

Системы случайных величин.

В практических применениях теории вероятностей приходится иметь дело с задачами, в которых результат опыта (стохастического эксперимента) описывается не одной случайной величиной, а двумя или более, образующими систему или случайный вектор.

Например, расходы случайно выбранной семьи зависят от затрат на питание, обувь, одежду, транспорт, удовлетворение духовных потребностей; на урожайность данной культуры влияют погодные условия, применяемые удобрения, характер почвы, качество посевного материала и т.д.

Случайным вектором или **n -мерной случайной величиной** называют упорядоченный набор из n случайных величин (X_1, X_2, \dots, X_n) .

Многомерная случайная величина, случайный вектор, система случайных величин – всё это различные интерпретации одного и того же математического объекта, каждой из которых можно пользоваться для удобства изложения.

Многомерные случайные величины так же, как и одномерные, могут быть дискретными и непрерывными.

Наиболее простой случай – двумерная случайная величина.

Функцией распределения двумерной случайной величины (X, Y) называют вероятность совместного выполнения двух неравенств:

$$F(x, y) = P\{X < x; Y < y\}.$$

Используя геометрическую интерпретацию двумерной случайной величины как случайной точки в декартовой системе координат, можно сказать, что функция распределения $F(x, y)$ есть вероятность попадания случайной точки (X, Y) в бесконечный квадрант с вершиной в точке (x, y) , лежащий левее и ниже её (границы не входят).

Свойства функции распределения.

1) функция распределения есть неубывающая функция по каждому из своих аргументов;

2) повсюду на $-\infty$ функция распределения равна 0 ;

3) если оба аргумента $= +\infty$, то функция распределения $= 1$;

4) при одном из аргументов, $= +\infty$, функция распределения двумерной случайной величины превращается в функцию распределения компоненты, соответствующей другому аргументу.

Функция распределения существует как для непрерывных, так и для дискретных случайных величин.

Распределение многомерных непрерывных случайных величин обычно характеризуют плотностью распределения.

Плотность распределения двумерной непрерывной случайной величины называют предел отношения вероятности попадания случайной величины в малый прямоугольник к площади этого прямоугольника, когда оба его раздела стремятся к нулю:

$$f(x, y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P\{(X, Y) \in R_{\Delta}\}}{\Delta x \cdot \Delta y}.$$

Если известна функция распределения, то плотность можно вычислить как вторую смешенную частную производную:

$$f(x, y) = F''_{x,y}(x, y).$$

Геометрически плотность распределения можно изобразить некоторой поверхностью, так называемой **поверхностью распределения**. Эта поверхность аналогична кривой распределения для одномерной случайной величины.

Свойства плотности распределения.

- 1) плотность распределения – неотрицательная функция;
- 2) двойной интеграл в бесконечных пределах от плотности двумерного распределения равен 1 (действительно, данный интеграл есть не что иное, как вероятность попадания во всю плоскость xOy);
- 3) плотности распределения компонент могут быть получены по формулам:

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

Последнее свойство следует из соответствующего свойства для функций распределения и формулы, связывающей функцию распределения функцию распределения случайного вектора и плотность.

Обратимся к дискретным двумерным случайным векторам.

Закон распределения дискретного случайного вектора (X, Y) – это совокупность всех возможных значений $\{x_i, y_j\}_{i,j}$ случайного вектора и их вероятностей $p_{ij} = P\{X=x_i, Y=y_j\}$.

Сумма вероятностей для всех возможных значений равна 1:

$$\sum_{i,j} p_{ij} = 1.$$

Так как и в непрерывном случае, зная закон распределения случайного вектора, можно найти распределение координат, просуммировав по соответствующему индексу:

$$p_{i\bullet} = P\{X = x_i\} = \sum_j p_{ij}; \quad p_{\bullet j} = P\{Y = y_j\} = \sum_i p_{ij}.$$

Пример 6.1. Качество продукции характеризуется двумя случайными величинами. Закон распределения случайного вектора (X, Y) задан таблицей:

	$x_1 = 0$	$x_2 = 0$	$x_3 = 0,1$	$x_4 = 0,2$	$p_{i\bullet}$
$y_1 = 5$	0,2	0,1	0,05	0,05	0,4
$y_2 = 6$	0	0,15	0,15	0,1	0,4
$y_3 = 7$	0	0	0,1	0,1	0,2
$p_{\bullet j}$	0,2	0,25	0,3	0,25	1

На пересечении i -ой строки и j -го столбца стоят вероятности $p_{ij} = P\{X=x_i, Y=y_j\}$.

