

Содержание

I	Теория вероятности	2
1	Вероятностное пространство	3
2	Условная вероятность. Независимость событий	6
3	Формула полной вероятности. Формула Байеса	8
4	Случайные величины	9
5	Понятие меры и интеграла Лебега	12

Часть I

Теория вероятности

1 Вероятностное пространство

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

1. Непредсказуемость
2. Воспроизводимость
3. Устойчивость частот

Для описания стохастических 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 .

Множество событий, обозначаемое \mathcal{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 событие («Что-нибудь да произойдет»)

Формализуя эти свойства, получаем определение алгебры.

Определение 1.1 Семейство \mathcal{A} подмножеств множества Ω называется алгеброй, если

1. $\forall A \in \mathcal{A}, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$
2. $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$
3. $\Omega \in \mathcal{A}$

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

Наименьшей возможной алгеброй является $\{\Omega, \emptyset\}$

Определение 1.2 Семейство \mathcal{A} подмножеств множества Ω называется σ -алгеброй, если

1. $\forall A_1, \dots, A_n, \dots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$
2. $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$
3. $\Omega \in \mathcal{A}$

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 $\sigma(A) = \{\emptyset, A, \bar{A}, \Omega\}$.

σ -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: $\Omega = \bar{\emptyset} \in \mathcal{A}$, 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 $\mathbb{N} = \bigcup_{i=1}^{\infty} \{i\}$, то есть не выполнено свойство счетной аддитивности из определения σ -алгебры.

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


Определение 1.5 Вероятностью называется функция $\mathbb{P}: \mathcal{A} \rightarrow \mathbb{R}_+$, удовлетворяющая свойствам

1. $\forall A \in \mathcal{A} \quad \mathbb{P}(A) \geq 0$

2. $\forall A_1, \dots, A_n, \dots \in \mathcal{A} \quad A_i \cap A_j = \emptyset \ (i \neq j) \Rightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

3. $\mathbb{P}(\Omega) = 1$

Определение 1.6 Вероятностным пространством $(\Omega, \mathcal{A}, \mathbb{P})$ называется измеримое пространство (Ω, \mathcal{A}) , снабженное вероятностью \mathbb{P} .

 Кому вообще нужна σ -алгебра событий, и зачем весь этот огород, если можно рассматривать множество всех подмножеств множества Ω ? Когда-то 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.

Свойства вероятности

1. $\mathbb{P}(\emptyset) = 0$

2. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

3. $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

4. $\mathbb{P}(A) \leq 1$

$$5. \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$$

$$6. \mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$$

$$7. \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n \sum_{i_1 < \dots < i_k} (-1)^{k+1} \mathbb{P}(A_{i_1} A_{i_2} \dots A_{i_k})$$

$$8. \mathbb{P}\left(\bigcap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n \mathbb{P}(\overline{A_i}) \text{ - неравенство Бонферрони}$$

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

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

, то они вместе будут эквивалентны 2. из определения вероятности.

Утверждение 1.1 $\forall A_1, \dots, A_n, \dots \in \mathcal{A} \quad A_i \cap A_j = \emptyset \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \Leftrightarrow$
 $(\forall A_1, \dots, A_n \in \mathcal{A} \quad A_i \cap A_j = \emptyset \ (i \neq j)) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) \wedge (\forall B_1, B_2, \dots \in \mathcal{A} \quad B_{n+1} \subseteq B_n \Rightarrow \lim_{n \rightarrow \infty} \mathbb{P}(B_n) = \mathbb{P}(B))$

Доказательство

\Rightarrow

Обозначим $C_n = B_n \setminus B_{n+1}$. Множества B, C_1, C_2, \dots не имеют общих точек.

$\forall n \quad B_n = \bigcup_{k=n}^{\infty} C_k \cup B$. Тогда $\mathbb{P}(B_1) = \mathbb{P}(B) + \sum_{k=1}^{\infty} \mathbb{P}(C_k)$. Отсюда следует, что ряд в

правой части сходится, так как имеет конечную сумму. $\mathbb{P}(B_n) = \mathbb{P}(B) + \sum_{k=n}^{\infty} \mathbb{P}(C_k)$. При $n \rightarrow \infty$ сумма ряда стремится к нулю как остаточный член ряда из предыдущего выражения. В предельном переходе получаем свойство непрерывности.

\Leftarrow

Рассмотрим произвольный набор $A_1, A_2, \dots \in \mathcal{A} \quad A_i A_j = \emptyset$.

