



ВЕРОЯТНОСТЕЙ ТЕОРИЯ

Авторы: Ю. В. Прохоров, Б. А. Севастьянов

ВЕРОЯТНОСТЕЙ ТЕОРИЯ, раздел математики, изучающий математич. модели случайных явлений. В. т. является основой мн. математич. дисциплин, напр. математической статистики, теории массового обслуживания, теории надёжности, финансовой и актуарной математики.

Экспериментальные основы теории вероятностей

Возможность изучения случайных событий основана на том, что массовые случайные явления в неизменных условиях обладают закономерностью, называемой статистич. устойчивостью частот, которая заключается в следующем. Пусть случайное событие A может произойти или не произойти при осуществлении некоторого комплекса условий S . Если условия S реализуются N раз, то говорят, что произведено N испытаний. Отношение $N(A)/N$, где $N(A)$ – число появлений события A при N испытаниях, называется относительной частотой события A . С ростом N относительная частота $N(A)/N$ колеблется около некоторого числа, называемого вероятностью события A и обычно обозначаемого $\mathsf{P}(A)$. Так, при большом числе бросаний монеты орёл появляется примерно в половине случаев, поэтому вероятность появления орла можно считать равной $1/2$. Статистика рождений показывает, что мальчиков рождается больше, чем девочек, причём наблюдаемая доля рождений мальчиков равна 0,51–0,52; поэтому вероятность рождения мальчика несколько больше $1/2$. См. также [Вероятность](#).

Основные понятия теории вероятностей

Исходя из данных событий A_1, \dots, A_r , можно определить их объединение и пересечение. Объединением событий A_1, \dots, A_r называют событие B , которое происходит тогда и только тогда, когда в данном испытании наступает хотя бы одно из событий A_1, \dots, A_r . Пересечением (или произведением, или совмещением) событий A_1, \dots, A_r называется событие C , которое происходит тогда и только тогда, когда в данном испытании наступают все события A_1, \dots, A_r . Для каждого события A вводится противоположное событие \overline{A} , которое происходит тогда и только тогда, когда A не происходит.

Для объединения B событий A_1, \dots, A_r обычно используются обозначения $B = A_1 \cup A_2 \cup \dots \cup A_r = \bigcup_{i=1}^r A_i$, а для пересечения $C = A_1 \cap A_2 \cap \dots \cap A_r = \bigcap_{i=1}^r A_i$. Иногда пишут $C = A_1 \dots A_r$.

В т. н. урновой схеме предполагается, что в урне содержатся шары, которые обозначаются элементами ω некоторого конечного множества Ω . Из урны случайным образом извлекается один шар ω . Если $\omega \in A$, где A – подмножество Ω , то говорят, что произошло событие A . Всё множество Ω называется достоверным событием, т. к. всегда $\omega \in \Omega$, а пустое множество \emptyset – невозможным событием, т. к. всегда $\omega \notin \emptyset$.

В В. т. вероятность P вводится аксиоматически. Предполагается, что события A образуют класс

подмножеств некоторого пространства элементарных исходов (элементарных событий) $\Omega = \{\omega\}$, этот класс Σ подмножеств является σ -алгеброй, т. е. Σ содержит невозможное \emptyset и достоверное Ω события, а также замкнут относительно образования разностей двух событий и объединения и пересечения событий в конечном или счётном числе. Вероятность P определена на всех множествах $A \in \Sigma$ и удовлетворяет следующим аксиомам:

A1. $\mathsf{P}(A) \geq 0$ (неотрицательность),

A2. $\mathsf{P}(\Omega) = 1$ (нормированность),

A3. $\mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i)$, если $A_i \cap A_j = \emptyset$ при $i \neq j$ (счётная аддитивность).

Тройка $(\Omega, \Sigma, \mathbf{P})$, в которой \mathbf{P} удовлетворяет аксиомам A1, A2, A3, называется вероятностным пространством.

Эта аксиоматика была предложена в 1933 А. Н. [Колмогоровым](#) и является наиболее распространённой логич. основой построения В. т. Свойствам неотрицательности, нормированности и конечной аддитивности удовлетворяют относит. частоты $N(A)/N$ реальных случайных событий, поэтому естественно было потребовать, чтобы этим же свойствам удовлетворяли и вероятности $\mathbf{P}(A)$, к которым близки относительные частоты. Требование счётной аддитивности вероятности \mathbf{P} необходимо для создания полноценной математич. теории. При построении вероятностных пространств вероятность \mathbf{P} может задаваться разными способами. Напр., если Ω – конечное множество, вероятностное пространство называется конечным, и в этом случае вероятность \mathbf{P} можно задать с помощью вероятностей $p(\omega)$ элементарных исходов $\omega \in \Omega$, эти вероятности удовлетворяют условиям

$p(\omega) \geq 0, \omega \in \Omega, \sum_{\omega \in \Omega} p(\omega) = 1$ а значение вероятности $\mathsf{P}(A)$ события A задаётся формулой $\mathsf{P}(A) = \sum_{\omega \in A} p(\omega) \leq 1$ Часто элементарные исходы $\omega \in A$ называются исходами, благоприятствующими событию A .

