

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Кафедра вычислительных методов и программирования

А.И. Волковец, А.Б. Гуринович

***ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА***

Конспект лекций
для студентов всех специальностей и форм обучения БГУИР

Минск 2003

УДК 519.2 (075.8)
ББК 22.171+22.172 я 73
В 67

Волковец А.И.

В 67 Теория вероятностей и математическая статистика: Конспект лекций для студ. всех спец. и форм обучения БГУИР / А.И. Волковец, А.Б. Гуринович. – Мн.: БГУИР, 2003. – 84 с.: ил.

Конспект лекций по курсу «Теория вероятностей и математическая статистика» включает в себя 17 лекций по темам, определенным типовой рабочей программой изучения данной дисциплины. Целью изучения является усвоение основных методов формализованного описания и анализа случайных явлений, обработки и анализа результатов физических и численных экспериментов. Для изучения данной дисциплины студенту необходимы знания, полученные при изучении разделов «Ряды», «Множества и операции над ними», «Дифференциальное и интегральное исчисления» курса высшей математики.

УДК 519.2 (075.8)
ББК 22.171+22.172 я 73

© Волковец А.И., Гуринович А.Б., 2003
© БГУИР, 2003

СОДЕРЖАНИЕ

ЛЕКЦИЯ 1	6
ВВЕДЕНИЕ	6
ОСНОВНЫЕ ПОНЯТИЯ.....	9
АКСИОМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ	9
НЕПОСРЕДСТВЕННЫЙ ПОДСЧЕТ ВЕРОЯТНОСТЕЙ.....	11
ОСНОВНЫЕ КОМБИНАТОРНЫЕ ФОРМУЛЫ.....	11
ЛЕКЦИЯ 2	13
ГЕОМЕТРИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТЕЙ.....	13
ТЕОРЕМЫ СЛОЖЕНИЯ ВЕРОЯТНОСТЕЙ	13
УСЛОВНАЯ ВЕРОЯТНОСТЬ	14
ЗАВИСИМЫЕ И НЕЗАВИСИМЫЕ СОБЫТИЯ	14
ТЕОРЕМЫ УМНОЖЕНИЯ ВЕРОЯТНОСТЕЙ.....	15
ВЕРОЯТНОСТЬ БЕЗОТКАЗНОЙ РАБОТЫ СЕТИ	16
ЛЕКЦИЯ 3	17
ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ.....	17
ФОРМУЛА БАЙЕСА.....	17
ТЕОРЕМА О ПОВТОРЕНИИ ОПЫТОВ	18
ЛЕКЦИЯ 4	21
СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. ЗАКОН РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ.....	21
ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ	21
РЯД РАСПРЕДЕЛЕНИЯ	22
ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ	23
ЛЕКЦИЯ 5	25
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНОЙ ВЕЛИЧИНЫ.....	25
<i>Математическое ожидание</i>	25
<i>Начальный момент</i>	25
<i>Центральный момент</i>	26
<i>Дисперсия</i>	26
<i>Среднее квадратическое отклонение</i>	27
<i>Мода</i>	27
<i>Медиана</i>	28
<i>Квантиль</i>	28
ЛЕКЦИЯ 6	29
ТИПОВЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ	29
<i>Индикатор случайного события</i>	29
<i>Геометрическое распределение</i>	29
<i>Биномиальное распределение</i>	29
<i>Распределение Пуассона</i>	29

<i>Равномерное распределение</i>	30
<i>Экспоненциальное распределение</i>	31
<i>Нормальное распределение</i>	32
ЛЕКЦИЯ 7	34
ФУНКЦИИ ОДНОГО СЛУЧАЙНОГО АРГУМЕНТА	34
ЗАКОН РАСПРЕДЕЛЕНИЯ ФУНКЦИИ СЛУЧАЙНОГО АРГУМЕНТА	34
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ФУНКЦИИ СЛУЧАЙНОГО АРГУМЕНТА	37
ХАРАКТЕРИСТИЧЕСКАЯ ФУНКЦИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ	38
ЛЕКЦИЯ 8	39
ДВУХМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. ДВУХМЕРНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ	39
ДВУХМЕРНАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ	39
МАТРИЦА РАСПРЕДЕЛЕНИЯ.	40
ДВУХМЕРНАЯ ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ	41
ЗАВИСИМЫЕ И НЕЗАВИСИМЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ	42
УСЛОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ	42
ЛЕКЦИЯ 9	44
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ДВУХМЕРНЫХ ВЕЛИЧИН	44
<i>Смешанный начальный момент</i>	44
<i>Смешанный центральный момент</i>	44
<i>Корреляционный момент</i>	44
<i>Коэффициент корреляции</i>	45
УСЛОВНЫЕ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ	46
ЛЕКЦИЯ 10	48
НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ НА ПЛОСКОСТИ	48
ЗАКОН РАСПРЕДЕЛЕНИЯ ФУНКЦИИ ДВУХ СЛУЧАЙНЫХ ВЕЛИЧИН	48
МНОГОМЕРНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ	49
ЛЕКЦИЯ 11	52
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ	52
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СУММЫ СЛУЧАЙНЫХ ВЕЛИЧИН	52
<i>Теорема о математическом ожидании суммы</i>	52
<i>Теорема о дисперсии суммы</i>	53
ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ПРОИЗВЕДЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН	54
<i>Теорема о математическом ожидании произведения</i>	54
<i>Теорема о дисперсии произведения</i>	54
ЛЕКЦИЯ 12	55
ЗАКОН БОЛЬШИХ ЧИСЕЛ	55
<i>Неравенство Чебышева</i>	55
<i>Теорема Чебышева</i>	56
<i>Теорема Бернулли</i>	56
ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА.	57

ЛЕКЦИЯ 13	61
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА. ОСНОВНЫЕ ПОНЯТИЯ	61
ОЦЕНКА ЗАКОНА РАСПРЕДЕЛЕНИЯ	61
<i>Эмпирическая функция распределения</i>	61
<i>Статистический ряд распределения</i>	62
<i>Интервальный статистический ряд</i>	62
<i>Гистограмма</i>	63
ЛЕКЦИЯ 14	65
ТОЧЕЧНЫЕ ОЦЕНКИ ЧИСЛОВЫХ ХАРАКТЕРИСТИК	65
<i>Оценка математического ожидания</i>	65
<i>Оценка дисперсии</i>	66
<i>Оценка вероятности</i>	67
ОЦЕНКА ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ.....	67
<i>Метод моментов</i>	67
<i>Метод максимального правдоподобия</i>	67
ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ЧИСЛОВЫХ ХАРАКТЕРИСТИК.....	68
<i>Доверительный интервал для математического ожидания</i>	69
<i>Доверительный интервал для дисперсии</i>	69
<i>Доверительный интервал для вероятности</i>	70
ЛЕКЦИЯ 15	71
ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ	71
<i>Проверка гипотезы о равенстве вероятностей</i>	71
КРИТЕРИИ СОГЛАСИЯ	72
<i>Критерий согласия Пирсона</i>	72
<i>Критерий согласия Колмогорова</i>	74
ЛЕКЦИЯ 16	76
СТАТИСТИЧЕСКАЯ ОБРАБОТКА ДВУХМЕРНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН	76
<i>Оценка корреляционного момента</i>	76
<i>Оценка коэффициента корреляции</i>	76
<i>Доверительный интервал для коэффициента корреляции</i>	76
СТАТИСТИЧЕСКИЕ КРИТЕРИИ ДВУХМЕРНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН	77
<i>Гипотеза об отсутствии корреляционной зависимости</i>	77
<i>t-критерий</i>	78
<i>F-критерий</i>	78
<i>Критерий Уилкоксона</i>	78
ЛЕКЦИЯ 17	80
ОЦЕНКА РЕГРЕССИОННЫХ ХАРАКТЕРИСТИК	80
МЕТОД НАИМЕНЬШИХ КВАДРАТОВ	80
ЛИТЕРАТУРА	83

ЛЕКЦИЯ 1

Введение

Теория вероятностей – раздел высшей математики, изучающий закономерности массовых случайных явлений.

Приведем несколько примеров случайных явлений.

1. Производится ряд испытаний заводских изделий определенного типа, например реле, на длительность безотказной работы. Результат испытания от одного раза к другому не остается постоянным, меняется. Эти изменения обусловлены влиянием ряда малозначительных, трудноуловимых факторов, таких, например, как микродефекты в металле; разные температурные условия; разные условия хранения и транспортировки изделий; отклонения напряжения от номинала и т.д.

2. Самолет совершает полет на заданной высоте; теоретически он летит горизонтально, равномерно и прямолинейно. Фактически полет сопровождается отклонениями центра массы самолета от теоретической траектории и колебаниями самолета около центра массы. Эти отклонения и колебания являются случайными и связаны с турбулентностью атмосферы; от одного раза к другому они не повторяются.

3. Рассматривается непрерывная работа ЭВМ между двумя очередными сбоями в решении задачи. Все контролируемые условия работы ЭВМ: температура, влажность, напряжение питания, характер решаемой задачи – остаются неизменными. Повторяя такой опыт несколько раз, мы убеждаемся, что время работы ЭВМ между двумя очередными сбоями будет разным (случайным). Это объясняется тем, что различные элементы ЭВМ подвергаются незначительным, неконтролируемым изменениям.

Все приведенные примеры рассмотрены здесь под одним и тем же углом зрения: подчеркнуты случайные вариации, неодинаковые результаты ряда опытов, основные условия которых остаются неизменными. Эти вариации всегда связаны с наличием каких-то второстепенных факторов, влияющих на исход опыта, но не заданных в числе его основных условий. Основные условия опыта, определяющие в общих и грубых чертах его протекание, сохраняются неизменными; второстепенные – меняются от опыта к опыту и вносят случайные различия в их результаты.

Совершенно очевидно, что в природе нет ни одного физического явления, в котором не присутствовали бы в той или иной мере элементы случайности. Как бы точно и подробно ни были фиксированы условия опыта, невозможно достигнуть того, чтобы при повторении опыта результаты полностью и в точности совпадали. Случайные отклонения неизбежно сопутствуют любому закономерному явлению. Тем не менее в ряде практических задач этими случайными элементами можно пренебречь, рассматривая вместо реального явления его упрощенную схему, «модель», и предполагая, что в данных условиях опыта явление протекает вполне определенным образом. При этом из бесчисленного множества факторов, влияющих на данное явление, выделяются самые главные, решающие; влиянием остальных, второстепенных факторов

просто пренебрегают. Такая схема изучения явлений постоянно применяется в физике, механике, технике. При использовании этой схемы для решения любой задачи прежде всего выделяется основной круг учитываемых условий и выясняется, на какие параметры задачи они влияют; затем применяется тот или иной математический аппарат (например, составляются и интегрируются дифференциальные уравнения, описывающие явление); таким образом выявляется основная закономерность, свойственная данному явлению и дающая возможность предсказать результат опыта по его заданным условиям. По мере развития науки число учитываемых факторов становится все больше; явление исследуется подробнее; научный прогноз становится точнее.

Однако для решения ряда вопросов описанная схема – классическая схема так называемых «точных наук» – оказывается плохо приспособленной. Существуют такие задачи, где интересующий нас исход опыта зависит от столь большого числа факторов, что практически невозможно зарегистрировать и учесть все эти факторы. Это задачи, в которых многочисленные второстепенные, тесно переплетающиеся между собой случайные факторы играют заметную роль, а вместе с тем число их так велико и влияние столь сложно, что применение классических методов исследования себя не оправдывает.

Рассмотрим следующий пример. Некоторое техническое устройство, например система автоматического управления, решает определенную задачу в условиях, когда на систему непрерывно воздействуют случайные помехи. Наличие помех приводит к тому, что система решает задачу с некоторой ошибкой, в ряде случаев выходящей за пределы допустимой. Возникают вопросы: как часто будут появляться такие ошибки? какие следует принять меры для того, чтобы практически исключить их возможность?

Чтобы ответить на такие вопросы, необходимо исследовать природу и структуру случайных возмущений, воздействующих на систему, изучить реакцию системы на такие возмущения, выяснить влияние конструктивных параметров системы на вид этой реакции. Все подобные задачи, число которых в физике и технике чрезвычайно велико, требуют изучения не только основных, главных закономерностей, определяющих явление в общих чертах, но и анализа случайных возмущений и искажений, связанных с наличием второстепенных факторов и придающих исходу опыта при заданных условиях элемент неопределенности.

Какие же существуют пути и методы для исследования случайных явлений? С чисто теоретической точки зрения те факторы, которые мы условно назвали «случайными», в принципе ничем не отличаются от других, которые мы выделили в качестве «основных». Теоретически можно неограниченно повышать точность решения каждой задачи, учитывая все новые и новые группы факторов: от самых существенных до самых ничтожных. Однако практически такая попытка одинаково подробно и тщательно проанализировать влияние решительно всех факторов, от которых зависит явление, привела бы только к тому, что решение задачи, в силу непомерной громоздкости и сложности, оказалось бы практически неосуществимым и к тому же не имело

бы никакой познавательной ценности. Очевидно, должна существовать принципиальная разница в методах учета основных, решающих факторов, определяющих в главных чертах течение явления, и вторичных, второстепенных факторов, влияющих на течение явления в качестве «погрешностей» или «возмущений». Элемент неопределенности, сложности, многопричинности, присущий случайным явлениям, требует создания специальных методов для изучения этих явлений.

Такие методы и разрабатываются в теории вероятностей. Ее предметом являются специфические закономерности, наблюдаемые в случайных явлениях. Практика показывает, что, наблюдая в совокупности массы однородных случайных явлений, мы обычно обнаруживаем в них вполне определенные закономерности, своего рода устойчивости, свойственные именно массовым случайным явлениям. Например, если много раз подряд бросать монету, частота появления герба (отношение числа появившихся гербов к общему числу бросаний) постепенно стабилизируется, приближаясь к вполне определенному числу, а именно к $1/2$. Такое же свойство «устойчивости частот» обнаруживается и при многократном повторении любого другого опыта, исход которого представляется заранее не определенным, случайным. Отметим, что именно массовость случайных явлений обеспечивает выполнение этой закономерности.

Подобного рода закономерности (их называют «статистическими») возникают, когда мы наблюдаем в совокупности массивы однородных случайных явлений. Они оказываются практически независимыми от индивидуальных особенностей отдельных случайных явлений, входящих в массив: эти особенности как бы взаимно погашаются, нивелируются; выражаясь образно, «из множества беспорядков возникает порядок». Средний массовый результат множества случайных явлений оказывается практически уже не случайным, предсказуемым. Это и является базой для практического применения вероятностных (статистических) методов исследования.

Методы теории вероятностей не отменяют и не упраздняют случайности, непредсказуемости исхода отдельного опыта, дают возможность предсказать, с каким-то приближением, средний суммарный результат массы однородных случайных явлений.

Цель вероятностных (статистических) методов – в том, чтобы, минуя слишком сложное (и зачастую практически невозможное) исследование отдельного случайного явления, обратиться непосредственно к законам, управляющим массами таких явлений. Изучение этих законов позволяет не только осуществлять прогноз в области случайных явлений, но и целенаправленно влиять на ход этих явлений, контролировать их, ограничивать сферу действия случайности, сужать ее влияние на практику.

В настоящее время нет практически ни одной области науки, в которой в той или иной степени не применялись бы вероятностные методы. В одних науках, в силу специфики предмета и исторических условий, эти методы находят применение раньше, в других – позднее. Исторически первые зачатки вероятностных методов с довольно примитивным математическим аппаратом

возникли в XVII в. при разработке теории азартных игр с целью дать рекомендации игрокам. Затем эти методы стали применяться в практике страховых компаний для установления разумных размеров страховых премий. Постепенно область применения вероятностных методов расширялась. Сегодня эти методы распространяются все шире и шире. Целые разделы современной физики (в частности, ядерная физика) базируются на математическом аппарате теории вероятностей. Широко применяются вероятностные методы в современных электронике, радиотехнике, теории связи, теории автоматического регулирования, кибернетике, вычислительной технике, теории автоматизированных систем управления. Это и естественно, так как работа современных радиотехнических, электронных систем протекает в условиях случайных воздействий, без учета которых невозможны разумное проектирование подобных систем, выбор их конструктивных параметров. Любая процедура управления чем бы то ни было (техническим устройством, группой устройств, человеко-машинным комплексом) протекает в заранее не известных, случайных условиях, неизбежно сопровождается случайными ошибками измерения тех или других параметров, ошибками выполнения команд и т. д.; анализ работы такой системы практически невозможен без учета случайных факторов.

Знакомство с методами теории вероятностей и математической статистики необходимо сегодня каждому грамотному инженеру. И не только инженеру. Биология, физиология, медицина, социология все шире применяют вероятностные методы. Не чуждаются их и такие «исконно гуманитарные» науки, как психология, лингвистика, литературоведение, даже эстетика.

Основные понятия

Случайное явление – это явление, которое при неоднократности воспроизведения одного и того же опыта протекает каждый раз по-иному, непредсказуемым образом.

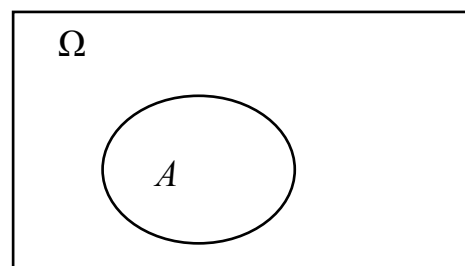
Опыт – воспроизводимая совокупность условий, в которых фиксируется тот или иной результат.

Случайное событие – всякий факт, который в опыте со случайным исходом может произойти или не произойти. Обозначение: A, B, C, \dots

Вероятность случайного события – количественная мера объективной возможности его осуществления.

Аксиомы теории вероятностей

Рассмотрим некоторый опыт. Каждый исход опыта обозначим элементарным событием $w_i, i = 1, 2, \dots, n$, где n – число исходов данного опыта. Множество всех возможных исходов опыта образуют $\Omega = \{w_1, w_2, \dots, w_n\}$ – универсальное множество опыта или пространство элементарных событий.



Тогда любое случайное событие A , возможное в данном опыте, есть некоторое подмножество универсального множества $A \in \Omega$:

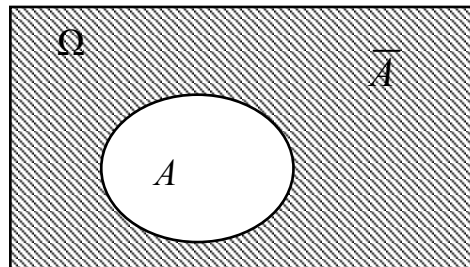
$$A = \{w_1, w_2, \dots, w_m\}, 0 \leq m \leq n,$$

где m – число исходов, благоприятных событию A .

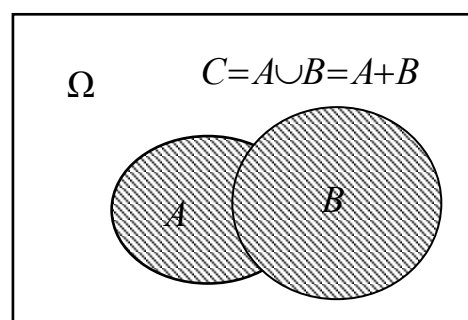
Событие A называется **достоверным**, если $A = \Omega$, т.е. происходит в каждом опыте.

Событие A называется **невозможным**, если $A = \emptyset$, т.е. никогда не происходит в данном опыте.

Противоположным к событию A называют событие \bar{A} , состоящее в невыполнении A , т.е. оно происходит всегда, когда не происходит A .



Событие C называется **суммой событий** A и B , $C = A \cup B = A + B$, если оно происходит тогда, когда происходит либо A , либо B , либо оба одновременно (хотя бы одно событие).



Событие C называется **произведением событий** A и B , $C = A \cap B = A \cdot B$, если C происходит тогда, когда происходят и A и B одновременно.

События A и B **несовместны**, если они не могут произойти одновременно, т.е. $A \cdot B = \emptyset$.

События A_i ($i = 1, 2, \dots, n$) образуют **полную группу**, если они попарно несовместны и в сумме образуют достоверное событие

$$\sum_{i=1}^n A_i = \Omega.$$

При преобразовании выражений можно пользоваться следующими тождествами:

$$A + \bar{A} = \Omega; \quad A \cdot \bar{A} = \emptyset; \quad A + \Omega = \Omega; \quad A \cdot \Omega = A; \quad A \cdot \emptyset = \emptyset;$$

$$A + \emptyset = A; \quad \overline{A + B} = \bar{A} \cdot \bar{B}; \quad \overline{A \cdot B} = \bar{A} + \bar{B}; \quad A + \bar{A} \cdot B = A + B.$$

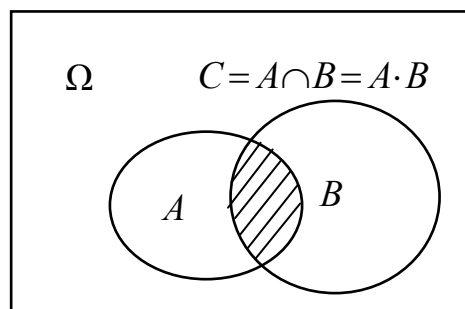
Аксиома 1. Вероятность $p(A)$ случайного события A есть функция множества элементарных исходов, благоприятных событию A , и вероятность любого события принимает значения:

$$0 \leq p(A) \leq 1, \quad (1.1)$$

причем $p(\emptyset) = 0, p(\Omega) = 1$.

Аксиома 2. Вероятность суммы несовместных случайных событий равна сумме вероятностей этих событий:

$$p\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n p(A_i), \quad A_i \cdot A_j = \emptyset, \quad \forall i \neq j. \quad (1.2)$$



Следствие аксиом 1 и 2:

$$1 = p(\Omega) = p(A + \bar{A}) = p(A) + p(\bar{A}) \Rightarrow p(\bar{A}) = 1 - p(A).$$

Непосредственный подсчет вероятностей

События $A_1 \dots A_n$ называются случаями, если они обладают следующими свойствами:

- события $A_1 \dots A_n$ несовместны, $A_i \cdot A_j = \emptyset, \forall i \neq j$;
- события $A_1 \dots A_n$ образуют полную группу, $\sum_{i=1}^n A_i = \Omega$;
- события $A_1 \dots A_n$ равновозможны, $p(A_i) = p, \forall i$.

Пусть некоторый опыт сводится к схеме случаев, тогда вероятность события A в этом опыте равна отношению числа благоприятных случаев к общему числу случаев:

$$p(A) = \frac{m}{n}, \quad (1.3)$$

где m – число случаев A_i , благоприятных событию A , т.е. входящих в множество $A = \{A_1 \dots A_m\}$;

n – число всех возможных случаев.

Доказательство. Очевидно, что $A = A_1 + A_2 + \dots + A_m$. Так как A_i несовместимы, то определим вероятность события A по второй аксиоме:

$$p(A) = p\left(\sum_{i=1}^m A_i\right) = \sum_{i=1}^m p(A_i) = m \cdot p,$$

$$p(\Omega) = p\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n p(A_i) = n \cdot p = 1 \Rightarrow p = \frac{1}{n} \Rightarrow p(A) = \frac{m}{n}.$$

Формула (1.3) называется **классическим определением вероятности** и использовалась как определение вероятности с XVII по XIX в. При определении значений m, n в (1.3) могут оказаться полезными следующие формулы из комбинаторики.

Основные комбинаторные формулы

Пусть имеется множество $X = \{x_1, x_2, \dots, x_n\}$, состоящее из n различных элементов. (n, r) -**выборкой** называется множество, состоящее из r элементов, взятых из множества X .

Упорядоченной называется выборка, для которой важен порядок следования элементов. Если каждый элемент множества X может извлекаться несколько раз, то выборка называется **выборкой с повторениями**.

Число упорядоченных (n, r) -выборок (**размещений**) с повторениями $\hat{A}(n, r)$ и без повторений $A(n, r)$ равно

$$\hat{A}_n^r = n^r, \quad (1.4)$$

$$A_n^r = \frac{n!}{(n-r)!}. \quad (1.5)$$

Если $r = n$, то размещения без повторений называются **перестановками**, т.е. это расположение элементов исходного множества в определенном порядке. Число перестановок из n элементов равно

$$P_n = n! = 1 \cdot 2 \cdot \dots \cdot n. \quad (1.6)$$

Пустое множество можно упорядочить только одним способом: $P_0 = 0! = 1$.

