогда и только тогда, когда ξ $\mathcal{G}$-измерима (т.е. $\{\omega \in \Omega | \xi(\omega) = x\} \in \mathcal{G}$ для всех x).

15^0 $\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$, если ξ независима относительно $\mathcal{G}$;

в частности, $\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$ для $\mathcal{G} = \{\Omega, \varnothing\}$.

16^0 $\mathbb{E}(\xi|\mathcal{H}) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{G})|\mathcal{H})$ при $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$;

в частности, $\mathbb{E}(\xi|\eta) = \mathbb{E}(\mathbb{E}(\xi|\eta, \zeta)|\eta)$,

$\mathbb{E}\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$,

$\mathbb{P}(A) = \mathbb{E}\mathbb{P}(A|\mathcal{G})$ для всех $A \in \mathcal{F}$

(формула полной вероятности),

$\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}(\xi|\mathbb{E}(\xi|\mathcal{G}))$ (свойство Дуба).

[С-но; 0,5 баллов] $\mathbb{E}(\xi\mathbb{E}(\eta|\mathcal{G})) = \mathbb{E}(\eta\mathbb{E}(\xi|\mathcal{G}))$.

Условная дисперсия и не только

Пусть ξ, η — дискретные случайные величины, и ξ интегрируема с квадратом (существует $\mathbb{E}\xi^2$).
Введем условную дисперсию $\mathbb{D}(\xi|\eta) : \Omega \to \mathbb{R}$ правилом:

$$\mathbb{D}(\xi|\eta) \triangleq \mathbb{E}\Big((\xi - \mathbb{E}(\xi|\eta))^2|\eta\Big) = \mathbb{E}\Big(\xi^2|\eta\Big) - \Big(\mathbb{E}(\xi|\eta)\Big)^2.$$

Считаем:

$$\begin{aligned}
\mathbb{E}(\mathbb{D}(\xi|\eta)) &= \mathbb{E}(\xi - \mathbb{E}(\xi|\eta))^2 \\
&= \mathbb{E}\Big((\xi - E\xi) - \mathbb{E}(\xi - (\mathbb{E}\xi|\eta))\Big)^2 \\
&= \mathbb{D}\xi - 2\mathbb{E}(\xi - E\xi)\mathbb{E}(\xi - E(\xi|\eta)) + \mathbb{E}(\mathbb{E}(\xi|\eta) - \mathbb{E}(\mathbb{E}(\xi|\eta))\Big)^2 \\
&= \mathbb{D}\xi - \mathbb{D}\mathbb{E}(\xi|\eta).
\end{aligned}$$

Условная дисперсия в качестве “теоремы Пифагора”

17^0 Для интегрируемой с квадратом дискретной случайной величины ξ

$$\mathbb{D}\xi = \mathbb{E}(\mathbb{D}(\xi|\eta)) + \mathbb{D}(\mathbb{E}(\xi|\eta));$$

в частности, $\mathbb{D}\xi = \mathbb{D}(\mathbb{E}(\xi|\eta)), \mathbb{E}(\mathbb{D}(\xi|\eta)) = 0$, в точности если $\sigma(\xi) \subset \sigma(\eta)$, то есть $\xi = \mathbb{E}(\xi|\eta) = h(\eta)$ для какой-то h; более того, $\mathbb{D}\xi = \mathbb{E}(\mathbb{D}(\xi|\eta)), \mathbb{D}(\mathbb{E}(\xi|\eta)) = 0$ для независимых ξ, η.

Подумать: теперь несложно доказать, что для любых интегрируемых с квадратом случайных величин η_1, η_2, ξ выполнено

$$\mathbb{D}(\mathbb{E}(\xi|\eta_1, \eta_2)) = \mathbb{D}(\mathbb{E}(\xi|\eta_1)) + \mathbb{D}(\mathbb{E}(\xi|\eta_2)) \leq \mathbb{D}\xi$$

для независимых η_1, η_2; но ведь независимых величин можно взять больше, чем две...

Вступайте в ряды Фурье!

Что разобрали:

- распределение случайных величин
- медиана и математическое ожидание
- независимость случайных величин
- производящие функции как матожидание
- дисперсия, ковариация и корреляция
- векторные случайные величины и их распределения
- энтропия и условная энтропия
- условное математическое ожидание

Мы научились находить зависимость одной случайной величины с помощью другой случайной величины:
1) линейную (как и любую наперед заданную двухпараметрическую) — надо рассчитать соответствующие дисперсии и ковариацию, для чего достаточно знать матожидания;
2) произвольного вида — надо рассчитать условное матожидание одной случайной величины относительно другой, для чего достаточно знать их совместное распределение.

