ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального образования
«ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Ю.И. Галанов

КОНСПЕКТ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета

Издательство ТПУ

Томск 2016

УДК 519.21(0.75,8)
ББК 22 171 Я73
Б24

Галанов Ю.И. Конспект по теории вероятностей: Учебное пособие. — Томск: Изд–во ТПУ, 2019. — 87 с.

Пособие содержит конспект лекций по теории вероятностей, которые автор читал для студентов ГРФ (ИПР) с 2000 г. по 2018 г.

УДК 519.21(0.75,8)
ББК 22 171 Я73

Рекомендовано к печати Редакционно-издательским советом
Томского политехнического университета

Рецензенты

Кандидат физико-математических наук, доцент Кандидат технических наук,

Введение

Теория вероятностей — математическая дисциплина, занимающейся разработкой и анализом математических моделей случайного эксперимента.

Случайный эксперимент (опыт. испытание. наблюдение и т.п.) характеризуется тем, что его исход не вполне определяется заданием начальных условий. Не всякий случайный эксперимент изучается теорией вероятностей, а только тот, который может быть (хотя бы мысленно) многократно воспроизведен и, при этом, обладает свойством «статистической устойчивости», заключающейся в том, что относительные частоты появления любого события A, связанного с изучаемым явлением, в сериях из n испытаний n_A/n, приближается к некоторому числу $P(A)$, называемому вероятностью события A.

Событие и вероятность являются первичными понятиями теории вероятностей. События, как математические объекты, можно комбинировать с помощью логических операций «нет», «и», «или», составляя из простых более сложные события, а вероятность есть числовая характеристика класса событий, свойства которой по определению аналогичны относительной частоте событий. Основная задача теории вероятностей заключается в том. чтобы найти вероятность некоторого, интересующего нас события, по известным вероятностям заданных событий, через которые оно выражается.

Исторически первым полем приложения статистических методов анализа явились азартные игры. Схемы азартных игр явились первыми четкими моделями случайных явлений (Кардано, Галилей, Паскаль, Ферма, Гюйгенс, Байес, Я. Бернулли).

В связи с серьезными потребностями естествознания (теория ошибок, теория стрельбы, статистика народонаселения) в XVIII веке создавался математический аппарат, развитие которого связано с именами Муавра, Лапласа, Гаусса, Пуассона.

Окончательно теория вероятностей как математическая наука оформилась в 30-х годах XX века, когда А. Н. Колмогоровым было предложено *аксиоматическое определение* вероятности.

Во второй половине XX века характерно проникновение статистических методов во все отрасли человеческих знаний. Это теоретическая физика, кибернетика, теория информации, теория массового обслуживания, теория надежности, математическая теория игр, теория операций и др.

Теория вероятностей, как прикладная наука, стала одним из надежных, точных и эффективных способов познания реальной действительности.

Часть I

Случайные события и их вероятности

Лекция 1.

Математическая модель случайного эксперимента

> Виды событий, операции над событиями, поле событий. Подходы к определению вероятности. Относительная частота события и её свойства. Аксиомы вероятности.
> Элементы комбинаторики. «Урновые схемы». Алгоритм решения задач на классическую вероятность

1.1. Случайный эксперимент и случайное событие

Существенными составляющими случайного явления, подлежащего изучению, выступают понятия *опыта* и *события*.

Под опытом (*экспериментом*, *испытанием*) будем понимать воспроизведение некоторого конкретного комплекса условий. Всякое испытание заканчивается одним и только одним из множества возможных исходов. Всякий исход опыта будем называть *событием*. Каждому исходу ставится в соответствие одно или несколько событий.

Опыт называется *случайным*, если его результат не вполне определяется заданием начальных условий опыта. Исходы случайного эксперимента — *случайные события*.

События будем обозначать большими буквами латинского алфавита $A, B, C, D, ...$. Некоторые специальные события обозначают греческими буквами.

Различают *элементарные* и *составные* события. *Элементарное* в условиях данного опыта событие является результатом одного и только одного исхода. Все остальные события называются *составными*, т.е. рассматриваются как логические объединения нескольких элементарных событий.

Примеры событий:

- Появление герба при бросании монеты.
- Попадание в цель при выстреле.
- Обнаружение объекта при одном цикле обзора радиолокационной станции.

- Обрыв нити в течение часа работы ткацкого станка.

Рассмотрим сначала два специальных события: *достоверное* и *невозможное*.

Событие называется *достоверным*, если оно происходит при любом исходе данного опыта, и обозначается Ω.

Событие называется *невозможным*, если оно не происходит ни при каком исходе данного опыта, и обозначается $\varnothing$.

Каждому событию A поставим в соответствие *противоположное* (дополнительное) событие, обозначается $\overline{A}$. По определению, событие $\overline{A}$ реализуется тогда и только тогда, когда событие A не происходит.

Возьмём, например, опыт: подбрасывание двух монет. Рассмотрим события: $A = \{RR\}$ — выпадение двух решек (R), $B = \{OO, OR, RO, RR\}$ — выпадение двух или менее решек, $C = \{\}$ — выпадение трёх решек, $D = \{OO, OR, RO\}$ — выпадение менее двух решек. Здесь A — случайное событие, B — достоверное событие, C — невозможное событие, A и D — противоположные события.

Два события называются *несовместными*, если их совместное осуществление — событие невозможное ни при каком исходе в данном опыте. Если события не являются несовместными, то они *совместны*.

Говорят, что несколько событий в условиях данного опыта образуют *полную группу событий*, если в результате опыта обязательно произойдет одно и только одно из них.

Основой для построения математической модели случайного эксперимента служит множество всех возможных взаимоисключающих (*элементарных*) в условиях данного опыта событий $(\omega_1,\ \omega_2, \ldots, \omega_n)$, составляющих *пространство элементарных событий*, которое обозначается символом Ω, а сами эти события называются точками этого пространства. Это множество составляет, в то же время, достоверное событие.

Случайное событие A в данной модели — это подмножество множества Ω: $A \subset \Omega$.

С помощью логических операций «нет», «и», «или» над подмножествами можно получить новые подмножества, которые также являются событиями. Исторически сложилось так, то указанные операции над событиями и их результаты в теории вероятностей имеют свои специальные названия.

Операция теории множеств	Операция в теории вероятностей
«нет» – дополнение	противоположное событие
«и» – пересечение	произведение событий
«или» – объединение	сумма событий

Множество элементов с операциями, установленными на данном множестве называется алгеброй. Далее мы рассмотрим алгебру событий.

1.2. Операции над событиями

Будем изображать пространство элементарных событий Ω в виде прямоугольника, а любое событие — в виде произвольной геометрической фигуры (схема Эйлера – Венна, см. рис. 1.1).

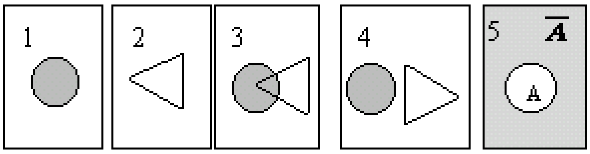


Рис. 1.1. Отношения между событиями A — 1 и B — 2: A и B совместны — 3; A и B несовместны — 4; событие A и ему противоположное — 5.

Для того чтобы математически выражать отношения между элементарными и составными (простыми или сложными) событиями достаточно ввести две основные операции: взятие противоположного события и одну из двух операций — сложение или умножение событий. Все другие операции выражаются через основные.

Событие C называется *суммой* событий A и B, обозначается $C = A + B$, если один и тот же исход испытания приводит к наступлению хотя бы одного из этих событий, то есть, событие C наступает, если в результате испытания наступает **ИЛИ** одно событие A, **ИЛИ** одно событие B, **ИЛИ** оба события вместе.

Произведением событий A и B называется событие $C = AB$, $(C = A \cap B)$, состоящее из всех элементарных событий, принадлежащих одновременно и событию A и событию B, т.е. когда в результате отдельного испытания происходят вместе **И** событие A, **И** событие B.

Введённые операции удобно изобразить на схеме Эйлера — Венна (1.2, 1.3), где результаты операций изображены в виде заштрихованных областей.

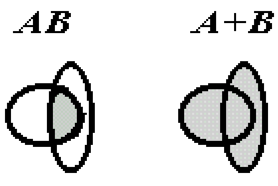


Рис. 1.2. Произведение и сумма совместных событий

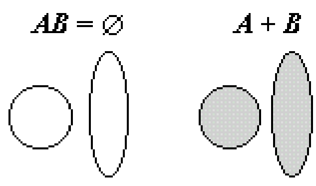


Рис. 1.3. Произведение и сумма несовместных событий

События называются *равными*, если они состоят из одних и тех же элементарных событий (эквивалентные, или равносильные события). Введённые операции над событиями подчиняются аксиомам:

$$
\begin{array}{ll}
A+A=A & A\cdot A=A\\
A+B=B+A & (A+B)+C=A+(B+C)\\
A+\Omega=\Omega & A\cdot\Omega=A\\
AB=BA & A\cdot\varnothing=\varnothing\\
A(B+C)=AB+AC & (A+B)(A+C)=A+(BC)\\
A+\overline{A}=\Omega & A+\varnothing=A\\
\overline{AB}=\overline{A}+\overline{B} & \overline{A+B}=\overline{A}\cdot\overline{B}
\end{array}
$$

Разностью событий A и B называется событие $C=A-B$ или ($C=A\backslash B=A\cdot\overline{B}$), состоящее из элементарных событий, которые принадлежат A, но не принадлежат B, т.е. событие A происходит без события B. *Симметрической разностью* двух событий A,B называют событие $A\vartriangle B=(A-B)+(B-A)$, которое происходит тогда и только тогда, когда когда происходит лишь одно из событий A,B.

Говорят, что событие A *влечёт* событие B ($A\subset B$), если при совершении события A событие B обязательно произойдет (событие A содержится в B). (Рис. 1.4)

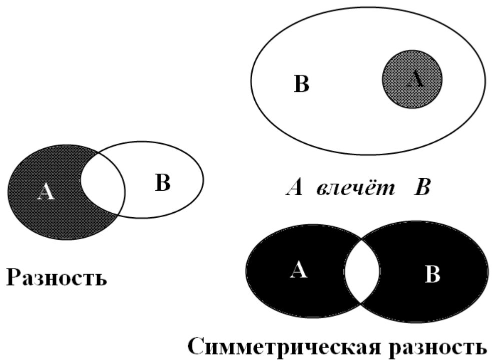


Рис. 1.4. Схема Эйлера–Венна для разности, симметрической разности *и* включения

Из изложенного выше мы выделяем две составляющие математической модели случайного эксперимента:

а) множество Ω, (пространство элементарных событий);

б) *булеву алгебру* подмножеств Ω, (события). *По определению булева алгебра подмножеств Ω есть класс подмножеств Ω, содержащий $\varnothing,\Omega$*

и замкнутый относительно операций образования дополнений, конечных объединений и конечных пересечений.

Третьей составляющей математической модели случайного эксперимента является *вероятность событий*. Она вводится как числовая функция на множестве элементарных событий.

Вероятность наделяет события свойством, подобным тому, что присуще материальным объектам: массой, длиной, площадью, объёмом — т.е. свойством *меры*.

1.3. Вероятность события

Вероятность события позволяет сравнивать события по степени возможности их наступления в ходе испытания.

Понятие вероятности события является первичным, не сводимым к другим понятиям. Имеется несколько подходов, поясняющих понятие вероятности.

Статистический подход к понятию вероятности. Пусть при проведении серии из n_1 испытаний событие A наступило m_1 раз $(m_1 \leqslant n_1)$. Число $\dfrac{m_1}{n_1} = P^*_{n_1}(A)$ называется относительной частотой появления события A в данной серии испытаний. Если провести другую серию из n_2 опытов, то получим другое число $P^*_{n_2}(A) = \dfrac{m_2}{n_2}$ и т.д..

Если в различных сериях испытаний относительные частоты наступления события A незначительно отличаются друг от друга, то говорят, что частота обладает свойством устойчивости. См. рис.(1.5) .

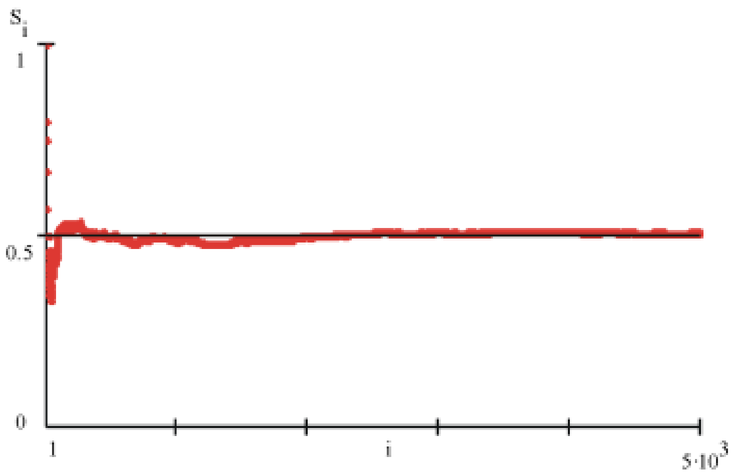


Рис. 1.5. Пример статистической устойчивости относительной частоты появления герба при подбрасывании монеты. Компьютерное моделирование.

В качестве вероятности события A принимают число $P_A = P(A)$, к которому стремится определённым образом частота события при неограниченном увеличении числа испытания в серии.

Замечание. Практическая польза от знания вероятности появления события состоит в следующем: зная вероятность p некоторого события, мы можем предсказать, что после проведении N испытаний в неизменных условиях *ожидаемое* число появлений регистрируемого события будет приблизительно равно $N \cdot p$.

Чтобы повысить шансы выигрыша в лотерею, надо купить больше билетов.

Более строго это свойство описывается следующим образом:

$$\lim_{n_k \to \infty} P\left\{ \mid P^*_{n_k} - P(A) \mid < \varepsilon \right\} = 1, \ \forall \varepsilon > 0. \qquad (1.1.)$$

Говорят, что частота события **сходится по вероятности к вероятности события** (Теорема Бернулли). Это свойство частоты мы докажем в разделе «Предельные теоремы теории вероятностей».

Относительная частота события обладает очевидными свойствами, не требующими доказательств:

а) неотрицательность;

б) принимает значения между нулём и единицей;

в) для несовместных событий относительная частота появления хотя бы одного из них равна сумме их относительных частот.

Аксиомы вероятности. В математической модели случайного эксперимента вероятность вводится как числовая функция, заданная на множестве событий, связанных с данным экспериментом. При этом вероятность события обладает всеми свойствами относительной частоты появления события. Эти свойства считаются аксиомами.

Аксиома 1.1. (Аксиома неотрицательности)

$$0 \leqslant P(A) \qquad \textit{для всех событий, определённых на} \qquad \Omega.$$

Аксиома 1.2. (Аксиома нормировки)

$$P(\Omega) = 1$$

Аксиома 1.3. (Аксиома сложения.) *Если $AB = \varnothing$, то*

$$P(A + B) = P(A) + P(B).$$

Аксиома сложения распространяется на любое конечное семейство попарно непересекающихся событий:

Аксиома 1.4. (Расширенная аксиома сложения) .

Если события $A_1, \ldots, A_n$ попарно несовместны, то

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

Следующие свойства вероятности выводятся как следствия из данных аксиом:

1. $P(\varnothing) = 0$;
2. $P(A) \leqslant P(B)$ если $A \subseteq B$;
3. Если $\overline{A}$ — событие, противоположное событию A, то

$$P\left(\bar{A}\right) = 1 - P\left(A\right).$$

Общее определение вероятности события. Аксиомы вероятности позволяют сформулировать *общее определение* вероятности события: вероятность события A равна сумме вероятностей элементарных событий ω_i, его составляющих.

$$P(A) = \sum_{\omega_i \in A} P(\omega_i) \tag{1.2.}$$

Классическое определение вероятности. Классическими называют опыты (испытания), обладающие двумя существенными признаками:

1. число исходов данного опыта конечно;
2. все исходы равновозможны, т.е. $P(\omega_i) = 1/n$.

Об опыте, исходы которого образуют полную группу попарно несовместных, равновозможных событий, говорят, что он укладывается в классическую схему.

Равновозможные события, составляющие полную группу, называют случаями.

По отношению к каждому событию случаи делятся на благоприятные, при которых происходит событие, и неблагоприятные, при которых событие не происходит.

Вероятностью появления некоторого события A в *классическом* случае называется отношение числа случаев, благоприятствующих появлению этого события, к общему числу равновозможных в данном опыте случаев:

$$P(A) = \sum_{\omega_i \in A} P(\omega_i) = \sum_{\omega_i \in A} 1/n = \frac{m}{n}, \tag{1.3.}$$

где

m — число исходов, благоприятствующих событию A,

n — общее число исходов опыта.

Достоинство определения – вероятность события можно определить до опыта. Недостаток – вероятность можно определить только для равновозможных исходов опыта.

Геометрическая вероятность. В задачах, сводящихся к случайному бросанию точки на конечный участок прямой, плоскости, трёх- или многомерного пространства, когда возможность появления точки внутри некоторой области в пространстве определяется не положением этой области и её границами, а только её мерой (мера обозначается как mes), т.е. длиной, площадью, объёмом, то вероятность появления случайной точки внутри некоторой области находится как отношение меры этой области к мере всей области, в которой может появиться данная точка:

$$P\left(A\right) = \frac{mes\left(A\right)}{mes\left(\Omega\right)} \tag{1.4.}$$

или

$$P(A) = \frac{l_A}{l}; \quad P(A) = \frac{S_A}{S}; \quad P(A) = \frac{V_A}{V}, \tag{1.5.}$$

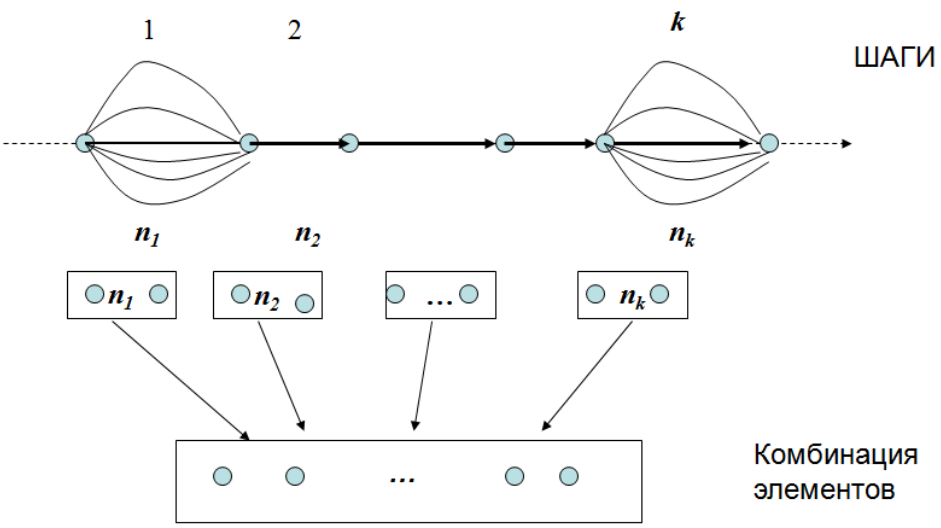
соответственно, для отрезка l прямой, области S на плоскости, области V трёхмерного пространства. Это определение вероятности называется *геометрическим*.