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbb{P}\left(\bigcup_{i=n+1}^{\infty} A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) + \mathbb{P}\left(\bigcup_{i=n+1}^{\infty} A_i\right).$$

$$\text{Обозначим } B_n = \bigcup_{i=n+1}^{\infty} A_i, \quad B_{n+1} \subseteq B_n \quad \forall n, \quad \bigcap_{n=1}^{\infty} B_n = \emptyset$$

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \lim_{n \rightarrow \infty} (\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) - \mathbb{P}(B_n)) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) - \lim_{n \rightarrow \infty} \mathbb{P}(B_n) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \blacksquare$$

Теорема 1.1 (Каратеодори) Пусть (Ω, \mathcal{A}) — измеримое пространство в широком смысле, а некоторая функция \mathbb{P} обладает свойствами вероятностной меры. Тогда на измеримом пространстве $(\Omega, \sigma(\mathcal{A})) \exists! \mathbb{P}': \forall A \in \mathcal{A} \quad \mathbb{P}(A) = \mathbb{P}'(A)$

Доказательство отсутствует \blacksquare



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.

2 Условная вероятность. Независимость событий

Рассмотрим произвольное $B \in \mathcal{A}$: $\mathbb{P}(B) > 0$.

Определение 2.1 Условной вероятностью события $A \in \mathcal{A}$ при условии B называется $\frac{\mathbb{P}(AB)}{\mathbb{P}(B)} =: \mathbb{P}(A|B) = \mathbb{P}_B(A)$

Что это означает на пальцах? Условная вероятность $\mathbb{P}(A|B)$ — это вероятность того, что произойдет событие A , если мы точно знаем, что произошло событие B .



Графически это означает, что, когда произошло событие B , мы оказались в круге B . Тогда формула $\frac{\mathbb{P}(AB)}{\mathbb{P}(B)}$ есть просто вероятность попасть в AB .

Из определения следует так называемый «Закон умножения вероятностей»:

$$\mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(AB)$$

Легко проверяется, что $(B, \mathcal{A}_B, \mathbb{P}_B)$, где $\mathcal{A}_B = \{A \cap B | A \in \mathcal{A}\}$, так же является вероятностным пространством.

Зачем нужно требование $\mathbb{P}(B) > 0$, если можно в случае $\mathbb{P}(B) = 0$ доопределить условную вероятность нулем как вероятность при условии невозможного события? При таком доопределении нарушится аксиома 3. вероятности \mathbb{P}_B , поскольку $\mathbb{P}_B(B)$ по доопределению будет равно 0.

Определение 2.2 События $A, B \in \mathcal{A}$ называются независимыми, если

$$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$$

Для независимых событий

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

Пример Являются ли несовместные события ($AB = \emptyset$) независимыми? Нет, пусть $A, B \in \mathcal{A}$: $\mathbb{P}(A) > 0$, $\mathbb{P}(B) > 0$. Тогда $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B) = 0$, что является противоречием. По-простому, если произошло одно из несовместных событий, то второе уже не может произойти, и его условная вероятность равна 0, а не вероятности самого события, что требуется для независимости.

Следующее определение обобщает понятие независимости на произвольное количество событий.

Определение 2.3 События A_1, A_2, \dots, A_n называются независимыми в совокупности, если

$$\forall m = 2, \dots, n \quad \forall 1 \leq j_1 < \dots < j_m \leq n \quad \mathbb{P}\left(\bigcap_{k=1}^m A_{j_k}\right) = \prod_{k=1}^m \mathbb{P}(A_{j_k})$$

Пример На примере тетраэдра Бернштейна можно убедиться в том, что попарной независимости событий недостаточно для независимости в совокупности. Рассмотрим тетраэдр, у которого три стороны покрашены в красный, синий и зеленый, а четвертая содержит все три цвета. События $\{\text{выпадет красный}\}=\{K\}$, $\{\text{выпадет синий}\}=\{C\}$, $\{\text{выпадет зеленый}\}=\{Z\}$ попарно независимы (например, вероятность события $\{C\} \cap \{K\}$ равна вероятности выпадения четвертой грани, т. е. $\frac{1}{4}$, в то время как выпадения каждого цвета равна $\frac{1}{2}$). Однако $\mathbb{P}(\{C\} \cap \{K\} \cap \{Z\}) = \frac{1}{4} \neq (\frac{1}{2})^3$.

3 Формула полной вероятности. Формула Байеса

Определение 3.1 B_1, \dots, B_n образуют полную группу, если выполнены следующие условия:

1. $\mathbb{P}(B_i) > 0 \quad \forall i = 1, \dots, n$
2. $B_i B_j = \emptyset \quad (i \neq j)$
3. $\bigcup_{i=1}^n B_i = \Omega$

Теорема 3.1 Пусть B_1, \dots, B_n образуют полную группу. Вероятность события $A \in \mathcal{A}$ можно вычислить по формуле полной вероятности:

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

Доказательство

$$A = \bigcup_{i=1}^n AB_i, \quad AB_i \cap AB_j = \emptyset \quad (i \neq j)$$

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(AB_i) = \sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

Последний переход следует из закона умножения вероятностей. ■

Первое требование определения полной группы необходимо для возможности определить условную вероятность, второе позволяет разбить множество A на непересекающиеся части. Третье требование, вообще говоря, можно ослабить, потребовав, чтобы $A \subseteq \bigcup_{i=1}^n B_i$. Доказательство при этом не изменится.