Число неупорядоченных (n, r) -выборок (**сочетаний**) с повторениями \hat{C}_n^r и без повторений C_n^r равно

$$\hat{C}_n^r = \frac{(n+r-1)!}{r!(n-1)!}, \quad (1.7)$$

$$C_n^r = \frac{A_n^r}{P_r} = \frac{n!}{r!(n-r)!}. \quad (1.8)$$

Число различных разбиений множества из n элементов на k непересекающихся подмножеств, причем в 1-м подмножестве r_1 элементов, во 2-м r_2 элементов и т.д., а $n = r_1 + r_2 + \dots + r_k$, равно

$$P_n(r_1, r_2, \dots, r_k) = \frac{n!}{r_1! r_2! \dots r_k!}. \quad (1.9)$$

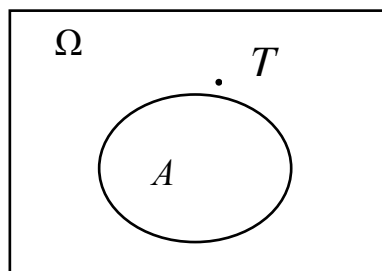
ЛЕКЦИЯ 2

Геометрическое определение вероятностей

Классическое определение вероятности предполагает, что число элементарных исходов конечно. На практике встречаются опыты, для которых множество таких исходов бесконечно. Чтобы преодолеть недостаток классического определения вероятности, состоящий в том, что оно неприменимо к испытаниям с бесконечным числом исходов, вводят геометрические вероятности – вероятности попадания точки в область.

Пусть в некоторую область случайным образом бросается точка T , причем все точки области Ω равноправны в отношении попадания точки T . Тогда за вероятность попадания точки T в область A принимается отношение

$$p(A) = \frac{S(A)}{S(\Omega)}, \quad (2.1)$$



где $S(A)$ и $S(\Omega)$ – геометрические меры (длина, площадь, объем и т.д.) областей A и Ω соответственно.

Теоремы сложения вероятностей

Теорема сложения двух случайных событий. Вероятность суммы случайных событий A и B равна сумме вероятностей этих событий минус вероятность их совместного появления:

$$p(A + B) = p(A) + p(B) - p(AB). \quad (2.2)$$

Доказательство. Представим событие $A + B$ в виде суммы трех несовместимых событий

$$A + B = A \cdot \bar{B} + AB + \bar{A} \cdot B.$$

Тогда на основании второй аксиомы

$$p(A + B) = p(A \cdot \bar{B}) + p(AB) + p(\bar{A} \cdot B).$$

Представим события A и B в виде суммы несовместимых событий:

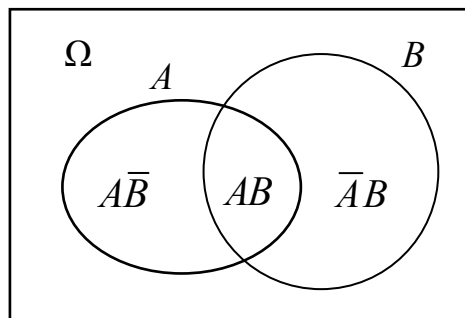
$$A = A \cdot \bar{B} + AB, \quad p(A) = p(A \cdot \bar{B}) + p(AB) \Rightarrow p(A \cdot \bar{B}) = p(A) - p(AB),$$

$$B = \bar{A} \cdot B + AB, \quad p(B) = p(\bar{A} \cdot B) + p(AB) \Rightarrow p(\bar{A} \cdot B) = p(B) - p(AB),$$

Подставим $p(A \cdot \bar{B})$ и $p(\bar{A} \cdot B)$ в выражение $p(A + B)$ и после преобразований получим: $p(A + B) = p(A) + p(B) - p(AB)$.

Теорема сложения для n случайных событий. Вероятность суммы n событий A_1, \dots, A_n равна

$$\begin{aligned} p\left(\sum_{i=1}^n A_i\right) &= \sum_{i_1=1}^n p(A_{i_1}) - \sum_{i_1, i_2} p(A_{i_1} A_{i_2}) + \dots \\ &\vdots \\ &\dots + (-1)^{k+1} \sum_{i_1, i_2, \dots, i_k} p(A_{i_1} A_{i_2} \dots A_{i_k}) + \dots + (-1)^{n+1} p(A_1 A_2 \dots A_n), \end{aligned} \quad (2.3)$$



где число слагаемых в k -й сумме равно C_n^k , т.е. перебираются все возможные сочетания из k слагаемых.

Доказательство. Используем метод математической индукции. Однако для экономии времени и места докажем переход от m слагаемых к $m+1$ для случая $m = 2$. Докажем, что

$$p(A_1 + A_2 + A_3) = p(A_1) + p(A_2) + p(A_3) + p(A_1 A_2) + p(A_1 A_3) + p(A_2 A_3) - p(A_1 A_2 A_3),$$

если $p(A_1 + A_2) = p(A_1) + p(A_2) - p(A_1 A_2)$.

Обозначим $B = A_2 + A_3$,

$$\begin{aligned} p(A_1 + A_2 + A_3) &= p(A_1 + B) = p(A_1) + p(B) - p(A_1 B) = p(A_1) + p(A_2 + A_3) - p(A_1 A_2 + A_1 A_3) = \\ &= p(A_1) + p(A_2) + p(A_3) - p(A_2 A_3) - p(A_1 A_2) - p(A_1 A_3) + p(A_1 A_2 A_3), \end{aligned}$$

что и требовалось доказать.

На практике, с учетом того что $p(A) = 1 - p(\bar{A})$, вероятность суммы n событий (если $n > 2$) удобнее вычислять по формуле

$$p(A_1 + A_2 + \dots + A_n) = 1 - p(\overline{A_1 + A_2 + \dots + A_n}) = 1 - p(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_n). \quad (2.4)$$

Условная вероятность

Ранее случайное событие определялось как событие, которое при осуществлении совокупности условий (опыта) может произойти или не произойти. Если при вычислении вероятности события никаких других ограничений, кроме этих условий, не налагается, то такую вероятность называют **безусловной**. Если же налагаются и другие дополнительные условия, то вероятность события называется **условной**.

Проводится опыт со случайным исходом, в результате которого возможны два события A и B . **Условной вероятностью** $p(B/A)$ называется вероятность события B , вычисленная при условии (в предположении), что событие A произошло.

Зависимые и независимые события

Событие A называется независимым от события B , если его вероятность не зависит от того, произошло B или нет, т.е. критерий независимости

$$p(A) = p(A/B) = p(A/\bar{B}). \quad (2.5)$$

В противном случае, т.е. когда критерий не выполняется, событие A зависит от события B .

Зависимость и независимость всегда взаимны, т.е. если событие A не зависит от события B (см. (2.5)), то и событие B не зависит от события A :

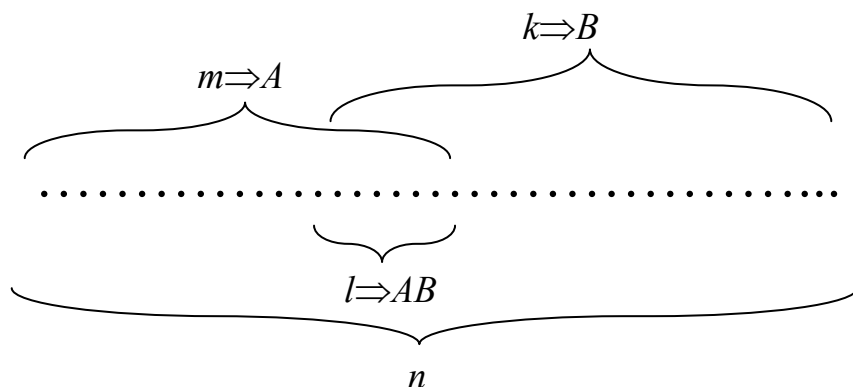
$$p(B) = p(B/A) = p(B/\bar{A}). \quad (2.6)$$

Теоремы умножения вероятностей

Теорема умножения вероятностей для двух событий. Вероятность произведения двух событий равна вероятности одного из них, умноженной на условную вероятность второго при наличии первого:

$$p(AB) = p(A)p(B/A) = p(B)p(A/B). \quad (2.7)$$

Доказательство. Докажем (2.7) для схемы случаев. Пусть в опыте возможны n несовместимых и равновозможных исходов. Событию A соответствует m исходов, событию B – k исходов. В l исходах события A и B происходят одновременно.



Очевидно, что $p(A) = \frac{m}{n}$, $p(B) = \frac{k}{n}$, $p(AB) = \frac{l}{n}$ (см. (1.1)). Вычислим условную вероятность $p(B/A)$, т.е. вероятность события B в предположении, что A произошло. Если известно, что событие A произошло, то из ранее возможных n случаев остаются возможными только те m , которые благоприятствовали событию A . Из них l благоприятны событию $B \Rightarrow p(B/A) = \frac{l}{m}$. Аналогично вычислим условную вероятность $p(A/B)$, т.е. вероятность события A в предположении, что B произошло: $p(A/B) = \frac{l}{k}$. Подставим найденные вероятности в (2.7): $\frac{l}{n} = \frac{m}{n} \cdot \frac{l}{m} = \frac{k}{n} \cdot \frac{l}{k}$, что и требовалось доказать. Очевидно, что безразлично, какое из событий считать первым, а какое вторым.

Теорема умножения вероятностей для n событий. Вероятность произведения n событий $A_1 \dots A_n$ равна

$$p(A_1 \cdot A_2 \cdot \dots \cdot A_n) = p(A_1) \cdot p(A_2/A_1) \cdot p(A_3/A_1 \cdot A_2) \cdot \dots \cdot p(A_n/A_1 \cdot A_2 \cdot \dots \cdot A_{n-1}), \quad (2.8)$$

где $p(A_k/A_1 \cdot \dots \cdot A_{k-1})$ – вероятность появления события A_k , при условии что события A_1, A_2, \dots, A_{k-1} в данном опыте произошли.

Доказательство. Используем метод математической индукции. Однако для экономии времени и места докажем переход от m сомножителей к $m+1$ для случая $m = 2$. Докажем, что $p(A_1 A_2 A_3) = p(A_1) p(A_2/A_1) p(A_3/A_1 A_2)$,

если $p(A_1 A_2) = p(A_1) p(A_2 / A_1)$. Обозначим $B = A_1 A_2$, тогда

$$p(A_1 A_2 A_3) = p(B A_3) = p(B) p(A_3 / B) = p(A_1 A_2) p(A_3 / A_1 A_2) = p(A_1) p(A_2 / A_1) p(A_3 / A_1 A_2),$$

что и требовалось доказать.

Если события $A_1 \dots A_n$ независимы, то вероятность произведения равна произведению вероятностей этих событий:

$$p(A_1 \cdot A_2 \cdot \dots \cdot A_n) = p(A_1) \cdot p(A_2) \cdot \dots \cdot p(A_n), \quad (2.9)$$

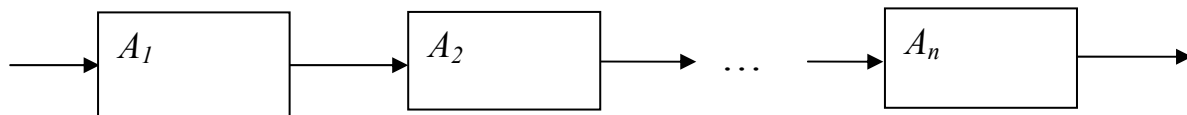
а вероятность $p(A_1 + A_2 + \dots + A_n)$ появления хотя бы одного события A_1, A_2, \dots, A_n равна (см. (2. 4))

$$p(A_1 + A_2 + \dots + A_n) = 1 - p(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_n) = 1 - p(\bar{A}_1) p(\bar{A}_2) \dots p(\bar{A}_n). \quad (2.10)$$

Вероятность безотказной работы сети

Событие B – безотказная работа сети, состоящей из n независимо работающих элементов A_i . Надежность $p(A_i) = p_i$ (вероятность безотказной работы) каждого элемента известна. Необходимо определить вероятность безотказной работы сети в целом.

Рассмотрим последовательное соединение элементов:



Очевидно, что $B = A_1 \cdot A_2 \cdot \dots \cdot A_n$, а с учетом (2.9)

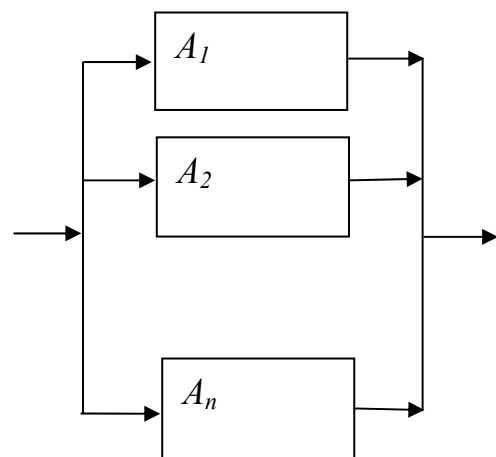
$$p(B) = p(A_1) \cdot p(A_2) \cdot \dots \cdot p(A_n) = p_1 \cdot p_2 \cdot \dots \cdot p_n. \quad (2.11)$$

Для параллельного соединения $B = A_1 + A_2 + \dots + A_n$, а с учетом (2.10)

$$p(B) = 1 - p(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_n) = 1 - q_1 \cdot q_2 \cdot \dots \cdot q_n, \quad (2.12)$$

где $q_i = 1 - p_i$.

Сети с любой другой схемой соединения всегда можно представить в виде участков либо с последовательным, либо с параллельным соединением и вероятность безотказной работы сети определить, последовательно применяя формулы (2.11) и (2.12).



ЛЕКЦИЯ 3

Формула полной вероятности

Следствием обеих теорем вероятности: теоремы сложения и теоремы умножения – является формула полной вероятности.

Пусть проводится опыт, об условиях которого можно сделать n исключающих друг друга предположений (гипотез), образующих полную группу:

$$H_1, H_2, \dots, H_n, \quad H_i H_j = \emptyset, \quad i \neq j, \quad \sum_{i=1}^n H_i = \Omega.$$

Каждая из гипотез осуществляется случайным образом и представляет собой случайное событие. Вероятности гипотез известны и равны:

$$p(H_1), p(H_2), \dots, p(H_n), \quad \sum_{i=1}^n p(H_i) = 1.$$

Рассмотрим некоторое событие A , которое может появиться только вместе с одной из гипотез. Известны условные вероятности события A для каждой из гипотез:

$$P(A/H_1), p(A/H_2), \dots, p(A/H_n).$$

Требуется определить полную (безусловную) $p(A)$ вероятности события A . Представим событие A как сумму из n несовместимых вариантов:

$$A = A \cdot \Omega = A(H_1 + H_2 + \dots + H_n) = A \cdot H_1 + A/H_2 + \dots + A/H_n.$$

На основании второй аксиомы

$$p(A) = \sum_{i=1}^n p(H_i A).$$

С учетом теоремы умножения вероятностей $p(H_i A) = p(H_i)p(A/H_i)$, тогда

$$p(A) = \sum_{i=1}^n p(H_i) \cdot p(A/H_i). \quad (3.1)$$

Формула Байеса

Базируется на формуле полной вероятности и теореме умножения вероятностей.

Пусть до проведения некоторого опыта об его условиях n можно сделать n исключающих друг друга предположений (гипотез), образующих полную группу:

$$H_1, H_2, \dots, H_n, \quad H_i H_j = \emptyset, \quad i \neq j, \quad \sum_{i=1}^n H_i = \Omega.$$

Вероятности гипотез до опыта (априорные вероятности) известны:

$$p(H_1), p(H_2), \dots, p(H_n).$$

Опыт произведен, и произошло некоторое событие A . Требуется определить вероятности гипотез с учетом того, что произошло событие A , т.е. определить апостериорные вероятности: $p(H_1/A), p(H_2/A), \dots, p(H_n/A)$.

Вероятность того, что событие A произошло совместно с H_i , на основании теоремы умножения вероятностей равна $p(H_i A) = p(H_i)p(A/H_i) = p(A)p(H_i/A)$. Отбросим левую часть равенства и выразим $p(H_i/A)$:

$$p(H_i / A) = \frac{p(H_i)p(A / H_i)}{p(A)}.$$

Раскроем $p(A)$ по формуле полной вероятности (3.1) и получим формулу Байеса

$$p(H_i / A) = \frac{p(H_i)p(A / H_i)}{\sum_{j=1}^n p(H_j)p(A / H_j)} \quad (3.2)$$

Формула Байеса позволяет пересчитать априорные вероятности гипотез с учетом того, что опыт завершился событием A .

Теорема о повторении опытов

Пусть проводятся n независимых **одинаковых** опытов, в каждом из которых событие A появляется с вероятностью p . Вероятность $P(n, k)$ того, что событие A произойдет ровно в k опытах, равна (**формула Бернулли**):

$$P(n, k) = C_n^k \cdot p^k \cdot q^{n-k} = \frac{n!}{k!(n-k)!} p^k \cdot q^{n-k}, 0 \leq k \leq n, \quad (3.3)$$

где $q = 1 - p$ – вероятность того, что A не появится в одном опыте.

Доказательство. Обозначим через B_k появление события A в k опытах и появление \bar{A} в $n - k$ опытах. Событие B_k представляет собой сумму несовместимых событий:

$$B_k = \underbrace{A_1 A_2 \dots A_k}_{k} \bar{A}_{k+1} \bar{A}_{k+2} \dots \bar{A}_n + \dots + \underbrace{\bar{A}_1 \bar{A}_2 \dots \bar{A}_{n-k}}_{n-k} A_{n-k+1} A_{n-k+2} \dots A_n$$

где A_i \bar{A}_i – появление и не появление событие A в i -м опыте.

Определим вероятность одного из слагаемых. Так как все опыты одинаковы, то вероятности всех вариантов одинаковы и равны

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_k \bar{A}_{k+1} \cdot \bar{A}_{k+2} \cdot \dots \cdot \bar{A}_n) = \underbrace{p \cdot p \cdot \dots \cdot p}_k \cdot \underbrace{q \cdot q \cdot \dots \cdot q}_{n-k} = p^k \cdot q^{n-k}.$$

Количество вариантов таких сложных событий равно числу выборов k номеров опытов из n возможных, в которых произойдут события A , т.е. равно числу сочетаний без повторения элементов C_n^k . Так как эти события несовместимы, то на основании второй аксиомы: $p(B_k) = P(n, k) = C_n^k \cdot p^k \cdot q^{n-k}$.

Свойства формулы Бернулли:

1. Правая часть формулы (3.3) представляет собой общий член разложения бинома Ньютона:

$$(q + p)^n = \sum_{k=0}^n P(n, k) = \sum_{k=0}^n C_n^k p^k q^{n-k} = 1. \quad (3.4)$$

2. Рекуррентная формула $P(n, k)$ имеет вид

$$P(n, k+1) = \frac{n-k}{k+1} \cdot \frac{p}{q} P(n, k). \quad (3.5)$$

3. Число k_0 , которому соответствует максимальная вероятность $P(n, k_0)$, называется *наивероятнейшим числом* появления события A и определяется неравенствами

$$np - q \leq k_0 \leq np + p. \quad (3.6)$$

Доказательство:

$$\frac{P_n(k+1)}{P_n(k)} = \frac{n-k}{k+1} \cdot \frac{p}{q} \Rightarrow P_n(k+1) \geq P_n(k) \Leftrightarrow \frac{n-k}{k+1} \cdot \frac{p}{q} \geq 1 \Leftrightarrow k \leq np - q,$$

$$\text{а } P_n(k+1) \leq P_n(k) \Leftrightarrow \frac{n-k}{k+1} \cdot \frac{p}{q} \leq 1 \Leftrightarrow k \geq np - q.$$

Итак, при $k < np - q$ функция $P(n, k)$ возрастает, а при $k > np - q$ — убывает. Тогда существует точка k_0 , в которой $P(n, k)$ достигает максимума, т.е.

$$\begin{cases} P(n, k_0) \geq P(n, k_0 - 1) \\ P(n, k_0) \leq P(n, k_0 + 1) \end{cases}.$$

Решив данную систему неравенств относительно k_0 , получим (3.6).

4. Вероятность $P_n(k_1 \leq k \leq k_2)$ того, что в n опытах схемы Бернулли событие A появится от k_1 до k_2 раз ($0 \leq k_1 \leq k_2 \leq n$), равна:

$$P(n, k_1 \leq k \leq k_2) = \sum_{k=k_1}^{k_2} P(n, k) = \sum_{k=k_1}^{k_2} C_n^k p^k q^{n-k}. \quad (3.7)$$

5 Вероятность $P(n, 1 \leq k \leq n)$ того, что в n опытах событие A появится хотя бы один раз, равна

$$P(n, 1 \leq k \leq n) = 1 - P(n, 0) = 1 - q^n. \quad (3.8)$$

Пусть производится n независимых опытов, каждый из которых имеет r ($r > 2$) попарно несовместных и единственно возможных исходов A_1, A_2, \dots, A_r с вероятностями $p_1 = p(A_1), p_2 = p(A_2), \dots, p_r = p(A_r)$.

Требуется определить вероятность того, что из серии n независимых опытов исход A_1 наступит k_1 раз, $A_2 - k_2, \dots, A_r - k_r$ ($k_1 + k_2 + \dots + k_r = n$), то

$$P(n, k_1, k_2, \dots, k_r) = \frac{(k_1 + k_2 + \dots + k_r)!}{k_1! k_2! \dots k_r!} \cdot p_1^{k_1} \cdot p_2^{k_2} \dots p_r^{k_r}. \quad (3.9)$$

Вычисление вероятностей $P(n, k)$ при больших значениях n по формуле Бернулли проблематично. Поэтому вычисление соответствующих вероятностей проводится с помощью следующих приближенных формул.

Если количество испытаний велико $n \rightarrow \infty$, а вероятность события мала $p \rightarrow 0$, так что $np \rightarrow a, 0 < a < \infty$ и $p \ll \frac{1}{\sqrt{n}}$, то используется **формула Пуассона**:

$$P(n, k) \approx \frac{a^k}{k!} \cdot e^{-a}, k = \overline{0, n}. \quad (3.10)$$

Если количество испытаний n велико, вероятности p и q не малы, так что выполняются следующие условия:

$$0 < np - 3\sqrt{npq}, np + 3\sqrt{npq} < n,$$

то применяются приближенные **формулы Муавра–Лапласа**:

$$\text{– локальная} \quad P(n, k) \approx \frac{\varphi(x)}{\sqrt{npq}}, \quad (3.11)$$

$$\text{где } \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x = \frac{k - np}{\sqrt{npq}};$$

$$\text{– интегральная} \quad P(n, k_1 \leq k \leq k_2) \approx \Phi(x_2) - \Phi(x_1), \quad (3.12)$$

$$\text{где } x_1 = \frac{(k_1 - np)}{\sqrt{npq}}, \quad x_2 = \frac{(k_2 - np)}{\sqrt{npq}},$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{x^2}{2}\right) dx \text{ – функция Лапласа.}$$

Функции $\varphi(x)$ и $\Phi(x)$ табулированы. При использовании таблиц следует помнить, что $\varphi(x)$ является четной ($\varphi(-x) = \varphi(x)$), а функция Лапласа – нечетной ($\Phi(-x) = -\Phi(x)$). Доказательство формул (3.11) и (3.12) будет приведено в лекции 12 (**Центральная предельная теорема**).

ЛЕКЦИЯ 4

Случайные величины. Закон распределения вероятностей

Под **случайной величиной** (СВ) понимается величина, которая в результате опыта со случайным исходом принимает то или иное значение, причем заранее, до опыта, неизвестно, какое именно. Случайные величины будем обозначать большими буквами: X, Y, Z ; их значения – соответствующими малыми буквами: x, y, z , а Ω_X – множество возможных значений величины X .

Примеры случайных величин:

1. Опыт – бросок одной игральной кости; случайные величины X – число выпавших очков; $\Omega_X = \{0, 1, 2, 3, 4, 5, 6\}$.

2. Опыт – работа ЭВМ до первого отказа; случайные величины X – время наработки на отказ; $\Omega_X = (0, \infty]$.

В зависимости от вида множества Ω_X случайные величины могут быть дискретными и непрерывными.

Случайная величина (СВ) X называется **дискретной**, если множество Ω_X – счетное, т.е. его элементы можно расположить в определенном порядке и пронумеровать.

Случайная величина X называется **непрерывной** (недискретной), если множество Ω_X – несчетное.