1.4. Решение задач на классическую вероятность

При решении задач на классическую вероятность приходится подсчитывать число способов (*комбинаций*), с помощью которых может осуществиться некоторое событие (действие). Задачи такого рода называют *комбинаторными*. При подсчёте числа комбинаций руководствуются принципами сложения и произведения комбинаций.

Принцип сложения комбинаций состоит в том, что если некоторое действие может осуществиться параллельно несколькими независимыми способами, то общее число способов осуществления этого действия равно числу таких способов.

Принцип произведения комбинаций заключается в следующем. Если какое-либо действие осуществляется за k последовательных шагов, при этом первый шаг может быть реализован n_1 числом способов, второй шаг n_2 числом способов, k-й шаг — n_k способами, то общее число способов реализации действия равно произведению $n_1 \cdot n_2 \cdot \ldots \cdot n_k$. Данный принцип можно пояснить также подсчётом числа способов выбора по одному элементу из каждого из k ящиков. содержащих n_k элементов (рис. 1.6). Пусть мы имеем конечное



$$N = n_1 \cdot n_2 \cdot \ldots \cdot n_k$$

Рис. 1.6. Принцип произведения комбинаций

множество элементов. Тогда из элементов данного множества можно составить различные соединения (комбинации, подмножества), отличающиеся либо своим составом, либо порядком взаимного расположения.

Перестановками из n элементов называют всевозможные упорядоченные соединения из данных элементов. Число таких перестановок P_n подсчитывается с помощью принципа произведения комбинаций: первый элемент в перестановке можно выбрать n числом способов, второй — $(n-1)$ числом, третий — $(n-2)$ числом и т.д.

Общее число комбинаций равно:

$$P_n = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1 = n! \tag{1.6.}$$

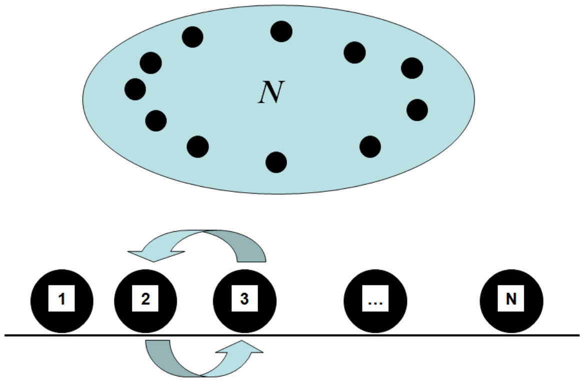


Рис. 1.7. Перестановки

Размещениями из n элементов по m называются всевозможные упорядоченные соединения (комбинации, подмножества) m элементов из n данных элементов. Число размещений A_n^m (от французского *arrangement* — размещение) подсчитывается по тому же принципу, что и число перестановок:

$$A_n^m = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-m+1) = \frac{n!}{(n-m)!}. \tag{1.7.}$$

Сочетанием из n элементов по m называется любое неупорядоченное подмножество m элементов из n. Число сочетаний C_n^m (от латинского *combinare* — соединить) подсчитывается следующим образом. Если во всех сочетаниях произвести всевозможные перестановки, то мы получим всевозможные размещения. Следовательно, имеет место формула

$$C_n^m = \frac{A_n^m}{m!} = \frac{n!}{m! \cdot (n-m)!} = \frac{n(n-1)(n-2)\ldots(n-m+1)}{m!}. \tag{1.8.}$$

Основное свойство сочетаний:

$$C_n^m = C_n^{n-m}. \tag{1.9.}$$

Образование сочетаний связано с задачей разбиения множества N элементов на два подмножества так, что одно из них содержит m элементов, а другое – оставшиеся $(N-m)$ элементов и является простейшим случаем более общей задачи о разбиении множества на k неупорядоченных подмножеств, содержащих $n_1, n_2, \ldots, n_k$ элементов, причём $n_1 + n_2 + \ldots + n_k = N$. Число таких комбинаций равно

$$\frac{N!}{n_1! \cdot n_2! \cdots n_k!}$$

1.5. «Урновые» схемы случайных экспериментов

Большое число опытов с дискретным пространством элементарных событий сводится к одной из четырёх схем выбора, т.н. «урновым» схемам выбора (см. рис. 1.8). Подсчитаем число исходов опыта для каждой схемы. Пусть урна содержит N *пронумерованных* элементов.

Выбор без возвращения с учётом порядка. Число исходов данного опыта равно числу упорядоченных подмножеств составленных из N элементов по m, т.е. числу размещений A_n^m.

Выбор без возвращения без учёта порядка. Число исходов данного опыта равно числу неупорядоченных подмножеств составленных из N элементов по m, т.е. числу сочетаний C_n^m.

Выбор с возвращением с учётом порядка. Поскольку каждый выбранный элемент возвращается в урну, то число вариантов выбора на каждом шаге остаётся неизменным. Тогда число исходов данного опыта равно N^m.

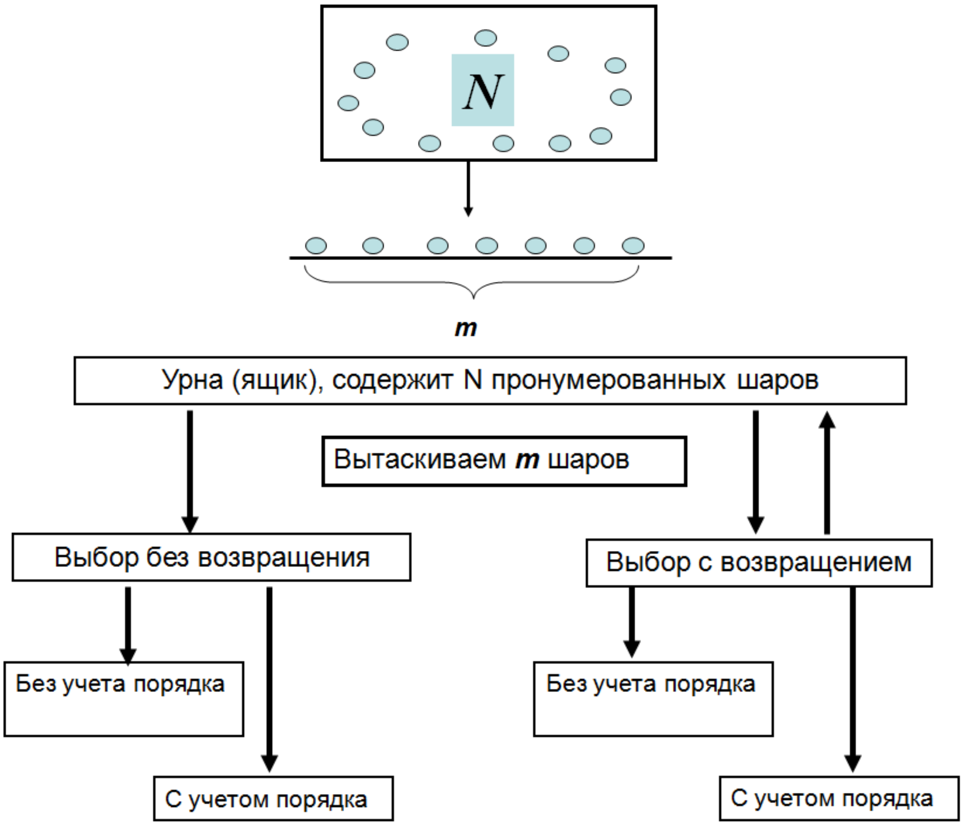


Рис. 1.8. Урновые схемы выбора

Выбор с возвращением без учёта порядка. Исход данного опыта – неупорядоченное подмножество содержащее m элементов, причём каждый элемент может неоднократно участвовать в выборе. Подсчитаем, сколько раз каждый элемент появится в выборке. Для этого отведём для каждого элементы персональную пронумерованную ячейку. При выборе некоторого элемента будем в его ячейку добавлять камешек. Тогда число всевозможных исходов данного опыта будет равно числу способов размещения m камешков по N ячейкам.

На рис. 1.9 показано несколько возможных исходов данного опыта.

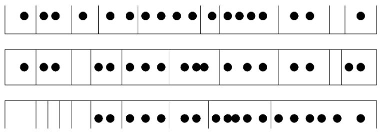


Рис. 1.9. Выбор с возвращением без учёта порядка или размещение m элементов по N ячейкам

Видно, что всевозможные комбинации можно получить меняя местами перегородки и камешки. Так как две внешние перегородки остаются на месте, то общее число подвижных элементов равно числу подвижных перегородок $(N + 1 - 2 = N - 1)$ плюс число камешков m. Число перестановок подвижных элементов равно $(N - 1 + m)!$. Надо ещё учесть, что $m!$ перестановок среди одинаковых камешков и $(N - 1)!$ перестановок среди одинаковых перегородок не меняют распределения камешков по ячейкам. Окончательно получаем, что число исходов данного опыта равно

$$\frac{(N-1+m)!}{(N-1)!\cdot(m!)} = C^{m}_{(N-1+m)} = C^{(N-1)}_{(N-1+m)} \qquad (1.10.)$$

Замечание. Обратите внимание на то, что опыт по выбору с возвращением без учёта порядка m элементов из N элементов эквивалентен размещению m элементов по N ячейкам.

Алгоритм решения задач на классическую вероятность заключает в последовательном выполнении следующих действий:

1. Уясняем, что является исходом данного опыта.

2. Убеждаемся, что число исходов конечно и они равновозможны.

3. Подсчитываем число исходов n.

4. Подсчитываем число исходов, благоприятствующих нашему событию m.

5. Рассчитываем вероятность события как отношение m/n.

Вопросы для самопроверки.

1. Какие события называются случайными? Приведите примеры случайных событий.
2. Какие события образуют полную группу несовместных событий?
3. Приведите примеры полных групп событий.
4. Какое событие называется суммой (объединением) нескольких событий?
5. Какое событие называется произведением (пересечением, совмещением) событий?
6. Что называется частотой события и каковы её свойства?
7. Сформулируйте классическое определение вероятности события. В каких пределах изменяется вероятность события?

Лекция 2.

Основные теоремы и формулы теории вероятностей

Как следствие из аксиом, выводится формула сложения вероятностей. Вводится понятие зависимых событий и условной вероятности. Выводится формула произведения событий. Приводится достаточное условие независимости событий. Используя понятие полной группы событий, выводятся формулы полной вероятности и формула Байеса.Рассмотрена специальная схема организации опытов — схема Бернулли, присущие ей типовые задачи и предельные случаи.

2.1. Сложение и умножение вероятностей

Любые два события A и B, возникающие в ходе случайного эксперимента, могут находиться между собой в некоторых отношениях. Например, можно говорить об их эквивалентности или включении одного в другое, о совместности или несовместности.

Особый интерес представляет отношение зависимости и независимости. Если, например, событие B уже наступило, то в некоторых случаях это может дать дополнительную информацию о событии A, а поэтому изменить его вероятность, а в других случаях эта информация не оказывает никакого влияния на вероятность события A.

Определение 2.1. *События A и B называются зависимыми, если вероятность появления одного из них зависит от того, произошло другое или нет. В противном случае события называются независимыми.*

Определение 2.2. *Вероятность события A, вычисленная при условии, что событие B произошло, называется условной вероятностью события A и обозначается $P(A \mid B)$.*

Условная вероятность. Если в ходе испытания становится известно, что произошло событие B, то эта дополнительная информация приводит к изменению пространства элементарных событий так, что именно событие B теперь играет роль достоверного события. При этом условии вероятность события A будет определяться вероятностной мерой пересечения A и B.

Действительно, если события несовместны, т.е. $P(AB) = 0$, то и $P(A \mid B) = 0$, а, если $A = B$, то $P(A \mid B) = 1$.

$$P(A|B) = \frac{mes(AB)}{mes(B)} = \frac{mes(AB)/mes\,(\Omega)}{mes\,(B)\,/mes\,(\Omega)} = \frac{P\,(AB)}{P\,(B)}. \qquad (2.1.)$$

Уравнение (2.1.) удобно переписать в такой форме, чтобы выразить вероятность произведения двух зависимых событий:

$$P\,(AB) = P\,(A)\,P\,(B|A) = P\,(B)\,P\,(A|B) = P\,(A)\,P\,(B|A)\,. \qquad (2.2.)$$

Теорема 2.1. *Вероятность произведения двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое произошло.*

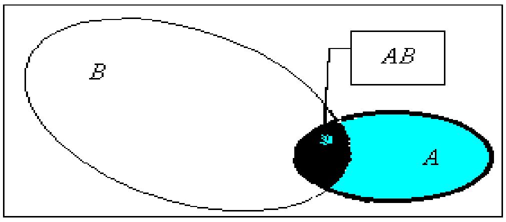


Рис. 2.1. К понятию условной вероятности

Следствие: все события на пространстве элементарных событий Ω для единичного эксперимента являются зависимыми.

Произведение большого числа событий удобно представить в виде упорядоченной последовательности событий $A_1,\ A_2,\ A_3, \ldots,\ A_n$. Тогда можно рассматривать условные вероятности типа

$$P(A_2 \mid A_1),\ P(A_3 \mid A_1A_2), P(A_4 \mid A_1A_2A_3), \ldots,$$
$$\ldots, P(A_n \mid A_1A_2 \cdots A_{n-1}),$$

и т. д., смысл которых ясен из определения условной вероятности. Например, для трёх событий получим:

$$P(A_1A_2A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)$$

$$P(A_1A_2A_3A_4) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)P(A_4 \mid A_1A_2A_3).$$

Замечание 1.1. Эта формула справедлива, если указанная последовательность событий по условиям задачи формируется единственным образом. В противном случае надо учесть число эквивалентных последовательностей (умножить на число перестановок сомножителей).

Теорема 2.2. *(Произведения вероятностей). Если события A и B независимы, то*

$$P(A \mid B) = P(A);\ \ P(B \mid A) = P(B) \text{ и } P(AB) = P(A)P(B).$$

Обобщение на случай n независимых событий $A_1,\ A_2,\ A_3,\ ...,\ A_n$:

$$P(A_1 \cdot A_2 \cdot A_3 \cdot \ldots \cdot A_n) = P(A_1) \cdot P(A_2) \cdot \ldots \cdot P(A_n).$$

Теорема сложения вероятностей совместных событий

Теорема 2.3. *(Сложения вероятностей). Пусть A и B — совместные события. Тогда вероятность появления* ***хотя бы одного из этих событий*** *(т.е. вероятность суммы событий A и B) равна сумме вероятностей этих событий без вероятности произведения этих событий*

$$P(A + B) = P(A) + P(B) - P(AB). \tag{2.3.}$$

Доказательство. Любое событие можно представить в виде суммы двух несовместных частей частей: $A = (A-B)+A\cdot B$. Тогда из аксиомы сложения следует, что $P(A-B) = P(A) - P(A\cdot B)$. Сумму двух событий представим в виде суммы трёх несовместных событий: $A + B = (A - B) + A \cdot B + (B - A)$. Тогда, согласно аксиоме сложения

$$P(A+B) = P(A) - P(A\cdot B) + P(A\cdot B) + P(B) - P(A\cdot B) = P(A) + P(B) - P(AB),$$

что и требовалось доказать.

Теорема может быть обобщена на любое конечное число совместных событий

$$\begin{aligned} &P(A_1 + A_2 + \cdots + A_n) = \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) - P(A_1A_2) - \\ &\cdots - P(A_1A_3) - \cdots - P(A_{n-1}A_n) - P(A_1A_2A_3) - \\ &\qquad\qquad - P(A_{n-2}A_{n-1}A_n) - \cdots - P(A_1A_2\cdots A_n). \end{aligned} \tag{2.4.}$$

При вычислении вероятности суммы большого числа событий $A = A_1 + A_2 + A_3 + \cdots + A_n$ часто бывает проще перейти к вычислению вероятности противоположного события. Для *независимых* событий получим формулу:[1]

[1]Предполагается справедливым утверждение: если два события независимы, то независимы и их противоположные события.

$$\begin{aligned} P(A) &= 1 - P(\overline{A_1 + A_2 + A_3 + \dots + A_n}) = \\ &= 1 - P(\overline{A}_1 \cdot \overline{A}_2 \cdot \overline{A}_3 \cdot \ldots \cdot \overline{A}_n) = \\ &= 1 - P\left(\overline{A}_1\right) \cdot P\left(\overline{A}_2\right) \cdot P\left(\overline{A}_3\right) \cdot \ldots \cdot P\left(\overline{A}_n\right). \end{aligned} \qquad (2.5.)$$

Или

$$P(A) = 1 - q_1 q_2 \cdots q_n, \qquad (2.6.)$$

где $q_i = 1 - P_i$.

Частный случай: если $P_1 = P_2 = \ldots = P_n = P,\ 1 - P = q$, то

$$P(A) = 1 - q^n.$$

Формула полной вероятности. Пусть событие A может произойти совместно с одним из событий H_1, H_2, $\ldots$, H_n, которые попарно несовместны и образуют полную группу событий, т.е. $\sum_{i=1}^{n} H_i = \Omega$, а, $H_i \cap H_i = \varnothing,\ \ i \neq j$ и $P(H_1) + P(H_2) + \ldots + P(H_n) = 1$.

