ВМК МГУ & ФУПМ МФТИ

Теория вероятности. Математическая статистика

Игорь Тао, Дима Сотников, Дима Гущин

7 августа 2019 г.

Содержание

Часть № I. Теория вероятности

Ты спросишь меня, кого я люблю больше: тебя или теорию вероятности. Я отвечу, что почти наверное я больше люблю тебя, и ты уйдешь, так и не узнав, что почти наверное значит с вероятностью 1.

Игорян

§1. Вероятностное пространство

1.1. Стохастические ситуации

Часть информации, изложенной дальше, взято из книги [1] Королева, так что любознательный читатель может ознакомиться с первоисточником при особом желании. Следуя Королеву, будем говорить о *случайности* как о принципиально неустранимой неопределенности. Чтобы понять, как с ней же работать, необходимо выделить те случаи, когда мы можем сказать: «А здесь нам поможет теория вероятностей,» — и не попасть впросак.

Методы теории вероятности работают в ситуациях, называемых *стохастическими*. Для них характерны три свойства:

- **1: Непредсказуемость:** исход ситуации нельзя предсказать абсолютно, то есть без какой-либо погрешности.
- **2: Воспроизводимость:** у испытателей результатов есть возможность (хотя бы теоретическая) повторить ситуацию сколь угодно много раз при неизменных условиях.
- **3: Устойчивость частот:** при многократном повторении ситуации частота исхода, то есть отношение числа опытов, в которых мы получили этот исход, к их общему числу, должна колебаться около некоторого значения, приближаясь к нему все ближе и ближе. Другими словами, должна иметь частота исходов должна иметь предел при стремлении числа опытов к ∞.

Для описания стохастических ситуаций ситуаций необходимо определить функцию вероятности. Её область определения называется *множеством событий*. В свою очередь событие (такое как, например, выпадание чётного числа на кубике) могут являться совокупностью неких более простых событий, описывающих стохастическую ситуацию (число, выпавшее на кубике). Последнее множество называется *множеством элементарных исходов* и обозначается Ω.

Есть разные теории описывающие вероятностные пространства, но самой используемой является теория, созданная Колмогоровым в начале прошлого века. Перейдем к ее описанию.

1.2. Измеримое пространство

Здесь по сути быстрым темпом рассказывается о началах теории меры Лебега, которая и приведет нас к определению вероятностного пространства. При должном желании читатель может ознакомится с целостным и строго математическим путем повествования меры Лебега, например, здесь — [2].

Вероятностное пространство, которое мы дальше формализуем, должны содержать в себе описание трех позиций, важных в эксперименте: во-первых, оно должно содержать множество элементарных исходов Ω, чтобы было предельно понятно, что является исходом; во-вторых, на этом множестве мы будем рассматривать некоторую структуру подмножеств, которая будет множеством событий, которые имеют для нас смысл (т.е. если мы бросаем кубик и нам важна четность выпавшего числа, то не имеет смысл рассматривать исход «выпало 1 очко» в отдельности от исходов «выпало 3 очка» и «выпало 5 очков»); в-третьих, на этом множестве мы зададим функцию, которая и будет мерой исхода, то есть некоторой характеристикой, описывающей частоту того или иного исхода.

Множество событий, обозначаемое $\mathcal{A}$ (используют наравне с этим еще обозначение $\mathscr{A}$), должно обладать следующими интуитивными свойствами:

1. Отрицание события есть событие (если «пойдет дождь» событие, то «не пойдет дождь» также событие).

2. Объединение событий есть событие («пойдет дождь» или «пойдет снег»).

3. Все множество элементарных исходов является событием («Что-нибудь да произойдет»).

Формализуя эти свойства, получаем определение алгебры.

Определение 1.1. Семейство $\mathcal{A}$ подмножеств множества Ω называется *алгеброй,* если верны три аксиомы:

(i) $\forall A \in \mathcal{A}, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$,

(ii) $\forall A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$,

(iii) $\Omega \in \mathcal{A}$.

↳ Из аксиом алгебры и формулы $A \cap B = \overline{\overline{A} \cup \overline{B}}$ следует, что пересечений событий является событием.

Пример. Наименьшей алгеброй является $\{\Omega, \varnothing\}$. А наибольшей — $\left\{\Omega, 2^{\Omega}\right\}$

Минус определения, которое мы только что дали, в том, что часто встречаются не конечные множества элементарных исходов, поэтому полезно, чтобы множество событий было замкнуто не только относительно объединения, но и относительно счетного объединения.

Определение 1.2. Семейство $\mathcal{A}$ подмножеств множества Ω называется *σ-алгеброй* (читается как сигма-алгебра), если

(i) $\forall A_1, \ldots, A_n, \ldots \in \mathcal{A} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{A}$,

(ii) $\forall A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$,

(iii) $\Omega \in \mathcal{A}$.