В том случае, когда есть основания считать элементарные исходы равновероятными, все $p(\omega)$ считают равными друг другу и получают в качестве частного случая (1) т. н. классич. определение вероятности $\mathsf{P}(A) = |A| / |\Omega|$, где $|A|$ – число элементов множества A , т. е. вероятность события A равна отношению числа исходов, благоприятствующих событию A , к общему числу элементарных исходов. Этот подход широко используется в вероятностной комбинаторике и вопросах защиты информации.

Другой важный случай, когда вероятность P задаётся исходя из обобщения понятия равновероятности, может быть описан следующим образом. Пусть Ω – некоторое ограниченное множество евклидова пространства, имеющее объём $V(\Omega)$ (соответственно длину или площадь в одномерном и двумерном случаях). Пусть ω – случайно взятая в Ω точка; полагая, что вероятность попасть точке ω в множество $A \subset \Omega$ пропорциональна его объёму $V(A)$, получают т. н. геометрич. определение вероятности $\mathsf{P}(A) = V(A)/V(\Omega)$. Это определение используется в интегральной геометрии.

Вероятность P , удовлетворяющая аксиомам A1 – A3, является нормированной мерой на σ -алгебре Σ подмножеств Ω (см. [Мера множества](#)). Таким образом, В. т. может с формальной точки зрения рассматриваться

как часть теории меры. Однако осн. проблемы В. т. и теории меры различны, что во многом связано со специфическим для В. т. понятием независимости.

Условную вероятность $P(A|B)$ события A при условии B определяют формулой $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $\text{если } P(B) \neq 0$. Событие A называется независимым (стохастически независимым) от события B , если $P(A|B) = P(A)$.
(3) Условие (3) можно записать в симметричной форме: $P(A \cap B) = P(A)P(B)$ (4)

В более общем случае σ -алгебры событий $\mathcal{A}_1, \dots, \mathcal{A}_r \subset \mathcal{A}$ называются независимыми, если для любых $A_i \in \mathcal{A}_i, i=1, \dots, r$, справедливо равенство $P(A_1 \cap \dots \cap A_r) = P(A_1)P(A_2) \dots P(A_r)$. События из различных независимых σ -алгебр называются независимыми.

Понятие независимости и условные вероятности оказываются особенно полезными при рассмотрении составных испытаний. В простых случаях испытание – это осуществление некоторых условий, при которых происходит одно и только одно из событий $\{A_i\}$, называемых исходами испытания. В вероятностном пространстве (Ω, \mathcal{A}, P) испытанию соответствует разбиение, где A_i попарно несовместимы (несовместны), т. е. $A_i \cap A_j = \emptyset$ при $i \neq j$. Говорят, что испытание T составлено из испытаний $T_1, T_2, \dots, T_{n-1}, T_n$, если каждый исход испытания T есть совмещение некоторых исходов $A_i, B_j, \dots, U_k, V_l$ соответствующих испытаний $T_1, T_2, \dots, T_{n-1}, T_n$. Из тех или иных соображений часто бывают известны вероятности $P(A_i), P(B_j|A_i), \dots, P(V_l|A_i \cap B_j \cap \dots \cap U_k)$. (5)

По вероятностям (5) с помощью (2) могут быть определены вероятности $P(E)$ для любого события вида $E = A_i \cap B_j \cap \dots \cap U_k \cap V_l$.

Наиболее значительными с практич. точки зрения представляются два типа составных испытаний, в первом из которых испытания T_1, T_2, \dots, T_n независимы, т. е. вероятности (5) равны безусловным вероятностям $P(A_i), P(B_j), \dots, P(V_l)$, а во втором на вероятности исходов к.-л. испытания влияют результаты лишь непосредственно предшествующего испытания, т. е. вероятности (5) равны соответственно $P(A_i), P(B_j|A_i), \dots, P(V_l|A_i \cap B_j \cap \dots \cap U_k)$. В этом случае говорят об испытаниях, связанных в [Маркова цепь](#); вероятности всех событий, связанных с составным испытанием, определяются здесь начальными вероятностями $P(A_i)$ и т. н. переходными вероятностями $P(B_j|A_i), \dots, P(V_l|A_i \cap B_j \cap \dots \cap U_k)$.