Ограничения, от которых мы будем избавляться

Для подсчета $\mathbb{E}(X|Y)$:
1. использовалось суммирование по значениям X, то есть предполагалось, что значений у X не более чем счетное число;
2. предполагалось, что вероятность каждого значения Y не равна нулю, то есть значений у Y тоже не более чем счетное число.
О том, что делать в общем случае — вся оставшаяся часть семестра.
Первый шаг здесь — ввести σ-алгебру над $\Omega = \mathbb{R}$. С этого и начнем...

- борелевские множества и мера Лебега
- случайные величины и измеримые отображения
- функции распределения
- абсолютно случайные величины и не только
- математическое ожидание как интеграл
- совместные функции распределения
- условное математическое ожидание как интеграл

Геометрическая вероятность. Вопросы

Пример 1. Алгебра ли $\mathcal{F} = \{[a,b] \subset [0,1]\}$?
Нет, надо рассмотреть всевозможные конечные объединения всевозможных промежутков из $[0,1]$.

Подумать: как задать на $\alpha(\mathcal{F})$ длину.

Пример 2. Алгебра ли множество Π всех прямоугольников на $[0,1]^2$?
Нет, не годятся даже всевозможные конечные объединения прямоугольников $\alpha(\Pi)$. Проще взять $\alpha(\mathcal{F}) \otimes \alpha(\mathcal{F})$.

Подумать: имеется ли в $\alpha(\mathcal{F}) \otimes \alpha(\mathcal{F})$ какой-нибудь треугольник?
Подумать: как задать на $\alpha(\mathcal{F}) \otimes \alpha(\mathcal{F})$ площадь.

Геометрическая вероятность. Решения

Пример 3. Борелевские множества на открытом вправо промежутке. Рассмотрим открытый вправо промежуток $I \subset \mathbb{R}$ и полукольцо $\mathcal{I}$ всех открытых вправо полуинтервалов $[a, b) \subset I$. Ее можно продолжить до σ-алгебры $\sigma(\mathcal{I})$, σ-алгебры борелевских множеств. А как выглядят $\sigma(\mathbb{R}), \mathcal{R}(\mathbb{R})$?

Подумать. А что будет, если начать со всех открытых влево промежутков, отрезков, интервалов?

Подумать. А как вероятность определить на таких множествах?

Пример 4. Борелевские множества на плоскости. Рассмотрим два открытых вправо промежутка $I_1, I_2 \subset \mathbb{R}$ и множества $\mathcal{I}_1, \mathcal{I}_2$ всех открытых вправо полуинтервалов в I_1, I_2. Элементы $\sigma(\mathcal{I}_1) \otimes \sigma(\mathcal{I}_2)$ называют борелевскими множествами на $I_1 \times I_2$.

Подумать: воспользовавшись этим примером, покажите, почему σ-алгебра удобнее, как измеримая структура, чем алгебра.

Два определения борелевских множеств

Определение 1.1. [очень умное, пользоваться им мы почти не будем]
Для топологического пространства (Ω, τ), где τ - совокупность всех открытых подмножеств множества Ω, под σ-алгеброй борелевских множеств понимается $\sigma(\tau)$,

$$\mathcal{B}_\Omega \triangleq \sigma(\tau).$$

Определение 1.2. [рабочее]
Минимальную σ-алгебру, содержащую все n-мерные параллелепипеды $\{[a,b] \times \ldots \times [c,d] \subset \mathbb{R}^n\}$, называют борелевской σ-алгеброй над $\mathbb{R}^n$, а ее элементы — борелевскими множествами множества $\mathbb{R}^n$,

$$\mathcal{B}_{\mathbb{R}^n} \triangleq \sigma\{[a,b] \times \ldots \times [c,d] \subset \mathbb{R}^n\}.$$

Подумать: покажите, почему эти определения эквивалентны для $\mathbb{R}^n$.
Подумать: покажите, что эти определения совпадают с определениями предыдущего слайда для $I = I_1 = I_2 = \mathbb{R}$.

Мера Лебега и не только

Борелевской мерой каждого параллелепипеда $[a,b]\times\ldots\times[c,d]$ назовем

$$\lambda\Big([a,b]\times\ldots\times[c,d]\Big)=(b-a)\cdot\ldots\cdot(d-c).$$

Эта функция счетно-аддитивна на полукольце всех параллелепипедов.
По теореме Каратеодори мы можем ее продолжить на все элементы из $\mathcal{B}\stackrel{\triangle}{=}\sigma(\tau)$, то есть на все борелевские множества.
Определение 2. Так доопределенную меру λ называют мерой Лебега.