Пример Проиллюстрировать формулу полной вероятности можно обычным экзаменом: A - {студент сдал экзамен}, B_i - {студент попал к преподавателю i }. Как и в любой другой лотерее, можно оценить вероятность попадания к преподавателю i , то есть $\mathbb{P}(B_i)$, а трезво оценивая свои силы можно прикинуть и вероятность сдать тому или иному преподавателю $\mathbb{P}(A|B_i)$. Зная все вышеперечисленное, несложно по формуле вычислить вероятность успешной сдачи.

Формула полной вероятности используется для вычисления априорной вероятности, т.е. вероятности события, которое еще не произошло. Пусть теперь $\mathbb{P}(A) > 0$. Тогда $\mathbb{P}(B_i|A) = \frac{\mathbb{P}(AB_i)}{\mathbb{P}(A)}$. Используя формулу полной вероятности, получаем формулу Байеса:

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^n \mathbb{P}(A|B_j)\mathbb{P}(B_j)}$$

Формула Байеса используется для вычисления апостериорной вероятности. То есть уже известно, что произошло некоторое событие A , и нужно вычислить вероятность того, что произошло некоторое B_i . В примере с экзаменом, например, может быть известно, что студент не сдал экзамен, и хочется вычислить вероятность того, что он сдавал преподавателю «Р».

4 Случайные величины

Случайные события — это хорошо, но с события типа «на монетке выпал герб» плохо формализуемы, а мы хотим формальности и математичности. Поэтому вместо всяких событий хочется работать с числами. Вот этим и займемся. При рассмотрении случайных событий мы ввели вероятностное пространство, которое выглядит так:


$$(\Omega, \mathcal{A}, \mathbb{P}),$$

где Ω — множество элементарных событий, \mathcal{A} — σ -алгебра подмножеств множества элементарных событий, а \mathbb{P} — вероятность. Мы же будем рассматривать теперь тройку

$$(\mathbb{R}, \mathcal{B}, \mathbb{P}),$$

где \mathbb{R} — действительная прямая, \mathcal{B} — борелевская σ -алгебра, а \mathbb{P} — вероятность. Поясним.

Определение 4.1 *Борелевской σ -алгеброй называется минимальная σ -алгебра, содержащая все открытые подмножества топологического пространства. Элементы борелевской σ -алгебры называются борелевскими множествами.*

 Мы будем рассматривать только топологическое пространство \mathbb{R} , так что это стремное словосочетание можно прямо сейчас забыть и понимать открытое множество как открытое множество из матана (все точки внутренние).

Пример Покажем, что все «хорошие» множества являются Борелевскими.

1. Все открытые интервалы входят по определению.
2. Отрезок вида $[a, b]$ входит как $\overline{(-\infty, a) \cup (b, +\infty)}$.
3. Точка ходит как вырожденный отрезок $[a, a]$.
4. Счетное объединение таких множеств входит по определению.


Теперь формально введем понятие случайной величины (может использоваться сокращение с.в.).

Определение 4.2 *Пусть $(\Omega, \mathcal{A}, \mathbb{P})$ — вероятностное пространство.*

Тогда случайной величиной ξ называется функция $\xi : \Omega \rightarrow \mathbb{R}$, измеримая относительно \mathcal{A} и \mathcal{B} . По-другому, ξ — случайная величина, если

$$\forall B \in \mathcal{B} \quad \xi^{-1}(B) = \{\omega : \xi(\omega) \in B\} \in \mathcal{A}$$

.

 Таким финтом ушами мы, по сути, сопоставили каждому событию какое-то «хорошее» множество на числовой прямой, и можем рассматривать не вероятности событий, а вероятности попадания в эти «хорошие» подмножества числовой прямой.

Введем еще несколько бесполезных определений, которые в дальнейшем использоваться не будут, но знать их не вредно.

Определение 4.3 С каждой случайной величиной свяжем два вероятностных пространства: первое — $(\Omega, \mathcal{A}_\xi, \mathbb{P})$ — вероятностное пространство, порожденное ξ . Здесь \mathcal{A}_ξ — наименьшая σ -алгебра, для которой выполняется свойство измеримости. Второе — $(\mathbb{R}, \mathcal{B}, \mathbb{P}_\xi)$, где $\mathbb{P}_\xi(B) = \mathbb{P}(\xi^{-1}(B)) \quad \forall B \in \mathcal{B}$ и называется распределением вероятностей ξ .