Вероятности $p_{i\bullet} = P\{X = x_i\}$ есть сумма вероятностей, стоящих в i -ой строке, а $p_{\bullet j} = P\{Y = y_j\}$ - в j -ом столбце.

Зависимые и независимые случайные величины.

Для того, чтобы полностью описать систему случайных величин, недостаточно знать распределение каждой из величин, входящих в эту систему; нужно знать еще зависимость между величинами, входящими в систему. Эта зависимость характеризуется с помощью условных законов распределения.

Для непрерывных величин большую роль играют **условные плотности распределения** – это плотности распределения одной случайной величины, вычисленные при условии, что другая случайная величина приняла определённое значение:

$$f_X(x | y_0) = \frac{f(x, y_0)}{\int_{-\infty}^{+\infty} f(x, y_0) dx};$$

$$f_Y(y | x_0) = \frac{f(y, x_0)}{\int_{-\infty}^{+\infty} f(y, x_0) dy}$$

Условное распределение компонент дискретного случайного вектора – это ряд распределения одной случайной величины, вычисленный при условии, что другая случайная величина приняла определённое значение:

$$p_X(x_i | y_j) = P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{\sum_{i=1}^n p_{ij}};$$

$$p_Y(y_j | x_i) = P\{Y = y_j | X = x_i\} = \frac{p_{ij}}{\sum_{j=1}^m p_{ij}}.$$

Здесь $\{p_{ij}\}_{i=1, \dots, n; j=1, \dots, m}$ - закон распределения дискретного случайного вектора (X, Y) ;

n и m могут быть конечны или бесконечны.

Пример 6.2. Закон распределения дискретного случайного вектора (X, Y) задан таблицей:

	$x_1 = 0$	$x_2 = 0$	$x_3 = 0,2$	$x_4 = 0,3$	$p_{i\cdot}$
$y_1 = 5$	0,2	0,1	0,05	0,05	0,4
$y_2 = 6$	0	0,15	0,15	0,1	0,4
$y_3 = 7$	0	0	0,1	0,1	0,2
$p_{\cdot j}$	0,2	0,25	0,3	0,25	1

Найдем условное распределение случайной величины X при условии, что случайная величина Y приняла значение $y_2 = 0,1$. Выбрав значения p_{i2} из столбца таблицы, соответствующего значению $y_2 = 0,1$, и разделив их на $p_{\cdot 2} = 0,25$, получаем условное распределение X при условии, что $Y = 0,1$:

x_i	5	6
$p_X(x_i/y_2)$	0,4	0,6

При изучении систем случайных величин всегда следует обращать внимание на степень и характер их зависимости. Эта зависимость может быть более или менее ярко выражена, более или менее тесная. В некоторых случаях зависимость между случайными величинами может быть настолько тесной, что зная значение одной величины, можно в точности указать значение другой. В другом крайнем случае зависимость между случайными величинами является настолько слабой, что их можно считать практически независимыми.

Перейдем к точным формулировкам.

Случайные величины X и Y называют **независимыми**, если

$$F(x,y)=F_X(x) \cdot F_Y(y),$$

т.е. события $\{X < x\}$ и $\{Y < y\}$ независимы при любых x и y .

Понятие независимости легко обобщить на случай n величин.

Случайные величины X_1, X_2, \dots, X_n называют **независимыми в совокупности**, если события $\{X_1 < x_1\}, \{X_2 < x_2\}, \dots, \{X_n < x_n\}$ независимы при любых x_1, x_2, \dots, x_n .

Многие результаты, относящиеся к предельным теоремам теории вероятностей, получены в предположении, что исходные случайные величины независимы в совокупности.

Обратимся к зависимым величинам.

Вероятностная зависимость между случайными величинами очень часто встречается на практике. Если случайные величины X и Y находятся в вероятностной зависимости, это не означает, что изменением величины одной величины, другая тоже изменится; это лишь означает, что изменением величины X величина Y имеет тенденцию также изменяться (например, возрастая или убывая с ростом X). Эта тенденция соблюдается лишь в общих чертах, и в каком-то отдельном случае от неё возможны отступления.