Законом распределения случайной величины X называется любая функция (правило, таблица и т.п.), устанавливающая соответствие между значениями случайной величины и вероятностями их наступления и позволяющая находить вероятности всевозможных событий $p\{a \leq X < b\}$, $\forall a, b$, связанных со случайной величиной.

Функция распределения

Функцией распределения $F(x)$ случайной величины X называется вероятность того, что она примет значение меньшее, чем аргумент функции x :

$$F(x) = p\{X < x\}. \quad (4.1)$$

Свойства функции распределения:

1. $F(-\infty) = 0$.

2. $F(+\infty) = 1$.

3. $F(x_1) \leq F(x_2)$, при $x_1 < x_2$.

Доказательство.

$$A = \{X < x_1\}, B = \{x_1 \leq X < x_2\}, C = \{X < x_2\},$$

тогда

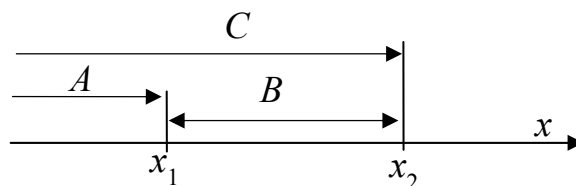
$$C = A + B, p(C) = p(A) + p(B), p(C) = F(x_2), p(A) = F(x_1),$$

$$F(x_2) = F(x_1) + p(B), p(B) \geq 0 \Rightarrow F(x_1) \leq F(x_2).$$

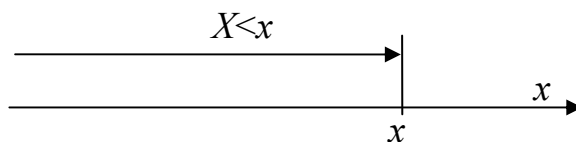
4. $p(x_1 \leq X < x_2) = F(x_1) - F(x_2)$. (4.2)

Доказательство.

$$p(x_1 \leq X < x_2) = p(B) = p(C) - p(A) = F(x_1) - F(x_2).$$



Проиллюстрируем эти свойства с помощью наглядной геометрической интерпретации. Для этого рассмотрим случайную величину как случайную точку X на оси OX , которая в результате опыта может занять то или иное положение. Тогда функция распределения $F(x)$ есть вероятность того, что случайная точка X в результате опыта попадет левее точки x . Увеличиваем x , перемещая точку вправо по оси абсцисс очевидно, что при этом вероятность выполнения неравенства $X < x$ убывать не может (свойство 3). При уменьшении x до $-\infty$ событие $X < x$ становится невозможным, т.е. $F(-\infty) = 0$ (свойство 1), при увеличении x до $+\infty$ – достоверным, т.е. $F(+\infty) = 1$ (свойство 2).



Функция распределения используется при рассмотрении как дискретных, так и непрерывных случайных величин.

Ряд распределения

Для описания дискретных случайных величин наряду с функцией распределения $F(x)$ используется ряд распределения вероятностей.

Рядом распределения дискретной СВ X называется таблица, в верхней строке которой перечислены все возможные значения СВ x_1, x_2, \dots, x_n ($x_{i-1} < x_i$), а в нижней – вероятности их появления p_1, p_2, \dots, p_n , где $p_i = p\{X = x_i\}$.

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

Так как события $\{X = x_1\}, \dots, \{X = x_n\}$ несовместны и образуют полную группу, то справедливо контрольное соотношение

$$p_1 + p_2 + \dots + p_n = 1. \quad (4.3)$$

Многоугольник вероятностей есть графическое изображение ряда распределения вероятностей. По оси абсцисс откладываются возможные значения случайной величины, а по оси ординат – вероятности этих значений. Для наглядности полученные точки соединяются отрезками прямых. Многоугольник распределения, так же как и ряд распределения, полностью характеризует случайную величину и является одной из форм закона распределения.

Функция распределения любой дискретной случайной величины есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины, и равны вероятностям этих значений:

$$F(x) = \sum_{x_i < x} p(X = x_i), \quad (4.5)$$

где суммирование распространяется на все значения x_i , которые меньше x .

Плотность распределения

Случайная величина X называется **непрерывной**, если ее функция распределения $F(x)$ – непрерывная и дифференцируемая функция для всех значений аргумента.

Для непрерывной функции распределения $F(x)$ вероятность любого отдельного значения случайной величины должна быть равна нулю, т.е. не должно быть скачков ни в одной точке. Такие события – возможные, но с нулевой вероятностью – появляются только при рассмотрении опытов, не сводящихся к схеме случаев. Это аналогично телу, имеющему определенную массу, но ни одна из точек внутри тела конечной массой не обладает. Малый объем обладает конечной массой, но она приближается к нулю по мере уменьшения объема и в пределе равна нулю для точки. То есть при непрерывном распределении вероятностей вероятность попадания на сколь угодно малый участок отлична от нуля, тогда вероятность попадания в строго определенную точку в точности равна нулю.

Вероятность попадания непрерывной случайной величины X на участок от x до $x + \Delta x$ равна приращению функции распределения на этом участке: $p\{x \leq X < x + \Delta x\} = F(x + \Delta x) - F(x)$. Тогда плотность вероятности на этом участке равна $\frac{p\{x \leq X < x + \Delta x\}}{\Delta x}$. Переходя к пределу при $\Delta x \rightarrow 0$, получим плотность вероятности в точке x

$$\lim_{\Delta x \rightarrow 0} \frac{p\{x \leq X < x + \Delta x\}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx} = F'(x) = f(x).$$

Полученная функция является одной из форм закона распределения непрерывных случайных величин.

Плотностью распределения (плотностью вероятности) $f(x)$ непрерывной случайной величины X называется производная ее функции распределения

$$f(x) = \frac{dF(x)}{dx} = F'(x), \quad (4.6)$$

а график плотности распределения называется кривой распределения.

Пусть имеется точка x и прилегающий к ней отрезок dx . Вероятность попадания случайной величины X на этот интервал равна $f(x)dx$. Эта величина называется **элементом вероятности**. Вероятность попадания случайной величины X на произвольный участок $[a, b]$ равна сумме элементов вероятности на этом участке:

$$p\{a \leq X < b\} = \int_a^b f(x)dx. \quad (4.7)$$

В геометрической интерпретации $p\{a \leq X < b\}$ равна площади, ограниченной сверху кривой плотности распределения $f(x)$ и участком $[a, b]$.

Соотношение (4.7) позволяет выразить функцию распределения $F(x)$ случайной величины X через ее плотность:

$$F(x) = p\{X < x\} = p\{-\infty < X < x\} = \int_{-\infty}^x f(x)dx. \quad (4.8)$$

Основные свойства плотности распределения:

1. Плотность распределения неотрицательна $f(x) \geq 0$, так как ее первообразная $F(x)$ является неубывающей функцией (см. свойство 3 $F(x)$).

2. Условие нормировки:
$$\int_{-\infty}^{\infty} f(x)dx = p(-\infty \leq X < +\infty) = 1. \quad (4.9)$$

Полная площадь, ограниченная кривой распределения и осью абсцисс, равна 1.

ЛЕКЦИЯ 5

Числовые характеристики случайной величины

Закон распределения случайной величины является исчерпывающей характеристикой, которая полностью описывает случайную величину с вероятностной точки зрения. Однако во многих практических задачах нет надобности в таком полном описании и достаточно указать только отдельные числовые параметры, характеризующие существенные черты распределения. Такие числа называются **числовыми характеристиками случайной величины**.

Математическое ожидание характеризует среднее значение случайной величины и определяется по формулам:

$$m_X = M[X] = \begin{cases} \sum_{i=1}^N x_i \cdot p_i & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} x \cdot f(x) dx & \text{для НСВ.} \end{cases} \quad (5.1)$$

где m_X обозначает число, полученное после вычислений по формуле (5.1);

$M[X]$ – оператор математического ожидания.

Как видно из (5.1), в качестве математического ожидания используется «среднее взвешенное значение», причем каждое из значений случайной величины учитывается с «весом», пропорциональным вероятности этого значения.

Физический смысл математического ожидания – среднее значение случайной величины, т.е. то значение, которое может быть использовано вместо случайной величины в приближительных расчетах или оценках.

Математическое ожидание обладает следующими свойствами:

1. $M[c] = c$.

Доказательство. Рассмотрим константу c как случайную дискретную величину, которая принимает одно значение c с вероятностью $p = 1$.

2. $M[X + c] = M[X] + c = m_X + c$.

Доказательство: $M[X + c] = \int_{-\infty}^{\infty} (x + c) \cdot f(x) dx = \int_{-\infty}^{\infty} x \cdot f(x) dx + \int_{-\infty}^{\infty} c \cdot f(x) dx = m_X + c$.

3. $M[c \cdot X] = c \cdot M[X] = c \cdot m_X$.

Доказательство: $M[cX] = \int_{-\infty}^{\infty} cx \cdot f(x) dx = c \int_{-\infty}^{\infty} x \cdot f(x) dx = c \cdot m_X$.

Начальный момент k -го порядка случайной величины X есть математическое ожидание k -й степени этой случайной величины:

$$\alpha_k(x) = M[X^k] = \begin{cases} \sum_{i=1}^N x_i^k \cdot p_i & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} x^k \cdot f(x) dx & \text{для НСВ.} \end{cases} \quad (5.2)$$

При $k = 0$ $\alpha_0(x) = M[X^0] = M[1] = 1$;

$k = 1$ $\alpha_1(x) = M[X^1] = M[X] = m_X$ – математическое ожидание;

$k = 2$ $\alpha_2(x) = M[X^2]$.

Центрированной случайной величиной $\overset{\circ}{X}$ называется случайная величина, математическое ожидание которой находится в начале координат (в центре числовой оси), т.е. $M[\overset{\circ}{X}] = 0$.

Операция центрирования (переход от нецентрированной величины X к центрированной $\overset{\circ}{X}$) имеет вид

$$\overset{\circ}{X} = X - m_X.$$

Центральный момент порядка k случайной величины X есть математическое ожидание k -й степени центрированной случайной величины $\overset{\circ}{X}$:

$$\mu_k(x) = M[\overset{\circ}{X}^k] = \begin{cases} \sum_{i=1}^N (x_i - m_X)^k \cdot p_i & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} (x - m_X)^k \cdot f(x) dx & \text{для НСВ.} \end{cases} \quad (5.3)$$

При $k = 0$ $\mu_0(x) = M[\overset{\circ}{X}^0] = M[1] = 1$;

$k = 1$ $\mu_1(x) = M[\overset{\circ}{X}^1] = M[\overset{\circ}{X}] = 0$;

$k = 2$ $\mu_2(x) = M[\overset{\circ}{X}^2] = M[(X - m_X)^2] = M[X^2] - 2m_X M[X] + m_X^2 = \alpha_2 - m_X^2 = D_X$ – дисперсия.

Дисперсия случайной величины характеризует степень рассеивания (разброса) значений случайной величины относительно ее математического ожидания и определяется по формулам:

$$D_x = D[X] = \mu_2(x) = \alpha_2(x) - m_X^2 = \begin{cases} \sum_{i=1}^N (x_i - m_X)^2 p_i = \sum_{i=1}^N x_i^2 p_i - m_X^2 & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} (x - m_X)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - m_X^2 & \text{для НСВ.} \end{cases} \quad (5.4)$$

Свойства дисперсии:

$$1. D[c] = 0.$$

$$\text{Доказательство: } D[c] = M[(c - M[c])^2] = M[(c - c)^2] = M[0] = 0.$$

$$2. D[X + c] = D_X.$$

Доказательство:

$$D[X + c] = M[(X + c - M[X + c])^2] = M[(X + c - m_X - c)^2] = M[(X - m_X)^2] = D_X$$

вытекает из свойства 3 математического ожидания. Оно становится понятным, если учесть, что величины X и $X + c$ отличаются лишь началом отсчета и рассеяны вокруг своих математических ожиданий одинаково. Очевидно, что операция центрирования не изменяет дисперсию случайной величины:

$$D[\overset{\circ}{X}] = D[X - m_X] = D[X].$$

$$3. D[c \cdot X] = c^2 \cdot D_X.$$

$$\text{Доказательство: } D[cX] = M[c^2 X^2] - (M[cX])^2 = c^2 (M[X^2] - m_X^2) = c^2 D_X.$$

Дисперсия случайной величины имеет размерность квадрата случайной величины, поэтому для анализа диапазона значений величины X дисперсия не совсем удобна. Этому недостатка лишено среднее квадратическое отклонение (СКО), размерность которого совпадает с размерностью случайной величины.

Среднее квадратическое отклонение случайной величины X характеризует ширину диапазона значений X и равно:

$$\sigma_X = \sigma[X] = +\sqrt{D[X]}. \quad (5.5)$$

СКО измеряется в тех же физических единицах, что и случайная величина.

Правило 3σ. Практически все значения случайной величины находятся в интервале

$$[m_X - 3\sigma_X; m_X + 3\sigma_X]. \quad (5.6)$$

Математическое ожидание и дисперсия (или СКО) – наиболее часто применяемые характеристики случайной величины. Они характеризуют наиболее важные черты распределения: его положение и степень разбросанности значений. Для более подробного описания используются начальные и центральные моменты высших порядков. Кроме математического ожидания на практике часто применяются и другие характеристики положения распределения значений.

Мода случайной величины равна ее наиболее вероятному значению, т.е. то значение, для которого вероятность p_i (для дискретной случайной величины) или $f(x)$ (для непрерывных случайной величины) достигает максимума:

$$f(Mo) = \max, \quad p(X = Mo) = \max.$$

Распределение с одним максимумом плотности распределения называется «унимодальным». Если многоугольник распределения или кривая распределения имеют более одного максимума, распределение называют

«полимодальным». Если распределение обладает посередине не максимумом, а минимумом, то оно называется «антимодальным».

Медиана случайной величины X равна такому ее значению, для которого выполняется условие $p\{X < Me\} = p\{X \geq Me\}$. Медиана, как правило, существует только для непрерывных случайных величин. Значение Me может быть определено как решение одного из следующих уравнений:

$$\int_{-\infty}^{Me} f(x)dx = 0,5; \int_{Me}^{+\infty} f(x)dx = 0,5; F(Me) = 0,5. \quad (5.7)$$

В точке Me площадь, ограниченная кривой распределения, делится пополам.

Квантиль χ_p случайной величины X – это такое ее значение, для которого выполняется условие

$$p\{X < \chi_p\} = F(\chi_p) = p. \quad (5.7)$$

Очевидно, что медиана – это квантиль $\chi_{0,5}$.

ЛЕКЦИЯ 6

Типовые законы распределения

Индикатор случайного события A – это дискретная случайная величина X , которая равна 1 при осуществлении события A и 0 при осуществлении \bar{A} :

$$X = \begin{cases} 1, & A \\ 0, & \bar{A} \end{cases}.$$

Ряд распределения вероятностей индикатора случайного события:

x_i	0	1
p_i	q	p

где p – вероятность осуществления A ;

$q = 1 - p$ – вероятность осуществления \bar{A} .

Числовые характеристики индикатора случайного события:

$$m_X = p, D_X = qp. \quad (6.1)$$

Геометрическое распределение имеет дискретная случайная величина X , если она принимает значения $0, 1, \dots, \infty$ с вероятностями:

$$p(X = i) = p_i = q^i p, \quad (6.2)$$

где p – параметр распределения ($0 \leq p \leq 1$), $q = 1 - p$.

Числовые характеристики геометрического распределения:

$$m_X = q / p, D_X = q / p^2. \quad (6.3)$$

Условия возникновения. Проводится ряд одинаковых независимых опытов до первого появления некоторого события A . Случайная величина X – число проведенных безуспешных опытов до первого появления события A .

Биномиальное распределение имеет дискретная случайная величина X , если она принимает значения $0, 1, \dots, n$ со следующими вероятностями:

$$p(X = i) = p_i = \frac{n!}{i!(n-i)!} p^i q^{n-i}, \quad (6.4)$$

где n, p – параметры распределения ($0 \leq p \leq 1$), $q = 1 - p$.

Числовые характеристики биномиального распределения:

$$m_X = np, D_X = npq. \quad (6.5)$$

Условия возникновения. Проводится n одинаковых независимых, в каждом из которых событие A появляется с вероятностью p . Случайная величина X – число опытов, в которых произошло событие A (см. теорему о повторении опытов).

Распределение Пуассона имеет дискретная случайная величина X , если она принимает значения $0, 1, \dots, \infty$ со следующими вероятностями:

$$p(X=i) = p_i = \frac{a^i}{i!} e^{-a}, \quad (6.6)$$

где a – параметр распределения ($a > 0$).

Числовые характеристики пуассоновской случайной величины:

$$m_X = a, D_X = a. \quad (6.7)$$

Условия возникновения:

1. Распределение Пуассона является предельным случаем биномиального, когда число опытов n неограниченно увеличивается, а вероятность p события A в одном опыте стремится к 0, так что существует предел $\lim_{\substack{n \rightarrow \infty \\ p \rightarrow 0}} np = a$

(см. формулу (3.10)).

2. Случайная величина X – число событий пуассоновского потока поступивших в течение интервала τ , причем параметр $a = \tau\lambda$, где λ – интенсивность потока.

Рассмотрим временную ось, на которой будем отмечать моменты возникновения случайных событий (например, отказы компонентов в сложном техническом устройстве, заявки на обслуживание и т.п.). Последовательность таких моментов называется *поток случайных событий*.

Поток случайных событий называется *стационарным*, если число событий, приходящихся на интервал τ , в общем случае не зависит от расположения этого участка на временной оси и определяется только его длительностью, т.е. среднее число событий в единице времени λ (*интенсивность потока*) постоянно.

Поток случайных событий называется *ординарным*, если вероятность попадания в некоторый участок Δt двух и более случайных событий значительно меньше, чем вероятность попадания 1-го события.

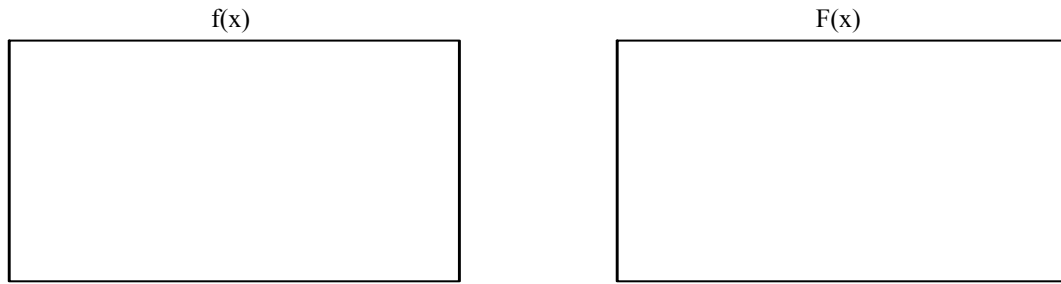
В потоке *отсутствует последствие*, если вероятность попадания событий на участок τ не зависит от того, сколько событий попало на другие участки, не пересекающиеся с данным.

Поток случайных событий называется *пуассоновским* или простейшим, если он является стационарным, ординарным и без последствия.

Равномерное распределение имеет непрерывная случайная величина X , если ее плотность вероятности в некотором интервале $[a; b]$ постоянна, т.е. если все значения X в этом интервале равновероятны:

$$f(x) = \begin{cases} 0, & x < a, \\ \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & x > b. \end{cases} \quad F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases} \quad (6.8)$$

Ниже приведены графики плотности и функции равномерного распределения при $b = 3$ и $a = 1$.



Числовые характеристики равномерно распределенной случайной величины:

$$m_X = \frac{a + b}{2}, D_X = \frac{(b - a)^2}{12}. \quad (6.9)$$

При необходимости определения параметров a и b по известным m_X , D_X используют следующие формулы:

$$a = m_X - \sigma_X \sqrt{3}, b = m_X + \sigma_X \sqrt{3}. \quad (6.10)$$

Условия возникновения:

1. Случайная величина X – ошибки округления при ограниченной разрядной сетке:

- округление до меньшего целого, $X \in [-1; 0]$, $m_X = -0,5$;
- округление до большего целого, $X \in [-0; 1]$, $m_X = 0,5$;
- округление до ближайшего целого, $X \in [-0,5; 0,5]$, $m_X = 0$,

где 1 – вес младшего разряда.

2. Случайная величина X – погрешность считывания значений с аналоговой шкалы измерительного прибора, $X \in [-0,5; 0,5]$, $m_X = 0$, где 1 – цена деления шкалы.

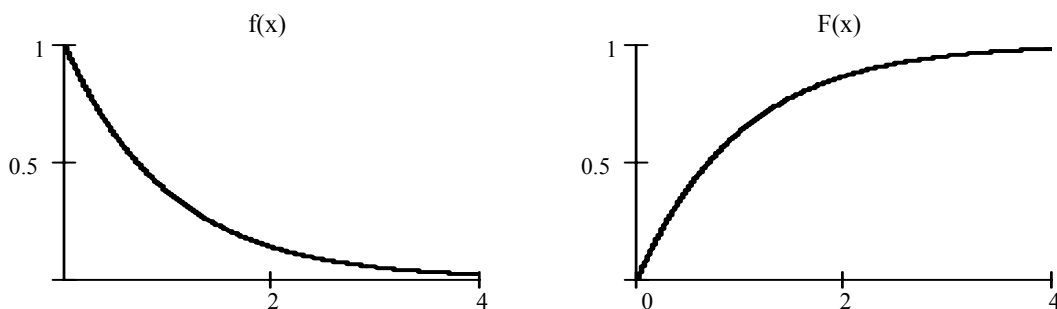
3. Генераторы псевдослучайных величин, например RANDOM, встроенные в языки программирования высокого уровня.

Экспоненциальное распределение имеет непрерывная случайная величина T , принимающая только положительные значения, если ее плотность вероятности и функция распределения равны:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases} \quad F(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases} \quad (6.11)$$

где λ – параметр распределения ($\lambda > 0$).

Ниже приведены графики плотности и функции экспоненциального распределения при $\lambda = 1$.



Числовые характеристики экспоненциальной случайной величины:

$$m_T = 1 / \lambda, D_T = 1 / \lambda^2. \quad (6.12)$$

Условия возникновения. Случайная величина T – интервал времени между двумя соседними событиями в простейшем или пуассоновском потоке случайных событий, причем параметр распределения λ – интенсивность потока.

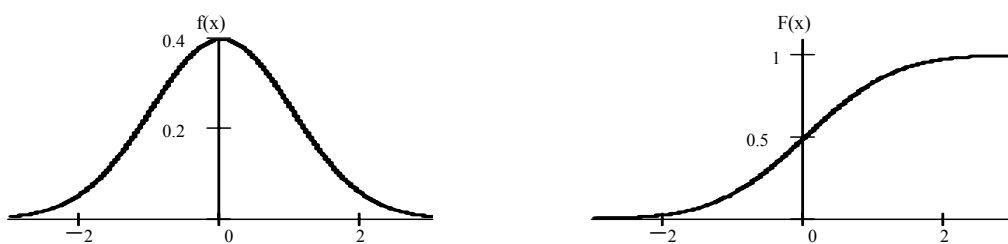
Нормальное распределение (распределение Гаусса) имеет непрерывная случайная величина X , если ее плотность вероятности и функция распределения равны:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, F(x) = 0.5 + \Phi\left(\frac{x-m}{\sigma}\right), \quad (6.13)$$

где m, σ – параметры распределения ($\sigma > 0$);

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \text{ — функция Лапласа.}$$

Ниже приведены графики плотности и функции нормального распределения при $m = 1, \sigma = 1$.



Так как первообразная для e^{-x^2} в аналитическом виде не существует, то для вычисления значений функции распределения и вероятностей событий, связанных с нормальной случайной величиной, используется табулированная функция Лапласа. При использовании таблицы значений функции Лапласа следует учитывать, что $\Phi(-x) = -\Phi(x)$, $\Phi(0) = 0$, $\Phi(\infty) = 0.5$.