Идем дальше в сторону упрощения работы со случайностями. Вместо того чтобы рассматривать произвольные борелевские множества, мы будем рассматривать только множества вида $(-\infty, x)$. Действительно, интервал (a, b) получается из полупрямых так: $(a, b) = (-\infty, b) \setminus (-\infty, a]$. Таким образом, мы можем рассматривать случайные величины только на таких множествах. Здесь имеется в виду, что для удовлетворения определению случайной величины достаточно измеримости только на полупрямых, что следует из следующих свойств полного прообраза: прообраз объединения есть объединение прообразов, прообраз пересечения есть пересечение прообразов, прообраз отрицания есть отрицание прообраза. Выше показано, что из полупрямых с помощью этих операций можно получить интервалы, а из интервалов и все \mathcal{B} .

Теперь несколько полезных утверждений. Пусть ξ — случайная величина. Тогда $-\xi$ также случайная величина, так как её прообраз от любой полупрямой является прообразом ξ от симметричной полупрямой, то есть лежит в \mathcal{A} . $\xi + c$ также будет случайной величиной, поскольку её прообразом для любой полупрямой будет прообраз ξ для полупрямой, сдвинутой на c , то есть будет лежать в \mathcal{A} .

Утверждение 4.1 Пусть ξ, η — случайные величины. Тогда множество $\{\omega | \xi(\omega) < \eta(\omega)\}$ является событием.

Доказательство $\{\omega | \xi(\omega) < \eta(\omega)\} = \bigcup_{r \in \mathbb{Q}} \{\omega | \xi(\omega) < r, \eta(\omega) > r\}$. Заметим, что $\{\omega | \xi(\omega) < r\}$ является событием. Аналогично для η . Выражение, написанное выше, является счетным объединением пересечений двух событий, то есть событием. ■

Похожими махинациями, а также с использованием этого утверждения, доказывается, что $\xi^2, \xi + \eta, \xi\eta$ являются случайными величинами. Более того, если ξ_1, \dots, ξ_n — с.в., а функция $\phi(x_1, \dots, x_n)$ является непрерывной на множестве их значений, то $\phi(\xi_1, \dots, \xi_n)$ будет случайной величиной. Это не доказывалось.

Определение 4.4 Рассмотрим вероятностное пространство $(\Omega, \mathcal{A}, \mathbb{P})$ и определенную на нем случайную величину ξ . Тогда её функцией распределения $F_\xi(x)$ называется функция $F_\xi : \mathbb{R} \rightarrow \mathbb{R}$

$$F_\xi(x) = \mathbb{P}(\omega : \xi(\omega) < x) = \mathbb{P}(\xi < x)$$

Запись $\mathbb{P}(\xi < x)$ является в некотором смысле жаргонной, так как аргументов вероятности должно быть событие из \mathcal{A} . Но $\xi < x$ мы в дальнейшем будем отождествлять с объединением элементарных событий, образ которых меньше x . Из определения случайной величины получаем, что это объединение является событием, поэтому применение к нему функции вероятности корректно.

Функция распределения (сокращение ф.р.) является очень полезной штукой, поскольку имеет достаточно простой вид и несет в себе всю информацию о распределении, то есть однозначно определяет \mathbb{P}_ξ .

Рассмотрим основные свойства функции распределения:

1. $F_\xi(x) \in [0, 1]$
2. $\lim_{x \rightarrow -\infty} F_\xi(x) = 0$
3. $\lim_{x \rightarrow +\infty} F_\xi(x) = 1$
4. $F_\xi(x)$ монотонно не убывает.
5. $F_\xi(x)$ непрерывна слева.

Вероятность попадания с.в. в полуинтервал $\mathbb{P}_\xi[a, b) = F_\xi(b) - F_\xi(a)$. При стремлении $b \rightarrow a$ получим $\mathbb{P}(\xi = a) = F_\xi(a+) - F_\xi(a)$, то есть вероятность попадания в точку равна скачку функции распределения в этой точке.

Определение 4.5 Точка x_0 называется точкой роста $F_\xi(x)$, если $\forall \epsilon > 0 \quad \mathbb{P}(x_0 - \epsilon \leq \xi < x_0 + \epsilon) > 0$

Пример Это очень полезный пример, который будет использоваться в матстате и который очень любят спрашивать. Пусть ξ — случайная величина. $\eta = F_\xi(\xi)$. Чему равна $F_\eta(x)$? По определению $F_\eta(x) = \mathbb{P}(\eta < x) = \mathbb{P}(F_\xi(\xi) < x) = \mathbb{P}(\xi < F_\xi^{-1}(x)) = F_\xi(F_\xi^{-1}(x)) = x$. Таким образом η имеет равномерное распределение.

5 Понятие меры и интеграла Лебега