Примерами случайных величин, находящихся в вероятностной зависимости, являются рост и возраст ребенка; затраты и прибыль при производстве определенной продукции; затраты на рекламу и объем продаваемой продукции и т.д.

Степень зависимости между случайными величинами обычно оценивают с помощью числовых характеристик зависимости.

Числовые характеристики зависимости.

Ковариация случайных величин X и Y – математическое ожидание произведения центрированных величин:

$$\text{Cov}(X, Y) = M[(X - m_X) \cdot (Y - m_Y)].$$

Если случайные величины дискретны, то их ковариацию вычисляют по формуле:

$$\text{cov}(X, Y) = \sum_{i,j} (x_i - m_X) \cdot (y_j - m_Y) \cdot p_{ij}.$$

Для непрерывных величин сумму заменяют на интеграл.

Ковариация описывает как рассеивание случайных величин, так и связь между ними.

Теорема. Если случайные величины независимы, то ковариация между ними равна 0 .

Доказательство проведем для дискретного случая (для непрерывного – аналогично).

Так как случайные величины X и Y - независимы, то

$$p_{ij} = P\{X=x_i, Y=y_j\} = P\{X=x_i\} \cdot P\{Y=y_j\} = p_i \cdot p_j.$$

Следовательно,

$$\begin{aligned} \text{cov}(X, Y) &= \sum_{i,j} (x_i - m_X) \cdot (y_j - m_Y) \cdot p_{ij} = \sum_{i,j} (x_i - m_X) \cdot (y_j - m_Y) \cdot p_i \cdot p_j = \\ &= \left[\sum_i (x_i - m_X) \cdot p_i \right] \cdot \left[\sum_j (y_j - m_Y) \cdot p_j \right] = \\ &= \left[\sum_i x_i \cdot p_i - \sum_i m_X \cdot p_i \right] \cdot \left[\sum_j y_j \cdot p_j - \sum_j m_Y \cdot p_j \right] = [m_X - m_X] \cdot [m_Y - m_Y] = 0. \end{aligned}$$

Здесь было использовано определение математического ожидания дискретной случайной величины и условие нормировки для вероятностей. \square

Для характеристики зависимости в чистом виде используют безразмерную величину – корреляцию.

Коэффициентом корреляции случайных величин называют их нормированную ковариацию:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)}.$$

Очевидно, что коэффициент корреляции обращается в ноль вместе с ковариацией случайных величин, т.е. для независимых случайных величин коэффициент корреляции равен 0 .

Случайные величины, имеющие нулевой коэффициент корреляции, называют **некоррелированными**.

Из независимости случайных величин следует их некоррелированность. Обратное неверно. Условие независимости более жесткое, чем некоррелированности.

Коэффициент корреляции характеризует линейную зависимость между случайными величинами.

Теорема. Если случайные величины связаны линейной зависимостью $Y=a \cdot X+b$, то коэффициент корреляции $\rho(X,Y)=|1|$.

Доказательство. Используем свойства математического ожидания и дисперсии:

$$\text{Cov}(X,Y)=M[(X-m_X) \cdot (a \cdot X+b-a \cdot m_X-b)]=a \cdot M[(X-m_X)^2]=a \cdot \sigma^2(X),$$

$$\sigma(Y)=|a| \cdot \sigma(X).$$

Можно показать, что обратное утверждение также верно.

Свойство коэффициента корреляции.

Величина коэффициента корреляции заключена в пределах $[-1;1]$.

Если коэффициент корреляции больше (меньше) нуля, то говорят о положительной (отрицательной) корреляции между случайными величинами, это означает, что при возрастании одной из случайных величин другая имеет тенденцию в среднем возрастать (убывать).

Например, рост и вес человека связаны положительной корреляцией; а время, потраченное на регулировку прибора и количество неисправностей, обнаруженных при работе прибора – отрицательной.

В качестве характеристики зависимости системы n случайных величин используют корреляционную матрицу $K=||\text{cov}(X_i, X_j)||_{i,j}$. Из определения ковариации следует, что эта матрица является симметричной.

Функции нескольких случайных аргументов.Свертка.

Рассмотрим важный частный случай: нахождение распределения суммы двух независимых случайных величин.