Числовые характеристики нормальной случайной величины:

$$m_X = m, D_X = \sigma^2; \quad (6.14)$$

$$\alpha_k(x) = k! \sum_{i=0}^{I[k/2]} \frac{m^{k-2i} (\sigma/2)^i}{(k-2i)! i!}; \quad (6.15)$$

$$\mu_k(x) = \begin{cases} 0, k - \text{нечетное}, \\ \frac{k!}{(k/2)!} \left(\frac{\sigma^2}{2} \right)^{k/2}, k - \text{четное}. \end{cases} \quad (6.16)$$

Условия возникновения. Это наиболее часто встречающийся на практике закон распределения – см. лекцию 12 (**Центральная предельная теорема**). Например, нормальный закон распределения имеют:

- погрешности измерительных приборов; при этом откалиброванный прибор не имеет систематической погрешности, т.е. $m = 0$, а величина σ определяется классом точности измерительного прибора;

- параметры радиоэлектронных компонентов (резисторов, конденсаторов и т.п.), причем m – номинальное значение, указанное на маркировке, а σ определяется классом точности.

ЛЕКЦИЯ 7

Функции одного случайного аргумента

Пусть некоторая случайная величина X подвергается детерминированному преобразованию φ , в результате которого получится величина Y , т.е. $Y = \varphi(x)$. Очевидно, что величина Y будет случайной, и, как правило, необходимо определить закон распределения и/или числовые характеристики случайной величины Y по известному закону распределения величины X и виду преобразования φ .

Закон распределения функции случайного аргумента

В случае, если X – дискретная случайная величина с известным рядом распределения вероятностей, определение ряда вероятностей Y не составит сложности.

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_i

y_i	$\varphi(x_1)$	$\varphi(x_2)$...	$\varphi(x_n)$
p_i	p_1	p_2	...	p_n

(*)

y_i	y_1	y_2	...	y_m
p_i	p_1	p_2	...	p_m

(**)

Из (*) путем упорядочивания и объединения одинаковых значений получаем ряд распределения случайной величины Y (**).

Если X – непрерывная случайная величина с известной плотностью вероятности $f(x)$, то алгоритм получения закона распределения $Y = \varphi(x)$ зависит от вида φ . Рассмотрим участок оси абсцисс $[a, b]$, на котором лежат все возможные значения величины X , т.е. $p(a \leq X \leq b) = 1$, в частном случае $a = -\infty, b = +\infty$. Способ решения поставленной задачи зависит от поведения функции φ на участке $[a, b]$: монотонна она на этом участке или нет. При этом отдельно проанализируем два случая: монотонного возрастания и монотонного убывания функции.

$Y = \varphi(x)$ – **монотонно возрастающая функция**. Определим функцию распределения $G(y)$ случайной величины Y . По определению она равна

$$G(y) = p(Y < y) = p(\varphi(x) < y) = p(X < \psi(y)) = \int_{-\infty}^{\psi(y)} f_X(x) dx,$$

где $\psi(y)$ – обратная функция $\varphi(x)$.

Для выполнения условия $Y < y$ необходимо и достаточно, чтобы случайная величина X попала на участок оси абсцисс от a до $\psi(y)$. Таким образом, функция распределения Y для аргумента X , распределенного в интервале $[a, b]$, равна

$$G(y) = \begin{cases} 0, & y < \psi(a), \\ \int_a^{\psi(y)} f_X(x) dx, & \psi(a) \leq y \leq \psi(b), \\ 1, & y > \psi(b). \end{cases}$$

$Y = \varphi(x)$ – **монотонно убывающая функция.** Определим функцию распределения $G(y)$ случайной величины Y . По определению она равна

$$G(y) = p(Y < y) = p(\varphi(x) < y) = p(X > \psi(y)) = \int_{\psi(y)}^{\infty} f_X(x) dx,$$

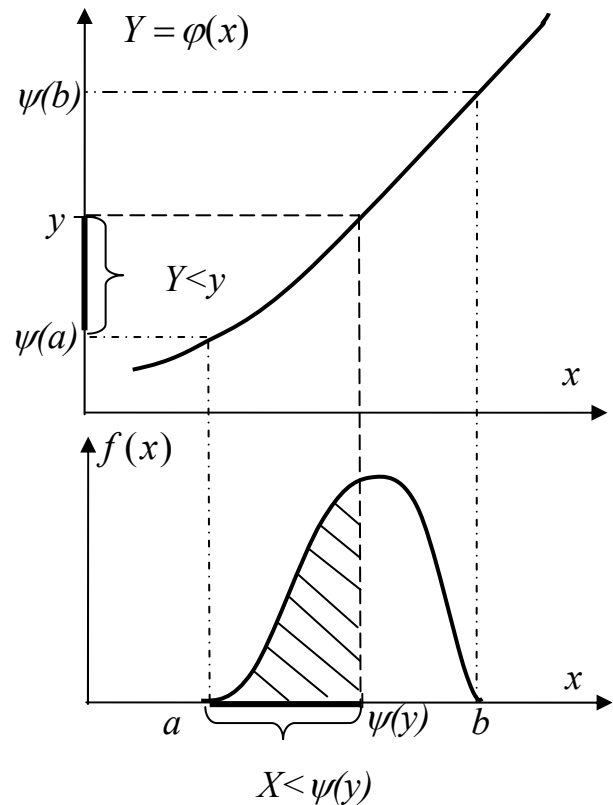
где $\psi(y)$ – обратная функция $\varphi(x)$.

Для выполнения условия $Y < y$ необходимо и достаточно, чтобы случайная величина X попала на участок оси абсцисс от $x = \psi(y)$ до b . Таким образом, функция распределения Y для аргумента X , распределенного в интервале $[a, b]$, равна:

$$G(y) = \begin{cases} 0, & y < \psi(b), \\ \int_{\psi(y)}^b f_X(x) dx, & \psi(b) \leq y \leq \psi(a), \\ 1, & y > \psi(a). \end{cases}$$

Плотность вероятностей случайной величины $Y = \varphi(x)$ для любого монотонного случая имеет следующий вид:

$$g(y) = G'(y) = \begin{cases} 0, & y < y_{\min}, \\ f_X(\psi(y)) \cdot |\psi'(y)|, & y_{\min} \leq y \leq y_{\max}, \\ 0, & y > y_{\max}. \end{cases} \quad (7.1)$$



Пример. Пусть случайная величина X имеет нормальный закон распределения $f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2}$, $Y = X^3$. Найти $g(y)$.

Функция $Y = \varphi(x)$ строго монотонна, дифференцируема и имеет обратную $X = \psi(y) = \sqrt[3]{Y}$. Воспользуемся формулой (7.1). Так как

$$f_X(\psi(y)) = f_X(y^{1/3}) = \frac{1}{\sigma \sqrt{2\pi}} e^{-y^{2/3}/2\sigma^2},$$

$$|\psi'(y)| = |(y^{1/3})'| = \frac{1}{3y^{2/3}},$$

то искомая плотность распределения функции $Y = X^3$:

$$g(y) = \frac{1}{3\sigma y^{2/3} \sqrt{2\pi}} e^{-y^{2/3}/2\sigma^2}.$$

$Y = \varphi(x)$ – **немонотонная функция**. Алгоритм получения закона распределения $Y = \varphi(x)$ приведен ниже.

1. Построить график $Y = \varphi(x)$ и определить диапазон значений $Y [y_{\min}, y_{\max}]$.
2. Диапазон Y разбить на M интервалов, в каждом из которых одинаковая степень неоднозначности $k_i, i = 1, 2, \dots, M$:

$$[y_{\min}, y_1), [y_1, y_2) \dots [y_{M-1}, y_{\max}].$$

Степень неоднозначности k_i – число значений X , соответствующих одному значению Y , или число обратных функций для данного интервала $\psi_j(y), j = 1, \dots, k_i$.

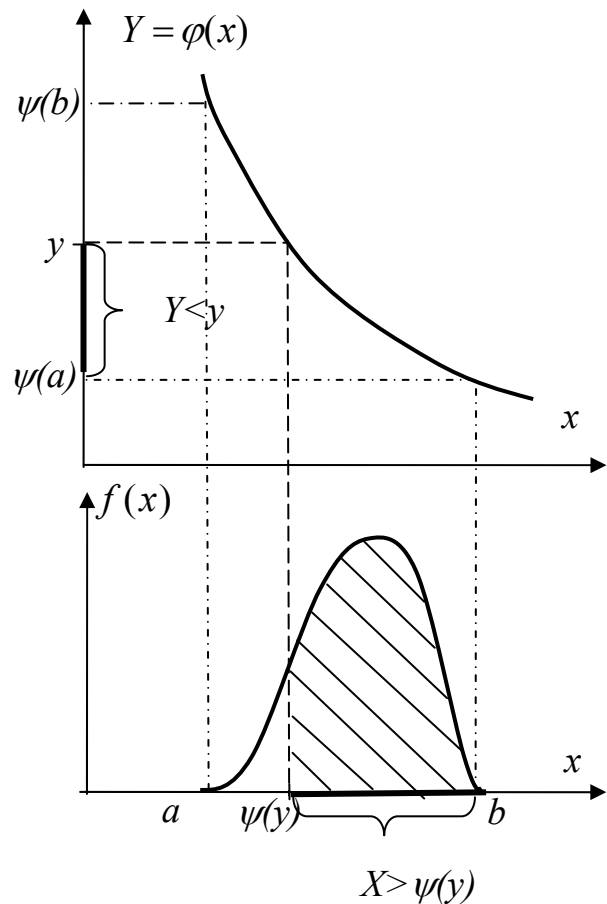
3. Определить обратные функции $\psi_j(y) = \varphi^{-1}(x)$ и вычислить $|\psi_j'(y)|$.

В общем случае число обратных функций $\psi_j(y)$ в i -м интервале равно k_i .

4. Определить плотность вероятностей $g(y)$ по следующей формуле:

$$g(y) = \begin{cases} 0, & y < y_{\min}, \\ M & \\ \sum_{j=1}^{k_i} f_X(\psi_j(y)) \cdot |\psi_j'(y)|, & y_{i-1} \leq y < y_i, \\ M & \\ 0, & y > y_{\max}. \end{cases} \quad (7.2)$$

В частном случае, когда обратные функций одинаковы для всех интервалов $\psi_j'(y) = \psi'(y), |\psi_j'(y)| = |\psi'(y)|$, формула (7.2) принимает вид



$$g(y) = \begin{cases} 0, y < y_{\min}, \\ M \\ k_i \cdot f_X(\psi(y)) \cdot |\psi'(y)|, y_{i-1} \leq y < y_i, \\ M \\ 0, y > y_{\max}. \end{cases} \quad (7.3)$$

а если величина X равномерно распределена в интервале $[a, b]$, т.е. ее плотность равна $f_X(x) = \begin{cases} 1/(b-a), a \leq x \leq b \\ 0, x < a, x > b \end{cases}$, то выражение для $g(y)$ можно представить как

$$g(y) = \begin{cases} 0, y < y_{\min}, \\ M \\ k_i \cdot \frac{1}{b-a} \cdot |\psi'(y)|, y_{i-1} \leq y < y_i, \\ M \\ 0, y > y_{\max}. \end{cases} \quad (7.4)$$

Числовые характеристики функции случайного аргумента

Пусть $Y = \varphi(x)$, где X – случайная величина с известным законом распределения, и необходимо определить числовые характеристики Y . В том случае, когда закон распределения Y определен (см. выражения (7.1 – 7.4)), то числовые характеристики Y легко вычислить по формулам (5.1 – 5.7). Однако, если закон распределения величины Y в явном виде не нужен, а необходимы только ее числовые характеристики, применимы следующие формулы.

Если X – дискретная случайная величина с известным рядом распределения вероятностей, то

$$m_Y = M[Y] = \sum_{i=1}^n \varphi(x_i) p_i; \quad (7.5)$$

$$D_Y = M[Y^2] - m_Y^2 = \sum_{i=1}^n \varphi^2(x_i) p_i - m_Y^2; \quad (7.6)$$

$$\alpha_k(y) = M[Y^k] = \sum_{i=1}^n \varphi^k(x_i) p_i; \quad (7.7)$$

$$\mu_k(y) = M[Y^{\circ k}] = \sum_{i=1}^n (\varphi(x_i) - m_Y)^k p_i. \quad (7.8)$$

Если X – непрерывная случайная величина с известной плотностью вероятностей $f(x)$, то формулы принимают вид

$$m_Y = M[Y] = \int_{-\infty}^{+\infty} \varphi(x) \cdot f(x) dx; \quad (7.9)$$

$$D_Y = M[Y^2] - m_Y^2 = \int_{-\infty}^{+\infty} \varphi^2(x) \cdot f(x) dx - m_Y^2; \quad (7.10)$$

$$\alpha_k(y) = M[Y^k] = \int_{-\infty}^{+\infty} \varphi^k(x) \cdot f(x) dx ; \quad (7.11)$$

$$\mu_k(y) = M[\overset{o}{Y}^k] = \int_{-\infty}^{+\infty} (\varphi(x) - m_Y)^k \cdot f(x) dx . \quad (7.12)$$

Характеристическая функция случайной величины

Пусть $Y = e^{itX}$, где X – случайная величина с известным законом распределения, t – параметр, $i = \sqrt{-1}$.

Характеристической функцией случайной величины X называется математическое ожидание функции $Y = e^{itX}$:

$$\nu_X(t) = M[e^{itX}] = \begin{cases} \sum_{k=1}^n e^{itx_k} p_k, & \text{для ДСВ,} \\ \int_{-\infty}^{+\infty} e^{itx} f(x) dx, & \text{для НСВ.} \end{cases} \quad (7.12)$$

Таким образом, характеристическая функция $\nu_X(t)$ и закон распределения случайной величины однозначно связаны *преобразованием Фурье*. Например, плотность распределения $f(x)$ случайной величины X однозначно выражается через ее характеристическую функцию при помощи *обратного преобразования Фурье*:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \nu(t) e^{-itx} dt . \quad (7.13)$$

Основные свойства характеристической функции:

1. Характеристическая функция величины $Z = aX + b$, где X – случайная величина с характеристической функцией $\nu_X(t)$, равна

$$\nu_Z(t) = M[e^{it(aX+b)}] = e^{itb} \nu_X(at) . \quad (7.14)$$

2. Начальный момент k -го порядка случайной величины X равен

$$\alpha_k(x) = \nu_X^{(k)}(0) i^{-k} , \quad (7.15)$$

где $\nu_X^{(k)}(0)$ – значение k -й производной характеристической функции при $t = 0$.

3. Характеристическая функция суммы $Y = \sum_{k=1}^n X_k$ независимых случайных величин равна произведению характеристических функций слагаемых:

$$\nu_Y(t) = \prod_{i=1}^n \nu_{X_i}(t) . \quad (7.16)$$

4. Характеристическая функция нормальной случайной величины с параметрами m и σ равна:

$$\nu_X(t) = e^{itm - \frac{t^2 \sigma^2}{2}} . \quad (7.17)$$

ЛЕКЦИЯ 8

Двухмерные случайные величины. Двухмерный закон распределения

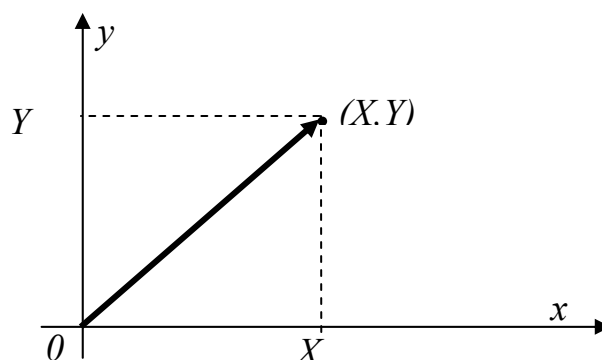
Двухмерная случайная величина (X, Y) – совокупность двух одномерных случайных величин, которые принимают значения в результате проведения одного и того же опыта.

Двухмерные случайные величины характеризуются множествами значений Ω_X, Ω_Y своих компонент и совместным (двухмерным) законом распределения. В зависимости от типа компонент X, Y различают дискретные, непрерывные и смешанные двухмерные случайные величины.

Двухмерную случайную величину (X, Y) геометрически можно представить как случайную точку (X, Y) на плоскости xOy либо как случайный вектор, направленный из начала координат в точку (X, Y) .

Двухмерный закон распределения вероятностей – функция, таблица, правило, позволяющие вычислить вероятности любых случайных событий, связанных двухмерной случайной величиной (X, Y) :

$$p(\alpha \leq X \leq \beta; \delta \leq Y \leq \gamma) \forall \alpha, \beta, \delta, \gamma.$$

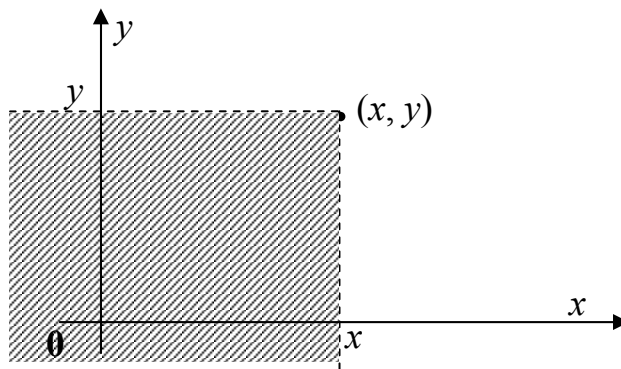


Двухмерная функция распределения

Двухмерная функция распределения двухмерной случайной величины (X, Y) равна вероятности совместного выполнения двух событий $\{X < x\}$ и $\{Y < y\}$:

$$F(x, y) = p(\{X < x\} \cdot \{Y < y\}). \quad (8.1)$$

Геометрически двухмерная функция распределения $F(x, y)$ – это вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной в точке (x, y) , лежащей левее и ниже ее. Компонента X приняла значения, меньшие действительного числа x , это функция распределения $F_X(x)$, а компонента Y – меньшие действительного числа y , это функция распределения $F_Y(y)$.



Свойства двухмерной функции распределения:

$$1. 0 \leq F(x, y) \leq 1.$$

Доказательство. Свойство вытекает из определения функции распределения как вероятности: вероятность – неотрицательное число, не превышающее 1.

$$2. F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0, F(+\infty, +\infty) = 1.$$

$$3. F(x_1, y) \leq F(x_2, y), \text{ если } x_2 > x_1; F(x, y_1) \leq F(x, y_2), \text{ если } y_2 > y_1.$$

Доказательство. Докажем, что $F(x, y)$ – неубывающая функция по переменной x . Рассмотрим вероятность

$$p(X < x_2, Y < y) = p(X < x_1, Y < y) + p(x_1 \leq X < x_2, Y < y).$$

Так как $p(X < x_2, Y < y) = F(x_2, y)$, а $p(X < x_1, Y < y) = F(x_1, y)$, то $F(x_2, y) - F(x_1, y) = p(x_1 \leq X < x_2, Y < y) \Rightarrow F(x_2, y) - F(x_1, y) \geq 0 \Rightarrow F(x_2, y) \geq F(x_1, y)$.

Аналогично и для y .

4. Переход к одномерным характеристикам:

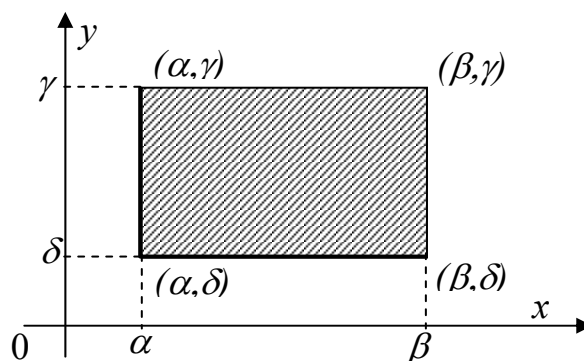
$$F(x, \infty) = p(X < x, Y < \infty) = p(X < x) = F_X(x); \quad (8.2)$$

$$F(\infty, y) = p(X < \infty, Y < y) = p(Y < y) = F_Y(y). \quad (8.3)$$

5. Вероятность попадания в прямоугольную область

$$p(\alpha \leq X \leq \beta, \delta \leq Y \leq \gamma) = F(\beta, \gamma) - F(\beta, \delta) - F(\alpha, \gamma) + F(\alpha, \delta).$$

Функция распределения – наиболее универсальная форма закона распределения и может быть использована для описания как непрерывных, так и дискретных двумерных случайных величин.



Матрица распределения

Двухмерная случайная величина (X, Y) является дискретной, если множества значений ее компонент Ω_X и Ω_Y представляют собой счетные множества. Для описания вероятностных характеристик таких величин используется двумерная функция распределения и матрица распределения.

Матрица распределения представляет собой прямоугольную таблицу, которая содержит значения компоненты $X - \Omega_X = \{x_1, x_2, \dots, x_n\}$, значения компоненты $Y - \Omega_Y = \{y_1, y_2, \dots, y_m\}$ и вероятности всевозможных пар значений $p_{ij} = p(X = x_i, Y = y_j), i = 1, \dots, n, j = 1, \dots, m$.

$x_i \setminus y_j$	y_1	y_2	...	y_m
x_1	p_{11}	p_{12}	...	p_{1m}
x_2	p_{21}	p_{22}	...	p_{2m}
...
x_n	p_{n1}	p_{n2}	...	p_{nm}

Свойства матрицы распределения вероятностей:

$$1. \sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1 .$$

2. Переход к ряду распределения вероятностей составляющей X :

$$p_i = p(X = x_i) = \sum_{j=1}^m p_{ij}, i = 1, \dots, n . \quad (8.3)$$

3. Переход к ряду распределения вероятностей составляющей Y :

$$p_j = p(Y = y_j) = \sum_{i=1}^n p_{ij}, j = 1, \dots, m . \quad (8.4)$$

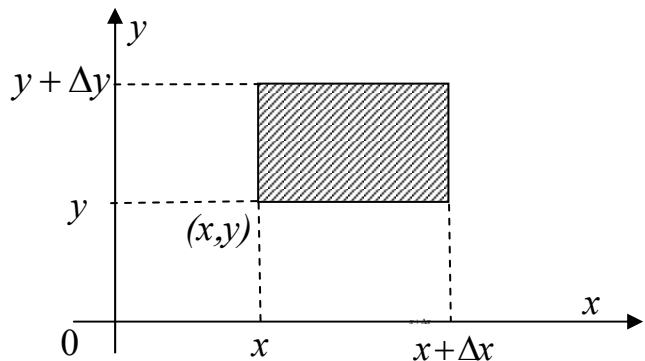
Двухмерная плотность распределения

Двухмерная случайная величина (X, Y) является непрерывной, если ее функция распределения $F(x, y)$ представляет собой непрерывную, дифференцируемую функцию по каждому из аргументов и существует вторая смешанная производная $\frac{\partial^2 F(x, y)}{\partial x \partial y}$.

Двухмерная плотность распределения $f(x, y)$ характеризует плотность вероятности в окрестности точки с координатами (x, y) и равна второй смешанной производной функции распределения:

$$f(x, y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{p(\{x \leq X < x + \Delta x\} \cap \{y \leq Y < y + \Delta y\})}{\Delta x \Delta y} = \frac{\partial^2 F(x, y)}{\partial x \partial y} . \quad (8.5)$$

Геометрически $f(x, y)$ – это некоторая *поверхность распределения*, она аналогична кривой распределения для одномерной случайной величины. Аналогично можно ввести понятие *элемента вероятности*: $f(x, y)dx dy$. Вероятность попадания значения двухмерной случайной величины (X, Y) в произвольную область D равна сумме всех элементов вероятности для этой области:



$$p\{(X, Y) \in D\} = \iint_{(D)} f(x, y) dx dy . \quad (8.6)$$

Свойства двухмерной плотности:

$$1. f(x, y) \geq 0 .$$

2. Условие нормировки:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1 . \quad (8.7)$$

Геометрически – объем тела, ограниченный поверхностью распределения и плоскостью xOy , равен единице.

3. Переход к функции распределения:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy. \quad (8.8)$$

4. Переход к одномерным характеристикам:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy; \quad (8.9)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx. \quad (8.10)$$

Зависимые и независимые случайные величины

Величина X **независима** от величины Y , если ее закон распределения не зависит от того, какое значение приняла величина Y . Для независимых величин выполняются следующие соотношения, т. е. критерии независимости:

$$1) F(x, y) = p(X < x, Y < y) = p(X < x)p(Y < y) = F_X(x)F_Y(y) \quad \forall x, y; \quad (8.11)$$

$$2) \text{ для непрерывных } - f(x, y) = f_X(x)f_Y(y) \quad \forall x, y; \quad (8.12)$$

$$3) \text{ для дискретных } - p_{ij} = p_i p_j, \text{ для } \forall i, j. \quad (8.13)$$

В том случае, если критерии не выполняются хотя бы в одной точке, величины X и Y являются зависимыми. Для независимых величин двумерные формы закона распределения не содержат никакой дополнительной информации, кроме той, которая содержится в двух одномерных законах. Таким образом, в случае зависимости величин X и Y переход от двух одномерных законов к двумерному закону осуществить невозможно. Для этого необходимо знать условные законы распределения.

