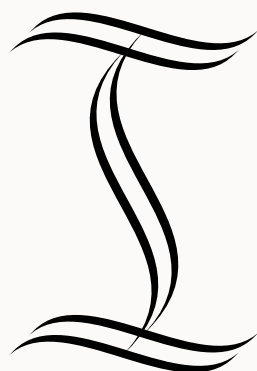


ВМК МГУ & ФУПМ МФТИ

Теория вероятности. Математическая статистика

Игорь Тао, Дима Сотников, Дима Гуцин



13 августа 2019 г.



Содержание

Часть I	Теория вероятности	2
1	Вероятностное пространство	3
1.1	Стохастические ситуации	3
1.2	Измеримое пространство	3
1.3	Вероятность. Вероятностное пространство	5
2	Условная вероятность и независимость событий	9
2.1	Условная вероятность	9
2.2	Независимость событий. Независимость в совокупности	10



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] В. Ю. Королева, так что любознательный читатель может ознакомиться с первоисточником при особом желании. Следуя Королеву, будем говорить о *случайности* как о принципиально неустрашимой неопределенности. Чтобы понять, как с ней же работать, необходимо выделить те случаи, когда мы можем сказать: «А здесь нам поможет теория вероятностей», — и не облажаться.

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. В ином случае использовать теорию вероятностей — что по воробьям из пушки палить!
- ❖ **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, частота исходов должна иметь предел при стремлении числа опытов к ∞ .

А где мы можем встретить такие ситуации, а? Правильно! В казино! Исторически теория вероятности создавалась, чтобы играть в азартные игры с опорой на науку, отсюда и типичные задачи: бросаем кости, гоняем шары в лототроне, делим ставку, и т. д.

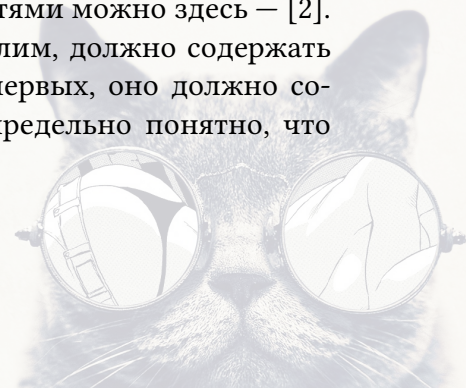
Для описания стохастических ситуаций необходимо определить функцию вероятности. Её область определения называется *множеством событий*. В свою очередь, событие (такое как, например, выпадание чётного числа на кубике) может являться совокупностью неких более простых событий, описывающих стохастическую 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 только суть. Быстрым темпом! И рассказываются начала теории меры, которые и приведут нас к определению вероятностного пространства. Строго излагать их мы, конечно, не будем, но обмазаться всякими формальностями можно здесь — [2].

Вероятностное пространство, которое мы дальше определим, должно 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 смысла 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 \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 событием («что-нибудь да произойдет»).

Если мы посадим эти три свойства в одну лодку, то получим следующее

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

- (i) $\forall A \in \mathcal{A}, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$ (если объединить два множества из семейства, то обязательно получим множество из семейства),
- (ii) $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$ (множество и его дополнение одновременно: либо лежат в нашем семействе, либо нет),
- (iii) $\Omega \in \mathcal{A}$ (все множество точно лежит в \mathcal{A}).

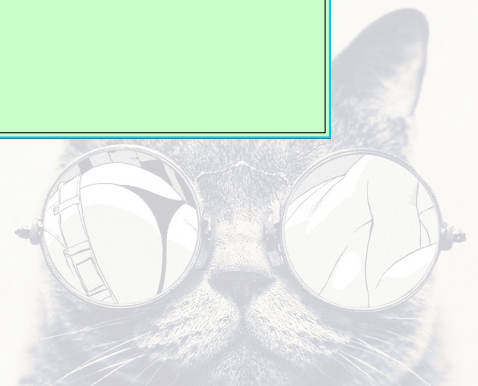
↳ Из аксиом алгебры и формулы $A \cap B = \overline{\bar{A} \cup \bar{B}}$ следует, что пересечение событий является событием. Подобным образом можно показать, что и разность, и симметрическая разность двух множеств из семейства \mathcal{A} тоже дадут множество из \mathcal{A} .

Пример. Наименьшей алгеброй является $\{\Omega, \emptyset\}$.

Косяк определения, которое мы только что дали, в том, что часто встречаются бесконечные 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) $\forall A_1, \dots, A_n, \dots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$,
- (ii) $\forall A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$,
- (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 $\mathcal{K} = A$ (класс состоит из одного элементарного события), будет являться $\sigma(A) = \{\emptyset, A, \bar{A}, \Omega\}$.