Пусть случайная величина $Z=X+Y$ (является суммой двух непрерывных случайных величин). Тогда её функция распределения :

$$G(z)=P\{Z<z\}=P\{X+Y<z\}=\int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{z-x} f(x, y)dy \right\} dx$$

Теперь рассмотрим плотность распределения случайной величины Z :

$$g(z) = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z-x)dx = \int_{-\infty}^{+\infty} f_X(z-y) \cdot f_Y(y)dy.$$

Для двух последних интегралов, выражающих плотность распределения суммы независимых случайных величин, используют специальное обозначение:

$$g = f_X * f_Y$$

и называют **композицией** (или **свёрткой**) плотностей распределения.

Пример 6.3. Если случайные величины X и Y равномерно распределены на $[a, b]$, то плотность распределения их суммы

$$g(z) = \left\{ \begin{array}{ll} \frac{z-2a}{(b-a)^2} & \text{при } 2a < z \leq a+b \\ \frac{2b-z}{(b-a)^2} & \text{при } a+b < z \leq 2b \\ 0 & \text{при } z \leq 2a, z > 2b \end{array} \right\},$$

функцию $g(z)$ называют плотностью распределения Симпсона.

Пример 6.4. Закон распределения суммы n независимых случайных величин, имеющих показательное распределение, называют законом Эрланга $(n-1)$ порядка. Показательное распределение является законом Эрланга нулевого порядка. Закон Эрланга был получен при моделировании работы телефонных сетей в первых работах по теории массового обслуживания. Плотность распределения Эрланга n -го порядка есть:

$$f(x) = \left\{ \begin{array}{ll} \frac{\lambda}{n!} (\lambda \cdot x)^n e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{array} \right\}.$$

Для того, что найти распределение суммы двух независимых дискретных величин, используют аналогичную формулу, но вместо интеграла пишут сумму.

Пример 6.5. Сумма независимых случайных величин, распределенных по закону Пуассона с параметрами λ_1 и λ_2 , также имеет распределение Пуассона с параметром $\lambda_1 + \lambda_2$.

Свойства числовых характеристик суммы случайных величин.

1) $M(X+Y) = M(X) + M(Y)$;

2) $D(X+Y) = D(X) + D(Y) + 2cov(X, Y)$;

3) Если случайные величины независимы, то $D(X+Y) = D(X) + D(Y)$;

4) Если случайные величины независимы, то $D(X-Y) = D(X) + D(Y)$;

5) Если случайные величины независимы, то

$$\sigma^2(X+Y) = \sqrt{\sigma^2(X) + \sigma^2(Y)}.$$

Многомерные распределения.

Двумерное нормальное распределение.

В теории вероятностей и её приложениях большую роль играет двумерное нормальное распределение. Плотность двумерной нормальной случайной величины имеет вид:

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

Это распределение зависит от 5 параметров:

m_X, m_Y – математические ожидания случайных величин;

σ_X, σ_Y – среднеквадратические отклонения;

r – коэффициент корреляции.

Если компоненты двумерной нормальной случайной величины не коррелированы, т.е. $r=0$, то

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

Таким образом, из некоррелированности нормально распределенных случайных величин следует их независимость. Так как для любых случайных величин из независимости следует некоррелированность, то термины «некоррелированные» и «независимые» величины для случая нормального распределения эквивалентны.

При $r \neq 0$ компоненты двумерного нормального вектора зависимы. Условные распределения имеют вид:

$$f(x|y) = \frac{1}{\sigma_X\sqrt{1-r^2}\sqrt{2\pi}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x-m_X}{\sigma_X} + r\frac{y-m_Y}{\sigma_Y}\right]^2\right\};$$

$$f(y|x) = \frac{1}{\sigma_Y\sqrt{1-r^2}\sqrt{2\pi}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{y-m_Y}{\sigma_Y} + r\frac{x-m_X}{\sigma_X}\right]^2\right\}.$$

Из последних формул следует, что условные распределения также являются нормальными с параметрами

$$m_{X|Y} = m_X + r \frac{\sigma_X}{\sigma_Y} (y - m_Y), \quad \sigma_{X|Y} = \sigma_X (1 - r);$$

$$m_{Y|X} = m_Y + r \frac{\sigma_Y}{\sigma_X} (x - m_X), \quad \sigma_{Y|X} = \sigma_Y (1 - r).$$

Величину $m_{X|Y}$ называют **условным математическим ожиданием** случайной величины X при условии, что $Y = y$ (или **регрессией** случайной величины X по величине Y).