События H_i называются гипотезами. Тогда любое событие можно представить в виде суммы непересекающихся составляющих: $A = \sum\limits_{i} A \cdot H_i$. (См. рис. 2.2)

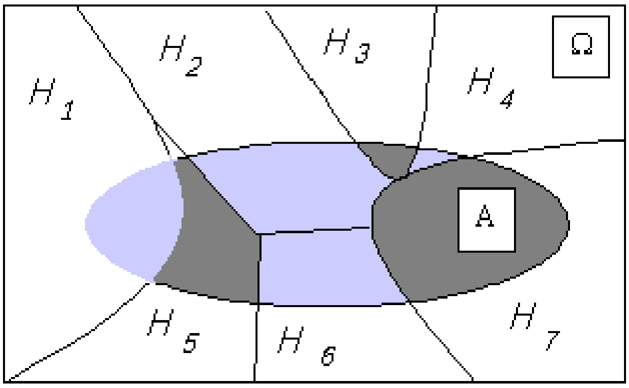


Рис. 2.2. Разложение события на составляющие

Пусть известны вероятности гипотез $P(H_1), P(H_2), \ldots, P(H_n)$ и условные вероятности $P(A \mid H_1),\ P(A \mid H_2),\ \ldots,\ P(A \mid H_n)$. Тогда вероятность события A равна сумме вероятностей составляющих его частей:

$$P\left(A\right) = \sum_{i=1}^{n} P\left(A \cdot H_i\right) = \sum_{i=1}^{n} P\left(H_i\right) \cdot P\left(A | H_i\right). \qquad (2.7.)$$

Формула (2.7.) называется формулой *полной вероятности*. Вероятности $P(H_i)$ называются доопытными (априорными) вероятностями гипотез.

Формула Байеса. Предположим, что произведен эксперимент, в результате которого наступило событие A. В связи с этим возникает вопрос: какова вероятность того, что данное событие произошло в результате реализации той или иной гипотезы? Эта вероятность рассчитывается как условная вероятность интересующей нас гипотезы при условии, что произошло событие A.

$$P\left(H_i|A\right) = \frac{P\left(H_i \cdot A\right)}{P\left(A\right)} = \frac{P\left(H_i\right) \cdot P\left(A|H_i\right)}{\sum\limits_{i=1}^{n} P\left(H_i\right) \cdot P\left(A|H_i\right)}. \tag{2.8.}$$

Формула (2.8.) называется *формулой Байеса*. Вероятности $P(H_i \mid A)$ называются послеопытными (апостериорными) вероятностями. Заметим, что знаменатель в формулах (2.8.) совпадает с правой частью формулы (2.7.). Формулы (2.8.) называют также формулами гипотез.

2.2. Повторные независимые испытания. Схема Бернулли.

Испытания, в которых вероятность появления события A в каждом последующем испытании не зависит от исходов других испытаний называются независимыми относительно события A. Очень большое число теоретических и практических задач сводится к схеме последовательных независимых испытаний, даже если они не последовательны во времени.

Определение 2.3. Схемой Бернулли *называется опыт, состоящий в проведении серии n независимых испытаний, в каждом из которых некоторое событие A может произойти с одной и той же вероятностью $P(A) = p$ и не произойти с вероятностью $q = 1 - p$.*

Подсчитаем вероятность $P_n(m)$ события: { Событие A в серии из n испытаний произошло ровно m раз}. Тогда вероятность элементарного события ω_i :

{ событие A произошло ровно m раз и не произошло $n - m$ раз}

будет равна произведению вероятностей соответствующих событий:

$$p \cdot p \cdot \ldots \cdot p \cdot (1-q) \cdot (1-q) \cdot \ldots \cdot (1-q) = p^m \cdot (1-p)^{n-m} = p^m \cdot q^{n-m}.$$

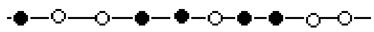


Рис. 2.3. Пример реализации серии независимых испытаний. Светлые точки — событие произошло; тёмные точки — событие не произошло

Число таких элементарных событий ω_i, составляющих интересующее нас событие, равно числу всевозможных реализаций последовательностей появлений и непоявлений (рис. 2.3) события A, т.е. числу сочетаний из n по m.

$$P_n(m) = C_n^m p^m (1-p)^{n-m}. \tag{2.9.}$$

Эта формула называется формулой Бернулли или биномиальной формулой, поскольку правая часть формулы Бернулли представляет собой общий член разложения бинома Ньютона:

$$(q+p)^n = \sum_{i=0}^{n} C_n^i p^i q^{n-i}. \tag{2.10.}$$

Если в формуле (2.10.) положить $p = q = 1/2$, то получим

$$\sum_{i=0}^{n} C_n^i = 2^n. \tag{2.11.}$$

Это выражение даёт число событий, которое можно построить на множестве Ω, содержащем n элементарных событий.

Формула Бернулли (2.9.) позволяет определить не только вероятность появления события A ровно m раз при n испытаниях, но и вероятность $P(m_1 \leqslant m \leqslant m_2)$ того, что число m появлений события A заключено на некотором отрезке $[m_1,\ m_2]$, $0 \leqslant m_1 < m_2 \leqslant n$. Искомая вероятность находится как сумма вероятностей несовместных событий:

$$P_n(m_1 \leqslant m \leqslant m_2) = \sum_{i=m_1}^{m_2} P_n(i). \qquad (2.12.)$$

Наиболее вероятное число появлений события. Если представить (2.9.) как функцию m, то получим зависимость с максимумом.

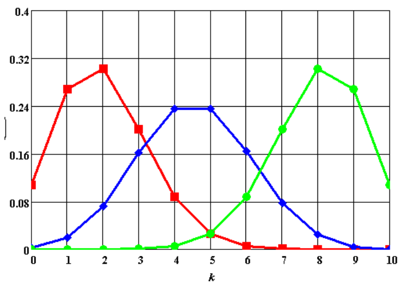


Рис. 2.4. График формулы Бернулли

Формула Бернулли позволяет установить, какое число появлений события A в серии из n испытаний наиболее вероятно.

Число m_0 называется наиболее вероятным, если $P_n(m_0) \geqslant P_n(m)$ при всех m, т.е. $P_n(m)$ достигает своего наибольшего значения при некотором m_0.

Из формулы Бернулли следует, что

$$\frac{P_n(m+1)}{P_n(m)} = \frac{n-m}{m+1} \cdot \frac{p}{q}.$$

Наиболее вероятное число m_0 определяется из двойного неравенства

$$np - q \leqslant m_0 \leqslant np + p, \qquad (2.13.)$$

устанавливающего для m_0 границы, которые отличаются на единицу.[2]

Если левое граничное значение $(np - q)$ — дробное, то дробным будет и правое граничное значение. Тогда существует одно число m_0; Если $(np - q)$ — целое, то $np - q + 1$ тоже целое. В этом случае существует два наиболее вероятнейших числа $m_0 = np - q$ и $m_0 + 1 = np + p$. См. рис.(2.4).

[2] Действительно, левая граница $np - q = np - (1 - p) = np + p - 1$ отличается от правой на единицу.

2.3. Предельные приближения в схеме Бернулли

При больших n вычисления по формуле Бернулли становятся проблематичными: с одной стороны, факториалы — очень быстро растущие функции[3], а, с другой стороны, возведение вероятностей в большие степени порождают очень маленькие числа. Т.е. при компьютерных расчётах возможны как переполнение разрядной сетки, так и получение аппаратного нуля.

Существует два предельных выражения для формулы Бернулли, снимающих эти проблемы.

Приближение Пуассона. Сформулируем предельное соотношение.

Теорема 2.4. ***Теорема Пуассона.*** *Если существует*

$$\lim_{\substack{n\to\infty\\ p\to 0}} np = \lambda, \tag{2.14.}$$

то справедливо приближение Пуассона:

$$P_n(m) = \lim_{\substack{n\to\infty\\ p\to 0}} C_n^m p^m (1-p)^{n-m} = \frac{\lambda^m}{m!}\cdot e^{-\lambda}, \tag{2.15.}$$

где $\lambda = np$ называется параметром Пуассона.

Практическое использование этой формулы допустимо при $\lambda \leqslant 12$. Формула (2.15.) табулирована. Из-за малости p *формулу Пуассона* или *распределение Пуассона* называют также законом редких явлений.

2.4. Локальная и интегральная теоремы Муавра—Лапласа

Если условие применимости формулы Пуассона (2.14.) нарушается, рассматриваются случаи, когда $p \neq 0$ и $p \neq 1$. При этом для подсчёта $P_n(m)$ пользуются локальной предельной теоремой Муавра—Лапласа.

Теорема 2.5. ***Локальная теорема Муавра—Лапласа.*** *Пусть вероятность события A в n независимых испытаниях равна p $(0 < p < 1)$. Тогда вероятность того, что в этих испытаниях событие A наступит ровно m раз, приближённо равна*

[3] При больших n факториал можно вычислить приближенно по формуле Стирлинга:

$$n! \approx \sqrt{2\cdot\pi\cdot n}\cdot(n/e)^n\cdot e^{\theta_n},\ \text{где } \theta_n \le \frac{1}{12n}$$

$$P_n(m) = \frac{1}{\sqrt{npq}}\varphi(x(m)), \tag{2.16.}$$

где

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), \quad x(m) = \frac{m-np}{\sqrt{npq}}, \quad q = 1-p.$$

Функция $\varphi(x)$ называется функцией Гаусса; она является чётной, т.е. $\varphi(-x) = \varphi(x)$; для неё составлены подробные таблицы.

На практике при большом числе испытаний n и не слишком малой вероятности p важно оценить вероятность того, что число появлений события A лежит в некоторых границах. Эту оценку устанавливает

Теорема 2.6. ***Интегральная теорема Муавра – Лапласа.*** *Если вероятность p наступления события A в каждом испытании постоянна, причём $0 < p < 1$, то вероятность того, что событие A появится в испытаниях от m_1 до m_2 раз, приближённо равна*

$$P_n(m_1 \leqslant m \leqslant m_2) \approx \Phi(x_2) - \Phi(x_1), \tag{2.17.}$$

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}}\int\limits_0^x exp\left(-\dfrac{t^2}{2}\right)dt, \quad x_1 = \dfrac{m_1-np}{\sqrt{npq}}, \quad x_2 = \dfrac{m_2-np}{\sqrt{npq}}, \quad q = 1-p.$

Функция $\Phi(x)$ называется функцией Лапласа. Она является нечётной, т.е. $\Phi(-x) = -\Phi(x)$; при $x \geqslant 0$ эта функция табулирована. Оценка погрешности при использовании формулы (2.16.) показывает, что эта формула обеспечивает хорошую точность уже при значениях $np \geqslant 12$.

Таким образом, для формулы Бернулли справедливы следующие приближения.

$$P_n(m) \overset{n\to\infty}{\longrightarrow} \begin{cases} \dfrac{\lambda^m}{m!}e^{-\lambda}, \ \lambda = n\cdot p = const < 12 \\ \dfrac{1}{\sqrt{npq}}\cdot\dfrac{1}{\sqrt{2\pi}}e^{-\dfrac{(m-np)^2}{2\cdot npq}}, n\cdot p\cdot q \geq 9. \end{cases}$$

2.5. Отклонение частоты появления события от его вероятности

Пусть n — число испытаний, p — вероятность появления события A в каждом испытании, $\dfrac{m}{n}$ — относительная частота появления события A. Тогда вероятность того, что отклонение частоты появления события при n испытаниях от его вероятности по абсолютной величине не превышает заданного числа ε, приближённо равна удвоенной функции Лапласа при $x = \varepsilon\sqrt{\frac{n}{pq}}$:

$$P\left(\left|\frac{m}{n} - p\right| \leqslant \varepsilon\right) = P\left(\left|\frac{m - np}{\sqrt{npq}}\right| \leqslant \varepsilon \cdot \sqrt{\frac{n}{pq}}\right) \approx \Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right) = \\ = 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) \qquad (2.18.)$$

Вопросы для самопроверки

1. Сформулируйте теорему сложения вероятностей для несовместных событий.
2. Чему равна сумма вероятностей несовместных событий, образующих полную группу?
3. Какая вероятность называется условной вероятностью?
4. Какие события называются независимыми?
5. Сформулируйте теорему умножения вероятностей и следствия из неё.
6. Как следует вычислять вероятность появления хотя бы одного из нескольких совместных событий?
7. Выведите формулу полной вероятности.
8. Выведите формулу вероятностей гипотез (Байеса).
9. При решении каких задач применяется формула полной вероятности?
10. При решении каких задач применяется формула Байеса?
11. При решении каких задач применяется формула Бернулли?
12. Какие изменения надо ввести в формулу Бернулли, если число исходов в испытаниях больше двух?
13. Дайте определение наиболее вероятного числа появления события при повторных испытаниях и приведите правило его вычисления.
14. Сформулируйте условия применимости приближения Пуассона в схеме Бернулли.
15. Когда следует применять локальную и интегральную теоремы Муавра-Лапласа?

Лекция 3.

Случайные величины и их законы распределения

3.1. Случайные величины

На практике, результаты случайного эксперимента чаще всего представляются в числовой форме. С другой стороны, результат случайного эксперимента — «случайное событие». Чтобы связать эти представления вводят новое понятие — «случайная величина». По сути дела — это числовая функция, заданная на множестве элементарных событий $\{\omega_i \in \Omega\}$ с областью значений в $\mathbb{R}$ или $\mathbb{R}_n$.

Полагают, что случайная величина в результате испытания принимает то или иное возможное значение, заранее неизвестное и зависящее от случайных обстоятельств.

С введением понятия «случайная величина» расширяется понятие «случайное событие» — теперь под случайным событием понимается событие, состоящее в том, что случайная величина в результате испытания приняла значение, принадлежащее некоторому конечному или бесконечному числовому множеству.[1]

Обычно рассматривают два вида случайных величин: *дискретные* и *непрерывные*.

Определение 3.1. *Случайная величина называется дискретной, если она принимает конечное или счётное множество значений.*

Дискретная случайная величина используется при описании измерений, принимающих целочисленные значения: число дефектных изделий, число телефонных вызовов, число неисправностей в приборе и т.д. и может быть записана в виде последовательности x_1, x_2, x_3, ..., x_n.

Для некоторых случайных величин число возможных значений, принимаемых этой величиной, бывает настолько велико, что удобнее представлять их в виде *непрерывных случайных величин*, которые принимают любое значение

[1] «случайная величина» — понятие более емкое, чем прежнее понятие «случайное событие». Её введение позволяет обойтись без описания Ω, отвечающего данному эксперименту. Ведь часто пространство элементарных событий описать очень сложно, а перечислить все ω_i не всегда и возможно.

в некотором интервале, например, продолжительность работы электрической лампы[2], дальность полета снаряда, уровень воды в половодье и т.д. Ниже мы дадим другое, более строгое, определение непрерывной случайной величины.

Закон распределения дискретной случайной величины. Для полного описания дискретной случайной величины необходимо:

- указать все её возможные значения;
- задать вероятности, с которыми принимаются эти значения.

Соотношение между возможными значениями дискретных случайных величин и их соответствующими вероятностями называется законом распределения дискретной случайной величины.

Удобен табличный способ задания закона распределения: в первой строке таблицы указывают значения случайной величины, во второй строке — вероятности этих значений.

Таблица 3.1.

Ряд распределения дискретной случайной величины

X	x_1	x_2	$\ldots$	x_i	$\ldots$	x_n	$\ldots$
P	P_1	P_2	$\ldots$	P_i	$\ldots$	P_n	$\ldots$

Эту таблицу называют рядом распределения дискретной случайной величины. Так как дискретная случайная величина обязательно примет одно из своих значений x_i, то события $\{X = x_i\}$ образуют полную группу событий, поэтому справедливо *условие нормировки*

$$\sum_{i=1}^{\infty} p_i = 1. \tag{3.1.}$$

Полагают, что $x_1 < x_2 < x_3 < \cdots < x_i < x_{i+1} \cdots$.

Функция распределения случайной величины. Дадим следующее определение.

Определение 3.2. *Функцией распределения, или интегральной функцией распределения случайной величины X называется вероятность того, что случайная величина X примет значения, меньшие заданного значения x, где x — любое действительное число:*

$$F(x) = P(X < x). \tag{3.2.}$$

[2] Срок службы электрической лампочки — пример смешанной случайной величины: её значения могут принимать одно дискретное, равное нулю значение, и любые значения из промежутка $(0,\ \infty)$.

Данное определение подходит как для дискретных, так и для непрерывных случайных величин.

Свойства $F(x)$:

1. Значения $F(x)$ принадлежат отрезку [0, 1], т.е.
$$0 \leqslant F(x) \leqslant 1.$$

2. $F(x)$ — неубывающая функция, т.е.
$$(x_1 < x_2) \Rightarrow F(x_1) \leqslant F(x_2).$$

3. $F(x)$ — непрерывная слева в каждой точке x_0, т.е. существует $F(x_0)$ и существует левосторонний предел:
$$F(x_0) = \lim_{x \to x_0 - 0} F(x).$$

4. При любом x_0 существует правосторонний предел, необязательно совпадающий с левосторонним:
$$\lim_{x \to x_0 + 0} F(x) = P(x \leqslant x_0).$$
Функция $F(x)$ может иметь разрывы только первого рода, причём в силу монотонности $F(x)$ и неравенства 0$\leqslant F(x) \leqslant$1 таких скачков конечное или счётное множество.

5. $$\lim_{x \to -\infty} F(x) = 0 \Longleftrightarrow \{ \text{ Невозможное событие}\},$$
$$\lim_{x \to +\infty} F(x) = 1 \Longleftrightarrow \{ \text{ Достоверное событие}\} \qquad (3.3.)$$

6. Вероятность того, что случайная величина попадёт на полуинтервал $[a,\ b)$ равна разности значений функции распределения в точках b и a:
$$P(a \leqslant x < b) = F(b) - F(a) \qquad (3.4.)$$

Замечание. Если $F(x)$ непрерывна в точке a и $a = b$, то $P(X \in [a,\ b]) = P(X = a) = F(b) - F(a) = 0$. Следовательно, для непрерывной в точке функции вероятность попадания на отрезок равна вероятности попадания на интервал.

Пусть дана дискретная случайная величина
$$X = \{x_1,\ x_2, \ldots, x_n\}.$$

Используя свойства функции $F(x)$, получаем, что при $x_{i-1} < x \leqslant x_i$

$$F(x) = p_1 + p_2 + \cdots + p_{i-1} = \sum_{i=1}^{i-1} p_i. \tag{3.5.}$$

В точке x_i $F(x)$ имеет скачок

$$p_i = P(X = x_i) = F(x_i + 0) - F(x_i).$$

Таким образом, функция распределения дискретной случайной величины является кусочно-непрерывной, в точках разрыва x_i имеет скачки P_i и непрерывна слева в точках разрыва x_i.