Сигма-алгебра является более узким понятием, нежели алгебра, то есть любая σ -алгебра является алгеброй, а обратное, вообще говоря, неверно. То есть это такая «элитка» среди всевозможных алгебр: не всем алгебрам дано быть «элиткой».

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. Вот видите: это не достойное нашего внимания семейство, оно не настолько элитное, чтобы называться сигма-алгеброй.

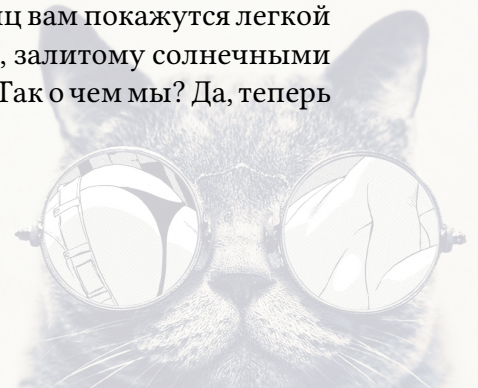
Вот мы и подошли к первому основополагающему определению.

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.

По сути измеримое пространство нам говорит о том, какие у нас есть элементарные исходы и что мы считаем за событие. Но если при бросании кости ты будешь использовать кость со смещенным центром тяжести, то игроки что-то заподозрят и популярно пояснят тебе, что ты неправ (поверь, тебе это не понравится), 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

Ну чё, народ, погнажи на!.. На новые баррикады! Они не такие высокие, как те, которые мы уже перепрыгнули, так что следующие пару страниц вам покажутся легкой прогулкой теплым осенним вечером по желтеющему бульвару, залитому солнечными лучами заходящего солнца. . . Кхм-Кхм. . . Это из другой пьесы. Так о чем мы? Да, теперь пора определить вероятность.



Определение 1.5. Пусть (Ω, \mathcal{A}) — измеримое пространство (\mathcal{A} — σ -алгебра). Вероятностью называется функция $\mathbb{P}: \mathcal{A} \rightarrow \mathbb{R}$, удовлетворяющая свойствам

1. $\forall A \in \mathcal{A} \mapsto \mathbb{P}(A) \geq 0$,
2. $\forall A_1, \dots, A_n, \dots \in \mathcal{A}, \quad A_i \cap A_j = \emptyset \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$,
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.

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

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

Утверждение 1.6.

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

$$\left[\forall A_1, \dots, A_n \in \mathcal{A} \quad A_i \cap A_j = \emptyset \ (i \neq j) \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) \right] \wedge$$

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

Доказательство.

\Rightarrow

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

$\forall n B_n = \bigcup_{k=n}^{\infty} C_k \cup B$. 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 $A_1, A_2, \dots \in \mathcal{A} \quad A_i \cap A_j = \emptyset$.

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

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



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

Определение 1.7. Вероятностным пространством $(\Omega, \mathcal{A}, \mathbb{P})$ называется измеримое пространство (Ω, \mathcal{A}) , с заданной на нём вероятностью \mathbb{P} .

Ничего сложного, если знать определения и уметь вертеть математическими понятиями и логикой. Главное не перепутайте: вертеть, а не класть.

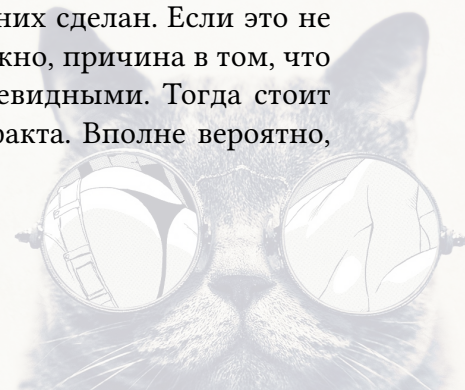
ЧО?! Кому вообще нужна σ -алгебра событий и зачем весь этот огород и маленькая тележка, если можно рассматривать множество всех подмножеств множества событий Ω ? Когда-то давно кто-то доказал, что в случае очень большого множества элементарных исходов, например, континуального (это много, поверьте наслову), множество 2^{Ω} будет иметь такую крокодильски, или алигаторски, 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. Да, реально тривиальны, а не как на матеше в школе.

Свойства вероятности:

1. $\mathbb{P}(\emptyset) = 0$,
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. $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$,
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. $\mathbb{P} \left(\bigcap_{i=1}^n A_i \right) \geq 1 - \sum_{i=1}^n \mathbb{P}(\overline{A_i})$ — неравенство Бонферрони.