Исходам испытаний могут соответствовать к.-л. числовые значения, в этом случае говорят о случайных величинах. Если задано вероятностное пространство (Ω, \mathcal{A}, P) , то случайная величина X – это функция $X(\omega)$ от элементарного исхода ω , для которой определена функция распределения $F_X(x) = P\{X \leq x\}$, $-\infty < x < \infty$.

Важный класс распределений составляют абсолютно непрерывные распределения, для которых существуют т. н. плотности вероятности $p_X(x)$, для этих распределений $F_X(x) = \int_{-\infty}^x p_X(u) du$.

Другой класс распределений – дискретные распределения; они задаются конечным или счётным числом точек x_k действительной прямой \mathbb{R} и вероятностями $P\{X = x_k\}$ так, что для любого $B \subseteq \mathbb{R}$ $P_X(B) = P\{X \in B\} = \sum_{x_k \in B} P\{X = x_k\}$.

Примерами абсолютно непрерывных распределений могут служить нормальное распределение, задаваемое плотностью

$$p_X(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (6)$$

где a и σ – параметры нормального распределения, $a \in \mathbb{R}$, $\sigma > 0$, а также показательное распределение, задаваемое плотностью $p_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$; $p_X(x) = 0$, $x < 0$,

где λ – параметр показательного распределения, $\lambda > 0$.

Примером дискретного распределения служит биномиальное распределение, задаваемое вероятностями

$$\mathbf{P}\{X=k\} = C_n^k p^k (1-p)^{n-k}, \quad k=0,1,\dots,n, \quad (7)$$

где X – число успехов в n испытаниях в [Бернулли схеме](#), p – вероятность успеха, $0 \leq p \leq 1$.

Часто вместо распределения вероятностей случайной величины можно ограничиться использованием небольшого количества числовых характеристик распределения. Из них наиболее употребительны [математическое ожидание](#) и [дисперсия](#).

При одновременном изучении нескольких случайных величин вводится совместное распределение, которое для случайных величин X_1, \dots, X_n задаётся функцией совместного распределения $F_{\{X_1, \dots, X_n\}}$

$(x_1, \dots, x_n) = \mathbf{P}\{X_1 < x_1, \dots, X_n < x_n\}$, где $-\infty < x_1, \dots, x_n < \infty$. Случайные величины X_1, \dots, X_n называются независимыми, если $F_{\{X_1, \dots, X_n\}}(x_1, \dots, x_n) = F_{\{X_1\}}(x_1) \dots F_{\{X_n\}}(x_n)$ для любых x_1, \dots, x_n , $-\infty < x_1, \dots, x_n < \infty$.

Предельные теоремы

С помощью совместного распределения случайных величин можно вычислить вероятность любого события, определяемого этими величинами, напр. события $a < X_1 + \dots + X_n < b$. Вычисление точных вероятностей таких событий, как правило, связано со значит. трудностями, поэтому обычно используются т. н. [предельные теоремы](#), которые позволяют получать приближённые значения таких вероятностей (с оценкой точности приближения).

Одним из примеров применения предельных теорем в В. т. может служить замена значения вероятности (7), трудно вычисляемой при больших n , приближённым значением $\frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{x^2}{2}}$, где $x = (k - np) / \sqrt{np(1-p)}$.

При формальном изложении В. т. предельные теоремы появляются в виде своего рода надстройки над её элементарными разделами, в которых все задачи имеют конечный, чисто арифметич. характер. Однако предельными теоремами раскрывается познавательная ценность В. т. Так, [Бернулли теорема](#) показывает, что при независимых испытаниях частота появления к.-л. события, как правило, мало отклоняется от его вероятности, а теорема Муавра – Лапласа указывает вероятности тех или иных отклонений. Смысл таких характеристик случайной величины, как её математич. ожидание и дисперсия, проявляется в [больших чисел законе](#) и [центральной предельной теореме](#).

Пусть X_1, X_2, \dots – независимые случайные величины, имеющие одно и то же распределение вероятностей с математич. ожиданием $\mathbf{E}X_k = a$ и дисперсией $\mathbf{D}X_k = \sigma^2$, и $S_n = (X_1 + \dots + X_n) / n$ – среднее арифметическое первых n величин этой последовательности.