Условные законы распределения

Условным законом распределения называется распределение одной случайной величины, найденное при условии, что другая случайная величина приняла определенное значение.

Условные ряды распределения для дискретных составляющих X и Y определяются по формулам:

$$p_{i/j} = p(X = x_i / Y = y_j) = p_{ij} / p(Y = y_j), i = 1, \dots, n; \quad (8.14)$$

$$p_{j/i} = p(Y = y_j / X = x_i) = p_{ij} / p(X = x_i), j = 1, \dots, m. \quad (8.15)$$

Матрица распределения вероятностей дискретной двумерной случайной величины (X, Y) , если ее компоненты зависимы, «порождает» два одномерных ряда вероятностей (см. (8.3, 8.4)) и два семейства условных рядов вероятностей (8.14, 8.15).

Условные плотности распределения для непрерывных составляющих X и Y определяются по формулам:

$$f(x/y) = f(x, y)/f_Y(y), \text{ для } f_Y(y) \neq 0; \quad (8.16)$$

$$f(y/x) = f(x, y)/f_X(x), \text{ для } f_X(x) \neq 0. \quad (8.17)$$

Условные законы распределения обладают всеми свойствами соответствующих им одномерных форм законов распределения.

Если величины X и Y независимы, то условные законы распределения равны соответствующим безусловным:

$$p_{i/j} = p_i, i = 1, \dots, n; \quad (8.18)$$

$$p_{j/i} = p_j, j = 1, \dots, m. \quad (8.19)$$

$$f(x/y) = f_X(x); \quad (8.20)$$

$$f(y/x) = f_Y(y). \quad (8.21)$$

Следует различать функциональную и статистическую (вероятностную) зависимости между случайными величинами. Если X и Y – случайные величины, которые связаны между собой функциональной зависимостью $y = \varphi(x)$, то, зная значение X , можно точно вычислить соответствующие значение Y , и наоборот.

Если между случайными величинами существует *статистическая зависимость* (величины X и Y зависимы – см. (8.11 – 8.13)), то по значению одной из них можно установить только условное распределение вероятностей другой, т.е. определить, с какой вероятностью появится то или иное значение другой величины.

Пример. Y – урожай зерна, X – количество удобрений на некотором участке земли. Очевидно, что между X и Y существует статистическая зависимость, так как значение Y (урожайность на участке) зависит и от многих других факторов.

ЛЕКЦИЯ 9

Числовые характеристики двумерных величин

Рассмотрим основные числовые характеристики двумерной случайной величины (X, Y) .

Смешанный начальный момент порядка $k + s$ равен математическому ожиданию произведения X^k и Y^s :

$$\alpha_{k,s}(x, y) = M[X^k Y^s] = \begin{cases} \sum_{i=1}^n \sum_{j=1}^m x_i^k y_j^s p_{i,j} & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^s f(x, y) dx dy & \text{для НСВ.} \end{cases} \quad (9.1)$$

Смешанный центральный момент порядка $k + s$ равен математическому ожиданию произведения центрированных величин $\overset{\circ}{X}^k$ и $\overset{\circ}{Y}^s$:

$$\mu_{k,s}(x, y) = M[(X - m_X)^k (Y - m_Y)^s] = \begin{cases} \sum_{i=1}^n \sum_{j=1}^m (x_i - m_X)^k (y_j - m_Y)^s p_{i,j} & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^k (y - m_Y)^s f(x, y) dx dy & \text{для НСВ,} \end{cases} \quad (9.2)$$

где p_{ij} – элементы матрицы распределения вероятностей дискретной случайной величины (ДСВ) (X, Y) ;

$f(x, y)$ – совместная плотность вероятности непрерывной случайной величины (НСВ) (X, Y) .

Рассмотрим наиболее часто используемые начальные и центральные моменты:

$$m_X = \alpha_{1,0}(x, y), \quad m_Y = \alpha_{0,1}(x, y); \quad (9.3)$$

$$D_X = \mu_{2,0}(x, y) = \alpha_{2,0}(x, y) - m_X^2, \quad D_Y = \mu_{0,2}(x, y) = \alpha_{0,2}(x, y) - m_Y^2. \quad (9.4)$$

Особую роль, как характеристика системы случайных величин, играет второй смешанный центральный момент порядка $1+1$ $\mu_{1,1}(x, y)$, который называется *корреляционным моментом* или *ковариацией* случайных величин X, Y .

Корреляционный момент K_{XY} характеризует степень тесноты линейной зависимости величин X и Y и рассеивание их значений относительно точки (m_X, m_Y) :

$$K_{XY} = \mu_{1,1}(x, y) = \alpha_{1,1}(x, y) - m_X m_Y. \quad (9.5)$$

Свойства ковариации K_{XY} :

1. $K_{XY} = K_{YX}$.
2. Корреляционный момент двух независимых случайных величин X и Y равен нулю.

Доказательство. $K_{XY} = \alpha_{1,1}(x, y) - m_X m_Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f(x, y) dx dy - m_X m_Y.$

Для независимых величин (см. (8.12)) $f(x, y) = f_X(x) \cdot f_Y(y)$, \Rightarrow

$$\Rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f(x, y) dx dy = \int_{-\infty}^{\infty} x f_X(x) dx \cdot \int_{-\infty}^{\infty} y f_Y(y) dy = m_X m_Y \text{ и } K_{XY} = m_X m_Y - m_X m_Y = 0.$$

3. Абсолютная величина корреляционного момента двух случайных величин не превышает среднего геометрического их дисперсий

$$|K_{xy}| \leq \sqrt{D_x \cdot D_y} \text{ или } |K_{xy}| \leq \sigma_x \cdot \sigma_y. \quad (9.6)$$

Доказательство. Приведено в лекции 11.

Если $K_{XY} < 0$, то между величинами X и Y существует отрицательная корреляционная зависимость, т.е. чем больше значение одной величины, тем более вероятны меньшие значения у другой (см. статистическую зависимость в лекции 8). *Пример.* X – число пропусков занятий студента, Y – оценка на экзамене.

Если $K_{XY} > 0$, то между величинами X и Y существует положительная корреляционная зависимость, т.е. чем больше значение одной величины, тем более вероятны большие значения у другой. *Пример.* X и Y – рост и вес наугад взятого студента.

Если $K_{XY} = 0$, то величины X и Y называются корреляционно независимыми или *некоррелированными*, т.е. между ними отсутствует зависимость *линейного* характера.

Если $K_{XY} \neq 0$, то величины X и Y называются *коррелированными*.

Итак, из коррелированности двух случайных величин следует их зависимость, но из зависимости еще не вытекает их коррелированность, так как зависимость может иметь и нелинейный характер. Из независимости случайных величин (см. критерии независимости (8.11 – 8.13)) обязательно следует их некоррелированность, но из некоррелированности не всегда следует независимость этих величин.

Величина ковариации K_{XY} зависит от дисперсии случайных величин X , Y , т.е. от рассеивания их значений относительно точки (m_X, m_Y) , поэтому для того, чтобы получить характеристику только степени тесноты линейной зависимости, корреляционный момент нормируется. Эта числовая характеристика называется коэффициентом корреляции.

Коэффициент корреляции R_{XY} характеризует степень линейной зависимости величин и равен:

$$R_{XY} = \frac{K_{XY}}{\sqrt{D_X D_Y}} = \frac{K_{XY}}{\sigma_X \sigma_Y}. \quad (9.7)$$

Свойства коэффициента корреляции:

1. Абсолютная величина коэффициента корреляции двух случайных величин не превышает единицы: $|R_{XY}| \leq 1$.

Доказательство. Разделим обе части двойного неравенства (9.6) $-\sigma_X\sigma_Y \leq K_{XY} \leq \sigma_X\sigma_Y$ на произведение положительных чисел $\sigma_X\sigma_Y \geq 0$ и получим $-1 \leq R_{XY} \leq 1 \Rightarrow |R_{XY}| \leq 1$.

2. $|R_{XY}| = 1$, если величины X, Y связаны линейной функциональной зависимостью $Y = aX + b$.

Доказательство:

$$K_{XY} = \mu_{1,1}(x, y) = M[(X - M[X])(aX + b - M[aX + b])] = \\ = M[(X - m_X)(aX + b - am_X - b)] = aM[(X - m_X)^2] = aD_X.$$

Найдем дисперсию Y : $D[Y] = D[aX + b] = a^2 D_X$, т.е. коэффициент корреляции: $R_{XY} = \frac{K_{XY}}{\sqrt{D_X D_Y}} = \frac{aD_X}{\sqrt{D_X a^2 D_X}} = \frac{a}{|a|} \Rightarrow |R_{XY}| = 1$.

Чем больше абсолютная величина коэффициента корреляции, тем ближе статистическая зависимость величины X, Y к линейной функциональной зависимости.

3. Если величины X и Y независимы, то $R_{XY} = 0$.

Условные числовые характеристики

Для зависимых двумерных величин могут быть определены условные законы распределения (см. (8.14 – 8.17)). Эти законы распределения обладают всеми свойствами безусловных законов, и на их основе по известным формулам (5.1 – 5.4) (после замены в них безусловных законов на условные) могут быть вычислены числовые характеристики, которые называются условными. Наибольшее практическое значение имеют условные математические ожидания.

Условным математическим ожиданием случайной величины X называется ее математическое ожидание, вычисленное при условии, что случайная величина Y приняла определенное значение $Y = y$:

$$M[X / y] = m_{X/y} = \begin{cases} \sum_{i=1}^n x_i \cdot p_{i/j}, & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} x \cdot f(x / y) dx, & \text{для НСВ.} \end{cases} \quad (9.8)$$

Аналогично и для

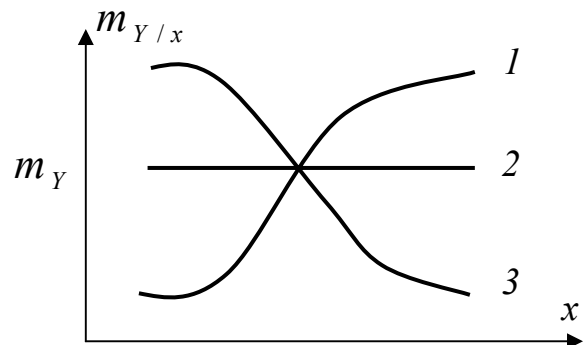
$$M[Y / x] = m_{Y/x} = \begin{cases} \sum_{j=1}^n y_j \cdot p_{j/i}, & \text{для ДСВ,} \\ \int_{-\infty}^{\infty} y \cdot f(y / x) dy, & \text{для НСВ.} \end{cases} \quad (9.9)$$

Условное математическое ожидание $m_{X/y}$ называется регрессией X на y , а условное математическое ожидание $m_{Y/x}$ – регрессией Y на x . Очевидно, что условные математические ожидания представляют собой некоторые функции, которых зависят от значения, взятого в условии, т.е. $m_{X/y} = \psi(y)$, а $m_{Y/x} = \varphi(x)$.

Графики этих зависимостей называются линиями регрессии (см. рисунок).

Линия регрессии 1 указывает, что между величинами X , Y существует положительная корреляционная зависимость, так как при увеличении значения x более вероятны большие значения Y (среднее значение Y увеличивается), т.е. $K_{XY} > 0$. Линия регрессии 2 указывает, что величины X , Y независимы, а линия регрессии 3 – что между величинами X , Y существует отрицательная корреляционная зависимость, т.е. $K_{XY} < 0$.

Регрессионный анализ позволяет выявить характер связи между величинами X , Y . Величины X , Y называются *линейно коррелированными*, если линии регрессии являются прямыми. Уравнения прямых регрессии имеют вид:



$$m_{Y/x} = m_Y + R_{XY} \frac{\sigma_Y}{\sigma_X} (x - m_X), \quad m_{X/y} = m_X + R_{XY} \frac{\sigma_X}{\sigma_Y} (y - m_Y). \quad (9.10)$$

Обе прямые проходят через точку (m_X, m_Y) , которую называют центром совместного распределения величин X и Y .

ЛЕКЦИЯ 10

Нормальный закон распределения на плоскости

Непрерывная двумерная случайная величина (X, Y) имеет **нормальное распределение**, если ее плотность вероятности равна:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-R_{XY}^2}} e^{-\frac{1}{2(1-R_{XY}^2)}\left[\frac{(x-m_X)^2}{2\sigma_X^2} - \frac{2R_{XY}(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{2\sigma_Y^2}\right]}, \quad (10.1)$$

где $m_X, m_Y, \sigma_X, \sigma_Y, R_{XY}$ – параметры распределения.

Если составляющие X, Y двумерной нормально распределенной случайной величины некоррелированы, то они и независимы, т.е. при $R_{XY} = 0$

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} e^{-\frac{1}{2}\left[\frac{(x-m_X)^2}{\sigma_X^2} + \frac{(y-m_Y)^2}{\sigma_Y^2}\right]} = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} \cdot \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{(y-m_Y)^2}{2\sigma_Y^2}} = f_X(x)f_Y(y).$$

Итак, для нормальных случайных величин понятия независимости и некоррелированности равносильны.

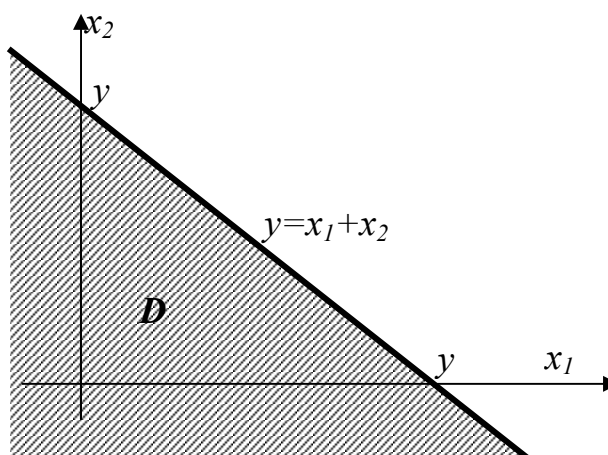
Закон распределения функции двух случайных величин

Рассмотрим функцию двух случайных аргументов $Y = \varphi(X_1, X_2)$. Необходимо определить закон распределения случайной величины Y по известному закону распределения двумерной случайной величины (X_1, X_2) и виду преобразования φ . Функция распределения $G(y)$ величины Y определяется по формуле

$$G(y) = p(Y < y) = p(\varphi(x_1, x_2) < y) = \iint_{(D)} f(x_1, x_2) dx_1 dx_2, \quad (10.2)$$

где $f(x_1, x_2)$ – совместная плотность вероятности величин X_1 и X_2 .

В формуле (10.2) интегрирование производится по области D , которая определяется из условия $\varphi(X_1, X_2) < y$. Форма области D зависит от вида функции $\varphi(X_1, X_2)$. В случае, когда $Y = X_1 + X_2$, область интегрирования имеет вид, показанный на рисунке, и функция распределения суммы двух случайных величин определяется по формуле



$$G(y) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{y-x_1} f(x_1, x_2) dx_2 \right\} dx_1 = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{y-x_2} f(x_1, x_2) dx_1 \right\} dx_2. \quad (10.3)$$

Дифференцируя это выражение по y , получим плотность распределения величины Y :

$$g(y) = \int_{-\infty}^{\infty} f(x_1, y - x_1) dx_1 = \int_{-\infty}^{\infty} f(y - x_2, x_2) dx_2. \quad (10.4)$$

Если величины X_1 и X_2 независимы, то

$$g(y) = \int_{-\infty}^{\infty} f_1(x_1) f_2(y - x_1) dx_1 = \int_{-\infty}^{\infty} f_1(y - x_2) f_2(x_2) dx_2. \quad (10.5)$$

В случае, когда складываются независимые случайные величины, говорят о *композиции законов распределения*. Произвести композицию двух законов распределения — это значит найти закон распределения суммы двух независимых случайных величин, распределенных по этим законам (см. (10.5)).

Многомерные случайные величины

Совокупность произвольного числа n одномерных случайных величин X_i , $i = 1, \dots, n$, которые принимают значение в результате проведения одного и того же опыта, называется n -мерной случайной величиной (X_1, X_2, \dots, X_n) . Ее можно интерпретировать как случайную точку или случайный вектор в n -мерном пространстве.

Полной характеристикой n -мерной случайной величины (X_1, X_2, \dots, X_n) является n -мерный закон распределения, который может быть задан функцией распределения или плотностью вероятности.

Функцией распределения n -мерной случайной величиной (X_1, X_2, \dots, X_n) называется вероятность выполнения n неравенств вида $X_i < x_i$:

$$F(x_1, x_2, \dots, x_n) = P\{(X_1 < x_1)(X_2 < x_2) \dots (X_n < x_n)\}. \quad (10.6)$$

Функцию распределения любой частной системы из величин, входящих в систему, можно получить, если положить все остальные аргументы n -мерной функции распределения равными бесконечности.

Плотностью распределения n -мерной случайной величиной (X_1, X_2, \dots, X_n) называется n -я смешанная частная производная функции $F(x_1, x_2, \dots, x_n)$, взятая один раз по каждому аргументу:

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (10.7)$$

Она обладает следующими свойствами:

$$1. f(x_1, \dots, x_n) \geq 0.$$

2. Условие нормировки:

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1 \quad (10.8)$$

3. Плотности распределения меньшего порядка могут быть получены путем интегрирования n -мерной плотности распределения по ненужным переменным. Например, одномерная плотность распределения величины X_k равна

$$f_k(x_k) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_1(x_1, \dots, x_n) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n. \quad (10.9)$$

4. Вероятность попадания случайной точки (X_1, X_2, \dots, X_n) в пределы n -мерной области D равна n -кратному интегралу по этой области:

$$p\{(X_1, X_2, \dots, X_n) \in D\} = \int \dots \int_{(D)} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n. \quad (10.10)$$

Случайные величины (X_1, X_2, \dots, X_n) называются *независимыми*, если закон распределения каждой частной системы, выделенной из системы (X_1, X_2, \dots, X_n) , не зависит от того, какие значения приняли остальные случайные величины.

Плотность распределения системы независимых случайных величин равна произведению плотностей распределения отдельных величин, входящих в систему: $f(x_1, x_2, \dots, x_n) = f_1(x_1) \cdot f_2(x_2) \dots f_n(x_n)$.

Основные *числовые характеристики* n -мерной случайной величиной (X_1, X_2, \dots, X_n) следующие.

1. Вектор математических ожиданий $\mathbf{M} = (m_1, m_2, \dots, m_n)$:

$$m_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i f(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (10.11)$$

2. Вектор дисперсий $\mathbf{D} = (D_1, D_2, \dots, D_n)$:

$$D_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - m_i)^2 f(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (10.12)$$

3. Корреляционная матрица, характеризующая попарную корреляцию всех величин, входящих в систему:

$$\|K_{ij}\| = \begin{vmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ K_{21} & K_{22} & \dots & K_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{vmatrix},$$

где

$$K_{ij} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - m_i)(x_j - m_j) f(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (10.13)$$

Данная матрица является симметричной ($K_{ij} = K_{ji}$) и включает в себя вектор дисперсий, так как $K_{ii} = D_i$

4. Матрица коэффициентов корреляции:

$$\|R_{ij}\| = \begin{vmatrix} 1 & R_{12} & \dots & R_{1n} \\ R_{21} & 1 & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & 1 \end{vmatrix},$$

где

$$R_{ij} = \frac{K_{ij}}{\sqrt{D_i D_j}}. \quad (10.14)$$

Матрица квадратная и симметричная.

ЛЕКЦИЯ 11

Числовые характеристики функции многих переменных

Пусть $Y = \varphi(x_1, x_2, \dots, x_n)$, где X_1, X_2, \dots, X_n – случайные величины с известной совместной n -мерной плотностью вероятностей $f(x_1, x_2, \dots, x_n)$.

Начальные моменты величины Y определяются по формуле

$$\alpha_k(y) = M[Y^k] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi^k(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n, \quad (11.1)$$

а центральные моменты – по формуле

$$\mu_k(y) = M[(Y - m_Y)^k] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\varphi(x_1, \dots, x_n) - m_Y)^k f(x_1, \dots, x_n) dx_1 \dots dx_n, \quad (11.2)$$

причем

$$m_Y = M[Y] = M[\varphi(x_1, \dots, x_n)] = \alpha_1(y), \quad (11.3)$$

$$D_Y = \mu_2(y) = \alpha_2(y) - m_Y^2. \quad (11.4)$$

В случае, когда совместная плотность вероятности аргументов $f(x_1, x_2, \dots, x_n)$ неизвестна, а известны числовые характеристики аргументов, то задача определения числовых характеристик Y разрешима только для определения классов функций φ .

Числовые характеристики суммы случайных величин

Пусть $Y = \sum_{i=1}^n X_i$, где X_1, X_2, \dots, X_n – случайные величины с известными числовыми характеристиками:

- вектор математических ожиданий $\mathbf{M} = (m_1, m_2, \dots, m_n)$;
- вектор дисперсий $\mathbf{D} = (D_1, D_2, \dots, D_n)$;
- корреляционная матрица $\|K_{ij}\|$.

Теорема о математическом ожидании суммы случайных величин. Математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых:

$$M[Y] = M\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n m_i. \quad (11.5)$$

Доказательство. Пусть $n = 2$, т.е. $Y = X_1 + X_2$, и предположим, что слагаемые – непрерывные случайные величины с некоторой совместной плотностью распределения $f(x_1, x_2)$. Тогда

$$\begin{aligned} m_Y &= M[X_1 + X_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f(x_1, x_2) dx_1 dx_2 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f(x_1, x_2) dx_1 dx_2 = m_1 + m_2. \end{aligned}$$

Аналогично и для дискретных слагаемых. Используя метод математической индукции, легко доказать, что теорема справедлива для любого n .

Теорема о дисперсии суммы случайных величин. Дисперсия суммы случайных величин равна сумме всех элементов корреляционной матрицы слагаемых:

$$D[Y] = D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \sum_{j=1}^n K_{ij} = \sum_{i=1}^n D_i + 2 \sum_{i=1}^n \sum_{j=i+1}^n K_{ij}. \quad (11.6)$$

Доказательство:

$$\begin{aligned} D[Y] &= M[(Y - m_Y)^2] = M\left[\left(\sum_{i=1}^n X_i - \sum_{i=1}^n m_i\right)^2\right] = M\left[\left(\sum_{i=1}^n \overset{\circ}{X}_i\right)^2\right] = \\ &= M\left[\sum_{i=1}^n \sum_{j=1}^n \overset{\circ}{X}_i \overset{\circ}{X}_j\right] = \sum_{i=1}^n \sum_{j=1}^n M\left[\overset{\circ}{X}_i \overset{\circ}{X}_j\right] = \sum_{i=1}^n \sum_{j=1}^n K_{ij} = \sum_{i=1}^n D_i + 2 \sum_{i=1}^n \sum_{j=i+1}^n K_{ij}. \end{aligned}$$

Следствие. Дисперсия суммы некоррелированных случайных величин равна сумме дисперсий этих величин, так как $K_{ij} = 0$, $\forall i, j$:

$$D\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n D_i. \quad (11.7)$$

Если $Y = a_0 + \sum_{i=1}^n a_i X_i$, a_i – неслучайные коэффициенты, то математическое ожидание и дисперсия Y равны:

$$m_Y = M\left[a_0 + \sum_{i=1}^n a_i X_i\right] = a_0 + \sum_{i=1}^n a_i m_i; \quad (11.8)$$

$$D_Y = D\left[a_0 + \sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 D_i + 2 \sum_{i=1}^n \sum_{j=i+1}^n a_i a_j K_{ij}. \quad (11.9)$$

Это легко доказать, используя (11.6), (11.7) и свойства математического ожидания ($M[c] = c$, $M[X + c] = m_X + c$, $M[c \cdot X] = c \cdot m_X$) и дисперсии ($D[c] = 0$, $D[X + c] = D_X$, $D[c \cdot X] = c^2 \cdot D_X$).

