Содержание

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 ситуаций необходимо определить функцию вероятности. Её область определения назывется 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 множетсвом элементарных исходов и обозначается Ω.

Множество событий, обозначаемое $\mathscr{A}$, 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 событие («Что-нибудь да произойдет»)

Формализуя эти свойства, получаем определение алгебры.

Определение 1.1 *Семейство $\mathscr{A}$ подмножеств множества Ω называется <u>алгеброй</u>, если*

1. $\forall A \in \mathscr{A}, B \in \mathscr{A} \Rightarrow A \bigcup B \in \mathscr{A}$
2. $\forall A \in \mathscr{A} \Rightarrow \overline{A} \in \mathscr{A}$
3. $\Omega \in \mathscr{A}$

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

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

Определение 1.2 *Семейство $\mathscr{A}$ подмножеств множества Ω называется <u>σ-алгеброй</u>, если*

1. $\forall A_1, \dots, A_n, \dots \in \mathscr{A} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{A}$
2. $\forall A \in \mathscr{A} \Rightarrow \overline{A} \in \mathscr{A}$
3. $\Omega \in \mathscr{A}$

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

Пример σ-алгеброй, порожденной $\mathscr{K} = A$, будет являться $\sigma(A) = \Big\{\varnothing, A, \bar(A), \Omega\Big\}$.

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

Пример Пусть $\Omega = \mathbb{R}$, $\mathscr{A}$ содержит конечные подмножества Ω и их дополнения. Для такого множества выполнены все аксиомы алгебры: $\Omega = \overline{\varnothing} \in \mathscr{A}$, 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, $\mathscr{A}$ является алгеброй. Все элементы $\mathscr{A}$ либо конечны, либо континуальны, поэтому $\mathscr{A}$ не содержит $\mathbb{N}$. Но $\mathbb{N} = \bigcup\limits_{i=1}^{\infty}\{i\}$, то есть не выполнено свойство счетной аддитивности из определения σ-алгебры.

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

Определение 1.5 *Вероятностью называется функция $\mathbb{P}\colon \mathscr{A} \to \mathscr{R}_+$, удовлетворяющая свойстам*

1. $\forall A \in \mathscr{A} \quad \mathbb{P}(A) \geq 0$
2. $\forall A_1, \ldots, A_n, \ldots \in \mathscr{A} \quad A_i \bigcap A_j = \varnothing \ (i \neq j) \Rightarrow \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} A_i$
3. $\mathbb{P}(\Omega) = 1$

Определение 1.6 *Вероятностным пространством $(\Omega, \mathscr{A}, \mathbb{P})$ называется измеримое пространство $(\Omega, \mathscr{A})$, 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. $\mathbb{P}(\varnothing) = 0$
2. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
3. $A \subseteq B \quad \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
4. $\mathbb{P}(A) \leq 1$

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

6. $\mathbb{P}(A \bigcup B) \le \mathbb{P}(A) + \mathbb{P}(B)$

7. $\mathbb{P}(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{k=1}^{n} \sum\limits_{i_1 < \dots < i_k} (-1)^{k+1} \mathbb{P}(A_{i_1} A_{i_2} \dots A_{i_k})$

8. $\mathbb{P}(\bigcap\limits_{i=1}^{n} A_i) \ge 1 - \sum\limits_{i=1}^{n} \mathbb{P}(\overline{A_i})$ - 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

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

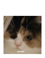
, то они вместе будут эквивалентны 2. из определения вероятности.

Утверждение 1.1 $\forall A_1, \ldots, A_n, \ldots \in \mathscr{A} \quad A_i \bigcap A_j = \varnothing \ (i \ne j) \Rightarrow \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} A_i \Leftrightarrow$ $(\forall A_1, \ldots, A_n \in \mathscr{A} \quad A_i \bigcap A_j = \varnothing \ (i \ne j) \Rightarrow \mathbb{P}(\bigcup\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} A_i) \wedge (\forall B_1, B_2, \ldots \in \mathscr{A} \quad B_{n+1} \subseteq B_n \Rightarrow \lim\limits_{n \to \infty} \mathbb{P}(B_n) = \mathbb{P}(B))$