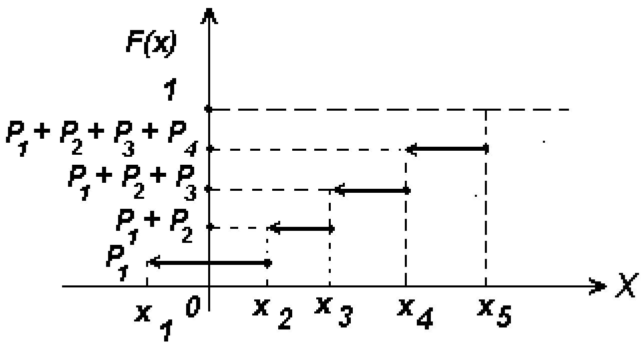


Рис. 3.1. Функция распределения дискретной случайной величины

Непрерывные случайные величины. Дадим необходимые определения.

Определение 3.3. *Случайная величина X называется непрерывной, если её функция распределения непрерывна.*

Определение 3.4. *Пусть F(x) — дифференцируемая функция. Производная от функции распределения $F(x)$ называется плотностью распределения вероятности или дифференциальной функцией распределения случайной величины*

$$f(x) = \frac{dF}{dx} = F'(x). \tag{3.6.}$$

Укажем вероятностный смысл $dF(x)$:

$$dF(x) = f(x)dx \approx F(x + dx) - F(x) = P(x \in (x,\ x + dx)).$$

Таким образом, дифференциал функции распределения есть вероятность попадания случайной величины на бесконечно малый промежуток от x до $x + dx$.

Свойства $f(x)$:

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt; \tag{3.7.}$$

$$f(x) \geqslant 0, \quad \text{т.е. не отрицательная функция}; \tag{3.8.}$$

$$P(a \leqslant x < b) = F(b) - F(a) = \int\limits_{a}^{b} f(x)dx; \tag{3.9.}$$

$$\int\limits_{-\infty}^{+\infty} f(x)dx = 1; \quad \text{условие нормировки}. \tag{3.10.}$$

3.2. Числовые характеристики случайных величин

Числовые характеристики в сжатой форме выражают наиболее существенные особенности распределения случайной величины.

Числовые характеристики широко используются в теории вероятностей и её многочисленных приложениях на практике. С их помощью в значительной степени облегчается решение вероятностных задач. Рассмотрим лишь важнейшие числовые характеристики, которые характеризуют форму распределения и положение случайной величины на числовой прямой.

Математическое ожидание. Пусть дискретная случайная величина задана своим *рядом распределения*:

X	x_1	x_2	x_3	$\ldots$	x_n	$\ldots$
p	p_1	p_2	p_3	$\ldots$	p_n	$\ldots$

Математическим ожиданием дискретной случайной величины называется число

$$m_X = M[X] = \sum_{i=1}^{\infty} x_i p_i, \tag{3.11.}$$

если числовой ряд сходится абсолютно. Если ряд расходится, то говорят, что дискретная случайная величина X не имеет конечного математического ожидания.

Если число значений случайной величины конечно и равно n, то математическое ожидание равно конечной сумме:

$$m_X = M[X] = \sum_{i=1}^{n} x_i p_i. \tag{3.12.}$$

Выясним вероятностный смысл математического ожидания дискретной случайной величины.

Пусть проведено n опытов, в которых случайная величина X приняла m_1 раз значение x_1, m_2 раз значение x_2, ..., m_k раз значение x_k, причём $m_1 + m_2 + \cdots + m_k = n$. Тогда среднее арифметическое значение $\overline{x}$ всех значений равно

$$\overline{x} = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_k m_k}{n} = x_1 \frac{m_1}{n} + x_2 \frac{m_2}{n} + \ldots + x_k \frac{m_k}{n} = \\ = x_1 p_1 * + x_2 p_2 * + \ldots + x_k p_k *,$$

где $p_i * = \dfrac{m_i}{n}$ — относительная частота значения x_i.

Если число опытов достаточно велико, то относительная частота приближённо равна вероятности события. Таким образом, математическое ожидание приближённо равно среднему арифметическому наблюдаемых значений случайной величины.

Выясним механический смысл математического ожидания дискретной случайной величины.

Пусть возможные значения дискретной случайной величины $x_1, x_2, \ldots, x_n$ являются координатами материальных точек с массами $p_1, p_2, \ldots, p_n$, расположенных на числовой прямой, причём $\sum\limits_{i=1}^{n} p_i = 1$.

Тогда

$$M[X] = \sum_{i=1}^{n} x_i p_i = \frac{\sum\limits_{i=1}^{n} x_i p_i}{\sum\limits_{i=1}^{n} p_i},$$

т.е. математическое ожидание есть абсцисса центра масс данной системы материальных точек.

Пусть X — непрерывная случайная величина и $f(x)$ — её дифференциальная функция распределения или плотность вероятности.

Математическим ожиданием непрерывной случайной величины X называется число

$$M[X] = m_x = \int\limits_{-\infty}^{\infty} x f(x) dx, \qquad (3.13.)$$

если этот несобственный интеграл сходится. Если он расходится, то непрерывная случайная величина X математического ожидания не имеет.

Математическое ожидание функции $\varphi(X)$ от случайной величины можно рассчитать без введения её плотности распределения по формуле:

$$M[\varphi(X)] = m_\varphi = \int\limits_{-\infty}^{\infty} \varphi(x) f(x) dx. \qquad (3.14.)$$

Свойства математического ожидания:

1. M[C]=C, т.е. математическое ожидание постоянной равно самой постоянной;

2. $M[C \cdot X] = C \cdot M[X]$ для любой случайной величины X и произвольного числа C;

3. $M[X + Y] = M[X] + M[Y]$ для произвольных случайных величин X, Y;

4. $M[X_1 \cdot X_2 \cdots X_n] = M[X_1] \cdot M[X_2] \cdots M[X_n]$ для n независимых случайных величин $X_1,\ X_2, \ldots, X_n$.

Размерность математического ожидания равна размерности случайной величины X.

Дисперсия. Дисперсией или рассеянием случайной величины X называется математическое ожидание квадрата отклонения случайной величины X от своего математического ожидания:

$$D[x] = M\left[(X - M[X])^2\right]. \tag{3.15.}$$

Если случайная величина X дискретна и задана своим рядом распределения, то

$$D[X] = \sum_{i=1}^{\infty} (x_i - m_X)^2 \, p_i. \tag{3.16.}$$

Если случайная величина X непрерывна и задана на $[a, b]$, то

$$D[X] = \int\limits_a^b (x_i - m_X)^2 \, f(x) dx, \tag{3.17.}$$

где $f(x)$ — функция плотности вероятности.

Из свойств математического ожидания и определения дисперсии имеем

$$\begin{aligned} D[X] &= M\left[(X - m_X)^2\right] = M\left[X^2 - 2Xm_X + m_X^2\right] = \\ &= M[X^2] - 2m_X M[X] + m_X^2 = M[X^2] - m_X^2. \end{aligned} \tag{3.18.}$$

Итак,

$$D[X] = M[X^2] - m_X^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx - m_X^2, \tag{3.19.}$$

то есть дисперсия случайной величины X равна разности математического ожидания квадрата случайной величины и квадрата математического ожидания.

Формула (3.19.) удобна для практического вычисления дисперсии. Размерность дисперсии равна квадрату размерности величины X. Величина $\sigma[X] = \sqrt{D[X]}$ называется средним квадратичным отклонением случайной величины.

Её размерность совпадает с размерностью величины X.

Вероятностный смысл дисперсии: дисперсия измеряет меру рассеивания значений случайной величины X относительно своего математического ожидания.

Механическим аналогом дисперсии служит момент инерции системы материальных точек, распределённых на прямой, относительно своего центра масс.

Свойства дисперсии:

1. $D[C] = 0$, где $C = const$;
2. $D[X] \geqslant 0$;
3. $D[C \cdot X] = C^2 \cdot D[X]$ для любой случайной величины X и произвольного числа C ;
4. $D[X \pm Y] = D[X] + D[Y]$ для независимых случайных величин X и Y.

Свойства среднего квадратического отклонения:

1. $\sigma[C] = 0$, где $C = Const$;
2. $\sigma[C \cdot X] = C \cdot \sigma[X]$;
3. $\sigma[X + Y] = \sqrt{\sigma^2[X] + \sigma^2[Y]}$ для независимых случайных величин X и Y.

Дополнение. *Математическое ожидание* и *дисперсия* являются частными случаями более общих понятий *моментов* распределения (по аналогии с моментами распределений материальных точек).

Различают *начальные моменты* $k-$го порядка

$$\alpha_k = M[x^k] \qquad (3.20.)$$

и *центральные* моменты $k-$го порядка

$$\mu_k = M[(x - M[x])^k]. \qquad (3.21.)$$

Выведем полезное соотношение между центральными и начальными моментами. Положим $(M[x] = m)$. Разложим $(x - m]^k$ по формуле бинома Ньютона:

$$(x - M[x])^k = ((-m) + x)^k = \sum_{i=0}^{k} (-1)^i \cdot C_k^i \cdot m^i \cdot x^{k-i}.$$

Взяв математическое ожидание от левой и правой частей данного выражения, получим:

$$\mu_k = M\left[(x - M[x])^k\right] = \sum_{i=0}^{k} (-1)^i \cdot C_k^i \cdot m^i \cdot \alpha_{k-i} \qquad (3.22.)$$

Используя (3.22.), вычислим несколько моментов.

Таблица 3.2.

Центральные моменты, как функции начальных моментов

k	μ_k	α_k
0	1	1
1	0	$M[X]$
2	$D[X] = \alpha_2 - \alpha_1^2$	$M[X^2]$
3	$\alpha_3 - 3\alpha_1 \cdot \alpha_3 - 2\alpha_1^2$	$M[X^3]$
4	$\alpha_4 - 4\alpha_1\alpha_3 + 6\alpha_1\alpha_2 - 3\alpha_4$	$M[X^4]$
...	...	...

Для более «тонкой» характеристики формы распределения вводятся коэффициенты ***асимметрии*** γ_1 и ***эксцесса*** γ_2:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \tag{3.23.}$$

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \tag{3.24.}$$

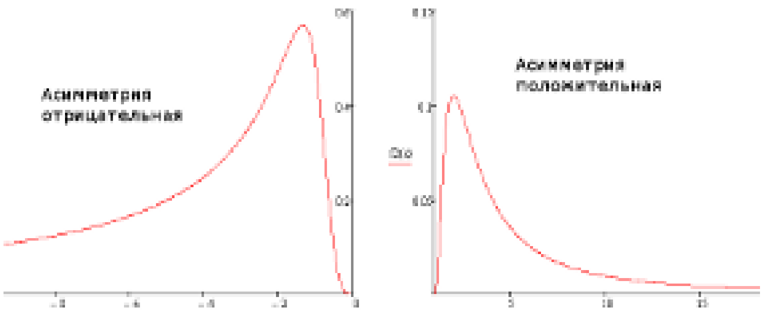


Рис. 3.2. Примеры несимметричных распределений

Набор начальных моментов индивидуален для каждого распределения и служит для его идентификации как отпечатки пальцев для идентификации человека.

Примеры законов распределения случайных величин.

Биномиальное распределение. Пусть производятся испытания по схеме Бернулли:

1. Опыты независимы, т.е. результат каждого опыта не оказывает влияния на другие;

2. Вероятность $P(A) = p$ наступления события A в каждом опыте одна и та же.

Через X обозначим число наступлений события A в серии из n опытов.

Определение 3.5. *Дискретная случайная величина X называется распределённой по биномиальному закону, если свои возможные значения она принимает с вероятностями*

$$P_n(k) = C_n^k p^k (1-p)^{n-k}, \ (k = 0, 1, 2, \ldots, n). \tag{3.25.}$$

Случайная величина X имеет ряд распределения

x	0	1	2	3	$\ldots$	n
p	$(1-p)^n$	$np(1-p)^{n-1}$	$C_n^2 p^2 (1-p)^{n-2}$	$C_n^3 p^3 (1-p)^{n-3}$	$\ldots$	p^n

Числовые характеристики для биномиального распределения:

Математическое ожидание

$$M[X] = n \cdot p. \tag{3.26.}$$

Дисперсия

$$D[X] = n \cdot p \cdot (1-p) \tag{3.27.}$$

Среднее квадратичное отклонение

$$\sigma[X] = \sqrt{np(1-p)}. \tag{3.28.}$$

Функция распределения $F(x)$ имеет вид ступенчатой функции с разрывами в точках $x = 0, 1, 2, \ldots, n$, причём величина скачка в точке $x = m$ равна вероятности $P_n(m) = C_n^m p^m (1-p)^{n-m}$.

Распределение Пуассона. Определение 3.6. *Случайная величина X называется распределённой по закону Пуассона, если свои возможные значения x_k, $k = 0,\ 1,\ 2, \ldots, n, \ldots$ она принимает с вероятностями*

$$P_n(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{3.29.}$$

где λ — параметр распределения (однопараметрическое распределение).

Замечание. Пуассоновское распределение является предельным для биномиального распределения. Дискретная случайная величина X распределена по закону Пуассона, если число испытаний велико, а вероятность p появления события в каждом испытании в схеме Бернулли очень мала и $\lambda = n{\cdot}p$ — среднее число появлений события в n испытаниях. Случайная величина X имеет ряд распределения:

x	0	1	2	3	$\ldots$	m	$\ldots$
p	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$	$\frac{\lambda^3}{3!}e^{-\lambda}$	$\ldots$	$\frac{\lambda^m}{m!}e^{-\lambda}$	$\ldots$

Числовые характеристики распределения Пуассона:

$$M[x] = \lambda;\ D[x] = \lambda;\ \sigma[X] = \sqrt{\lambda}. \tag{3.30.}$$

Отсюда следует смысл параметра λ.

Равномерное распределение Определение 3.7. *Непрерывная случайная величина X называется равномерно распределённой на $[a,\ b]$, если её дифференциальная функция распределения имеет вид:*

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leqslant x \leqslant b \\ 0, & x < a,\ x > b. \end{cases} \tag{3.31.}$$

Интегральная функция распределения равномерно распределённой случайной величины X имеет вид

$$F(x) = \begin{cases} 0, & x \leqslant a \\ \dfrac{x-a}{b-a}, & a < x \leqslant b \\ 1, & x > b \end{cases}. \tag{3.32.}$$

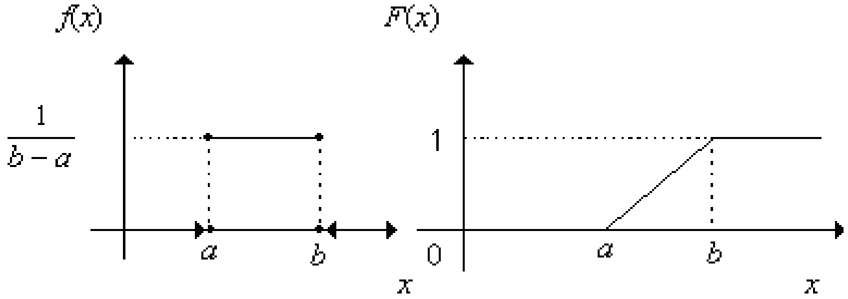


Рис. 3.3. Графики функций $f(x)$ и $F(x)$ равномерного распределения

Вероятность того, что случайная величина, равномерно распределённая в интервале (α, β), принадлежащем [a, b], выражается формулой

$$P(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{\beta - \alpha}{b - a}. \qquad (3.33.)$$

Для равномерного распределения случайной величины X :

$$M[x] = \frac{a+b}{2}, \qquad (3.34.)$$

т.е. математическое ожидание является серединой промежутка $[a, b]$;

$$D[x] = \frac{(b-a)^2}{12}, \ \sigma[X] = \frac{b-a}{2\sqrt{3}}. \qquad (3.35.)$$

Пример. Ошибка отсчёта показаний стрелочного прибора распределена равномерно на отрезке, равном цене деления.

Экспоненциальное распределение. Определение 3.8. *Непрерывная случайная величина X называется распределённой по экспоненциальному (показательному) закону, если дифференциальная функция распределения имеет вид:*

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geqslant 0 \end{cases}, \qquad (3.36.)$$

где λ — параметр распределения. Интегральная функция распределения имеет вид:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geqslant 0 \end{cases}. \qquad (3.37.)$$

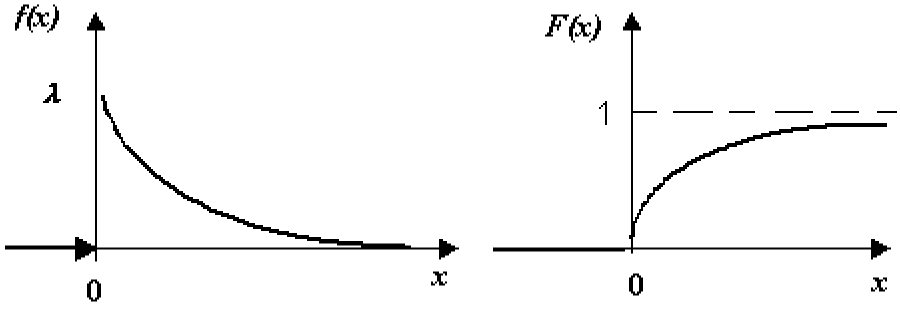


Рис. 3.4. Графики функций $f(x)$ и $F(x)$:

Числовые характеристики распределения:

$$M[x] = \frac{1}{\lambda}, \quad D[x] = \frac{1}{\lambda^2}, \quad \sigma[x] = \frac{1}{\lambda}. \tag{3.38.}$$

Эти формулы устанавливают вероятностный смысл параметра λ.

Вероятность попадания в интервал $(\alpha,\ \beta)$ даётся выражением:

$$P(\alpha \leqslant x < \beta) = \int\limits_{\alpha}^{\beta} f(x)dx = e^{-\lambda\alpha} - e^{-\lambda\beta}. \tag{3.39.}$$

Примеры непрерывных случайных величин, распределённых по показательному закону:

- продолжительность телефонного разговора;
- срок службы радиоэлектронной аппаратуры;
- время ожидания при техническом обслуживании;
- длина пути молекулы между двумя последовательными столкновениями с другими молекулами;
- время обнаружения цели локатором.