Прими к факту, что в силу простоты этих определений, математики используют их на уровне интуиции, то есть ты можешь не заметить, а при этом в паре строк уже будет использовано несколько этих свойств. В некотором роде это и достоинство теории вероятности, так как происходит подобное из-за ясной бытовой аналогией, связанной всем распространённостью азартных игр. С другой стороны, есть отличный способ избежать возможности быть посаженным на что-нибудь за отсутствие формализма: для этого стоит каждый раз, как перед тобой будут выписаны математические выкладки, пробовать осмыслить каждый переход, который в них сделан. Если это не получается сделать, то не стремись хвататься за голову: возможно, причина в том, что используются какие-то свойства, которые автор посчитал очевидными. Тогда стоит поискать другие в других источниках доказательства этого факта. Вполне вероятно,



что это поможет найти такое доказательство, в котором будет все предельно понятно или в тяжелых моментах будет указано, что стоит перечитать и повторить.

В заключение вырежем клинышком в этой глиняной массе

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

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

Доказательство. Ушло и не вернулось. Последний раз его видели здесь [2]. ■

Зочем? 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 и независимость событий

2.1. Условная вероятность

Условно мы будем тут работать с вероятностью нескольких событий. Ведь бывают разные игры: и в карты, и в кости. Есть и такие, в которых придется не единожды просить судьбу «выдать козырь», поэтому выгодно понимать, сколько в действительности у тебя шансов дёрнуть судьбу так, как тебя тогда на рынке дёрнули.

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

Определение 2.1. Условной вероятностью события $A \in \mathcal{A}$ при условии B называется

$$\frac{\mathbb{P}(AB)}{\mathbb{P}(B)} \stackrel{\text{def}}{=} \mathbb{P}(A|B) = \mathbb{P}_B(A).$$

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

Пример. Представь, что ты бросаешь 2 кубика. Событие A — у тебя выпало суммарно меньше 5 очков. А B , например, означает, что у тебя выпало 6 очков на первом кубике. Очевидно, что если на одном выпало 6, то суммарно у тебя меньше 6 быть не может, так что вероятность тоже должна быть равна нулю, что согласуется с определением:

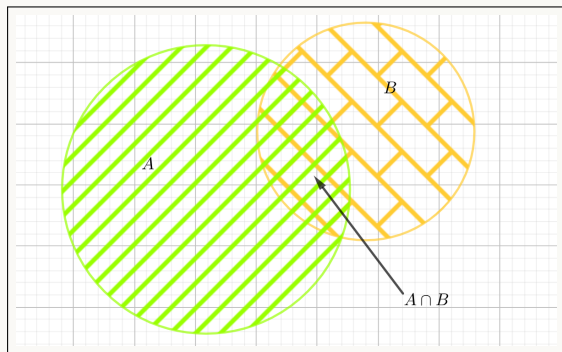
$$P_B(A) = \frac{P(AB)}{P(B)} = \frac{0}{P(B)} = 0.$$

Допустим теперь, что B — у тебя выпало на первом кубике 1 очко. Тогда понятно, что вероятность события A — будет равна $1/2$, так как при выпадении 1, 2, 3 на втором кубике событие произойдет, а при выпадении других чисел — нет. Вычислим теперь по определению.

Вероятность выпадения 1 на первом кубике, притом чтобы сумма была меньше пяти, равна $1/12$, так как всего 36 комбинаций на кубиках, а подходит только 3: $(1, 1)$, $(1, 2)$, $(1, 3)$. Вероятность выпадения любого из чисел на кубике — $1/6$, так что

$$P_B(A) = \frac{P(AB)}{P(B)} = \frac{1/12}{1/6} = \frac{1}{2}.$$

Опа! Америка-Европа! Все сошлось с интуитивными предположениями.



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 $A \cap B$.



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.

ЧО?! 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.

2.2. Независимость событий. Независимость в совокупности

Да, и в теории вероятности есть независимость! Но это уже другая независимость, а не та, про которую так много новостей нынче.

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 ($A \cap B = \emptyset$) независимыми? Ответ: 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.

Обобщим понятие независимости на произвольное количество событий.

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

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

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 «выпадет красный» = {K}, «выпадет синий» = {C}, «выпадет зеленый» = {3} попарно независимы (например, вероятность события $\{C\} \cap \{K\}$ равна вероятности выпадения четвертой грани, т. е. $1/4$, в то время как вероятность выпадения каждого цвета равна $1/2$). Однако

$$\mathbb{P}(\{C\} \cap \{K\} \cap \{3\}) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3.$$



Список литературы

- [1] Королев В. Ю. Теория вероятностей и математическая статистика: учеб. М.: ТК Вебли, Изд-во Проспект, 2006. с. 160.
- [2] Гусев Н. А. Заметки к курсу: мера и интеграл Лебега. 2019. <https://www.ngusev.ru/mt/mt-lectures.pdf>.