Геометрическая вероятность. Итог

Эксперимент удовлетворяет геометрическому определению вероятности, если все его исходы можно изобразить в виде некоторого борелевского множества Ω, причем $0 < \lambda(\Omega) < +\infty$, а вероятность любого его борелевского подмножества зависит только от лебеговской меры (длины, площади, объема, ... n-мерного объема) этого подмножества и задается по формуле

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}, \forall A \in \mathcal{B}.$$

Лебеговские множества. Пополнение меры

Итак, имея меру на всех борелевских подмножествах, мы можем свести геометрическую вероятность к общей схеме. Полученную нами меру (как и всякую другую меру) можно попробовать распространить и дальше. Рассмотрим алгебру нулевых (пренебрежимых) множеств:

$$\mathcal{N} = \{N \mid \exists A \in \mathcal{B}, N \subset A, \lambda(A) = 0\},$$

а с ней — σ-кольцо лебеговских множеств $\sigma(\mathcal{B} \cup \mathcal{N})$.

Замечание. Так пополненное σ-кольцо уже учитывает исходную меру λ, в отличие от конструкции Каратеодори.

Теперь продолжим меру на все лебеговские множества правилом
$\lambda(N) = 0 \Rightarrow \lambda|_{2^N} \equiv 0 \qquad \forall N \in \mathcal{B}$.

Замечание. Такую операцию пополнения можно делать с любой мерой.

Определение 3. [на вырост] Меру μ со свойством
$\mu(N) = 0 \Rightarrow \mu|_{2^N} \equiv 0$ для всех $N \in \mathcal{F}$ называют полной.

Случайные величины. Определения

Итак, задано некоторое вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$.

Определение 4.1. Случайной (скалярной) величиной ξ называется такая функция $\xi : \Omega \to \mathbb{R}$, что для любого борелевского множества $B \in \mathcal{B}$ выполнено

$$\xi^{-1}(B) = \{\omega \,|\, \xi(\omega) \in B\} \in \mathcal{F}.$$

Определение 4.2. Случайной (скалярной) величиной ξ называется такая функция $\xi : \Omega \to \mathbb{R}$, что для всех $a, b \in \mathbb{R}$ $(a < b)$ выполнено

$$\xi^{-1}((a, b]) = \{\omega \,|\, a < \xi(\omega) \le b\} \in \mathcal{F}.$$

Определение 4.3. Случайной (скалярной) величиной ξ называется такая функция $\xi : \Omega \to \mathbb{R}$, что для всех $b \in \mathbb{R}$ выполнено

$$\xi^{-1}((-\infty, b]) = \{\omega \,|\, \xi(\omega) \le b\} \in \mathcal{F}.$$

Случайные величины. Эквивалентность

Теорема 1. Определения 4.1-4.3 эквивалентны.

Доказательство.

Очевидно, что из 4.1 следует 4.2, а из 4.2 следует 4.3. Покажем, что из 4.3 следует 4.1.

Рассмотрим класс

$$\{B \subset \mathbb{R} \,|\, \xi^{-1}(B) \in \mathcal{B}\},$$

он замкнут относительно не более чем счетного применения теоретико-множественных операций. Следовательно, это σ-алгебра. Он заведомо содержит все $(-\infty, b]$. Следовательно, он содержит минимальную σ-алгебру, их содержащую, то есть $\mathcal{B}$.

О минимальности

Подумать. Следующие определения эквивалентны:

- ξ — случайная величина;
- $\sigma\{\{\omega \mid \xi(\omega) \in B\} \mid B \in \mathcal{B}\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid \xi(\omega) < x\} \mid x \in \mathbb{R}\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid \xi(\omega) \leq x\} \mid x \in \mathbb{R}\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid \xi(\omega) \geq x\} \mid x \in \mathbb{R}\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid \xi(\omega) > x\} \mid x \in \mathbb{R}\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid y < \xi(\omega) < x\} \mid x, y \in \mathbb{R}, y < x\} \subset \mathcal{B}$;
- $\sigma\{\{\omega \mid y \leq \xi(\omega) \leq x\} \mid x, y \in \mathbb{R}, y < x\} \subset \mathcal{B}$.

  Однако в этот список нельзя добавить
- $\sigma\{\{\omega \mid \xi(\omega) = x\} \mid x \in \mathbb{R}\} \subset \mathcal{B}$.

Сейчас мы дали много эквивалентных друг другу определений $\sigma(\xi)$. Для определенности примем:

$$\sigma(\xi) \overset{\triangle}{=} \sigma\{\{\omega \mid \xi(\omega) \in B\} \mid B \in \mathcal{B}\}.$$

Что получили в рамках избавления от ограничений

- борелевские множества и мера Лебега
- случайные величины и измеримые отображения
- функции распределения
- абсолютно случайные величины и не только
- математическое ожидание как интеграл
- совместные функции распределения
- условное математическое ожидание как интеграл

Мы нашли подходящую σ-алгебру на $\mathbb{R}$, теперь мы умеем работать хотя бы с геометрической вероятностью. Мы дали определение случайной величине без каких-либо требований на мощность ее значений и нашли какой σ-алгебре она соответствует. Теперь надо научиться работать с такими объектами...