Пример. Докажем, что абсолютная величина корреляционного момента двух случайных величин не превышает среднего геометрического их дисперсий:

$$|K_{xy}| \leq \sqrt{D_x \cdot D_y} \quad \text{или} \quad |K_{xy}| \leq \sigma_x \cdot \sigma_y.$$

Введем в рассмотрение случайные величины $Z_1 = \sigma_Y X - \sigma_X Y$, $Z_2 = \sigma_Y X + \sigma_X Y$ и вычислим их дисперсии по формуле (11.9):

$$D[Z_1] = 2\sigma_X^2 \sigma_Y^2 - 2\sigma_X \sigma_Y \cdot K_{XY}; \quad D[Z_2] = 2\sigma_X^2 \sigma_Y^2 + 2\sigma_X \sigma_Y \cdot K_{XY}.$$

Так как дисперсия всегда неотрицательна, то $2\sigma_X^2 \sigma_Y^2 - 2\sigma_X \sigma_Y \cdot K_{XY} \geq 0 \Rightarrow \sigma_X \sigma_Y \geq K_{XY}$ и $2\sigma_X^2 \sigma_Y^2 + 2\sigma_X \sigma_Y \cdot K_{XY} \geq 0 \Rightarrow -\sigma_X \sigma_Y \leq K_{XY}$. Таким образом, $-\sigma_X \sigma_Y \leq K_{XY} \leq \sigma_X \sigma_Y \Rightarrow |K_{XY}| \leq \sigma_X \sigma_Y$.

Числовые характеристики произведения случайных величин

Пусть $Y = \prod_{i=1}^n X_i$, где (X_1, X_2, \dots, X_n) – случайные величины с известными

числовыми характеристиками:

- вектор математических ожиданий $\mathbf{M} = (m_1, m_2, \dots, m_n)$;
- вектор дисперсий $\mathbf{D} = (D_1, D_2, \dots, D_n)$;
- корреляционная матрица $\|K_{ij}\|$.

Теорема о математическом ожидании произведения случайных величин. Математическое ожидание произведения двух случайных величин равно произведению их математических ожиданий плюс ковариация:

$$m_Y = M[X_1 X_2] = m_1 m_2 + K_{12}. \quad (11.10)$$

Доказательство. По определению ковариация равна

$$K_{12} = M[\overset{\circ}{X}_1 \overset{\circ}{X}_2] = M[(X_1 - m_1)(X_2 - m_2)] = M[X_1 X_2 - m_1 X_2 - m_2 X_1 + m_1 m_2] = M[X_1 X_2] - m_1 M[X_2] - m_2 M[X_1] + m_1 m_2 = M[X_1 X_2] - m_1 m_2.$$

Откуда следует формула (11.10).

Следствие. Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий:

$$m_Y = M\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n m_i. \quad (11.11)$$

Доказательство. Пусть $n = 2$. Для независимых случайных величин $K_{ij} = 0, \forall i, j$, тогда формула (11.10) примет вид $m_Y = M[X_1 X_2] = m_1 m_2$. Используя метод математической индукции, легко доказать, что (11.11) справедлива для любого n .

Теорема о дисперсии произведения случайных величин. Дисперсия произведения независимых случайных величин равна

$$D_Y = D\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n (D_i + m_i^2) - \prod_{i=1}^n m_i^2. \quad (11.12)$$

Доказательство. По определению дисперсия равна

$$D\left[\prod_{i=1}^n X_i\right] = M\left[\left(\prod_{i=1}^n X_i\right)^2\right] - \left(M\left[\prod_{i=1}^n X_i\right]\right)^2 = M\left[\prod_{i=1}^n X_i^2\right] - \prod_{i=1}^n m_i^2 = \prod_{i=1}^n (D_i + m_i^2) - \prod_{i=1}^n m_i^2.$$

Следствие. Дисперсия произведения независимых централизованных случайных величин равна произведению дисперсий этих величин:

$$D_Y = D\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n D_i. \quad (11.12)$$

ЛЕКЦИЯ 12

Закон больших чисел

Пусть проводится некоторый опыт, в котором нас интересует значение случайной величины X . При однократном проведении опыта нельзя заранее сказать, какое значение примет величина X . Но при n -кратном ($n > 100 \dots 1000$) повторении «среднее» (среднее арифметическое) значение величины X теряет случайный характер и становится близким к некоторой константе.

Закон больших чисел – совокупность теорем, определяющих условия стремления средних арифметических значений случайных величин к некоторой константе при проведении большого числа опытов.

Неравенство Чебышева. Для любой случайной величины X с математическим ожиданием m_X и дисперсией D_X выполняются следующие неравенства:

$$p(|X - m_X| \geq \varepsilon) \leq \frac{D_X}{\varepsilon^2}, \quad (12.1)$$

где $\varepsilon > 0$.

Доказательство. Рассмотрим вероятность $p(|X| \geq \varepsilon)$:

$$\begin{aligned} p(|X| \geq \varepsilon) &= \int_{|x| \geq \varepsilon} f(x) dx = \int_{|x| \geq \varepsilon} \frac{x^2}{\varepsilon^2} \cdot \frac{\varepsilon^2}{x^2} f(x) dx \leq \int_{|x| \geq \varepsilon} \frac{x^2}{\varepsilon^2} f(x) dx = \frac{1}{\varepsilon^2} \int_{|x| \geq \varepsilon} x^2 f(x) dx \leq \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} x^2 f(x) dx = \frac{M[X^2]}{\varepsilon^2}. \end{aligned}$$

Таким образом, $p(|X| \geq \varepsilon) \leq \frac{M[X^2]}{\varepsilon^2}$. Заменяя нецентрированную величину X на центрированную $\overset{\circ}{X} = X - m_X$, получим

$$p(|X - m_X| \geq \varepsilon) \leq \frac{M[(X - m_X)^2]}{\varepsilon^2} = \frac{D_X}{\varepsilon^2}.$$

Пример. Определим вероятность, что случайная величина примет значение за пределами интервала $3\sigma_X$. Полагаем в неравенстве Чебышева $\varepsilon = 3\sigma_X$, имеем:

$$p(|X - m_X| \geq 3\sigma_X) \leq \frac{D_X}{9\sigma_X^2} = \frac{1}{9}.$$

Неравенство Чебышева дает только верхнюю границу вероятности данного отклонения. Значение вероятности быть выше этой границы (1/9) не может *ни при каком* законе распределения. Таким образом, правило $3\sigma_X$ выполняется с вероятностью не меньшей, чем 8/9.

Сходимость по вероятности. Последовательность случайных величин X_n сходится по вероятности к величине a , $X_n \xrightarrow[n \rightarrow \infty]{p} a$, если при увеличении n

вероятность того, что X_n и a будут сколь угодно близки, неограниченно приближается к единице:

$$p(|X_n - a| < \varepsilon) > 1 - \delta,$$

где ε, δ – произвольно сколь угодно малые положительные числа.

Одна из наиболее важных форм закона больших чисел – теорема Чебышева, она устанавливает связь между средним арифметическим наблюдаемых значений случайной величины и ее математическим ожиданием.

Теорема Чебышева. Пусть произведены n одинаковых независимых опытов, в каждом из которых случайная величина X приняла значения X_1, X_2, \dots, X_n . При достаточно большом числе независимых опытов среднее арифметическое значений случайной величины X сходится по вероятности к ее математическому ожиданию:

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p} m_X. \quad (12.2)$$

Доказательство. Рассмотрим величину $Y = \frac{1}{n} \sum_{i=1}^n X_i$. Определим числовые характеристики Y (см. (11.5), (11.7)):

$$m_Y = M \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n} \sum_{i=1}^n M [X_i] = \frac{1}{n} n m_X = m_X;$$

$$D_Y = D \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n^2} \sum_{i=1}^n D [X_i] = \frac{1}{n^2} n D_X = \frac{D_X}{n}.$$

Запишем неравенство Чебышева для величины Y :

$$p(|Y - m_Y| \geq \varepsilon) = p\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - m_X\right| \geq \varepsilon\right) \leq \frac{D_Y}{\varepsilon^2} = \frac{D_X}{\varepsilon^2 n}.$$

Как бы ни было мало число ε , можно взять n таким большим, чтобы выполнялось неравенство $\frac{D_X}{\varepsilon^2 n} < \delta$, где δ – сколь угодно малое число. Тогда

$$p\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - m_X\right| \geq \varepsilon\right) < \delta. \quad \text{Переходя к противоположному событию}$$

$$p\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - m_X\right| < \varepsilon\right) > 1 - \delta, \text{ т.е. } Y \text{ сходится по вероятности к } m_X.$$

Теорема Бернулли. Пусть произведены n одинаковых независимых опытов, в каждом из которых возможно событие A с вероятностью p . Тогда частота появления события A в n опытах сходится по вероятности к вероятности появления A в одном опыте:

$$p^*(A) \xrightarrow[n \rightarrow \infty]{p} p(A), \quad (12.3)$$

где $p^*(A) = \frac{m}{n}$ – частота события A в n опытах;

m – число опытов в которых произошло событие A ;

n – число проведенных опытов.

Пусть случайная величина X – индикатор события A :

$$X = \begin{cases} 1, & A \\ 0, & \bar{A} \end{cases}$$

тогда X_i – индикатор события A в i -м опыте.

Числовые характеристики индикатора X случайного события (см. (6.1)):

$$m_X = p, D_X = qp,$$

где $q = 1 - p$ – вероятность осуществления \bar{A} .

Применим теорему Чебышева:

$$\frac{1}{n} \sum_{i=1}^n X_i = \frac{m}{n} = p^*(A) \xrightarrow[n \rightarrow \infty]{p} m_X = p = p(A).$$

Центральная предельная теорема

Данная теорема определяет условия, при которых возникает случайная величина с нормальным законом распределения. Различные формы центральной предельной теоремы различаются между собой условиями, накладываемыми на распределения образующих сумму случайных слагаемых X_1, X_2, \dots, X_n . Чем жестче эти условия, тем легче доказывается теорема; чем они шире, тем труднее доказательство. Здесь мы докажем одну из самых простых форм этой теоремы, а именно, центральную предельную теорему для одинаково распределенных слагаемых.

Теорема. Если X_1, X_2, \dots, X_n – независимые случайные величины, имеющие одно и то же распределение с математическим ожиданием m и дисперсией σ^2 , то при неограниченном увеличении n ($n \rightarrow \infty$) закон

распределения их суммы $Y = \sum_{i=1}^n X_i$ неограниченно приближается к нормальному закону с параметрами:

$$m_Y = n \cdot m, \sigma_Y = \sigma \sqrt{n}. \quad (12.4)$$

Доказательство. Проведем доказательство для случая непрерывные случайных величин (для дискретных оно будет аналогичным). Применим для этого аппарат характеристических функций. Случайные величины X_1, X_2, \dots, X_n имеют одну и ту же плотность $f(x)$, а значит, одну и ту же характеристическую функцию $\nu_X(t)$. Не нарушая общности, можно перенести начало отсчета всех случайных величин X_1, X_2, \dots, X_n в их общее математическое ожидание m ; это равносильно их центрированию и, значит, тому, что $m = 0$. Согласно свойству (7.16) характеристическая функция суммы равна произведению характеристических функций слагаемых:

$$\nu_Y(t) = [\nu_X(t)]^n. \quad (12.5)$$

Разложим функцию $\nu_X(t)$ в окрестности точки $t = 0$ в ряд Маклорена с тремя членами:

$$\nu_X(t) = \nu_X(0) + \nu'_X(0)t + [\nu''_X(0)/2 + \alpha(t)]t^2, \quad (12.6)$$

где производные берутся по t ; $\alpha(t) \rightarrow 0$ при $t \rightarrow 0$.

Используя свойство (7.15) характеристических функций определим значения

$$\begin{aligned} \nu_X(0) &= \nu_X^{(0)}(0) = \alpha_0(x)i^0 = 1, \\ \nu'_X(0) &= \nu_X^{(1)}(0) = \alpha_1(x)i^1 = m \cdot i = 0, \\ \nu''_X(0) &= \nu_X^{(2)}(0) = \alpha_2(x)i^2 = \sigma^2 \cdot i^2 = -\sigma^2. \end{aligned}$$

Подставив их в (12.5), получим

$$\nu_X(t) = 1 - [\sigma^2/2 - \alpha(t)]t^2. \quad (12.7)$$

Перейдем от Y к линейно связанной с ней «нормированной» случайной величине $Z = \frac{Y}{\sigma\sqrt{n}}$. Эта величина удобна тем, что ее дисперсия не зависит от n и равна единице при любом. Если мы докажем, что случайная величина Z имеет нормальное распределение, это будет означать, что и случайная величина Y , линейно связанная с Z , распределена нормально.

Вместо того чтобы доказывать, что закон распределения случайной величины Z при увеличении n приближается к нормальному, докажем, что ее характеристическая функция, однозначно определяющая плотность, приближается к характеристической функции нормального закона $e^{-\frac{t^2}{2}}$ с теми же, что у Z , параметрами $m_Z = 0$; $\sigma_Z = 1$ (см. (7.17)).

Найдем характеристическую функцию случайной величины Z . Из свойства (7.14) характеристической функции имеем:

$$\nu_Z(t) = \nu_Y\left(\frac{t}{\sigma\sqrt{n}}\right). \quad (12.8)$$

Подставив (12.5) и (12.7) в (12.8), получим:

$$\nu_Z(t) = \left\{ 1 - \left[\frac{\sigma^2}{2} - \alpha(t/\sigma\sqrt{n}) \right] t^2 / (n\sigma^2) \right\}^n. \quad (12.9)$$

Прологарифмируем это выражение:

$$\ln \nu_Z(t) = n \ln \left\{ 1 - \left[\frac{\sigma^2}{2} - \alpha(t/\sigma\sqrt{n}) \right] t^2 / (n\sigma^2) \right\}.$$

Введем обозначение $\left[\frac{\sigma^2}{2} - \alpha(t/\sigma\sqrt{n}) \right] t^2 / (n\sigma^2) = \varepsilon$, тогда $\ln \nu_Z(t) = n \ln(1 - \varepsilon)$.

Будем неограниченно увеличивать n ; при этом величина ε будет стремиться к нулю. Разложим $\ln(1 - \varepsilon)$ в ряд по степеням ε и ограничимся одним членом

разложения (остальные при $n \rightarrow \infty$ станут пренебрежимо малыми):
 $\ln(1 - \varepsilon) \approx -\varepsilon$.

Тогда

$$\begin{aligned} \lim_{n \rightarrow \infty} \ln v_Z(t) &= \lim_{n \rightarrow \infty} n \cdot (-\varepsilon) = \lim_{n \rightarrow \infty} \left\{ -\frac{t^2}{2} + \alpha(t / \sigma \sqrt{n}) t^2 / \sigma^2 \right\} = \\ &= -\frac{t^2}{2} + \lim_{n \rightarrow \infty} \alpha(t / \sigma \sqrt{n}) t^2 / \sigma^2 = -\frac{t^2}{2}. \end{aligned}$$

Откуда $\lim_{n \rightarrow \infty} v_Z(t) = e^{-\frac{t^2}{2}}$, а это есть не что иное, как характеристическая функция случайной величины, распределенной по нормальному закону с параметрами $m = 0, \sigma = 1$.

Более общую форму центральной предельной теоремы мы приведем без доказательства.

Теорема. Если X_1, X_2, \dots, X_n – независимые случайные величины, имеющие примерно одинаковые дисперсии $D_i \approx D$ для $\forall i$, то при неограниченном увеличении n ($n \rightarrow \infty$) закон распределения их суммы $Y = \sum_{i=1}^n X_i$ неограниченно приближается к нормальному закону с параметрами:

$$m_Y = \sum_{i=1}^n m_i, \quad \sigma_Y = \sqrt{\sum_{i=1}^n D_i}. \quad (12.10)$$

Требование $D_i \approx D, \forall i$ означает, что ни одно из слагаемых не носит доминирующего характера (влияние всех X_i на сумму Y приблизительно одинаково).

Таким образом, нормальное распределение возникает тогда, когда суммируется много независимых (или слабо зависимых) случайных величин, сравнимых по порядку своего влияния на рассеивание суммы. На практике такая обстановка встречается нередко. Пусть рассматривается отклонение Y какого-то параметра, например радиоэлектронного устройства от номинала. Это отклонение (при известных допущениях) может быть представлено как сумма n элементарных отклонений, связанных с отдельными причинами:

$$Y = \sum_{i=1}^n X_i$$

где, например:

X_1 – отклонение, вызванное влиянием температуры;

X_2 – отклонение, вызванное влиянием влажности воздуха;

X_3 – отклонение, вызванное недостаточной чистотой материала изделия;

...

X_n – отклонение, вызванное недостаточной чистотой материала изделия.

Число n этих элементарных отклонений весьма велико, как и число n причин, вызывающих суммарное отклонение Y . Обычно слагаемые X_1, X_2, \dots, X_n сравнимы по порядку своего влияния на рассеивание суммы. Действительно,

если бы какая-то из случайных величин X_1, X_2, \dots, X_n оказывала существенно большее влияние на рассеивание суммы, чем все остальные, было бы естественно принять специальные меры для того, чтобы устранить главную причину рассеивания; поскольку такие меры не предпринимаются, можно предположить, что оставшиеся случайные слагаемые сравнимы по порядку своего (равномерно малого) влияния на рассеивание суммы.

Нормальный закон широко распространен в технике; в большинстве случаев ошибки измерения параметров, ошибки выполнения команд, ошибки ввода различных величин в техническое устройство распределены по нормальному (или близкому к нормальному) закону; такая ошибка обычно может быть представлена в виде суммы многих «элементарных ошибок» X_i , каждая из которых связана с отдельной, практически независимой от других причиной. Именно в применении к теории ошибок был впервые обоснован Лапласом и Гауссом нормальный закон.

На практике при суммировании величин с одинаковым законом распределения закон распределения суммы можно считать нормальным, если $n > 10 \dots 20$.

Пример 1. Пусть X – случайная величина, равномерно распределенная на интервале $[0, 1]$, и формируется, например, генератором псевдослучайных величин. На основании центральной предельной теоремы величина

$$Y = \sigma \left(\sum_{i=1}^{12} x_i - 6 \right) + m \quad (12.11)$$

будет иметь практически нормальный закон распределения с параметрами m, σ .

Пример 2. Докажем теоремы Муавра–Лапласа (см. (3.11), (3.12)), т.е. что число появлений событий A в n независимых одинаковых опытах распределено по нормальному закону, если количество опытов n велико, а вероятности p и q не малы, так что выполняются следующие условия:

$$0 < np - 3\sqrt{npq}, np + 3\sqrt{npq} < n. \quad (12.12)$$

Пусть X_i – индикатор события A в i -м опыте, тогда число появлений событий A в n опытах равно $Y = \sum_{i=1}^n X_i$, причем $0 \leq Y \leq n$. На основании

центральной предельной теоремы величина Y будет иметь практически нормальный закон распределения с параметрами

$$m_Y = \sum_{i=1}^n m_i = np, \quad \sigma_Y = \sqrt{\sum_{i=1}^n D_i} = \sqrt{npq}. \quad \text{Условие (12.12) получено на основании}$$

правила $3\sigma_Y$ для величины Y , так чтобы практически все значения нормальной величины Y находились в интервале $[0; n]$.

ЛЕКЦИЯ 13

Математическая статистика. Основные понятия

Математической статистикой называется наука, занимающаяся методами обработки опытных данных, полученных в результате наблюдений над случайными явлениями. Любой такой результат можно представить как совокупность значений, принятых в результате n опытов случайной одномерной или многомерной величиной.

Генеральной совокупностью опыта называется множество объектов, из которых производится выборка. Каждый из объектов задает фиксированное значение случайной величины X . Количество (N) входящих в генеральную совокупность объектов называют *объемом генеральной совокупности*. Она может состоять из бесчисленного множества объектов.

Выборка – множество $\{x_1, x_2, \dots, x_n\}$ случайно отобранных объектов (значений) из генеральной совокупности. *Объемом* выборки n называется число входящих в нее объектов. К выборке предъявляется требование, чтобы она адекватно представляла генеральную совокупность, т.е. была *репрезентативной* (*представительной*). В силу закона больших чисел можно утверждать, что выборка будет *репрезентативной*, если ее осуществлять случайно, т.е. каждый из объектов генеральной совокупности имеет одинаковую вероятность попасть в выборку. Очевидно, что можно осуществить в одинаковых условиях k выборок объема n и получить различные совокупности значений случайной величины X : $\{x_1^{(1)}, \dots, x_n^{(1)}\}, \{x_1^{(2)}, \dots, x_n^{(2)}\}, \dots, \{x_1^{(k)}, \dots, x_n^{(k)}\}$. Пусть для генеральной совокупности опыта случайная величина X имеет функцию распределения $F(x)$, тогда каждую из выборок $\{x_1^{(1)}, \dots, x_n^{(1)}\}, \{x_1^{(2)}, \dots, x_n^{(2)}\}, \dots, \{x_1^{(k)}, \dots, x_n^{(k)}\}$ можно рассматривать, как реализацию n -мерной случайной величины (X_1, \dots, X_n) , где составляющая X_i , $i = 1, \dots, n$, есть значение величины X в i -м опыте. Очевидно, что все составляющие X_i будут иметь одинаковый закон распределения $F(x)$. Так как компоненты X_i независимы, то функция распределения n -мерной случайной величины (X_1, X_2, \dots, X_n) определяется формулой

$$F(x_1, x_2, \dots, x_n) = F(x_1) F(x_2) \dots F(x_n).$$

Вариационным рядом называется выборка $\{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n\}$, полученная в результате расположения значений исходной выборки в порядке возрастания. Значения \hat{x}_i называются вариантами.

Одной из главных задач математической статистики является определение закона распределения случайной величины X .

Оценка закона распределения

Эмпирическая функция распределения случайной величины X равна частоте того, что X примет значение меньшее, чем аргумент функции x , и определяется формулой

$$F^*(x) = p^*(X < x) = \begin{cases} 0, & x \leq \hat{x}_1, \\ \vdots \\ \frac{i}{n}, & \hat{x}_i < x \leq \hat{x}_{i+1}, \\ \vdots \\ 1, & x > \hat{x}_n. \end{cases} \quad (13.1)$$

При $n \rightarrow \infty$ эмпирическая функция распределения $F^*(x)$ сходится по вероятности к теоретической функции распределения $F(x)$.

Основные свойства функции $F^*(x)$.

1. $0 \leq F^*(x) \leq 1$.
2. $F^*(x)$ – неубывающая ступенчатая функция.
3. $F^*(x) = 0$, для $x \leq \hat{x}_1$.
4. $F^*(x) = 1$, для $x > \hat{x}_n$.

Эмпирическая функция распределения является наилучшей оценкой закона распределения (несмещенной, состоятельной, эффективной). Недостаток функции $F^*(x)$ заключается в ее невысокой наглядности: визуально сложно определить закон распределения случайной величины X .

Статистический ряд распределения вероятностей определяется по исходной выборке объемом n , если анализируемая случайная величина X является дискретной с известным множеством значений $\{x_1, x_2, \dots, x_m\}$:

x_j	x_1	x_2	x_m
p_j^*	k_1	k_2	k_m

Здесь $p_j^* = \frac{k_j}{n}$ – частота появления j -го значения;

k_j – число значений x_j в выборке.

Свойства статистического ряд распределения вероятностей:

1. $\sum_{j=1}^m k_j = n$.
2. $\sum_{j=1}^m p_j^* = 1$.