В соответствии с законом больших чисел, каково бы ни было число $\varepsilon > 0$, вероятность неравенства $|\bar{S}_n - a| \geq \varepsilon$ при $n \rightarrow \infty$ имеет пределом 1 и, таким образом, \bar{S}_n , как правило, мало отличается от a (это – аналог устойчивости частот). Центральная предельная теорема уточняет этот результат, показывая, что отклонения \bar{S}_n от a приближённо подчинены нормальному распределению со средним 0 и дисперсией σ^2/n . Т. о., для вычисления (в первом приближении) вероятностей тех или иных отклонений \bar{S}_n от a при больших n нет надобности знать во всех деталях распределение величин X_n ; достаточно знать лишь их дисперсию. Для оценки точности этого приближения необходимо привлекать моменты порядка, большего 2. Использование таких моментов позволяет также строить более точные приближения.

Эти утверждения могут быть с надлежащими изменениями распространены на различно распределённые слагаемые (см. [Ляпунова теорема](#)) и на случайные векторы (из конечномерных и некоторых бесконечномерных пространств). Условия независимости могут быть заменены условиями слабой (в том или ином смысле) зависимости случайных величин X_1, X_2, \dots .

В 1920-х гг. было обнаружено, что даже в схеме последовательности одинаково распределённых и независимых случайных величин появляются предельные распределения, отличные от нормального.

Случайные процессы

Механизм возникновения большинства предельных закономерностей может быть понят лишь в связи с теорией [случайных процессов](#). В ряде физич. и химич. исследований в сер. 20 в. возникла потребность наряду с одномерными и многомерными случайными величинами рассматривать случайные процессы. В В. т. случайный процесс рассматривают как параметрич. семейство случайных величин X_t . Примером случайного процесса может служить процесс X_t , где X_t – координата в момент t частицы, совершающей броуновское движение. Обычно в приложениях параметр t является временем, но этим параметром может быть, напр., произвольная независимая переменная, и тогда говорят о случайной функции (если t – точка пространства, то говорят о случайном поле). В том случае, когда параметр t пробегает целочисленные значения, случайная функция называется случайной последовательностью (или временным рядом). Подобно тому как случайная величина характеризуется законом распределения, случайный процесс может быть characterized т. н. конечномерными распределениями – совокупностью совместных законов распределения, где t_1, \dots, t_n – всевозможные моменты времени, $n=1, 2, \dots$. В теории случайных процессов наиболее изучены [марковские процессы](#), [стационарные случайные процессы](#), [ветвящиеся процессы](#), а также [мартингалы](#). Интенсивно развивается теория случайных процессов, происходящих в случайной среде.

Исторически первыми изучались марковские процессы. Случайный процесс X_t называется марковским, если для любых моментов времени t_0 и t_1 , $t_0 < t_1$ условное распределение вероятностей при условии, что заданы все значения X_t при $t \leq t_0$, зависит только от (в силу этого марковские случайные процессы иногда называются процессами без последствия). Марковские процессы являются естественным обобщением детерминированных процессов, рассматриваемых в классич. физике. В детерминированных процессах состояние системы в момент t_0 однозначно определяет ход процесса в будущем; в марковских процессах состояние системы в момент времени t_0 однозначно определяет распределение вероятностей процесса при $t > t_0$, причём никакие сведения о поведении процесса до момента времени t_0 не изменяют это распределение. Подобно тому как изучение непрерывных детерминированных процессов сводится к дифференциальным

уравнениям относительно функций, описывающих состояние системы, изучение непрерывных марковских процессов сводится к дифференциальным или интегродифференциальным уравнениям относительно распределений вероятностей процесса.

Другим крупным разделом в теории случайных процессов является теория стационарных случайных процессов. Стационарность процесса, т. е. неизменность во времени его вероятностных характеристик, налагает сильное ограничение на процесс и позволяет из одного этого допущения извлечь ряд важных следствий. Для б. ч. теории достаточно предположения о стационарности в широком смысле, т. е. требования независимости от t математич. ожиданий $\mathsf{E}X_t$ и $\mathsf{E}X_{t+\tau}$ для всех t .

Теория случайных процессов тесно связана с классич. проблематикой предельных теорем для сумм случайных величин. Те законы распределения, которые выступают при изучении сумм случайных величин как предельные, в теории случайных процессов являются точными законами распределения соответствующих характеристик. Этот факт позволяет доказывать многие предельные теоремы с помощью соответствующих случайных процессов.