Не всегда по задаче понятно, какую сигма-алгебру надо выбрать, однако большой бонус ее аксиом заключается в том, что мы можем пересекать различные сигма-алгебры и получать снова сигма-алгебру. Из-за этого если нам известно неполное множество событий $\mathcal{K}$, т.е. $\mathcal{K} \subseteq \mathcal{A}$, то мы можем восстановить сигма-алгебру с точностью до тех событий, которых нет в $\mathcal{K}$. Иными словами, мы можем найти такую минимальную сигма-алгебру, которая содержит в себе $\mathcal{K}$ как подмножество.

Определение 1.3. Пусть $\mathcal{K}$ — класс подмножеств Ω. σ-алгебра $\sigma(\mathcal{K})$, *порожденная классом $\mathcal{K}$* — наименьшая σ-алгебра, содержащая $\mathcal{K}$, то есть любая σ-алгебра, содержащая $\mathcal{K}$, содержит и $\sigma(\mathcal{K})$.

Пример. σ-алгеброй, порожденной $\mathcal{K} = A$, будет являться $\sigma(A) = \left\{\varnothing, A, \overline{A}, \Omega\right\}$.

Сигма-алгебра является более узким понятием, нежели алгебра, то есть любая σ-алгебра является алгеброй, а обратное, вообще говоря, неверно.

Пример. Пусть $\Omega = \mathbb{R}$, $\mathcal{A}$ содержит конечные подмножества Ω и их дополнения. Для такого множества выполнены все аксиомы алгебры: $\Omega = \overline{\varnothing} \in \mathcal{A}$, объединение конечных множеств есть конечное множество, объединение конечного множества с дополнением к конечному множеству так же является дополнением к некоторому множеству. То же можно сказать и об объединении двух дополнений. Таким образом, $\mathcal{A}$ является алгеброй. Все элементы $\mathcal{A}$ либо конечны, либо континуальны, поэтому $\mathcal{A}$ не содержит $\mathbb{N}$. Но $\mathbb{N} = \bigcup\limits_{i=1}^{\infty}\{i\}$, то есть не выполнено свойство счетной аддитивности из определения σ-алгебры.

Вот мы и подошли к первому основополагающему определению.

Определение 1.4. Пара $(\Omega, \mathcal{A})$ называется *измеримым пространством*, если $\mathcal{A}$ является σ-алгеброй. Если же $\mathcal{A}$ — алгебра, то $(\Omega, \mathcal{A})$ — *измеримое пространство в широком смысле*.

По сути измеримое пространство нам говорит о том, какие у нас есть элементарные исходы и что мы считаем за событие. Однако если бы в известной игре «Кости» при броске кубика каждое число выпадало бы не с равной вероятностью, то это была бы уже совершенно другая игра, так что имеет смысл определить не только множество событий, но и вероятность того или иного события. Мы этим и займемся дальше.

1.3. Вероятность. Вероятностное пространство

С корабля — на бал или, по-современному, не отходя от кассы, сразу определим, что такое вероятность.

Определение 1.5. *Вероятностью* называется функция $\mathbb{P}\colon \mathcal{A} \to \mathbb{R}$, удовлетворяющая свойствам

1. $\forall A \in \mathcal{A} \mapsto \mathbb{P}(A) \geqslant 0$,
2. $\forall A_1, \ldots, A_n, \ldots \in \mathcal{A}, \quad A_i \cap A_j = \varnothing \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} \mathbb{P}(A_i)$,
3. $\mathbb{P}(\Omega) = 1$.

Второй пункт в определении вероятностной меры нельзя заменить аналогичным с конечными объединением и суммой. Однако если добавить к данному требованию так называемое свойство непрерывности вероятностной меры, т.е.

$$\forall B_1, B_2, \ldots \in \mathcal{A}, \quad B_{n+1} \subseteq B_n \Rightarrow \lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B), B = \bigcap_{n=1}^{\infty} B_n$$

то они вместе будут эквивалентны 2 из определения вероятности. Покажем это.

Утверждение 1.6. $\forall A_1, \ldots, A_n, \ldots \in \mathcal{A}, \quad A_i \cap A_j = \varnothing \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} A_i \Leftrightarrow$

$$\Bigg[\forall A_1, \ldots, A_n \in \mathcal{A} \quad A_i \cap A_j = \varnothing \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} A_i \wedge \forall B_1, B_2, \ldots \in \mathcal{A},$$
$$B_{n+1} \subseteq B_n \Rightarrow \lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(B), B = \bigcap_{n=1}^{\infty} B_n\Bigg].$$

Доказательство.
($\Rightarrow$)
Обозначим $C_n = B_n \setminus B_{n+1}$. Множества $B, C_1, C_2, \ldots$ не имеют общих точек.
$\forall n \quad B_n = \bigcup\limits_{k=n}^{\infty} C_k \bigcup B$. Тогда $\mathbb{P}(B_1) = \mathbb{P}(B) + \sum\limits_{k=1}^{\infty} \mathbb{P}(C_k)$. Отсюда следует, что ряд в правой части сходится, так как имеет конечную сумму. $\mathbb{P}(B_n) = \mathbb{P}(B) + \sum\limits_{k=n}^{\infty} \mathbb{P}(C_k)$.
При $n \to \infty$ сумма ряда стремится к нулю как остаточный член ряда из предыдущего выражения. В предельном переходе получаем свойство непрерывности.