Интервальный статистический ряд вероятностей строится по исходной выборке, если анализируемая случайная величина X является непрерывной:

j	A_j	B_j	h_j	v_j	p_j^*	f_j^*
1	A_1	B_1	h_1	v_1	p_1^*	f_1^*
•						
M	A_M	B_M	h_M	v_M	p_M^*	f_M^*

Здесь j – номер интервала;

M – число непересекающихся и примыкающих друг к другу интервалов, на которые разбивается диапазон значений $[\hat{x}_1, \hat{x}_n]$:

$$M \approx \begin{cases} \text{int}(\sqrt{n}), n \leq 100, \\ \text{int}((2 \div 4) \cdot \lg(n)), n > 100, \end{cases} \quad (13.2)$$

где $\text{int}(x)$ – целая часть числа x . Желательно, чтобы n без остатка делилось на M ;

A_j, B_j – левая и правая границы j -го интервала ($A_{j+1} = B_j$), причем

$$A_1 = \hat{x}_1, B_M = \hat{x}_n;$$

$h_j = B_j - A_j$ – длина j -го интервала;

v_j – количество чисел в выборке, попадающих в j -й интервал;

$$p_j^* = \frac{v_j}{n} \text{ – частота попадания в } j\text{-й интервал;}$$

$$f_j^* = \frac{p_j^*}{h_j} = \frac{v_j}{nh_j} \text{ – статистическая плотность вероятности в } j\text{-м интервале.}$$

При построении интервального статистического ряда вероятностей используют следующие методы разбиения диапазона значений на интервалы:

1) **равноинтервальный**, т.е. все интервалы одинаковой длины:

$$h_j = h = \frac{\hat{x}_n - \hat{x}_1}{M}, \quad \forall j; \quad (13.3)$$

$$A_j = \hat{x}_1 + (j-1)h, \quad j = \overline{2, M}. \quad (13.4)$$

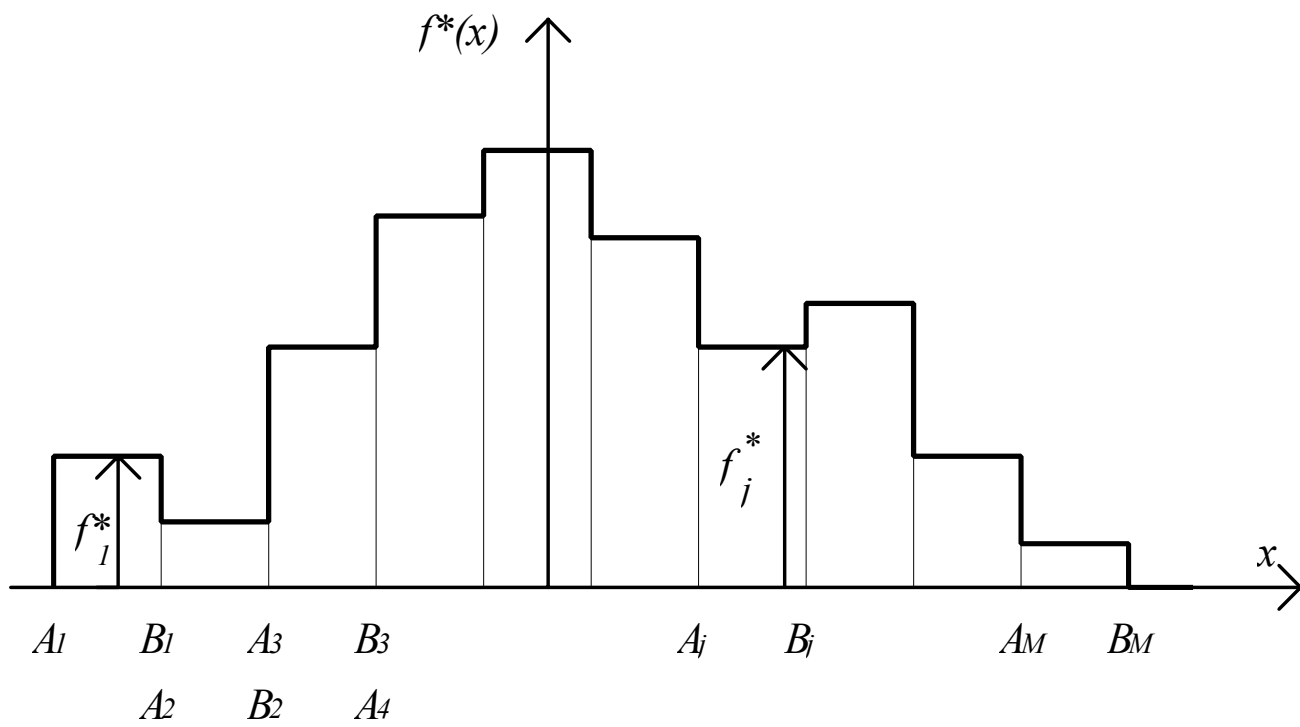
2) **равновероятностный**, т.е. границы интервалов выбирают так, чтобы в каждом интервале было одинаковое число выборочных значений (необходимо, чтобы n без остатка делилось на M):

$$v_j = v = \frac{n}{M}, \quad p_j^* = \frac{1}{M} \quad \forall j; \quad (13.5)$$

$$A_j = \frac{\hat{x}_{(j-1)v} + \hat{x}_{(j-1)v+1}}{2}, \quad j = \overline{2, M}. \quad (13.6)$$

Гистограмма – статистический аналог графика плотности вероятности $f^*(x)$ случайной величины, и она строится по интервальному статистическому ряду. Гистограмма представляет собой совокупность прямоугольников, построенных, как на основаниях, на интервалах h_j статистического ряда с высотой равной статистической плотности вероятности f_j^* в соответствующем интервале.

Для равноинтервального метода все прямоугольники гистограммы имеют одинаковую ширину, а для равновероятностного метода – одинаковую площадь. Сумма площадей всех прямоугольников гистограммы равна 1. Достоинства гистограммы: простота построения, высокая наглядность.



ЛЕКЦИЯ 14

Точечные оценки числовых характеристик

Статистической оценкой \hat{Q} параметра Q распределения называется приближенное значение параметра, вычисленное по результатам эксперимента (по выборке). Статистические оценки делятся на точечные и интервальные.

Точечной называется оценка, определяемая одним числом. Точечная оценка \hat{Q} параметра Q случайной величины X в общем случае равна

$$\hat{Q} = \varphi(x_1, x_2, \dots, x_n), \quad (14.1)$$

где x_i – значения выборки.

Очевидно, что оценка \hat{Q} – это случайная величина, так как она является функцией от n -мерной случайной величины (X_1, X_2, \dots, X_n) , где X_i – значение величины X в i -м опыте, и значения \hat{Q} будут изменяться от выборки к выборке случайным образом. К оценкам предъявляется ряд требований.

1. Оценка \hat{Q} называется **состоятельной**, если при увеличении объема выборки n она сходится по вероятности к значению параметра Q :

$$\hat{Q} \xrightarrow[n \rightarrow \infty]{p} Q \Rightarrow \lim_{n \rightarrow \infty} (P(|\hat{Q} - Q| < \varepsilon)) = 1, \forall \varepsilon > 0. \quad (14.2)$$

Состоятельность – это минимальное требование к оценкам.

2. Оценка \hat{Q} называется **несмещенной**, если ее математическое ожидание точно равно параметру Q для любого объема выборки:

$$M[\hat{Q}] = Q, \forall n. \quad (14.3)$$

Несмещенная оценка является состоятельной, если $\lim_{n \rightarrow \infty} D[\hat{Q}] = 0$.

3. Несмещенная оценка \hat{Q} является **эффективной**, если ее дисперсия минимальна по отношению к дисперсии любой другой оценки этого параметра:

$$D[\hat{Q}] = \min. \quad (14.4)$$

Оценка математического ожидания. На основании теоремы Чебышева в качестве состоятельной оценки математического ожидания может быть использовано среднее арифметическое значений выборки \bar{x} , называемое выборочным средним:

$$m_X^* = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (14.5)$$

Определим числовые характеристики оценки \bar{x} .

$$M[\bar{x}] = M\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n M[X_i] = \frac{1}{n} \sum_{i=1}^n m_X = m_X,$$

т.е. оценка несмещенная.

$$D[\bar{x}] = D\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n^2} \sum_{i=1}^n D[X_i] = \frac{1}{n^2} \sum_{i=1}^n D_X = \frac{1}{n} D_X. \quad (14.6)$$

Оценка (14.5) является эффективной, т.е. ее дисперсия минимальна, если величина X распределена по нормальному закону.

Состоятельная **оценка начального момента k -го** порядка определяется по формуле

$$\hat{\alpha}_k(x) = \frac{1}{n} \cdot \sum_{i=1}^n (x_i)^k. \quad (14.7)$$

Оценка дисперсии. В качестве состоятельной оценки дисперсии может быть использовано среднее арифметическое квадратов отклонений значений выборки от выборочного среднего:

$$S^2 = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2. \quad (14.8)$$

Определим математическое ожидание оценки S^2 . Так как дисперсия не зависит от того, где выбрать начало координат, выберем его в точке m_X , т.е. перейдем к центрированным величинам:

$$\begin{aligned} M[S^2] &= M\left[\frac{1}{n} \sum_{i=1}^n \overset{\circ}{X}_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \overset{\circ}{X}_i\right)^2\right] = M\left[\frac{n-1}{n^2} \sum_{i=1}^n \overset{\circ}{X}_i^2 - \frac{2}{n^2} \sum_{i<j} \overset{\circ}{X}_i \overset{\circ}{X}_j\right] = \\ &= \frac{n-1}{n^2} \sum_{i=1}^n M[\overset{\circ}{X}_i^2] - \frac{2}{n^2} \sum_{i<j} M[\overset{\circ}{X}_i \overset{\circ}{X}_j] = \frac{n-1}{n^2} \sum_{i=1}^n D_X - \frac{2}{n^2} \sum_{i<j} K_{ij} = \frac{n-1}{n} D_X. \end{aligned}$$

Ковариация $K_{ij}=0$, так как опыты, а следовательно, и X_i – значение величины X в i -м опыте – независимы. Таким образом, величина S^2 является смещенной оценкой дисперсии, а несмещенная состоятельная оценка дисперсии равна:

$$S_0^2 = \frac{n}{n-1} S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n x_i^2 - \frac{n}{n-1} \bar{x}^2. \quad (14.9)$$

Дисперсия величины S_0^2 равна:

$$D[S_0^2] = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)} D^2. \quad (14.10)$$

Для нормального закона распределения величины X формула (14.10) примет вид

$$D[S_0^2] = \frac{2}{(n-1)} D^2, \quad (14.11)$$

для равномерного закона распределения –

$$D[S_0^2] \approx \frac{0.8n + 1.2}{n(n-1)} D^2. \quad (14.12)$$

Состоятельная несмещенная оценка **среднеквадратического отклонения** определяется по формуле

$$S_0 = \sqrt{S_0^2}. \quad (14.13)$$

Состоятельная оценка **центрального момента k -го** порядка равна:

$$\hat{\mu}_k(x) = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^k. \quad (14.14)$$

Оценка вероятности. На основании теоремы Бернулли несмещенная состоятельная и эффективная оценка вероятности случайного события A в схеме независимых опытов равна частоте этого события:

$$p^*(A) = \frac{m}{n}, \quad (14.15)$$

где m – число опытов, в которых произошло событие A ;

n – число проведенных опытов.

Числовые характеристики оценки вероятности $p^*(A) = p^*$ равны:

$$M[p^*] = p(A) = p, \quad D[p^*] = \frac{p(1-p)}{n}. \quad (14.16)$$

Оценка параметров распределения

Для вычисления оценок параметров распределения чаще всего применяются методы моментов и максимального правдоподобия.

Метод моментов. Пусть имеется выборка $\{x_1, \dots, x_n\}$ независимых значений случайной величины с известным законом распределения $f(x, Q_1, \dots, Q_m)$ и m неизвестными параметрами Q_1, \dots, Q_m . Необходимо вычислить оценки $\hat{Q}_1, \dots, \hat{Q}_m$ параметров Q_1, \dots, Q_m . Последовательность вычислений следующая:

1. Вычислить значения m начальных и/или центральных теоретических моментов

$$\alpha_k(x) = M[X^k], \quad \mu_k(x) = M[(X - m_x)^k]. \quad (14.17)$$

2. Определить m соответствующих выборочных начальных $\hat{\alpha}_k(x)$ и/или центральных $\hat{\mu}_k(x)$ моментов по формулам (14.7, 14.14).

3. Составить и решить относительно неизвестных параметров Q_1, \dots, Q_m систему из m уравнений, в которых теоретические моменты приравниваются к выборочным моментам. Каждое уравнение имеет вид $\alpha_k(x) = \hat{\alpha}_k(x)$ или $\mu_k(x) = \hat{\mu}_k(x)$. Найденные корни являются оценками $\hat{Q}_1, \dots, \hat{Q}_m$ неизвестных параметров.

Замечание. Часть уравнений может содержать начальные моменты, а оставшаяся часть – центральные.

Метод максимального правдоподобия. Согласно данному методу оценки $\hat{Q}_1, \dots, \hat{Q}_m$ получаются из условия максимума по параметрам Q_1, \dots, Q_m положительной функции правдоподобия $L(x_1, \dots, x_n, Q_1, \dots, Q_m)$.

Если случайная величина X непрерывна, а значения x_i независимы, то

$$L(x_1, \dots, x_n, Q_1, \dots, Q_m) = \prod_{i=1}^n f(x_i, Q_1, \dots, Q_m).$$

Если случайная величина X дискретна и принимает независимые значения x_i с вероятностями $p(X = x_i) = p_i(x_i, Q_1, \dots, Q_m)$, то функция правдоподобия равна

$$L(x_1, \dots, x_n, Q_1, \dots, Q_m) = \prod_{i=1}^n p_i(x_i, Q_1, \dots, Q_m).$$

Система уравнений согласно этому методу может записываться в двух видах:

$$\frac{\partial L(x_1, \dots, x_n, Q_1, \dots, Q_m)}{\partial Q_i} = 0, \quad i = 1, 2, \dots, m, \quad (14.18)$$

или

$$\frac{\partial \ln(L(x_1, \dots, x_n, Q_1, \dots, Q_m))}{\partial Q_i} = 0, \quad i = 1, 2, \dots, m. \quad (14.19)$$

Найденные корни выбранной системы уравнений являются оценками $\hat{Q}_1, \dots, \hat{Q}_m$ неизвестных параметров Q_1, \dots, Q_m .

Интервальные оценки числовых характеристик

Пусть для параметра Q получена из опыта несмещенная оценка \hat{Q} . Оценим возможную ошибку, возникающую при замене параметра Q его оценкой \hat{Q} . Возьмем достаточно большую вероятность γ , такую, что событие с вероятностью γ можно считать практически достоверным, и найдем такое значение ε , для которого

$$P(|\hat{Q} - Q| < \varepsilon) = \gamma. \quad (14.20)$$

Тогда диапазон практически возможных значений ошибки, возникающей при замене Q на \hat{Q} , будет $\pm \varepsilon$; большие по абсолютной величине ошибки будут появляться только с малой вероятностью $\alpha = 1 - \gamma$. Равенство (14.19) означает, что с вероятностью γ неизвестное значение параметра Q попадает в интервал

$$I_\gamma = (\hat{Q} - \varepsilon; \hat{Q} + \varepsilon). \quad (14.21)$$

Доверительным называется интервал I_γ , в который с заданной вероятностью (*надежностью*) γ попадают значения параметра Q . Вероятность γ выбирается близкой к 1: 0,9; 0,95; 0,975; 0,99.

Очевидно, что для построения доверительного интервала должен быть известен закон распределения величины \hat{Q} . Затруднение состоит в том, что закон распределения оценки \hat{Q} зависит от закона распределения величины X и, следовательно, от его неизвестных параметров (в частности и от самого параметра Q). Для решения этой проблемы воспользуемся тем, что величина \hat{Q}

представляет собой, как правило, сумму n независимых одинаково распределенных случайных величин и, согласно центральной предельной теореме, при достаточно большом n ($n > 20 \dots 50$) ее закон распределения можно считать нормальным.

Доверительный интервал для математического ожидания. Интервал I_γ для математического ожидания случайной величины X с неизвестным законом распределения имеет вид

$$\bar{x} - \frac{S_0 \cdot z_\gamma}{\sqrt{n}} < m_X < \bar{x} + \frac{S_0 \cdot z_\gamma}{\sqrt{n}}, \quad (14.22)$$

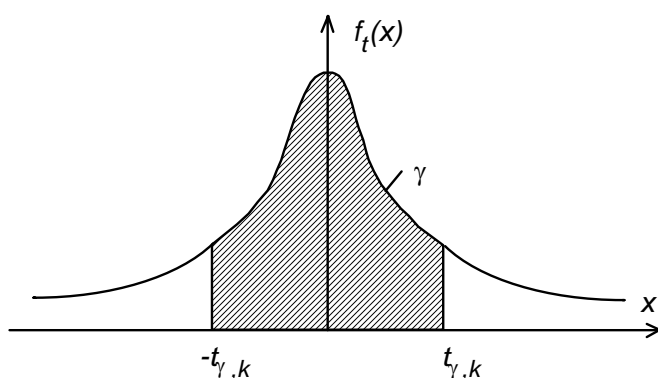
где $z_\gamma = \arg \Phi\left(\frac{\gamma}{2}\right)$ – значение аргумента функции Лапласа, т.е. $\Phi(z_\gamma) = \frac{\gamma}{2}$.

Если случайная величина X распределена по нормальному закону с параметрами m_X и σ_X , то величина

$T = \frac{(\bar{x} - m_X) \sqrt{n}}{S_0}$ распределена по

закону Стюдента с $(n - 1)$ степенью свободы.

Распределение Стюдента с k степенями свободы имеет следующую плотность распределения:



$$f_k(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \cdot \Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \quad (14.23)$$

где $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ – гамма-функция.

Доверительный интервал с надежностью γ для математического ожидания имеет вид

$$\bar{x} - \frac{S_0 \cdot t_{\gamma, n-1}}{\sqrt{n}} < m_X < \bar{x} + \frac{S_0 \cdot t_{\gamma, n-1}}{\sqrt{n}}, \quad (14.24)$$

где $t_{\gamma, n-1}$ – значение, взятое из таблицы распределения Стюдента.

Доверительный интервал для дисперсии. Интервал I_γ для дисперсии случайной величины X с неизвестным законом распределения имеет вид

$$S_0^2 - z_\gamma \sqrt{\frac{2}{n-1}} S_0^2 < D_X < S_0^2 + z_\gamma \sqrt{\frac{2}{n-1}} S_0^2, \quad (14.25)$$

где $z_\gamma = \arg \Phi\left(\frac{\gamma}{2}\right)$ – значение аргумента функции Лапласа, т.е. $\Phi(z_\gamma) = \frac{\gamma}{2}$.

Если случайная величина X распределена по нормальному закону с параметрами m_x и σ_x , то величина $v = \frac{(n-1)S_0^2}{\sigma_x^2}$ распределена по закону χ^2 с $(n-1)$ степенью свободы и доверительный интервал с надежностью γ для дисперсии имеет вид

$$\frac{(n-1)S_0^2}{\chi_{\frac{1-\gamma}{2}, n-1}^2} < D_X < \frac{(n-1)S_0^2}{\chi_{\frac{1+\gamma}{2}, n-1}^2}, \quad (14.26)$$

где $\chi_{\frac{1-\gamma}{2}, n-1}^2, \chi_{\frac{1+\gamma}{2}, n-1}^2$ – значения, взятые из таблицы распределения χ^2 .

Формулы (14.24, 14.26) можно использовать при любом объеме выборки n , так как эти интервалы I_γ построены на основе знания точных законов распределения величин, связывающих Q и \hat{Q} . Кроме этого, если случайная величина X распределена по нормальному закону и ее дисперсия σ_x^2 известна, то точный интервал I_γ для математического ожидания при любом объеме выборки n определяют по формуле (14.22), заменив в ней оценку S_0 СКО его точным значением σ_x .

Доверительный интервал для вероятности. Интервал I_γ для вероятности события A в схеме независимых опытов Бернулли имеет вид

$$p^* - z_\gamma \cdot \sqrt{\frac{p^*(1-p^*)}{n}} < p(A) < p^* + z_\gamma \cdot \sqrt{\frac{p^*(1-p^*)}{n}}, \quad (14.25)$$

где $p^* = p^*(A) = \frac{m}{n}$ – частота появления события A в n опытах;

m – число опытов, в которых произошло событие A ;

n – число проведенных опытов;

$z_\gamma = \arg \Phi\left(\frac{\gamma}{2}\right)$ – значение аргумента функции Лапласа, т.е. $\Phi(z_\gamma) = \frac{\gamma}{2}$.

ЛЕКЦИЯ 15

Проверка статистических гипотез

Статистической гипотезой называется всякое непротиворечивое множество утверждений $\{H_0, H_1, \dots, H_{k-1}\}$ относительно свойств распределения случайной величины. Любое из утверждений H_i называется альтернативой гипотезы. Простейшей гипотезой является двухальтернативная: $\{H_0, H_1\}$. В этом случае альтернативу H_0 называют нулевой гипотезой, а H_1 – конкурирующей гипотезой.

Критерием называется случайная величина $U = \varphi(x_1, \dots, x_n)$, где x_i – значения выборки, которая позволяет принять или отклонить нулевую гипотезу H_0 . Значения критерия, при которых гипотеза H_0 отвергается, образуют критическую область проверяемой гипотезы, а значения критерия, при которых гипотезу принимают, область принятия гипотезы (область допустимых значений). *Критические точки* отделяют критическую область от области принятия гипотезы.

Ошибка первого рода состоит в том, что будет отклонена гипотеза H_0 , если она верна ("пропуск цели"). Вероятность совершить ошибку первого рода обозначается α и называется *уровнем значимости*. Наиболее часто на практике принимают, что $\alpha = 0,05$ или $\alpha = 0,01$.

Ошибка второго рода заключается в том, что гипотеза H_0 принимается, если она неверна ("ложное срабатывание"). Вероятность ошибки этого рода обозначается β . Вероятность не допустить ошибку второго рода ($1 - \beta$) называют *мощностью критерия*. Для нахождения мощности критерия необходимо знать плотность вероятности критерия при альтернативной гипотезе. Простые критерии с заданным уровнем значимости контролируют лишь ошибки первого рода и не учитывают мощность критерия.

Проверка гипотезы о равенстве вероятностей. Пусть произведено две серии опытов, состоящих соответственно из n_1 и n_2 опытов. В каждом из них регистрировалось появление одного и того же события A . В первой серии событие A появилось в k_1 опытах, во второй – в k_2 опытах, причем частота события A в первой серии получилась больше, чем во второй:

$$p_1^* = \frac{k_1}{n_1} > p_2^* = \frac{k_2}{n_2}. \text{ Разность между двумя частотами получилась равной}$$

$$U = p_1^* - p_2^*. \quad (15.1)$$

Спрашивается, значимо или не значимо это расхождение? Указывает ли оно на то, что в первой серии опытов событие A действительно вероятнее, чем во второй, или расхождение между частотами надо считать случайным?

Выдвинем двухальтернативную гипотезу $\{H_0, H_1\}$, где:

H_0 – различия в вероятностях не существует, т.е. обе серии опытов произведены в одинаковых условиях, а расхождение U объясняется случайными причинами,

H_1 – различие в вероятностях существует, т.е. обе серии опытов произведены не в одинаковых условиях.

В данном случае нуль-гипотеза H_0 состоит в том, что обе серии опытов однородны и что вероятность p появления события A в них одна и та же, приблизительно равная частоте, которая получится, если обе серии смешать в одну: $p \approx p^* = \frac{k_1 + k_2}{n_1 + n_2}$.

При достаточно больших n_1 и n_2 каждая из случайных величин p_1^* и p_2^* распределена практически нормально, с одним и тем же математическим ожиданием $m = p \approx p^*$. Что касается дисперсий D_1 и D_2 в первой и во второй сериях, то они различны и равны соответственно (см. (14.16))

$$D_1 \approx \frac{p_1^*(1-p_1^*)}{n_1}, \quad D_2 \approx \frac{p_2^*(1-p_2^*)}{n_2}.$$

В качестве критерия будем использовать случайную величину $U = p_1^* - p_2^*$, которая также имеет приблизительно нормальное распределение с математическим ожиданием $m_U = 0$ и дисперсией

$$D_U = D_1 + D_2 \approx \frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}, \text{ откуда } \sigma_U = \sqrt{D_U} \approx \sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}.$$

Определим критическую точку U_α для заданного уровня значимости α из уравнения:

$$\alpha = p(U \geq U_\alpha) = 0.5 - \Phi\left(\frac{U_\alpha}{\sigma_U}\right) \text{ т.е. } U_\alpha = \sigma_U \cdot \arg \Phi(0.5 - \alpha).$$

Если значение, вычисленное по формуле (15.1), больше, чем критическое значение, т.е. $U > U_\alpha$, то гипотеза H_0 отклоняется, в противном случае нет оснований ее отклонить.

Критерии согласия

Критериями согласия называются критерии, используемые для проверки гипотез о предполагаемом законе распределения.

Критерий согласия Пирсона (χ^2). Это один из наиболее часто применяемых критериев. Алгоритм проверки следующий.