Нормальное распределение (распределение Гаусса).

Определение 3.9. *Непрерывная случайная величина X называется распределённой по нормальному закону, или имеет гауссовское распределение, если дифференциальная функция распределения имеет вид:*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \tag{3.40.}$$

где a, σ — параметры распределения.

График функции $f(x)$ называется нормальной кривой. Функция $f(x)$ имеет единственную точку экстремума $x = a$, в которой функция принимает наибольшее значение $f(A) = \frac{1}{\sigma\sqrt{2\pi}}$. В точках $x = a \pm \sigma$ кривая имеет перегиб и $f(a \pm \sigma) = \frac{1}{\sigma\sqrt{2\pi e}}$. Так как $\lim\limits_{x\to\pm\infty} f(x) = 0$, то ось Ox — горизонтальная асимптота для нормальной кривой. Изменение параметра σ ведет к изменению формы кривой: чем меньше σ, тем кривая круче; при увеличении σ она становится более пологой.

Нормальная кривая симметрична относительно прямой $x = a$.

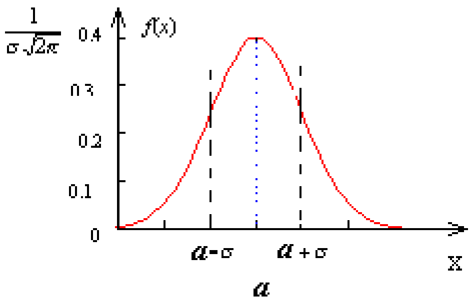


Рис. 3.5. Нормальная кривая

Вероятностный смысл параметров нормального распределения:

$$M[X] = a, \quad D[X] = \sigma^2, \quad \sigma[X] = \sigma, \tag{3.41.}$$

Функция нормального распределения

$$F(x) = \int\limits_{-\infty}^{x} f(\xi)d\xi = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{(\xi-a)^2}{2\sigma^2}} d\xi. \tag{3.42.}$$

Этот «неберущийся» интеграл удобно выразить через табулированную функцию Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-\frac{\xi^2}{2}} d\xi. \tag{3.43.}$$

Именно

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right). \tag{3.44.}$$

Функция $\Phi(x)$ — нечетная.

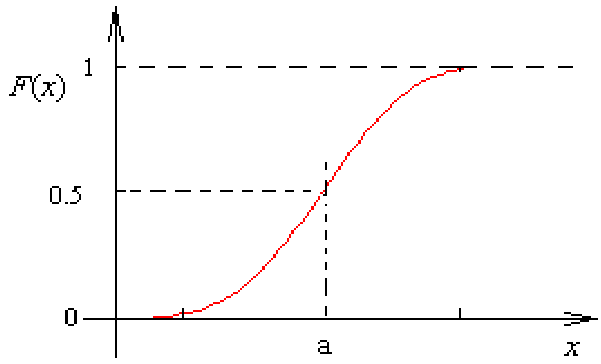


Рис. 3.6. График функции $F(x)$ нормального распределения

При $a = 0$, $\sigma = 1$ получаем стандартное (нормированное) нормальное распределение. Для него

$$f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \varphi(x). \tag{3.45.}$$

Функция $\varphi(x)$ — четная. Для нее составлены подробные таблицы.

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{\xi^2}{2}} d\xi \tag{3.46.}$$

Функция $F_0(x)$ — табулирована.
Вероятность попадания случайной величины в интервал $[\alpha, \beta)$:

$$P(\alpha \leqslant x < \beta) = F(\beta) - F(\alpha) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \tag{3.47.}$$

Если $\alpha = a - \delta$, $\beta = a + \delta$, где δ — произвольное число, то

$$P(a - \delta \leqslant x < a + \delta) = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(-\frac{\delta}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

При $\delta = 3\sigma$

$$P(a - 3\sigma \leqslant x < a + 3\sigma) = 2\Phi(3) = 2 \cdot 0.49865 = 0.9973.$$

Правило трех сигм. Если случайная величина подчинена нормальному закону, то вероятность ее отклонения от математического ожидания больше трех средних квадратичных ошибок, близка к нулю $(p = 0.0027)$.

Или практически достоверно, что нормальная случайная величина принимает значения на отрезке $[a-3\sigma,\ a+3\sigma]$, так как $p=0.9973$.

На практике многие случайные величины распределены нормально или почти нормально. Например: ошибки прямых измерений; ошибки стрельбы, наведения; отклонение напряжения в сети от номинала; суммарная выплата страхового общества за большой период; дальность полета снаряда; частота события при большом числе опытов; масса вылавливаемой рыбы одного вида; рост мужчин (женщин) одного возраста и национальности.

Вопросы для самопроверки

1. Какая величина называется случайной величиной?
2. Дайте определение дискретной и непрерывной случайных величин. Приведите примеры дискретной и непрерывной случайных величин.
3. Что называется законом распределения случайной величины?
4. Что называется рядом распределения дискретной случайной величины?
5. Дайте определение функции распределения вероятности. перечислите и докажите свойства функции распределения.
6. Как найти вероятность попадания случайной величины в заданный интервал?
7. В чем различаются графики функции распределения дискретной и непрерывной случайных величин?
8. Дайте определение плотности распределения вероятностей. Перечислите и докажите ее свойства. Пригодно ли понятие плотности распределения вероятностей для дискретных случайных величин?
9. Как определить вероятность попадания случайной величины в заданный интервал с помощью плотности распределения?
10. Что называется математическим ожиданием дискретной случайной величины?
11. Что называется математическим ожиданием непрерывной случайной величины?
12. Какова механическая интерпретация математического ожидания?
13. Что называется модой случайной величины? Что называется медианой случайной величины?
14. Дайте определение дисперсии случайной величины. Перечислите ее свойства.
15. Что называется средним квадратичным отклонением случайной величины?
16. Что называется начальным моментом $k-$го порядка случайной величины?
17. Что называется центральным моментом $k-$го порядка случайной величины?

18. Какое распределение случайной величины называется биномиальным? Чему равны математическое ожидание и дисперсия случайной величины, имеющей биномиальное распределение?
19. Какое распределение случайной величины называется распределением Пуассона? Чему равны математическое ожидание и дисперсия случайной величины, имеющей распределением Пуассона?
20. Какое распределение случайной величины называется равномерным?
21. Какое распределение случайной величины называется показательным распределением?
22. Какое распределение случайной величины называется нормальным распределением?
23. Как называется график плотности нормального распределения и каковы его свойства?
24. Что называется функцией Лапласа и каковы ее свойства?
25. В чем заключается сущность закона больших чисел?
26. Как записывается неравенство Чебышева?
27. Какое практическое и теоретическое значение имеет неравенство Чебышева?
28. Сформулируйте и докажите теорему Чебышева.
29. Какое практическое значение имеет теорема Чебышева?
30. Объясните, пользуясь теоремой Бернулли, свойство устойчивости относительных частот появления события в серии испытаний.
31. В чем заключается сущность центральной предельной теоремы?
32. Приведите примеры задач, в которых применяется теорема Муавра-Лапласа?

Лекция 4.

Многомерные случайные величины

На пространстве Ω элементарных событий каждому элементарному событию ω_i поставим в соответствие n случайных величин:

по закону φ_1 одномерную случайную величину X_1;

по закону φ_2 одномерную случайную величину X_2;

... ...

по закону φ_n одномерную случайную величину X_n.

В этом случае говорят, что, на Ω определена система случайных величин $(X_1,\ X_2,\ \ldots X_n)$. В дальнейшем будем использовать сокращённое обозначение для термина *«случайная величина»* – СВ.

Примеры:

1. Пусть пространство элементарных событий – множество студентов данной группы, участвующих в социологическом опросе. Введем случайные величины: X_1- рост человека, X_2- его масса. Тогда $X = (X_1,\ X_2)$ – двумерная случайная величина.

2. Пусть $X_1,\ X_2, X_3\quad -$ координаты центра тяжести в некоторой декартовой системе координат. Тогда случайное положение центра тяжести ракеты в пространстве – трёхмерная СВ $X = (X_1,\ X_2,\ X_3)$.

Для простоты рассмотрим двухмерный случай $(X,\ Y)$. Возможные значения двухмерной СВ $(X,\ Y)$ есть пары чисел $(x,\ y)$.

Геометрически двумерную СВ можно трактовать как случайную точку $M\,(X,\ Y)$ на плоскости xOy. Если составляющие СВ X и Y – дискретные (непрерывные) СВ, то и двумерная случайная величина называется дискретной (непрерывной).

4.1. Законы распределения двухмерной случайной величины

4.1.1. Матрица распределения двумерной дискретной случайной величины

Пусть $X\,(x_1,\ x_2,\ \ldots,\ x_n)$ и $Y\,(y_1,\ y_2,\ \ldots,\ y_m)$ — одномерные дискретные случайные величины. Тогда система $(X,\ Y)$ принимает значение $(x_i,\ y_j)$ $(i = \overline{1,n};\ j = \overline{1,m})$.

Обозначим $p_{ij} = P\,(X = x_i,\ Y = y_j)$ вероятность того, что СВ X принимает значение x_i, а СВ Y – значение y_j. Из чисел p_{ij} можно составить таблицу размера $m \times n$. Эта таблица называется законом распределения двумерной СВ или матрицей распределения системы $(X,\ Y)$.

Таблица 4.1.

Ряд распределения системы из двух дискретных случайных величин

Y	X			
	x_1	x_2	$\ldots$	x_n
y_1	p_{11}	p_{12}	$\ldots$	p_{1n}
y_2	p_{21}	p_{22}	$\ldots$	p_{2n}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
y_m	p_{m1}	p_{m2}	$\ldots$	p_{mn}

Так как события $(X = x_i,\ Y = y_j)$ образуют полную группу, то $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n} p_{ji} = 1$.

Зная матрицу распределения системы $(X,\ Y)$, можно найти *частные* законы распределения СВ X и Y, входящих в систему.

Так, СВ X принимает значение x_i в комбинациях

$$(x_i, y_1)\,, (x_i, y_2)\,,\ (x_i,\ y_3)\,,\ \ldots,\ (x_i,\ y_m)$$

с вероятностями $p_{1i},\ p_{2i},\ p_{3i},\ \ldots,\ p_{mi}$.

Событие $A = (X = x_i)$ (i – фиксировано) представлено в виде группы m несовместных событий.

$$\begin{aligned} A = (X = x_i) \quad &= \quad (X = x_i,\ Y = y_1) \cup (X = x_i,\ Y = y_2) \cup \ldots \\ &\ldots \quad \cup (X = x_i,\ Y = y_m) \end{aligned} \qquad (4.1.)$$

Следовательно, по теореме сложения вероятностей несовместных событий, имеем

$$P\left(X=x_i\right)=\sum_{j=1}^{m} P\left(X=x_i,\ Y=y_j\right)=\sum_{j=1}^{m} p_{ji}=p_{1i}+p_{2i}+\ldots+p_{mi}$$

(i– фиксировано).

Аналогично находится закон распределения СВ Y. В результате мы получаем следующие ряды распределения:

Таблица 4.2.
Ряд распределения СВ X

X	x_1	x_2	$\ldots$	x_n
p_x	$\sum\limits_{j=1}^{m} p_{j1}$	$\sum\limits_{j=1}^{m} p_{j2}$	$\ldots$	$\sum\limits_{j=1}^{m} p_{jm}$

Таблица 4.3.
Ряд распределения СВ Y

Y	y_1	y_2	$\ldots$	y_m
p_y	$\sum\limits_{i=1}^{n} p_{1i}$	$\sum\limits_{i=1}^{n} p_{2i}$	$\ldots$	$\sum\limits_{i=1}^{n} p_{mi}$

4.1.2. Функции от дискретных случайных величин

Пусть X – дискретная СВ с законом распределения

X	x_1	x_2	x_3	$\ldots$	x_n
p	p_1	p_2	p_3	$\ldots$	p_n

Рассмотрим СВ $Y=\varphi(X)$, где φ – заданная функция. Если СВ X принимает значения $x_1,\ x_2,\ \ldots,\ x_n$, то СВ $Y=\varphi(X)$ принимает значения $\varphi(x_1),\ \varphi(x_2),\ \ldots,\ \varphi(x_n)$ с вероятностями $p_1,\ p_2,\ \ldots, p_n$. Если значения $\varphi(x_i)$ различны, то закон распределения СВ $Y=\varphi(X)$ имеет вид

Таблица 4.4

X	x_1	x_2	$\ldots$	x_n
Y	$\varphi(x_1)$	$\varphi(x_2)$	$\ldots$	$\varphi(x_n)$
p	p_1	p_2	$\ldots$	p_n

Если $\varphi(X_i)=\varphi(X_j)$, $(i\neq j)$, то в силу несовместности событий $(X=x_i)$ и $(X=x_j)$ получаем $(X=x_j)\cup(X=x_i)=(Y=\varphi(x_i))$.

По теории сложения вероятностей несовместных событий при $i\neq j$

$$P\left(Y=\varphi\left(x_i\right)\right)=P\left(X=x_i\right)+P\left(X=x_j\right)=P_i+P_j$$

Пусть X и Y – дискретные независимые СВ с законами распределения

X	x_1	x_2	$\ldots$	x_n
p_x	p_1	p_2	$\ldots$	p_m

Y	y_1	y_2	$\ldots$	y_m
p_y	p_1	p_2	$\ldots$	p_m

Найдём закон распределения СВ $Z = X + Y$. Для этого надо найти все возможные значения Z и их вероятности. Возможные значения Z есть суммы каждого возможного значения X со всеми возможными значениями Y: $z_{ij} = x_i \cdot y_j$.

Найдём вероятности этих возможных значений.

Каждое событие $\{Z = z_{ij}\}$ есть произведение двух событий: $\{X = x_i\}$ и $\{Y = y_j\}$, вероятности которых равны p_i и p_j соответственно. Так как $\{X = x_i\}$ и $\{Y = y_j\}$ – независимые, то вероятности их совместного наступления по теории умножения равны $p_i \cdot p_j$.

Если значения $Z_{ij} = x_i + y_j$ различные, то закон распределения СВ Z имеет вид

$$P(Z_{ij}) = p_i \cdot p_j \tag{4.2.}$$

Если среди значений Z_{ij} есть одинаковые, то их вероятности надо сложить.

Чтобы составить закон распределения СВ $Z = XY$, надо найти все возможные значения Z и их вероятности. Возможные значения Z есть произведения каждого возможного значения X со всеми возможными значениями Y: $Z_{ij} = x_i + y_j$. Вероятности этих значений равны произведению вероятностей $p_i \cdot p_j$. Если значения Z_{ij} различные, то закон распределения Z имеет вид такой же как и для суммы (4.2.).

$$P(Z_{ij}) = p_i \cdot p_j$$

Если среди значений Z_{ij} есть одинаковые, то нужно сложить их вероятности.

4.1.3. Функция распределения двумерной случайной величины

Пусть $(X,\ Y)$ – двумерная СВ.

Введем вероятность совместного появления двух событий $(X < x)$ и $(Y < y)$

$$P\left(X < x,\ Y < y\right) \forall x,\ y \in R$$

Функция

$$F\left(x,\ y\right) = P\left(X < x,\ Y < y\right) \forall x,\ y \in R \tag{4.3.}$$

называется функцией распределения двумерной СВ $(X,\ Y)$.

Геометрически $F(x,\ y)$ задаёт вероятность попадания случайной точки $(X,\ Y)$ в бесконечный прямой угол с вершиной в точке $(x,\ y)$

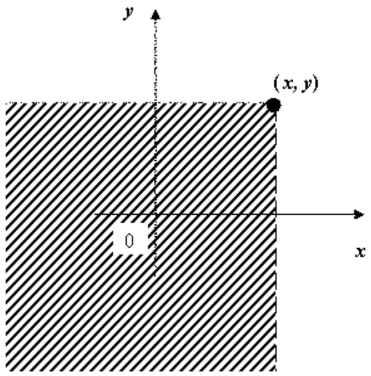


Рис. 4.1. Геометрическая интерпретация функции распределения системы двух СВ

Функция распределения $F(x,\ y)$ обладает следующими свойствами:

1. $0 \leqslant F(x,\ y) \leqslant 1, \quad \forall x, y \in R;$
2. $F(x,\ y)$ - неубывающая функция по каждому аргументу:
   если $x_2 > x_1$, y фиксировано, то $F(x_2, y) \geqslant F(x_1, y)$,
   если $y_2 > y_1$, x – фиксировано, то $F(x, y_2) \geqslant F(x, y_1)$.
3. $\lim\limits_{x \to -\infty} F(x, y) = 0,$
4. $\lim\limits_{y \to -\infty} F(x, y) = 0,$
5. $\lim\limits_{\substack{x \to -\infty \\ y \to -\infty}} F(x, y) = 0,$
6. $\lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y) = 1,$

Первые три предела соответствуют вероятностям невозможных событий а четвёртый – достоверного события.

Частные функции распределения для СВ X и Y $F_1(x)$ и $F_2(y)$ получим с помощью следующих преобразований:

$$F_1(x) = \lim_{y \to +\infty} F(x, y); \quad F_2(y) = \lim_{x \to +\infty} F(x, y) \tag{4.4.}$$

Из определения функции распределения следует, что

$$P\left(x_1 \leqslant X \leqslant x_2,\ y_1 \leqslant Y \leqslant y_2\right) = \\ = \left[F\left(x_2, y_2\right) - F\left(x_1, y_2\right)\right] - \left[F\left(x_2, y_1\right) - F\left(x_1, y_1\right)\right] = \\ = \left[F\left(x_1, y_1\right) + F\left(x_2, y_2\right)\right] - \left[F\left(x_1, y_2\right) + F\left(x_2, y_1\right)\right] \quad (4.5.)$$

4.1.4. Плотность распределения непрерывной двумерной случайной величины

Пусть функция распределения $F(x,\ y)$ системы $(X,\ Y)$ непрерывна и имеет непрерывные частные производные $F'_x,\ F'_y,\ F''_{xy}$. Функция

$$f(x,y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P\left(x \leqslant X \leqslant x + \Delta x,\ y \leqslant Y \leqslant y + \Delta y\right)}{\Delta x \cdot \Delta y} = \frac{\partial^2 F(x,y)}{\partial x \partial y} \quad (4.6.)$$

называется плотностью распределения вероятностей системы $(X,\ Y)$.