Доказательство
$\Rightarrow$
Обозначим $C_n = B_n \setminus B_n + 1$. Множества $B, C_1, C_2, \ldots$ 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 $A_1, A_2, \ldots \in \mathscr{A} \quad A_i A_j = \varnothing$.
$\mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) = \mathbb{P}(\bigcup\limits_{i=1}^{n} A_i) + \mathbb{P}(\bigcup\limits_{i=n+1}^{\infty} A_i) = \sum\limits_{i=1}^{n} \mathbb{P}(A_i) + \mathbb{P}(\bigcup\limits_{i=n+1}^{\infty} A_i)$.
Обозначим $B_n = \bigcup\limits_{i=n+1}^{\infty} A_i, \quad B_{n+1} \subseteq B_n \quad \forall n, \quad \bigcap\limits_{n=1}^{\infty} B_n = \varnothing$
$\sum\limits_{i=1}^{\infty} \mathbb{P}(A_i) = \lim\limits_{n \to \infty} (\mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) - \mathbb{P}(B_n)) = \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i) - \lim\limits_{n \to \infty} \mathbb{P}(B_n) = \mathbb{P}(\bigcup\limits_{i=1}^{\infty} A_i)$ ■

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

Доказательство отсутсвует ■

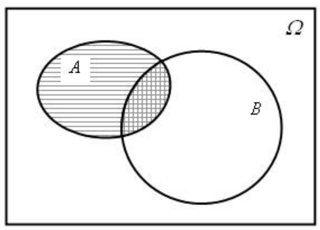
 Зачем это нужно? Теорема Каратеодори говорит о том, что любую вероятностную меру, заданную на алгебре, можно однозначно продолжить на σ-алгебру, то есть расширить область ее определения. При этом значения функции на алгебре не изменятся. Теорема будет использоваться при определении интеграла Лебега.

2 Условная вероятность. Независимость событий

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

Определение 2.1 *<u>Условной вероятностью</u> события $A \in \mathscr{A}$ 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 веротяность 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 $(B, \mathscr{A}_B, \mathbb{P}_B)$, где $\mathscr{A}_B = \{A \bigcap B | A \in \mathscr{A}\}$, 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 $A, B \in \mathscr{A}$ называются <u>независимыми</u>, 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 $(AB = \varnothing)$ независимыми? Нет, пусть $A, B \in \mathscr{A}: \quad \mathbb{P}(A) > 0,\ \mathbb{P}(B) > 0$. 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 веротяность равна 0, а не вероятности самого события, что требуется для независимости.

Следующее определение обобщает понятие независимости на произвольное количество событий.

Определение 2.3 *События $A_1, A_2, \ldots, A_n$ называются <u>независимыми в совокупности</u>, если*

$$\forall m = 2, \ldots, n \quad \forall 1 \le j_1 < \ldots < j_m \le n \quad \mathbb{P}\Big(\bigcap_{k=1}^{m} A_{j_k}\Big) = \prod_{k=1}^{m} \mathbb{P}(A_{j_k})$$

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

3 Формула полной веротяности. 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. $B_i B_j = \varnothing \quad (i \neq j)$
3. $\bigcup\limits_{i=1}^{n} B_i = \Omega$

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

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

Доказательство
$A = \bigcup\limits_{i=1}^{n} AB_i, \quad AB_i \bigcap AB_j = \varnothing \quad (i \neq j)$
$\mathbb{P}(A) = \sum\limits_{i=1}^{n} \mathbb{P}(AB_i) = \sum\limits_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$
Последний переход следует из закона умножения вероятностей. ■

Первое требование определения полной группы необходимо для возможности определить условную веротяность, второе позволяет разбить множество A на непересекающиеся части. Третье требование, вообще говоря, можно ослабить, потребовав, чтобы $A \subseteq \bigcup\limits_{i=1}^{n} B_i$. 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 вышеперечисленое, несложно по формуле вычислить вероятность успешной сдачи.

Формула полной веротяности 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:

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

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 Рублеву.