Исторический очерк

Первые работы по В. т., принадлежащие Б. [Паскалю](#), П. [Ферма](#) и Х. [Гюйгенсу](#), появились в сер. 17 в. и были связаны с подсчётом разл. вероятностей в азартных играх. Первый строго доказанный результат В. т. принадлежит Я. [Бернулли](#), установившему закон больших чисел для схемы независимых испытаний с двумя исходами (опубл. в 1713).

Второй период истории В. т. (18 – 1-я пол. 19 вв.) связан с именами А. де [Муавра](#), П. [Лапласа](#), К. [Гаусса](#) и С. [Пуассона](#). В этот период В. т. находит ряд актуальных применений в естествознании и технике, гл. обр. в теории ошибок, развившейся в связи с потребностями геодезии и астрономии, и в теории стрельбы. К этому периоду относятся доказательство первого варианта центральной предельной теоремы (А. де Муавр, 1733, П. Лаплас, 1812) и [Пуассона теоремы](#). А. [Лежандром](#) (1806) и К. Гауссом (1808) был разработан метод наименьших квадратов. В 18 в. ряд трудов по В. т. был написан работавшими в России Л. [Эйлером](#), Н. [Бернулли](#) и Д. [Бернулли](#); появились работы М. В. [Остроградского](#) по вопросам В. т., связанным с математич. статистикой, и В. Я. [Буняковского](#) по применениям В. т. к страховому делу, статистике и демографии.

Третий период истории В. т. (2-я пол. 19 в.) связан в осн. с именами П. Л. [Чебышева](#) и его учеников А. М. [Ляпунова](#) и А. А. [Маркова](#). Они поставили ряд общих задач, решение которых привело к обобщению теорем Бернулли и Муавра – Лапласа. Чебышев доказал (1867) закон больших чисел при весьма общих предположениях. Он же впервые сформулировал центральную предельную теорему для сумм независимых случайных величин и указал один из методов её доказательства (1887). Другим методом доказательство этой теоремы в условиях, близких к окончательным, получил А. М. Ляпунов (1901). А. А. Марков впервые рассмотрел (1907) один случай зависимых испытаний, который впоследствии получил название цепей Маркова. Со 2-й пол. 19 в. исследования по В. т. в России занимают ведущее место в мире. В Зап. Европе во 2-й пол. 19 в. получили большое развитие работы по математич. статистике (А. [Кетле](#), Ф. Гальтон) и статистич. физике (Л. [Больцман](#)), которые наряду с основными теоретич. работами П. Л. Чебышева, А. М. Ляпунова и А. А. Маркова создали основу для существенного расширения проблематики В. т. в совр. период её развития.

Четвёртый (современный) период истории В. т., начавшийся в 20 в., характеризуется существенным расширением круга её применений, созданием нескольких систем строгого математич. обоснования В. т., появлением новых мощных методов, требующих применения, помимо классич. анализа, средств теории множеств, теории функций действительного переменного и функционального анализа; в области В. т. плодотворно работали во Франции – Э. [Борель](#), П. [Леви](#), М. [Фреше](#), в Германии – Р. [Мизес](#), в США – Н. [Винер](#), У. [Феллер](#), Дж. Дуб, в Швеции – Г. [Крамер](#); отеч. наука продолжала занимать значительное, а в ряде направлений и ведущее положение. В нашей стране новый период развития В. т. открывается деятельностью С. Н. [Бернштейна](#), обобщившего классич. предельные теоремы П. Л. Чебышева, А. М. Ляпунова и А. А. Маркова и указавшего на ряд применений В. т. в естествознании. А. Я. Хинчин и А. Н. Колмогоров успешно применяли методы теории функций действительного переменного к В. т. В 1930-х гг. ими и Е. Е. [Слуцким](#) были заложены основы теории случайных процессов. В. И. [Романовский](#), Н. В. [Смирнов](#), Ю. В. [Линник](#) и Л. Н. Большев внесли большой вклад в развитие математич. статистики, применяя методы В. т. к статистич. задачам.

Литература

Лит.: История математики с древнейших времен до начала XIX столетия. М., 1970–1972. Т. 2–3; Колмогоров А. Н. Основные понятия теории вероятностей. 2-е изд. М., 1974; Феллер В. Введение в теорию вероятностей и ее приложения: В 2 т. М., 1984; Прохоров Ю. В., Розанов Ю. А. Теория вероятностей. 3-е изд. М., 1987; Боровков А. А. Теория вероятностей. 4-е изд. М., 2003; Севастьянов Б. А. Курс теории вероятностей и математической статистики. 3-е изд. М., 2004; Ширяев А. Н. Вероятность: В 2 т. 3-е изд. М., 2004; Гнеденко Б. В. Курс теории вероятностей. 8-е изд. М., 2005.