Свойства плотности вероятностей $f(x,y)$ системы $(X,\ Y)$.

Неотрицательность.

$$f(x,y) \geqslant 0,\ \forall x,\ y \leqslant R; \quad (4.7.)$$

Вероятность попадания в область. Если $f(x,y)$ непрерывна в произвольной замкнутой области D на плоскости, то

$$P\left((X,Y) \in D\right) = \iint\limits_D f(x,y)\,dxdy. \quad (4.8.)$$

Связь с функцией распределения.

$$F(x,\ y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(x,y)\ dxdy; \quad (4.9.)$$

Условие нормировки.

$$\int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(x,y)\,dxdy = 1; \quad (4.10.)$$

Частные функции распределения.

$$F_1(x) = \int_{-\infty}^{x} \left[\int_{-\infty}^{+\infty} f(x,y)\,dy \right] dx \tag{4.11.}$$

$$F_2(y) = \int_{-\infty}^{y} \left[\int_{-\infty}^{+\infty} f(x,y)\,dx \right] dy; \tag{4.12.}$$

Частные плотности распределения.

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, \quad f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx. \tag{4.13.}$$

4.1.5. Зависимые и независимые случайные величины. Условные законы распределения

Определение 4.1. *Две случайные величины являются независимыми, если закон распределения одной из них не зависит от того, какое значение приняла другая величина. В противном случае случайные величины считаются зависимыми.*

Определение 4.2. *Распределение одной случайной величины, входящей в систему, найденное при условии, что другая случайная величина, входящая в систему, приняла определённое значение, называется условным законом распределения.*

Система дискретных случайных величин

Событие $A = (X = x_i,\ Y = y_j)$ является произведением двух событий $B = (X = x_i)$ и $C = (Y = y_j)$. Если события B и C независимы, то и СВ X и Y, входящие в систему, называются независимыми. В противном случае они зависимы.

Тогда

$$P(A) = P(BC) = \\ = \begin{cases} P(B) \cdot P(C) & B \text{ и } C \text{ независимы} \\ P(B) \cdot P(C\,|B\,) = P(C) \cdot P(B\,|C\,) & B \text{ и } C \text{ зависимы} \end{cases} \tag{4.14.}$$

Условные вероятности

$$P\left(Y = y_j \middle| \, X = x_i\right) = \frac{P\left(X = x_i, \; Y = y_j\right)}{P\left(X = x_i\right)} = \frac{p_{ij}}{p_i} \tag{4.15.}$$

$$P\left(X = x_i \middle| \, Y = y_j\right) = \frac{P\left(X = x_i, \; Y = y_j\right)}{P\left(Y = y_j\right)} = \frac{p_{ij}}{p_j} \tag{4.16.}$$

и есть условные законы распределения двух дискретных случайных величин, входящих в систему.

Если случайные величины независимы, то из условия независимости (4.14.) следует, что $p_{ij} = p_j \cdot p_i$ и условные распределения совпадают с безусловными:

$$\begin{aligned} P(y \mid X = x_i) &= \frac{p_{ji}}{p_i} = \frac{p_j \cdot p_i}{p_i} = p_j \\ P(x \mid Y = y_j) &= \frac{p_{ji}}{p_j} = \frac{p_j \cdot p_i}{p_j} = p_i \end{aligned} \tag{4.17.}$$

Условные функции распределения дискретных СВ вводятся аналогично:

$$\begin{aligned} F(y \mid X = x_i) &= P(Y < y \mid X = x_i) = \frac{P(Y < y, X = x_i)}{P(X = x_i)} \\ F(x \mid Y = y_j) &= P(X < x \mid Y = y_j) = \frac{P(X < x, Y = y_j)}{P(Y = y_j)} \end{aligned} \tag{4.18.}$$

Непрерывные СВ

Для непрерывной случайной величины вероятность того, что случайная величина примет какое-либо наперед заданное значение равна нулю:

$$P(X = \overset{*}{x}) = F(\overset{*}{x}) - F(\overset{*}{x}) = 0,$$

поэтому определение условной функции распределения вида (4.18.) невозможно.

Условную функцию распределения мы получим совершив предельный переход в выражении

$$\frac{P(Y < y, x \leqslant X \leqslant X + \triangle\, x)}{P(x \leqslant X < x + \triangle\, x)}.$$

Разделим числитель и знаменатель на $\triangle\ x$ и заменим вероятности попадания в область соответствующими приращениями функций распределения, а затем перейдем к пределу при стремлении $\triangle\ x$ к нулю.

$$F(y \mid x) = \lim_{\triangle x \to 0} \frac{\dfrac{F(x + \triangle\ x, y) - F(x, y)}{\triangle\ x}}{\dfrac{F_1(x + \triangle\ x) - F_1(x)}{\triangle\ x}} =$$

$$= \frac{\dfrac{\partial F(x,y)}{\partial x}}{f_1(x)} = \int\limits_{-\infty}^{y} \frac{f(x,t)}{f_1(x)}\,dt = \int\limits_{-\infty}^{y} f(t \mid x)\,dt \qquad (4.19.)$$

Мы ввели плотность распределения СВ X $f_1(x)$ под знак интеграла, поскольку она от y не зависит.

Аналогично найдем условную функцию распределения для СВ X.

$$F(x \mid y) = \frac{\dfrac{\partial F(x,y)}{\partial y}}{f_2(y)} = \int\limits_{-\infty}^{x} \frac{f(t,y)}{f_2(y)}\,dt = \int\limits_{-\infty}^{x} f(t \mid y)\,dt \qquad (4.20.)$$

Введем плотности условных распределений

$$f(x \mid y) = \frac{\partial}{\partial x} F(x \mid y) = \frac{f(x,y)}{f_2(y)}$$
$$f(y \mid x) = \frac{\partial}{\partial y} F(y \mid x) = \frac{f(x,y)}{f_1(x)} \qquad (4.21.)$$

Теорема 4.1. *Для независимости СВ X и Y необходимо и достаточно, чтобы выполнялось условие*

$$F\left(x,y\right) = F_1\left(x\right) \cdot F_2\left(y\right) \qquad (4.22.)$$

или (для непрерывных случайных величин)

$$f\left(x,y\right) = f_1\left(x\right) \cdot f_2\left(y\right). \qquad (4.23.)$$

4.2. Числовые характеристики системы двух случайных величин

4.2.1. Математическое ожидание двумерной С.В.

Дискретные случайные величины

Пусть дискретная двумерная величина $(X,\ Y)$ задана матрицей распределения (3.1). Суммируя по столбцам, а затем по строкам элементы матрицы распределения, находим ряды распределений С.В. X и Y (3.2) и (3.3). Тогда математические ожидания С.В. X и Y можно рассчитать следующим образом:

$$M[x] = m_x = \sum_i x_i p_i = \sum_i x_i \sum_j p_{ij} = \sum_i \sum_j x_i p_{ij}.$$
$$M[y] = m_y = \sum_j y_j p_j = \sum_j y_j \sum_i p_{ij} = \sum_j \sum_i y_j p_{ij}.$$

Если $Z = \varphi\left(X\right)$, то

$$M[Z] = \sum_{i=1}^{n} p_i \varphi(x_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ji} \varphi(x_j).$$

Если $Z = \Psi\left(Y\right)$, то

$$M[Z] = \sum_{j=1}^{m} p_j \Psi(y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ji} \Psi\left(y_j\right).$$

Непрерывные случайные величины

Пусть система непрерывных величин (X, Y) распределена в области D плоскости $x0y$ с плотностью распределения $f\left(x,\ y\right)$. Тогда математические ожидания С.В. X и Y будут рассчитываться следующим образом:

$$M[X] = m_x = \iint\limits_D x \cdot f\left(x, y\right)\ dxdy = \int\limits_{-\infty}^{+\infty} x \cdot f_1\left(x\right)\ dx$$

$$M[Y] = m_y = \iint\limits_D y \cdot f\left(x, y\right) dxdy = \int\limits_{-\infty}^{+\infty} y \cdot f_2\left(y\right) dy.$$

Если $Z = \varphi(x)$, то

$$M[Z] = m_z = \iint\limits_D \varphi(x) \cdot f(x, y)\, dxdy$$

Если $Z = \varphi(X,\ Y)$, то

$$M[Z] = m_z = \iint\limits_D \varphi(x, y) \cdot f(x, y)\, dxdy.$$

4.2.2. Дисперсия двумерной С.В. $(X,\ Y)$.

Пусть С.В. X и Y – дискретные и имеют математические ожидания m_x и m_y.Тогда дисперсии находятся следующим образом:

$$D[X] = d_x = M\left[(X - m_x)^2\right] = \sum_{i=1}^{n} (x_i - m_x)^2 \cdot P_i = M\left[X^2\right] - m_x^2,$$

$$D[Y] = d_y = M\left[(Y - m_y)^2\right] = \sum_{j=1}^{m} (y_j - m_y)^2 \cdot P_j = M\left[Y^2\right] - m_y^2.$$

Средние квадратичные отклонения

$$\sigma[X] = \sigma_x = \sqrt{D[X]}; \quad \sigma[Y] = \sigma_y = \sqrt{D[Y]}.$$

Для непрерывных С.В. X и Y формулы принимают вид:

$$D[X] = \int\limits_{-\infty}^{+\infty} (x - m_x)^2 \cdot f_1(x)\, dx;$$

$$D[Y] = \int\limits_{-\infty}^{+\infty} (y - m_y)^2 \cdot f_2(y)\, dy;$$

или

$$D[X] = \int\limits_{-\infty}^{+\infty} x^2 \cdot f_1(x)\, dx - m_x^2 = M\left[X^2\right] - m_x^2$$

$$D[Y] = \int\limits_{-\infty}^{+\infty} y^2 \cdot f_2(y)\, dy - m_y^2 = M\left[Y^2\right] - m_y^2.$$

Иногда удобнее использовать формулы, содержащие дифференциальную функцию системы:

$$M[X] = \iint\limits_G x \cdot f\,(x, y)\,dx, \quad M[Y] = \iint\limits_G y \cdot f\,(x, y)\,dy \quad ,$$

$$D[X] = \iint\limits_G (x - m_x)^2 \cdot f\,(x, y)\,dxdy = \iint\limits_G x^2 \cdot f\,(x, y)\,dxdy - m_x^2,$$

$$D[Y] = \iint\limits_G (y - m_y)^2 \cdot f\,(x, y)\,dxdy = \iint\limits_G y^2 \cdot f\,(x, y)\,dxdy - m_y^2.$$

Здесь двойные интегралы берутся по области возможных значений системы G.

Ковариация и коэффициент корреляции

Ковариацией или *корреляционным моментом* С.В. X и Y называется величина

$$K_{x,y} = \operatorname{cov}(X,\ Y) = M[(X - m_x) \cdot (Y - m_y)].$$

Для дискретных С.В. X и Y

$$K_{x,y} = \operatorname{cov}(X,\ Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i - m_x)\,(y_j - m_y) \cdot P_{ij}$$

или

$$\operatorname{cov}(X,\ Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j \cdot P_{ij} - m_x \cdot m_y.$$

Формулы можно объединить в одну

$$\operatorname{cov}(X,\ Y) = M[X \cdot Y] - m_x m_y$$

Безразмерная величина

$$\rho_{x,y} = \frac{K_{x,y}}{\sigma_x \sigma_y} = \frac{cov\,(X,\ Y)}{\sigma_x \sigma_y}$$

называется *коэффициентом корреляции*.

Если $\rho_{x,y} = 0$, то С.В. X и Yназываются некоррелированными; если $\rho_{x,y} \neq 0$, то X и Y коррелированы.

4.3. Свойства числовых характеристик двумерной случайной величины

Свойства математического ожидания

1. $M[C] = C, \ C - const$;
2. Если m_x и m_y– конечны, и $\alpha, \ \beta$ – константы, то
$$M[\alpha X + \beta Y] = \alpha M[X] + \beta M[Y].$$
При $\alpha = \beta = 1 \quad M[X + Y] = M[X] + M[Y]$;
3. Если m_x и m_y– конечны, то
$$M[XY] = M[X] \cdot M[Y] + cov\,(X, \ Y)\,.$$
Если С.В. X и Y не коррелированы (независимы), то
$$M[XY] = M[X] \cdot M[Y] \tag{4.24.}$$
(т.к. $\operatorname{cov}(X, \ Y) = 0$).

Свойства дисперсии

1. $D[C] = 0 \quad C - const$
2. $D[CX] = C^2 D[X]$
3. Если d_x и d_y – конечны, то
$$D[\alpha X + \beta Y] = \alpha^2 d_x + \beta^2 d_y + 2\alpha\beta \operatorname{cov}(X, \ Y)\,.$$
Если С.В. X и Yне коррелированы (т.е. независимы), то $\operatorname{cov}(X, \ Y) = 0$ и
$$D[\alpha X + \beta Y] = \alpha^2 d_x + \beta^2 d_y.$$
Формулу можно обобщить на любое конечное число слагаемых
$$D\Big[\sum \alpha_i X_i\Big] = \sum_{i=1}^{n} \alpha_i^2 D[X_i] + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \alpha_i \alpha_j \operatorname{cov}(X_i, \ Y_j).$$
Для независимых С.В.
$$D\Big[\sum \alpha_i X_i\Big] = \sum_{i=1}^{n} \alpha_i^2 D[X_i].$$
4. Если С.В. X и Yнезависимы и d_x и d_y – конечны, то
$$D[X \cdot Y] = d_x \cdot m_y^2 + d_y \cdot m_x^2 + d_x \cdot d_y.$$

Свойства ковариации

1. Ковариация симметрична относительно своих аргументов: $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$ – это следует из определения.
2. Если С.В. X и Y независимы, то они некоррелированы.

   Действительно, для независимых С.В. X и Y $Y f_1(x) \cdot f_2(y) = f(x,y)$.

   Тогда

$$\operatorname{cov}(X,Y) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xy \cdot f_1(x) f_2(y)\, dxdy - m_x m_y =$$
$$= \int\limits_{-\infty}^{+\infty} x \cdot f_1(x)\, dx \cdot \int\limits_{-\infty}^{+\infty} y \cdot f_2(y)\, dy - m_x m_y = m_x m_y - m_x m_y = 0$$

3. Обратное утверждение, в общем случае, не верно, т.е. из равенства нулю ковариации не следует независимость X и Y.

   Только в случае нормального распределения, справедливо также и обратное утверждение.

Свойства коэффициента корреляции:

1. $\rho_{x,x} = 1$
2. $\rho_{x,y} = \rho_{y,x}$
3. Если С.В. X и Yнезависимы, то $\rho_{x,y} = 0$
4. $|\rho_{x,y}| \leqslant 1$
5. $|\rho_{x,y}| = 1 \Leftrightarrow Y = aX + b$, где $a,\ b$ – некоторые постоянные, т.е. между С.В. X и Yсуществует линейная функциональная зависимость.

Свойства 1–3 следуют из определений независимости (4.23.) и коэффициента корреляции.

Докажем свойство 4. Рассмотрим подробнее структуру выражения $\rho_{x,y}$:

$$\begin{aligned}
\rho_{x,y} &= \frac{M\big[(X - m_X)(Y - m_Y)\big]}{\sigma_X \cdot \sigma_Y} = \\
&= M\left[\left(\frac{X - m_X}{\sigma_X}\right) \cdot \left(\frac{Y - m_Y}{\sigma_Y}\right)\right] = \\
&= M\big[\widetilde{X} \cdot \widetilde{Y}\big]
\end{aligned} \tag{4.25.}$$

Случайные величины $\widetilde{X}$ и $\widetilde{Y}$ получены из исходных с помощью операций *центрирования* и *нормирования*. Центрирование – перенос начала координат в центр распределения $(m_X,\ m_Y)$, а нормирование – изменение масштаба, когда за единицу измерения берется среднее квадратичное отклонение. Такие СВ называют *стандартизированными*. Стандартные случайные величины имеют следующие значения числовых характеристик:

$$m_{\widetilde{X}} = M\Big[\frac{X - m_X}{\sigma_X}\Big] = \frac{M[X] - m_X}{\sigma_X} = 0; \qquad (4.26.)$$
$$m_{\widetilde{Y}} = 0$$

$$D_{\widetilde{X}} = M\big[\widetilde{X}^2\big] = D\Big[\frac{X - m_X}{\sigma_X}\Big] = \frac{M\big[(X - m_X)^2\big]}{\sigma_X^2} = 1; \qquad (4.27.)$$
$$D_{\widetilde{Y}} = M\big[\widetilde{Y}^2\big] = 1$$

Рассмотрим очевидное неравенство:

$$(\widetilde{X} \pm \widetilde{Y})^2 \geqslant 0$$

или, переходя к математическому ожиданию,

$$M\big[(\widetilde{X} \pm \widetilde{Y})^2\big] \geqslant 0 \qquad (4.28.)$$

Раскроем скобки и выделим произведение СВ:

$$M\Big[\frac{\widetilde{X}^2 + \widetilde{Y}^2}{2}\Big] \geqslant \mp M\big[\widetilde{X}\widetilde{Y}\big] \qquad (4.29.)$$

Найдем математическое ожидание от левой и правой частей этого неравенства:

$$\frac{M\big[\widetilde{X}^2\big] + M\big[\widetilde{Y}^2\big]}{2} \geqslant \mp M\big[\widetilde{X} \cdot \widetilde{Y}\big] = \mp\rho_{X,Y} \qquad (4.30.)$$

или, с учётом свойств стандартных величин

$$1 \geqslant \mp\rho_{X,Y} \Longrightarrow 1 \geqslant \mid \rho_{X,Y} \mid,$$

что и требовалось доказать.

Докажем свойство 5.