($\Leftarrow$)
Рассмотрим произвольный набор $A_1, A_2, \ldots \in \mathcal{A} \quad A_i A_j = \varnothing$.
$\mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) = \mathbb{P}(\bigcup\limits_{i=1}^{n} A_i) + \mathbb{P}(\bigcup\limits_{i=n+1}^{\infty} A_i) = \sum\limits_{i=1}^{n} \mathbb{P}(A_i) + \mathbb{P}(\bigcup\limits_{i=n+1}^{\infty} A_i)$.
Обозначим $B_n = \bigcup\limits_{i=n+1}^{\infty} A_i, \quad B_{n+1} \subseteq B_n \quad \forall n, \quad \bigcap\limits_{n=1}^{\infty} B_n = \varnothing$
$\sum\limits_{i=1}^{\infty} \mathbb{P}(A_i) = \lim\limits_{n\to\infty} (\mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) - \mathbb{P}(B_n)) = \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) - \lim\limits_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i)$ ■

Определение 1.7. *Вероятностным пространством* $(\Omega, \mathcal{A}, \mathbb{P})$ называется измеримое пространство $(\Omega, \mathcal{A})$, снабженное вероятностью $\mathbb{P}$.

Как видите: и здесь нет ничего сложного. Идет по сути просто жонглирование математическими понятиями и задание одних определений.
Рубрика «Зачем так сложно?» Кому вообще нужна σ-алгебра событий и зачем весь этот огород, если можно рассматривать множество всех подмножеств множества событий Ω? Когда-то давно кто-то доказал, что в случае очень большого множества элементарных исходов, например, континуального, множество 2^{Ω} будет иметь такую крокодильски большую мощность, что вся теория сломается. Таким образом, алгебры нужны для того, чтобы вероятность имела хорошую область определения.

Перечислим свойства вероятности. Доказательства их можно найти в любом из классических учебников по теории вероятностей или можно их придумать самому: большинство из них тривиальны.

Свойства вероятности:

1. $\mathbb{P}(\varnothing) = 0$,
2. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$,
3. $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$,
4. $\mathbb{P}(A) \leq 1$,
5. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$,
6. $\mathbb{P}(A \cup B) \leqslant \mathbb{P}(A) + \mathbb{P}(B)$,
7. $\mathbb{P}\left(\bigcup\limits_{i=1}^{n} A_i\right) = \sum\limits_{k=1}^{n} \sum\limits_{i_1 < \dots < i_k} (-1)^{k+1} \mathbb{P}(A_{i_1} A_{i_2} \dots A_{i_k})$,
8. $\mathbb{P}\left(\bigcap\limits_{i=1}^{n} A_i\right) \geqslant 1 - \sum\limits_{i=1}^{n} \mathbb{P}(\overline{A_i})$ — неравенство Бонферрони.

Если порождать сигма-алгебру алгеброй, то вероятность будет определена однозначно вероятностью на данной алгебре. Это и есть содержание следующей теоремы.

Теорема 1.8 (Каратеодори)**.** Пусть $(\Omega, \mathcal{A})$ — измеримое пространство в широком смысле, а некоторая функция $\mathbb{P}$ обладает свойствами вероятностной меры. Тогда на измеримом пространстве $(\Omega, \sigma(\mathcal{A}))$

$$\exists!\ \mathbb{P}' : \forall A \in \mathcal{A} \mapsto \mathbb{P}(A) = \mathbb{P}'(A).$$

Доказательство. См. [2]. ■

Why? Зачем это нужно? Теорема Каратеодори говорит о том, что любую вероятностную меру, заданную на алгебре, можно однозначно продолжить на σ-алгебру, то есть расширить область ее определения. При этом значения функции на алгебре не изменятся. Теорема будет использоваться при определении интеграла Лебега.

Список литературы

[1] Королев В. Ю. Теория вероятностей и математическая статистика: учеб. М.: ТК Вебли, Изд-во Проспект, 2006. 160 с.

[2] Гусев Н. А. Заметки к курсу: мера и интеграл Лебега. 2019. https://www.ngusev.ru/mt/mt-lectures.pdf.