1. Построить интервальный статистический ряд и гистограмму.
2. По виду гистограммы выдвинуть гипотезу:

H_0 – величина X распределена по такому-то закону: $f(x) = f_0(x)$,

H_1 – величина X не распределена по такому-то закону: $f(x) \neq f_0(x)$,

где $f_0(x)$, $F_0(x)$ – плотность и функция гипотетического закона распределения.

3. Используя метод моментов или максимального правдоподобия, определить оценки неизвестных параметров $\hat{Q}_1, \dots, \hat{Q}_m$ гипотетического закона распределения.

4. Вычислить значение критерия по формуле

$$\chi^2 = n \sum_{j=1}^M \frac{(p_j - p_j^*)^2}{p_j} = \sum_{j=1}^M \frac{(v_j - np_j)^2}{np_j}, \quad (15.2)$$

где p_j – теоретическая вероятность попадания случайной величины в j -й интервал при условии, что гипотеза H_0 верна:

$$p_j = p(A_j \leq X < B_j) = \int_{A_j}^{B_j} f_0(x) dx = F_0(B_j) - F_0(A_j). \quad (15.3)$$

Замечания. При расчете p_1 и p_M в качестве крайних границ первого и последнего интервалов A_1, B_M следует использовать теоретические границы гипотетического закона распределения. Например, для нормального закона $A_1 = -\infty, B_M = +\infty$. После вычисления всех вероятностей p_i проверить, выполняется ли контрольное соотношение

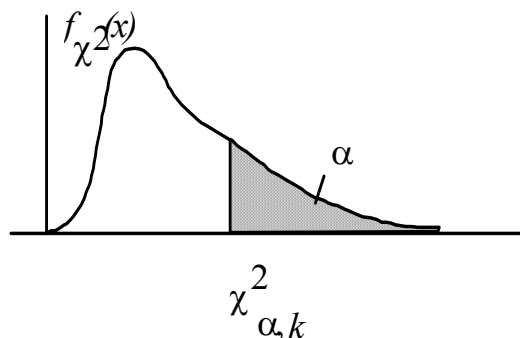
$$\left| 1 - \sum_{j=1}^M p_j \right| \leq 0,01.$$

Величина χ^2 распределена по закону, который называется распределением χ^2 . Данное распределение не зависит от закон распределения величины X , а зависит от параметра k , который называется числом степеней свободы:

$$f_k(u) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} u^{\frac{k}{2}-1} e^{-\frac{u}{2}}, & u \geq 0, \\ 0, & u < 0, \end{cases} \quad (15.4)$$

где $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ – гамма-функция.

Так как аналитическое выражение плотности распределения χ^2 является довольно сложным, то в практике используют таблицу значений $\chi_{\alpha,k}^2$, рассчитанных из уравнения $p(\chi^2 > \chi_{\alpha,k}^2) = \alpha$, для различных значений k .



5. Из таблицы распределения χ^2 выбирается значение $\chi_{\alpha,k}^2$, где α – заданный уровень значимости ($\alpha = 0,05$ или $\alpha = 0,01$), а k – число степеней свободы, которое определяется по формуле

$$k = M - 1 - s.$$

Здесь s – число неизвестных параметров гипотетического закона распределения, значения которых были определены в п. 3.

6. Если значение, вычисленное по формуле (15.2), больше, чем критическое значение, т.е. $\chi^2 > \chi_{\alpha,k}^2$, то гипотеза H_0 отклоняется, в противном случае нет оснований ее отклонить.

Критерий согласия Колмогорова. Алгоритм проверки следующий:

1. Построить вариационный ряд и график эмпирической функции распределения $F^*(x)$.

2. По виду графика $F^*(x)$ выдвинуть гипотезу:

$$H_0 : F(x) = F_0(x),$$

$$H_1 : F(x) \neq F_0(x),$$

где $F_0(x)$ – функция гипотетического закона распределения.

3. Используя метод моментов или максимального правдоподобия определить оценки неизвестных параметров $\hat{Q}_1, \dots, \hat{Q}_m$ гипотетического закона распределения.

4. Рассчитать 10...20 значений функции $F_0(x)$ и построить ее график в одной системе координат с функцией $F^*(x)$.

5. По графику определить максимальное по модулю отклонение между функциями $F^*(x)$ и $F_0(x)$.

$$Z = \max_{i=1}^n |F^*(x_i) - F_0(x_i)|. \quad (15.5)$$

6. Вычислить значение критерия Колмогорова

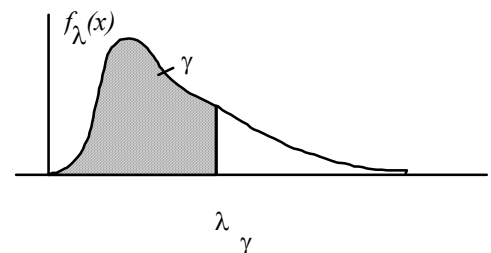
$$\lambda = \sqrt{n} \cdot Z. \quad (15.6)$$

Величина λ распределена по закону Колмогорова, который не зависит от закона распределения величины X :

$$F(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2}. \quad (15.7)$$

Так как аналитическое выражение функции распределения $F(\lambda)$ является довольно сложным, то в практике используют таблицу значений λ_γ , рассчитанных из уравнения $p(0 \leq \lambda < \lambda_\gamma) = \gamma$.

7. Из таблицы распределения Колмогорова выбрать критическое значение λ_γ , $\gamma = 1 - \alpha$. α – заданный уровень значимости ($\alpha = 0,05$ или $\alpha = 0,01$).



8. Если $\lambda > \lambda_\gamma$, то нулевая гипотеза H_0 отклоняется, в противном случае нет оснований ее отклонить.

Достоинствами критерия Колмогорова по сравнению с критерием χ^2 являются возможность его применения при очень маленьких объемах выборки ($n < 20$), более высокая "чувствительность", а следовательно, меньшая трудоемкость вычислений. Недостатком является то, что эмпирическая функция распределения $F^*(x)$ должна быть построена по несгруппированным выборочным данным, что затруднительно при больших объемах выборки. Кроме этого, следует отметить, что критерий Колмогорова можно применять только в случае, когда гипотетическое распределение полностью известно заранее из каких-либо теоретических соображений, т.е. когда известен не только вид функции распределения $F_0(x)$, но и все входящие в нее параметры Q_1, \dots, Q_k . Такой случай сравнительно редко встречается на практике. Обычно из теоретических соображений известен только общий вид функции $F_0(x)$, а входящие в нее числовые параметры определяются по данному статистическому материалу. При применении критерия χ^2 это обстоятельство учитывается соответствующим уменьшением числа степеней свободы распределения k . Критерий Колмогорова такого согласования не предусматривает. Если все же применять этот критерий в тех случаях, когда параметры теоретического распределения определяются по статистическим данным, критерий дает заведомо заниженные значения λ ; поэтому мы в ряде случаев рискуем принять как правдоподобную гипотезу, которая в действительности плохо согласуется с опытными данными.

ЛЕКЦИЯ 16

Статистическая обработка двумерных случайных величин

Пусть проводится n независимых опытов, в каждом из которых двумерная случайная величина (X, Y) принимает определенные значения и результаты опытов представляют собой двумерную выборку вида $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Статистическая обработка опытных данных включает в себя обработку и анализ составляющих X и Y , как одномерных величин (см. лекции 13–15), и вычисление оценок и анализ параметров, присущих только двумерным (многомерным) случайным величинам. Как правило, определяются следующие оценки числовых характеристик случайной величины (X, Y) :

оценки математических ожиданий:

$$m_X^* = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad (16.1)$$

$$m_Y^* = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i; \quad (16.2)$$

оценки дисперсии:

$$S_0^2(x) = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n x_i^2 - \frac{n}{n-1} \bar{x}^2; \quad (16.3)$$

$$S_0^2(y) = \frac{1}{n-1} \cdot \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} \sum_{i=1}^n y_i^2 - \frac{n}{n-1} \bar{y}^2. \quad (16.4)$$

Оценка корреляционного момента. Состоятельная несмещенная оценка корреляционного момента равна

$$K_{XY}^* = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad (16.5)$$

где x_i, y_i – значения, которые приняли случайные величины X, Y в i -м опыте;

\bar{x}, \bar{y} – средние значения случайных величин X и Y соответственно.

Оценка коэффициента корреляции. Состоятельная оценка коэффициента корреляции равна

$$R_{XY}^* = \frac{K_{XY}^*}{S_0(x)S_0(y)}, \quad (16.6)$$

где $S_0(x), S_0(y)$ – оценки среднеквадратического отклонения случайных величин X и Y соответственно.

Доверительный интервал для коэффициента корреляции с надежностью γ для случая двумерного нормального распределения имеет вид

$$\frac{e^{2a} - 1}{e^{2a} + 1} < R_{XY} < \frac{e^{2b} - 1}{e^{2b} + 1}, \quad (16.7)$$

где $a = 0,5 \cdot \ln \left(\frac{1 + R_{XY}^*}{1 - R_{XY}^*} \right) - \frac{z_\gamma}{\sqrt{n-3}}$; $b = 0,5 \cdot \ln \left(\frac{1 + R_{XY}^*}{1 - R_{XY}^*} \right) + \frac{z_\gamma}{\sqrt{n-3}}$;

$z_\gamma = \arg \Phi \left(\frac{\gamma}{2} \right)$ – значение аргумента функции Лапласа, т.е. $\Phi(z_\gamma) = \frac{\gamma}{2}$.

Статистические критерии двумерных случайных величин

Гипотеза об отсутствии корреляционной зависимости.

Предполагается, что двумерная случайная величина (X, Y) распределена по нормальному закону. Алгоритм проверки следующий.

1. Формулируется гипотеза:

$$H_0: R_{XY} = 0;$$

$$H_1: R_{XY} \neq 0.$$

Здесь R_{XY} – теоретический коэффициент корреляции.

2. Вычисляется оценка коэффициента корреляции R_{XY}^* по формуле (16.6)

3. Если объем выборки не велик ($n < 50$), определяется значение критерия

$$t = \frac{R_{XY}^* \sqrt{n-2}}{\sqrt{1 - (R_{XY}^*)^2}}, \quad (16.8)$$

который распределен по закону Стьюдента с $(n-2)$ степенями свободы, если гипотеза H_0 верна.

4. По заданному уровню значимости α вычисляется доверительная вероятность $\gamma = 1 - \alpha$ и из таблицы Стьюдента выбирается критическое значение $t_{\gamma, n-2}$.

5. Если $|t| > |t_{\gamma, n-2}|$, то гипотеза H_0 отклоняется, а следовательно, величины X, Y коррелированы. В противном случае гипотеза H_0 принимается.

3*. Если объем выборки велик ($n > 50$), то определяется значение критерия

$$Z = \frac{R_{XY}^* \sqrt{n}}{\sqrt{1 - (R_{XY}^*)^2}}, \quad (16.9)$$

который распределен практически по нормальному закону, если гипотеза H_0 верна.

4*. По заданному уровню значимости α из таблицы функции Лапласа определяется критическое значение $Z_\alpha = \arg \Phi \left(\frac{1-\alpha}{2} \right)$, т.е. $\Phi(Z_\alpha) = \frac{1-\alpha}{2}$.

5*. Если $|Z| > |Z_\alpha|$, то гипотеза H_0 отклоняется, а следовательно, величины X, Y коррелированы. В противном случае гипотеза H_0 принимается.

t-критерий. *t*-критерий служит для сравнения двух средних значений из нормально распределенных генеральных совокупностей в предположении, что дисперсии σ_X и σ_Y равны, хотя и неизвестны. Таким образом, проверяемая гипотеза H_0 утверждает, что $m_X = m_Y$. Пусть $\{x_1, x_2, \dots, x_{n_1}\}, \{y_1, y_2, \dots, y_{n_2}\}$ – независимые случайные выборки из обеих генеральных совокупностей; в общем случае они могут иметь совершенно разные объемы. В качестве критерия используем величину

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)S_0^2(x) + (n_2 - 1)S_0^2(y)}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}. \quad (16.10)$$

При сделанных предпосылках (нормальная распределенность X и Y и равенство дисперсий) и в предположении, что гипотеза H_0 верна, величина T удовлетворяет распределению Стьюдента с $k = n_1 + n_2 - 2$ степенями свободы.

Поэтому критическая область может быть установлена следующим образом. Для заданного уровня значимости α по таблице распределения Стьюдента определяем значение $t_{1-\alpha, n-1}$. Если вычисленное (согласно (16.10)) значение T удовлетворяет неравенству $|T| > t_{1-\alpha, n-1}$, то гипотезу H_0 отвергают.

По отношению к предпосылке «нормальной распределенности» *t*-критерий не очень чувствителен. Его можно применять, если статистические распределения обеих выборок не имеют нескольких вершин (т.е. унимодальные) и не слишком ассиметричны. Предпосылка $\sigma_X = \sigma_Y$ во многих случаях может быть обоснована на содержательном уровне; а гипотезу $\sigma_X = \sigma_Y$ можно проверить по *F*-критерию (см. ниже).

F-критерий. Гипотезы о дисперсии имеют в технике большое значение, так как σ_X^2 есть мера таких характеристик, как точность машин, ошибки измерительных приборов, точность технологических процессов и т. п.

F-критерий служит для проверки гипотезы о равенстве дисперсий при условии, что X и Y распределены нормально. Проверяемая гипотеза H_0 утверждает, что $\sigma_X = \sigma_Y$. Из каждой генеральной совокупности производятся выборки объема n_1 и n_2 . В качестве критерия используем величину

$$F = \frac{S_0^2(x)}{S_0^2(y)}, \text{ или } F = \frac{S_0^2(y)}{S_0^2(x)}, \quad (16.11)$$

причем, большую дисперсию выбирают в качестве числителя.

Величина F удовлетворяет *F*-распределению с $(n_1 - 1, n_2 - 1)$ степенями свободы. Критическая область выбирается следующим образом. Для уровня значимости α по таблице *F*-распределения определяем критическое значение $F_{\alpha/2; n_1-1, n_2-1}$. Если F , вычисленное по выборке, больше, чем это критическое значение $F_{\alpha/2; n_1-1, n_2-1}$, то гипотеза H_0 должна быть отклонена.

Критерий Уилкоксона. Данный критерий служит для проверки, относятся ли две выборки к одной и той же генеральной совокупности; другими словами, гипотеза H_0 утверждает, что $F_X(x) \equiv F_Y(y)$. Относительно закона

распределений величин X и Y никаких предположений не делается. Способы проверки, при которых не делается предположений о распределении в генеральной совокупности, называются способами, свободными от параметров, в противоположность рассматривавшимся выше параметрическим критериям, в которых предполагалась нормальная распределенность X и Y . Значения $\{x_1, x_2, \dots, x_{n_1}\}$ и $\{y_1, y_2, \dots, y_{n_2}\}$ обеих выборок упорядочиваются вместе в порядке их возрастания. Пара значений (x_i, y_j) образует *инверсию*, если $y_j < x_i$. Пусть, например, для $n_1 = 4$ и $n_2 = 5$ получилась такая последовательность: $y_5 x_3 x_4 y_1 y_2 x_2 y_4 y_3 x_1$. В нашем примере x_3 и x_4 образуют по одной инверсии (с y_5), x_2 образует три инверсии (с y_5, y_1, y_2), а x_1 образует пять инверсий (со всеми y).

В качестве критерия используется величина U – полное число инверсий. Если гипотеза верна, значение U не должно слишком сильно отклоняться от своего математического ожидания $M_U = \frac{n_1 n_2}{2}$. Данная величина распределена по закону Уилкоксона и от гипотезы H_0 отказываются, если U больше критического значения U_α , взятого из таблицы Уилкоксона для заданного уровня значимости α . Для больших объемов выборки (n_1 и n_2 больше 25) критическое значение U_α определяется по формуле

$$U_\alpha = Z_\alpha \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}, \quad (16.12)$$

где $Z_\alpha = \arg \Phi\left(\frac{1-\alpha}{2}\right)$ – значение аргумента функции Лапласа, т.е. $\Phi(Z_\alpha) = \frac{1-\alpha}{2}$

ЛЕКЦИЯ 17

Оценка регрессионных характеристик

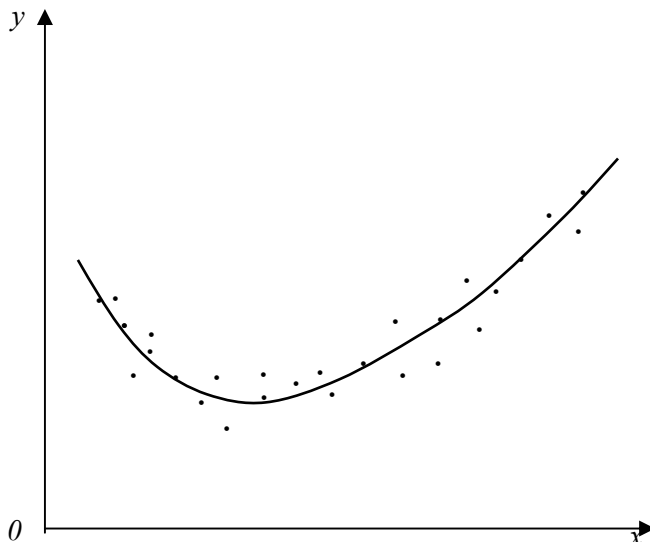
Пусть проводится n независимых опытов, в каждом из которых двумерная случайная величина (X, Y) принимает определенные значения и результаты опытов представляют собой двумерную выборку вида $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Необходимо на основании имеющейся выборки выявить характер связи между величинами X, Y , т.е. получить оценку условного математического ожидания $m_{Y/x}^*$ – оценку регрессии Y на x . Данная оценка представляет собой некоторую функцию:

$$m_{Y/x}^* = \bar{y}(x) = \varphi(x, a_0, a_1, \dots, a_m),$$

где a_0, a_1, \dots, a_m – неизвестные параметры.

Таким образом, во-первых, необходимо установить тип зависимости $\varphi(x, a_0, a_1, \dots, a_m)$ – т.е. является ли она линейной, квадратичной, показательной и т.д., во-вторых, определить значения неизвестных параметров a_0, a_1, \dots, a_m .

Для определения типа зависимости строится *диаграмма рассеивания* или *корреляционное поле*, которую можно получить, если результаты опытов изобразить в виде точек на плоскости в декартовой системе координат (см. рисунок).



На основании анализа корреляционного поля выбираем тип *эмпирической линии регрессии* $\bar{y}(x) = \varphi(x, a_0, a_1, \dots, a_m)$, которая должна проходить через точки $(x_1, y_1) \dots (x_n, y_n)$ так, чтобы ее график наилучшим образом соответствовал бы к неизвестной линии регрессии, т.е. ее значения должны быть приблизительно равны средним арифметическим значений Y для каждого значения $X=x$. Во многих случаях тип зависимости может быть выбран на основе теоретических или иных соображений.

Для определения значений параметров, при которых обеспечивается наилучшее согласования кривой $y = \varphi(x, a_0, a_1, \dots, a_m)$ и экспериментальных точек $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, используется метод наименьших квадратов.

Метод наименьших квадратов

Суть данного метода заключается в том, что значения параметров a_0, a_1, \dots, a_m необходимо выбрать так, чтобы *сумма квадратов отклонений экспериментальных точек от сглаживающей кривой обращалась в минимум*:

$$\sum_{i=1}^n [y_i - \varphi(x_i, a_0, \dots, a_m)]^2 = \min \quad (17.1)$$

Найдем значения $a_j, j = 1, \dots, m$, обращающие левую часть выражения (17.1) в минимум. Для этого продифференцируем его по $a_j, j = 1, \dots, m$, и приравняем производные к нулю (в точке экстремума производная равна нулю):

$$\sum_{i=1}^n [y_i - \varphi(x_i, a_0, \dots, a_m)] \frac{\partial \varphi(x_i)}{\partial a_j} = 0, j = 0, 1, \dots, m, \quad (17.2)$$

где $\frac{\partial \varphi(x_i)}{\partial a_j}$ – значение частной производной функции φ по параметру a_j в точке x_i .

Система уравнений (17.2) содержит столько же уравнений, сколько неизвестных параметров, т.е. $m+1$.

Решить систему (17.2) в общем виде нельзя; для этого необходимо задаться конкретным видом функции φ .

Пусть y представляет собой степенной ряд:

$$y = \varphi(x, a_0, \dots, a_m) = \sum_{j=0}^m a_j x^j. \quad (17.3)$$

Тогда (17.2) примет вид системы линейных уравнений (СЛУ):

$$\sum_{j=0}^m a_j \sum_{i=1}^n (x_i)^{j+k} = \sum_{i=1}^n y_i (x_i)^k, k = 0, 1, \dots, m \quad (17.4)$$

Поделим обе части уравнений на объем выборки n , система примет вид

$$\sum_{j=0}^m a_j \hat{\alpha}_{j+k}(x_i) = \hat{\alpha}_{k,1}(x_i, y_i), k = 0, 1, \dots, m \quad (17.5)$$

где $\hat{\alpha}_k(x) = \frac{1}{n} \cdot \sum_{i=1}^n (x_i)^k$ – оценка начального момента k -го порядка величины X ;

$\hat{\alpha}_{k,1}(x, y) = \frac{1}{n} \cdot \sum_{i=1}^n x_i^k y_i$ – оценка смешанного начального момента порядка $k+1$ величин X и Y .

Переменными в системе (17.4) являются $a_j, j = 1, \dots, m$, а вычисленные по исходной выборке оценки начальных моментов являются коэффициентами СЛУ. Решив данную систему, мы определим оценки параметров $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m$, обеспечивающие наилучшее согласование кривой $y = \varphi(x, a_0, a_1, \dots, a_m)$ и экспериментальных точек $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$.

Пример. Определим оценку линейной регрессии $m_{Y/X} = a_0 + a_1 x$.

Система (17.5) для $m=1$ имеет вид

$$\begin{cases} \hat{\alpha}_0(x) a_0 + \hat{\alpha}_1(x) a_1 = \hat{\alpha}_{0,1}(x, y) \\ \hat{\alpha}_1(x) a_0 + \hat{\alpha}_2(x) a_1 = \hat{\alpha}_{1,1}(x, y) \end{cases}$$

С учетом того, что $\hat{\alpha}_0(x) = 1$, $\hat{\alpha}_1(x) = \bar{x}$, $\hat{\alpha}_{0,1}(x, y) = \bar{y}$, получаем:

$$\begin{cases} a_0 + \bar{x}a_1 = \bar{y} \\ \bar{x}a_0 + \hat{\alpha}_2(x)a_1 = \hat{\alpha}_{1,1}(x, y) \end{cases}$$

Отсюда

$$\hat{a}_1 = \frac{\hat{\alpha}_{1,1}(x, y) - \bar{x} \cdot \bar{y}}{\hat{\alpha}_2(x) - \bar{x}^2} = \frac{K_{XY}^*}{S_0^2(x)}, \quad (17.6)$$

$$\hat{a}_0 = \bar{y} - \hat{a}_1 \cdot \bar{x}, \quad (17.7)$$

что соответствует уравнениям прямых регрессий (9.10) (см. лекцию 9).

ЛИТЕРАТУРА

1. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. – М.: Наука, 1988. – 416 с.
2. Вентцель Е.С. Теория вероятностей и математическая статистика: Учебник. – 5-е изд., стереотип. – М.: Высш. шк., 1999. – 576 с.
3. Герасимович А.И. Математическая статистика. – Мн.: Выш. шк., 1983. – 279 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика. – М.: Высш. шк., 1977. – 479 с.
5. Жевняк Р.М., Карпук А.А., Унукович В.Т. Теория вероятностей и математическая статистика: Учеб. пособие для студентов. инж.-экон. спец. – Мн.: Харвест, 2000.-384 с.

Учебное издание

**Волковец Александр Иванович,
Гуринович Алевтина Борисовна**

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Конспект лекций
для студентов всех специальностей и форм обучения БГУИР

Редактор Т.А. Лейко
Корректор Е.Н. Батурчик
Компьютерная верстка Т.В. Шестакова.

Подписано в печать .08.2003.
Печать ризографическая.
Уч.-изд. л. 4,0.

Формат 60x84 1/16.
Гарнитура «Таймс».
Тираж 300 экз.

Бумага офсетная.
Усл. печ. л. .
Заказ 210.

Издатель и полиграфическое исполнение:
Учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники».
Лицензия ЛП №156 от 30.12.2002.
Лицензия ЛВ №509 от 03.08.2001.
220013, Минск, П. Бровки, 6.