Пусть между случайными величинами существует функциональная линейная зависимость $Y = aX + b$. Тогда имеем:

$$\begin{aligned} m_Y &= a \cdot m_X + b \\ \sigma_Y^2 &= a^2 \cdot \sigma_X^2, \quad \sigma_Y = \mid a \mid \cdot \sigma_X \\ \operatorname{cov}(X,Y) &= M\big[(X - m_x)(aX + b - am_x - b)\big] = \\ &= M\big[a(X - m_x)^2\big] = a \cdot \sigma_X^2 \\ \rho_{x,y} &= \frac{a \cdot \sigma_X^2}{\mid a \mid \cdot \sigma_X^2} = \frac{a}{\mid a \mid} = \begin{cases} 1, & \textbf{при} \ \ a > 0 \\ -1, & \textbf{при} \ \ a < 0 \end{cases} . \end{aligned}$$

Пусть теперь $\rho_{x,y} = \pm 1$. Строгое равенство соответствует строгому равенству в выражении (4.28.).

$$M\big[(\widetilde{X} \pm \widetilde{Y})^2\big] = 0.$$

Но равенство нулю математического ожидания неотрицательной величины возможно лишь в том случае, если она тождественно равна нулю. Тогда

$$(\widetilde{X} \pm \widetilde{Y})^2 = 0 \Rightarrow (\widetilde{X} \pm \widetilde{Y}) = 0 \Rightarrow \pm\widetilde{X} = \widetilde{Y}. \tag{4.31.}$$

Отсюда получаем

$$Y = \pm\frac{\sigma_y}{\sigma_x} \cdot X \mp \frac{\sigma_y}{\sigma_x} \cdot m_x = a \cdot X + b, \tag{4.32.}$$

т.е случайные величины связаны функциональной зависимостью.

4.4. Числовые характеристики условных законов распределения

4.4.1. Условные математические ожидания. Линии регрессии

Для дискретной двумерной С.В. $(X,\ Y)$ находим ряд распределения X при $Y = y_j$ – фиксированным, используя формулы (4.15.,4.16.)

$$P\left(X = x_i|Y = y_j\right) = \frac{P\left(X = x_i,\ Y = y_j\right)}{P\left(Y = y_j\right)} = \frac{P_{ij}}{P_j}$$

и ряд распределения Y при $X = x_i$ – фиксированном по формуле

$$P\left(Y = y_j|X = x_i\right) = \frac{P\left(Y = y_j,\ X = x_i\right)}{P\left(X = x_i\right)} = \frac{P_{ji}}{P_i}.$$

Тогда условные математические ожидания находим по формулам:

$$M[X|Y = y_j] = x_1 \cdot \frac{P_{1j}}{P_j} + x_2 \cdot \frac{P_{2j}}{P_j} + \ldots + x_n \cdot \frac{P_{nj}}{P_j} = \sum_{i=1}^{n} x_i \cdot \frac{P_{ij}}{P_j},$$

$$M[Y|X = x_i] = y_1 \cdot \frac{P_{1i}}{P_i} + y_2 \cdot \frac{P_{2i}}{P_i} + \ldots + y_m \cdot \frac{P_{mj}}{P_j} = \sum_{j=1}^{m} y_j \cdot \frac{P_{ji}}{P_i}.$$

Для непрерывных С.В. X и Y

$$M[X|Y = y] = \int_{-\infty}^{+\infty} y \cdot f\left(x|y\right) dx, \tag{4.33.}$$

$$M[Y|X = x] = \int_{-\infty}^{+\infty} y \cdot f\left(y|x\right) dy \tag{4.34.}$$

Эти условные математические ожидания являются некоторыми функциями $\varphi\left(y\right)$ и $\psi\left(x\right)$ соответственно. Функция $M[X|Y = y] = \varphi\left(y\right)$ называется *функцией регрессии* X на Y.

График функции $x = \varphi\left(y\right)$ называется *кривой регрессии* X *на* Y.

Функция $\psi\left(x\right) = M[Y|X = x]$ называется *функцией регрессии* С.В. Y на С.В. X.

График функции $y = \psi\left(x\right)$ называется *кривой регрессии* Y *на* X.

Если С.В. X и Y независимы, то условные математические ожидания равны безусловным и кривые регрессии являются прямыми, параллельными осям координат.

Зависимости $\psi(x)$ и $\varphi(y)$, показывающие, как *среднее значение* одной величины изменяется при изменении другой величины, называются *корреляционными* (*случайными, статистическими, стохастическими*)[1].

В отличии от корреляционных, *функциональные* зависимости предполагают, что не среднее, а *каждое* значение одной величины есть однозначная функция другой величины.

4.4.2. Условные дисперсии

Условные дисперсии для дискретных С.В. X и Y:

$$D[X|Y=y_j]=\sum_{i=1}^{n}(x_i-M[X|Y=y_j])^2\cdot P(x_i|y_j);$$

$$D[Y|X=x_i]=\sum_{j=1}^{m}(y_j-M[Y|X=x_i])^2\cdot P(y_j|x_i).$$

Условные дисперсии для непрерывных С.В. X и Y:

$$D[X|Y=y]=\int_{-\infty}^{+\infty}(x-M[X|Y=y])^2\cdot f(x|y)dx, \tag{4.35.}$$

$$D[Y|X=x]=\int_{-\infty}^{+\infty}(y-M[Y|X=x])^2\cdot f(y|x)dy. \tag{4.36.}$$

Рассмотрим на примере случайной величины Y очень полезное соотношение, связывающее условную дисперсию $D[Y\mid X=x]$ и функцию регрессии $\psi(x)$.

Условные дисперсии характеризуют разброс случайной величины относительно соответствующей функции регрессии. Если усреднить $D[Y\mid X=x]$ по X, то полученная величина является обобщающей мерой разброса С.В. Y относительно линии регрессии.

[1] В общем случае, корреляционная зависимость предполагает, что изменение одной случайной величины приводит к изменению закона распределения другой случайной величины.

$$\int\limits_{-\infty}^{\infty} D\big[Y \mid X = x\big] \cdot f_1(x)\,dx = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\big(y-\psi(x)\big)^2 f(y \mid x)\cdot f_1(x)\,dy\,dx =$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\big(y-\psi(x)\big)^2 f(x,y)\,dy\,dx = D_\psi(Y) \qquad (4.37.)$$

Она является одной из составляющих безусловной дисперсии:

$$\begin{aligned} D[Y] &= \iint\limits_D (y-m_y)^2 \cdot f\,(x,y)\,dxdy = \\ &= \iint\limits_D [y-\psi(x)+\psi(x)-m_y]^2 \cdot f\,(x,y)\,dxdy = \\ &= \iint\limits_D [y-\psi(x)]^2 \cdot f\,(x,y)\,dxdy + \iint\limits_D [\psi(x)-m_y]^2 \cdot f\,(x,y)\,dxdy + \\ &+ 2\iint\limits_D [y-\psi(x)][\psi(x)-m_y] \cdot f\,(x,y)\,dxdy \end{aligned} \qquad (4.38.)$$

Последний интеграл обращается в нуль:

$$\iint\limits_D [y-\psi(x)][\psi(x)-m_y] \cdot f\,(x,y)\,dxdy =$$

$$= \int\limits_{-\infty}^{\infty} [\psi(x)-m_y]\left\{\int\limits_{-\infty}^{\infty} [y-\psi(x) f(y \mid x)\,dy]\right\} f_1(x)\,dx =$$

$$= \int\limits_{-\infty}^{\infty} [\psi(x)-m_y]\,\{\psi(x)-\psi(x)\}\,f_1(x)\,dx = 0 \qquad (4.39.)$$

Поэтому мы получаем:

$$D[Y] = D_\psi(y) + D_{m_y}(\psi) \qquad (4.40.)$$

Первое слагаемое есть дисперсия С.В. относительно линии регрессии (в регрессионном анализе она называется остаточной дисперсией), а второе слагаемое – дисперсия функции регрессии относительно безусловного среднего.

Покажем, что функция регрессии обеспечивает минимум остаточной дисперсии по сравнению с другими пробными функциями, подставляемыми вместо неё в (4.37.).

Возьмём произвольную функцию $g(x)$ и подставим её вместо функции регрессии в выражение для остаточной дисперсии:

$$\begin{aligned}
D_g(Y) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (y-g(x))^2 f(x,y)\,dy\,dx = \\
&= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (y-\psi(x)+\psi(x)-g(x))^2 f(x,y)\,dy\,dx = \\
&= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \left\{(y-\psi(x))^2+(\psi(x)-g(x))^2\right\} f(x,y)\,dy\,dx+ \\
&+ 2\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \{(y-\psi(x))(\psi(x)-g(x))\} f(x,y)\,dy\,dx = \\
&= D_Y(\psi)+M\left[(\psi(x)-g(x))^2\right]+ \\
&+ 2\int\limits_{-\infty}^{\infty}[\psi(x)-g(x)]\left\{\int\limits_{-\infty}^{\infty}(y-\psi(x))f(y\mid x)\,dy\right\} f_1(x),dx = \\
&= D_\psi(Y)+M\left[(\psi(x)-g(x))^2\right]
\end{aligned} \tag{4.41.}$$

Здесь последний интеграл, как и в (4.39.), обращается в нуль. Отсюда следует, что

$$D_g(Y) \geqslant D_\psi(Y), \tag{4.42.}$$

так как $M\left[(\psi(x)-g(x))^2\right]$ есть величина неотрицательная.

Свойство функции регрессии минимизировать остаточную дисперсию является теоретическим обоснованием метода выбора оптимальных значений параметров кривых, использующихся для аппроксимации регрессионных зависимостей.[2]

[2]Более общий метод нахождения оптимальных аппроксимирующих зависимостей известен как *метод наименьших квадратов*. Предложен и обоснован К.Гауссом и А.Лежандром, строгое математическое обоснование дано А.А. Марковым и А.Н. Колмогоровым

4.4.3. Линейная регрессия и корреляция

На практике особую важность представляют случаи, когда функцию регрессии можно аппроксимировать линейной зависимостью. Рассмотрим для примера случайную величину Y:

$$M[Y|X = x] = \psi(x) = A \cdot x + B \tag{4.43.}$$

Параметры линейной зависимости A, B подбираются так, чтобы остаточная дисперсия $D_Y(A \cdot x + B)$ достигала своего минимума.

Итак, мы будем искать минимум функции

$$D_{A \cdot x+B}(Y) = M\left[(y - A \cdot x - B)^2\right], \tag{4.44.}$$

рассматривая её как функцию параметров A и B.

Как известно из курса дифференциального исчисления, необходимым условием существования экстремума функции двух переменных является равенство нулю частных производных по этим переменным:

$$\begin{cases} \dfrac{\partial}{\partial A} M\left[(y - A \cdot x - B)^2\right] & = 0 \\ \dfrac{\partial}{\partial B} M\left[(y - A \cdot x - B)^2\right] & = 0. \end{cases} \tag{4.45.}$$

Внеся операцию дифференцирования под знак математического ожидания (это эквивалентно дифференцированию по параметру подинтегральной функции), получим:

$$\begin{cases} M\left[(Y - A \cdot X - B) \cdot X\right] & = 0 \\ M\left[(Y - A \cdot X - B)\right] & = 0. \end{cases} \tag{4.46.}$$

Из первого уравнения системы следует:

$$M[X \cdot Y] = A \cdot M[X^2] + B \cdot m_x,$$

а из второго

$$m_y = A \cdot m_x + B. \tag{4.47.}$$

Исключая B, получим:

$$M[X \cdot Y] - m_x \cdot m_y = A \cdot \left(M[X^2] - m_x^2\right).$$

Отсюда

$$
\begin{aligned}
A = {} & \frac{M[X \cdot Y] - m_x \cdot m_y}{M[X^2] - m_x^2} = \\
& = \frac{\operatorname{cov}\bigl(X \cdot Y\bigr)}{{\sigma_x}^2} = \frac{\operatorname{cov}\bigl(X \cdot Y\bigr)}{\sigma_x \cdot \sigma_y} \cdot \frac{\sigma_y}{\sigma_x} = \\
& = \rho_{x,y} \cdot \frac{\sigma_y}{\sigma_x},
\end{aligned} \tag{4.48.}
$$

$$
B = m_y - \rho_{x,y} \cdot \frac{\sigma_y}{\sigma_x} \cdot m_x. \tag{4.49.}
$$

Обозначив функцию регрессии $\psi(x)$ как y, запишем уравнение линейной регрессии в симметричной форме:

$$
\frac{y - m_y}{\sigma_y} = \rho_{x,y} \cdot \frac{x - m_x}{\sigma_x} \tag{4.50.}
$$

Коэффициент корреляции оценивает тесноту линейной связи между X и Y. Чем ближе абсолютная величина коэффициента корреляции к единице, тем связь сильнее, то есть ближе к функциональной; связь слабеет, когда абсолютная величина коэффициента корреляции стремится к нулю. В пределе, когда $\rho_{x,y} = 0$, линия регрессии вырождается в прямую, параллельную оси абсцисс: $y = m_Y$, что характерно для независимых случайных величин.

Найдем выражение для остаточной дисперсии (4.44.). Для этого подставим в него найденные значения параметров линейной регрессии:

$$
\begin{aligned}
D_\psi(Y) = {} & M\Bigl[\Bigl((y - m_y) - \rho\frac{\sigma_y}{\sigma_x}(x - m_x)\Bigr)^2\Bigr] = \\
= {} & \sigma_y^2 - 2\rho\frac{\sigma_y}{\sigma_x}\operatorname{cov}(X, Y) + \bigl(\rho\frac{\sigma_y}{\sigma_x}\bigr)^2\sigma_x^2 = \\
= {} & \sigma_y^2(1 - \rho^2)
\end{aligned} \tag{4.51.}
$$

Видно, что в отсутствии корреляции ($\rho = 0$) остаточная дисперсия принимает максимальное значение, равное безусловной дисперсии, а в случае функциональной линейной зависимости ($\rho = 1$) остаточная дисперсия равна нулю.

Вопросы для самопроверки

1. Что называется системой случайных величие?
2. Как можно трактовать систему случайных величин7
3. Дайте определение функции распределения системы двух случайных величин и укажите ее свойства.
4. Дайте определения плотности распределения вероятностей системы двух случайных величин. Перечислите и докажите ее свойства.
5. Как определить вероятность попадания в данную область?
6. Что называется условным законом распределения?
7. Как выражается плотность распределения каждой из величин, входящих в систему, через плотность распределения системы?
8. какие случайные величины называются зависимыми (независимыми)?
9. Что является необходимым и достаточным условием независимости случайных величин?
10. Что называется корреляционным моментом? коэффициентом корреляции?
11. Чему равен коэффициент корреляции для независимых случайных величин?
12. Какие случайные величины называются некоррелированными?
13. Следует ли из независимости случайных величин их некоррелированность и наоборот?
14. Равносильны ли понятия некоррелированности и независимости для нормально распределённой системы?

Лекция 5.

Предельные теоремы теории вероятностей

Рассматриваются две группы теорем. Первая группа называется «Закон больших чисел» (ЗБЧ). ЗБЧ утверждает, что среднее арифметическое большого числа случайных слагаемых «стабилизируется» с ростом этого числа. Как бы сильно каждая с. в. не отклонялась от своего среднего значения, при суммировании эти отклонения «взаимно гасятся», так что среднее арифметическое приближается к постоянной величине.
Вторая группа теорем называется «Центральная предельная теорема» (ЦПТ). ЦПТ утверждает. что распределение суммы случайных величин с ростом числа слагаемых при соблюдении некоторых условий приближается нормальному распределению.

5.1. Законы больших чисел

При рассмотрении предельных теорем нам придётся ввести понятие сходимости *последовательности случайных величин.*

5.1.1. Сходимость по вероятности

Мы определили случайную величину как числовую функцию, заданную на множестве элементарных событий Ω. И если мы хотим говорить о сходимости последовательности случайных величин, не будем забывать, что мы имеем дело не с последовательностью чисел, а с последовательностью функций, заданных на одном и том же множестве Ω.

Существуют разные виды сходимости последовательностей случайных величин (функций). Нам понадобится «сходимость по вероятности».

Определение 5.1. *Говорят, что последовательность случайных величин $\{X_n\}$ сходится по вероятности к случайной величине X при $n \to \infty$* (в краткой записи $X_n \xrightarrow{p} X$)*, если $P\left(|X_n - X| > \varepsilon\right) \to 0$ для любого $\varepsilon > 0$.*

Чтобы доказывать сходимость по вероятности, надо просто уметь вычислять $P\left(|X_n - X| > \varepsilon\right)$ при больших n. Но для этого необходимо знать распределение X_n, что не всегда возможно. Оценить эту вероятность сверху позволяют неравенства П. Л. Чебышёва.

Теорема 5.1. (первое неравенство Чебышёва) *Если неотрицательная случайная величина X имеет конечное математическое ожидание $\mathbf{M}\Big[X\Big]$, то для любого $\varepsilon > 0$ справедливо неравенство*

$$P(X \geq \varepsilon) \leq \frac{\mathbf{M}\Big[X\Big]}{\varepsilon} \ \text{ или } \ P(X < \varepsilon) \geq 1 - \frac{\mathbf{M}\Big[X\Big]}{\varepsilon} \tag{5.1.}$$

Доказательство. Обозначим $F(x)$ функцию распределения случайной величины X, тогда

$$P\Big\{X \geqslant \varepsilon\Big\} = \int\limits_{x \geqslant \varepsilon} dF(x).$$

Так как в области интегрирования $\frac{x}{\varepsilon} \geq 1$, то

$$P(X \geq \varepsilon) = \int\limits_{\varepsilon}^{\infty} dF(x) \leq \int\limits_{0}^{\infty} \frac{x}{\varepsilon} dF(x) = \frac{\mathbf{M}\Big[X\Big]}{\varepsilon.}$$

Первое неравенство Чебышёва доказано.

Неравенство Чебышёва. Докажем следующее неравенство[1]:

Теорема 5.2. *Для любой случайной величины X, имеющей конечную дисперсию $\mathbf{D}\Big[X\Big]$, при любом $\varepsilon > 0$. имеет место неравенство*

$$P\Big\{\Big|X - \mathbf{M}\Big[X\Big]\Big| \geqslant \varepsilon\Big\} \leqslant \frac{\mathbf{D}\Big[X\Big]}{\varepsilon^2} \tag{5.2.}$$

Доказательство. Обозначим $F(x)$ функцию распределения случайной величины X, тогда

$$P\Big\{\Big|X - \mathbf{M}\Big[X\Big]\Big| \geqslant \varepsilon\Big\} = \int\limits_{|x - M[x]| \geqslant \varepsilon} dF(x).$$

Так как в области интегрирования

$$\frac{|x - M[X]|}{\varepsilon} \geqslant 1,$$

[1] В 1853 г. И. Бьенеме (I. Bienayme) и в 1866 г., независимо от него, П. Л. Чебышёв прямыми методами доказали это неравенство.

то

$$\int\limits_{|x-M[x]|\geqslant\varepsilon} dF(x) \leqslant \frac{1}{\varepsilon^2} \int\limits_{|x-M[x]|\geqslant\varepsilon} \Big(x - \mathbf{M}\Big[X\Big]\Big)^2 dF(x).$$

Мы только усилим неравенство, распространив интегрирование на все x

$$\int\limits_{|x-M[x]|\geqslant\varepsilon} dF(x) \leqslant \frac{1}{\varepsilon^2} \int\limits_{-\infty}^{+\infty} \Big(x - \mathbf{M}\Big[X\Big]\Big)^2 dF(x) = \frac{\mathbf{D}\Big[X\Big]}{\varepsilon^2}.$$

Неравенство Чебышёва доказано.

Для противоположного события неравенство Чебышёва выглядит так:

$$P\left\{\left|X - M\left[X\right]\right| < \varepsilon\right\} \geq 1 - \frac{D\left[X\right]}{\varepsilon^2}. \qquad (5.3.)$$

В качестве следствия получим так называемое «правило трёх сигм», которое формулируют, например, так: вероятность случайной величине отличаться от своего математического ожидания более, чем на три корня из дисперсии, мала. Для каждого распределения величина этой вероятности своя: для нормального распределения, например, эта вероятность равна 0,0027. Мы получим верную для *всех* распределений с конечной дисперсией оценку сверху для «вероятности с. в. отличаться от своего математического ожидания более, чем на три корня из дисперсии».

$$P\left\{\left|X - M\left[X\right]\right| > 3\cdot\sqrt{\mathbf{D}\Big[X\Big]}\right\} \leqslant \frac{\mathbf{D}\Big[X\Big]}{\left(3\cdot\sqrt{\mathbf{D}\Big[X\Big]}\right)^2} = \frac{1}{9}$$

.

Определение 5.2. *Говорят. что последовательность случайных величин (СВ)* $\{X_n\}_{i=1}^{\infty}$ *с конечными первыми моментами удовлетворяет закону больших чисел (ЗБЧ), если*

$$\frac{X_1 + \dots + X_n}{n} - \frac{\mathbf{M}\Big[X_1\Big] + \dots + \mathbf{M}\Big[X_n\Big]}{n} \xrightarrow{p} 0 \qquad (5.4.)$$

Законами больших чисел принято называть утверждения об условиях, при которых последовательность СВ «удовлетворяет закону больших чисел».

Рассмотрим последовательность независимых одинаково распределённых СВ

Теорема 5.3. **(ЗБЧ в форме Чебышёва)** *Среднее арифметическое любой последовательности независимых одинаково распределённых СВ сходится по вероятности к своему математическому ожиданию.*

Доказательство. Поскольку СВ, входящие в последовательность одинаково распределены, то они имеют одинаковые математические ожидания $\mathbf{M}\Big[X_i\Big] = m$ и дисперсии $\mathbf{D}\Big[X_i\Big] = D$.

Введем случайную величину $Y_n = \dfrac{1}{n} \cdot \sum\limits_{i=1}^{n} X_i$ – среднее арифметическое последовательности СВ. Y_n имеет следующие числовые характеристики:

$$\mathbf{M}\Big[Y_n\Big] = \mathbf{M}\Big[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\Big] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{M}\Big[X_i\Big] = \frac{n \cdot m}{n} = m.$$

Из условия независимости имеем:

$$\mathbf{D}\Big[Y_n\Big] = \mathbf{D}\Big[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\Big] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbf{D}\Big[X_i\Big] = \frac{n \cdot D}{n^2} = \frac{D}{n}.$$

Применим к величине Y_n неравенство Чебышёва [5.2.].

$$P\left\{\big|Y_n - M\left[Y_n\right]\big| > \varepsilon\right\} \leqslant \frac{D\left[Y_n\right]}{\varepsilon^2} = \frac{D}{n \cdot \varepsilon^2} \xrightarrow{n \to \infty} 0.$$

Таким образом, мы доказали теорему Чебышёва.

Теорема Чебышёва обобщается и на случаи, когда случайные величины имеют различные математические ожидания и дисперсии, а также на случай зависимых СВ

Теорема Бернулли. Одним из важных частных случаев теоремы Чебышёва является теорема Бернулли, объясняющая факт устойчивости относительной частоты события в серии независимых испытаний и дающая теоретическое обоснование статистическому определению вероятности события.

Теорема 5.4. **(Бернулли)** *Пусть μ — число наступлений события A в n независимых испытаниях и p — вероятность наступления события A в каждом из испытаний. Тогда, каково бы ни было $\varepsilon > 0$,*

$$\lim_{n \to \infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1 \qquad (5.5.)$$

т.е. каким бы малым положительным числом ε ни было, вероятность события

$$\left|\frac{\mu}{n} - p\right| < \varepsilon$$

стремится к единице.

Доказательство. Введем случайную величину $I_i(A)$ — индикатор события, которая с вероятностью p принимает значение единица, если событие A произошло в i–м испытании и равное нулю в противном случае с вероятностью $q = 1 - p$. Числовые характеристики индикатора:

$$\mathbf{M}\Big[I\Big] = 0 \cdot q + 1 \cdot p = p, \ \mathbf{D}\Big[I\Big] = p - p^2 = p \cdot q \leqslant \frac{1}{4}. \tag{5.6.}$$

Представим число появления события через индикатор $\mu = \sum\limits_{i=1}^{n} I_i$, а относительную частоту появления события как $\frac{\mu}{n} = \frac{1}{n} \cdot \sum\limits_{i=1}^{n} I_i$, то есть как среднее арифметическое индикаторов.

Числовые характеристики:

$$\mathbf{M}\Big[\frac{\mu}{n}\Big] = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{M}\Big[I_i\Big] = \frac{n \cdot p}{n} = p, \ \mathbf{D}\Big[\frac{\mu}{n}\Big] = \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathbf{D}\Big[I_i\Big] = \frac{n \cdot p \cdot q}{n^2} = \frac{p \cdot q}{n}. \tag{5.7.}$$

Тогда, согласно теореме Чебышёва, среднее значение индикаторов (относительная частота события) сходится по вероятности к математическому ожиданию индикатора — вероятности события в отдельном испытании.

Замечание 1.1. Сходимость по вероятности в законах больших чисел можно пояснить следующим образом: последовательность СВ есть последовательность функций плотности. Тогда сходимость по вероятности означает «перетекание» с ростом n плотности распределения в наперед заданную как угодно малую ε–окрестность около предельной точки. В пределе СВ вырождается в константу.

Замечание 1.2. Последовательность среднего СВ сходится по вероятности к среднему математических ожиданий, если её дисперсия стремится к нулю.

5.2. Центральная предельная теорема

Существует известное замечание Липмана (цитируемое Пуанкаре), гласящее, что «"...каждый уверен в справедливости закона ошибок, экспериментаторы — потому, что они думают, что это математическая теорема, математики — потому, что они думают, что это экспериментальный факт». Стоит отметить, что обе стороны совершенно правы, если это их убеждение не слишком безусловно: математическое доказательство говорит нам, что при некоторых ограничительных условиях *мы вправе ожидать нормального распределения, а статистический опыт показывает, что в действительности распределения часто* приближенно нормальные.

Г. Крамер

Все формы ЦПТ устанавливают условия, при которых возникает нормальный закон распределения. При доказательстве ЦПТ используется математический аппарат — преобразование Фурье.

Характеристические функции. Характеристической функцией (ХФ) $g_X(t)$ случайной величины X называется математическое ожидание величины $Z = e^{itX}$, где t — некоторый вещественный параметр, а i — мнимая единица $(i^2 = -1)$.

Если X дискретная СВ, заданная рядом распределения, то

$$g_X(t) = \sum_{k=1}^{\infty} p_k \cdot e^{it \cdot x_k} \tag{5.8.}$$

Если же X непрерывна с плотностью $f(x)$, то

$$g_X(t) = \int_{-\infty}^{+\infty} e^{it \cdot x} \cdot f(x)dx \tag{5.9.}$$

Как видно, ХФ есть преобразование Фурье плотности распределения. По известной ХФ обратным преобразованием Фурье можно найти плотность (функцию) распределения.

$$f(x) = \frac{1}{2 \cdot \pi} \int_{-\infty}^{+\infty} e^{-itx} g_x(t)dt \tag{5.10.}$$

Доказано, что два соотношения: $F_n(x) \to F(x)$ и $g_n(t) \to g_X(t)$ равносильны. Причём сходимость ХФ доказывается с меньшими трудностями, чем сходимость функций распределения. Т.е. зная предельное значение ХФ, с помощью обратного преобразования можем найти предельное значение функции распределения.

При практическом использовании ХФ удобно её представить формулой Тейлора при $t = 0$ (ограничимся тремя членами разложения):

$$g_X(t) = g_X(0) + \frac{g'_X(0)}{1!} \cdot (t)^1 + \frac{g''_X(0)}{2!} \cdot (t)^2 + \frac{g'''_X(\xi)}{(3)!} \cdot (t)^3. \tag{5.11.}$$

Имеем:

$$g_X(0) = \int\limits_{-\infty}^{+\infty} e^0 f(x) dx = 1 \tag{5.12.a}$$

$$g'_X(0) = i \cdot \int\limits_{-\infty}^{+\infty} e^0 x \cdot f(x) dx = i \cdot \alpha_1 = i \cdot \mathbf{M}\Big[X\Big] \tag{5.12.b}$$

$$g''_X(0) = i^2 \cdot \int\limits_{-\infty}^{+\infty} e^0 x^2 \cdot f(x) dx = i^2 \cdot \alpha_2 \tag{5.12.c}$$

$$g^{(k)}_X(0) = (i)^k \cdot \int\limits_{-\infty}^{+\infty} e^0 x^k \cdot f(x) dx = (i)^k \cdot \alpha_k \tag{5.12.d}$$

Здесь α_k – начальные моменты $k-$го порядка.

В результате имеем

$$g_X(t) = 1 + \frac{\alpha_1}{1!} \cdot (it)^1 + \frac{\alpha_2}{2!} \cdot (it)^2 + \varepsilon(t) \cdot (t)^2. \tag{5.13.}$$

Где $\varepsilon(t)$ — бесконечно малая величины.

Пусть СВ Z есть сумма попарно независимых случайных величин: $Z = \sum\limits_{k=1}^{n} X_k$. Тогда её ХФ равна произведению ХФ случайных величин:

$$g_Z(x) = \mathbf{M}\Big[\exp\Big(it \cdot \sum_{k=1}^{n} X_k\Big)\Big] = \prod_{k=1}^{n} \mathbf{M}\Big[\exp\big(it \cdot X_k\big)\Big] = \prod_{k=1}^{n} g_{X_k}(x). \tag{5.14.}$$

Для удобства расчётов вводится *кумулятивная функция*, равная логарифму характеристической функции:

$$\varphi(t) = \ln(g(t)) \tag{5.15.}$$

Кумулятивная функция суммы независимых СВ равна сумме их кумулятивных функций. Ряд Маклорена кумулятивной функции:

$$\varphi(t) = \varphi(0) + \varphi'(0) \cdot t + \frac{\varphi''(t) \cdot t^2}{2!} + \varepsilon(t) \cdot t^2, \tag{5.16.}$$

где ε бесконечно малая величина.

$$\varphi'(t) = \frac{g'_x(t)}{g_x(t)},$$

$$\varphi''(t) = \frac{g''_x(t) \cdot g_x(t) - [g'_x(t)]^2}{[g_x(t)]^2}.$$

Учитывая, что $g_x(0) = 1$, находим

$$\varphi'(0) = i \cdot \mathbf{M}\Big[X\Big],\ \mathbf{M}\Big[X\Big] = m = -i \cdot \varphi'(0),$$

$$\varphi''(0) = g''_x(0) - [g'_x(0)]^2 = i^2 \cdot \left(\mathbf{M}\Big[X^2\Big] - \left(\cdot\, \mathbf{M}\Big[X\Big]\right)^2\right) =$$

$$= -\mathbf{M}\Big[X^2\Big] + \mathbf{M}\Big[X\Big]^2 = -D_X,\ D_X = \sigma^2 = -\varphi''(0).$$

Тогда

$$\varphi(t) = m \cdot it + \frac{\sigma^2 \cdot (it)^2}{2} + \varepsilon(t) \cdot t^2, \tag{5.17.}$$

Величины $k_j = i^j \cdot \varphi^{(j)}$ называются кумулянтами (семиинвариантами) $j-$го порядка случайной величины. При сложении независимых СВ их кумулянты складываются.

СВ $Y = aX + b$, полученная путём линейного преобразования СВ X, имеет ХФ

$$g_Y(t) = \int\limits_{-\infty}^{\infty} e^{it(ax+b)} f(x)dx = e^{itb} g_X(at) \tag{5.18.}$$

и кумулятивную функцию

$$\varphi(t) = itb + \varphi(at)$$

Характеристические функции нормального распределения. Рассмотрим вначале стандартное распределение. [2]

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx - \frac{x^2}{2}} dx = e^{-\frac{t^2}{2}} \tag{5.19.}$$

Для произвольного нормального распределения $N(a, \sigma^2)$ по формуле 5.18. получим

$$g_x(t) = e^{iat - \frac{\sigma^2 t^2}{2}}, \quad \varphi_x(t) = iat - \frac{\sigma^2 t^2}{2} \tag{5.20.}$$

А.М. Ляпунов показал, что нормальный закон возникает во всех случаях, когда исследуемая случайная величина формируется как сумма большого числа независимых или слабо зависимых СВ, каждая из которых в отдельности незначительно влияет на суммы (не доминирует).

Рассмотрим простейший случай ЦПТ.

Теорема 5.5. *Если* $X_1, X_2, \ldots, X_n$*– независимые одинаково распределённые СВ с* $\mathbf{M}\Big[X\Big] = m$ *и* $\mathbf{D}\Big[X\Big] = \sigma^2 < \infty$*, то при неограниченном увеличении* n *закон распределения СВ*

$$Y_n = \frac{\sum\limits_{i=1}^{n} X_i - n \cdot m}{\sqrt{n} \cdot \sigma}$$

неограниченно приближается к нормальному.

Доказательство. Заметим, что СВ Y_n– приведена к стандартному виду, т.е. имеет нулевое математическое ожидание и единичную дисперсию. Так как X_i одинаково распределены, то они имеют общую кумулятивную функцию. (5.17.). Кумулятивная функция суммы равна сумме кумулятивных функций:

$$\varphi_{\sum X_i(t)} = n \cdot \varphi_X(t) = n \cdot m \cdot it + \frac{n \cdot \sigma^2 \cdot (it)^2}{2} + n \cdot \varepsilon(t) \cdot t^2$$

Выражение для кумулятивной функции $\varphi_{Y_n}(t)$ получим используя формулу линейного преобразования ХФ (5.18.), положив $a = \frac{1}{\sqrt{n}\sigma}$, $b = -\frac{nm}{\sqrt{n}\sigma}$:

$$\varphi_{Y_n}(t) = -it\frac{nm}{\sqrt{n}\sigma} + it\frac{nm}{\sqrt{n}\sigma} - \frac{n\sigma^2}{2} \cdot \frac{t^2}{n\sigma^2} + n\varepsilon\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot \frac{t^2}{n\sigma^2} \tag{5.21.}$$

$$= -\frac{t^2}{2} + \varepsilon\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot \frac{t^2}{\sigma^2}. \tag{5.22.}$$

[2]Воспользуемся известной формулой $\int\limits_{-\infty}^{+\infty} \exp(-Ax^2 \pm 2Bx - C)dx = \sqrt{\frac{\pi}{A}} \exp\big(-\frac{AC-B^2}{A}\big)$

Величина ε стремится к нулю при $n \to \infty$. Таким образом, $\lim\limits_{n\to\infty} \varphi_{Y_n}(t) \to -\frac{t^2}{2}$.

Мы получили, что предельное значение кумулятивной функции СВ Y_n равно кумулятивной функции стандартного распределения. Значит СВ Y_n имеет плотность распределения

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

Теорема доказана.

Список литературы

1. Бочаров П.П., Печенкин А.В. Теория вероятностей. Математическая статистика. – М.: Гардарика, 1998. – 328 с.

2. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. – Спб.: Издательство «Лань», 2004. – 256 с.

3. Магазинников Л.И. Курс лекций по теории вероятностей. — Томск: Издательство ТГУ, 1989. — 212 с.

4. Гнеденко Б.В. Курс теории вероятностей. — М.: Наука, 1988. — 400 с.

5. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1972.

6. Чудесенко В.Ф. Сборник заданий по спецкурсам высшей математики: Типовые расчеты.– М.,1983.

7. Пестова Н.С., Самойлова М.В. Практические занятия по теории вероятностей.– Томск,1975.

8. Агапов Г.И. Задачник по теории вероятностей. – М.,1994.

9. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. — М.: Наука, 1984. — 245 с.

10. Сборник задач по математике для ВТУЗов. Специальные курсы./ Под редакцией А.В.Ефимова. — М.: Наука, 1984. — 250 с.

11. Сборник задач по теории вероятностей, математической статистике и теории случайных функций. / Под редакцией А.А.Свешникова. — М.: Наука, 1965. — 656 с.

12. Дж.Ламперти. Вероятность.

13. Тутубалин В.Н. Теория вероятностей.

14. Феллер В. Введение в теорию вероятностей и ее приложения// т.1,2.

15. Эмиль Борель. Вероятность и достоверность.

16. Д.Данин. Вероятностный мир.

Содержание

Учебное издание

ГАЛАНОВ Юрий Иванович

КОНСПЕКТ ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

Учебное пособие

Научный редактор
доктор физ.-мат. наук,
профессор К.П. Арефьев

Редактор Н.Я. Горбунова

Верстка в системе *MikTeX* 2.7
с использованием шрифтов
из пакета LHFONTS 3.5. Ю.И. Галанов

Дизайн обложки И.О. Фамилия

Подписано к печати 00.00.2008. Формат 60x84/8.
Бумага «Снегурочка». Печать XEROX.
Усл.печ.л. 000. Уч.-изд.л. 000. Заказ XXX. Тираж XXX экз.

Томский политехнический университет
Система менеджмента качества
Томского политехнического университета сертифицирована
NATIONAL QUALITY ASSURANCE по стандарту ISO
9001:2000

ИЗДАТЕЛЬСТВО ТПУ 634050, г. Томск, пр. Ленина, 30.