К числу немногих плоских фигур, вероятность попадания в которые может быть вычислена в явном виде, принадлежит эллипс равных вероятностей:

$$G_C = \left\{ (x, y) : \left[\frac{(x - m_X)^2}{\sigma_X^2} - \frac{2r(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} = C^2 \right] \right\}.$$

Вероятность попадания точки (X, Y) внутрь эллипса равной вероятности G_C есть

$$P(C) = 1 - \exp\left\{-\frac{C^2}{2(1 - r^2)}\right\}.$$

Пример 7.1. Случайные величины X и Y – независимы и стандартно нормально распределены. Найти вероятность того, что случайная точка (X, Y) попадёт в кольцо $K = \{(x, y) : 4 \leq x^2 + y^2 \leq 9\}$.

$$\begin{aligned} \text{Ясно, что } P\{(X, Y) \in K\} &= P\{(X, Y) \in G_3\} - P\{(X, Y) \in G_2\} = P(3) - P(2) = \\ &= e^{-2} - e^{-4,5} \approx 0,1242. \end{aligned}$$

Общий случай n -мерного нормального распределения.

Плотность нормального распределения в пространстве n измерений имеет вид:

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(K)}} \exp\left\{-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n K_{ij}^{-1} (x_i - m_i)(x_j - m_j)\right\},$$

где m_i – математическое ожидание i -ой компоненты нормального случайного вектора;

K_{ij}^{-1} – элементы матрицы, обратной по отношению к корреляционной матрице K .

Из общего выражения вытекают все формы нормального закона для любого числа измерений и любых видов зависимости между случайными величинами.

Функции от нормально распределенных случайных величин.

1. Линейная функция.

Линейная функция является наиболее простой. Плотность распределения суммы двух независимых случайных величин, подчинённых нормальному закону распределения, можно найти, используя формулу свёртки.

Правило. При композиции нормальных распределений получается снова нормальное распределение, причём математические ожидания и дисперсии суммируются. Т.е., если

$$X_1 \sim \mathcal{N}(m_1, \sigma_1^2), X_2 \sim \mathcal{N}(m_2, \sigma_2^2), \text{ то } X_1 + X_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2).$$

Если слагаемые зависимы, то сумма снова является нормально распределённой случайной величиной со средним $m_1 + m_2$ и дисперсией, равной $\sigma_1^2 + \sigma_2^2 + 2r\sigma_1\sigma_2$, где r - коэффициент корреляции между слагаемыми.

Рассмотрим общий случай. $Y = \sum_{i=1}^n a_i \cdot X_i + b$, где $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$. Случайная

величина Y также распределена нормально с параметрами $m_Y = \sum_{i=1}^n a_i \cdot m_i + b$;

$$\sigma_Y^2 = \sum_{i=1}^n a_i^2 \cdot \sigma_i^2 + 2 \sum_{i < j} a_i a_j r_{ij} \sigma_i \sigma_j.$$

Следующие распределения играют большую роль в статистике.

2. Распределение хи-квадрат с n степенями свободы $\chi^2(n)$.

Пусть X_1, X_2, \dots, X_n – независимые случайные величины, распределённые по одному и тому же нормальному закону с параметрами m и σ^2 . Функция распределения суммы

$$\chi^2(n) = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2$$

носит название хи-квадрат распределение с n степенями свободы. Плотность такого распределения для $x > 0$ имеет вид:

$$f(x) = \frac{x^{n/2-1} \cdot x^{-x/2}}{2^{n/2} \Gamma(n/2)}, \text{ где } \Gamma(\lambda) \text{ – гамма-функция.}$$

Распределение $\chi^2(n)$ табулировано.

$$M[\chi^2(n)] = n, D[\chi^2(n)] = 2n.$$

3. Распределение Стьюдента $St(n)$.

[Стьюдент – псевдоним английского статистика Уильяма Сили Госсета, впервые нашедшего этот закон в 1908 году эмпирическим путём.]

Распределение Стьюдента с n степенями свободы определяется как распределение отношения:

$$St = \frac{X}{\sqrt{\chi^2(n)/n}}, \text{ где } X \sim \mathcal{N}(0, 1).$$

В силу определения это распределение является симметричным.

Важная роль распределения Стьюдента в математической статистике объясняется следующим фактом: если случайные величины $X_0, X_1, X_2, \dots, X_n$ независимы и одинаково нормально распределены с параметрами m и σ^2 , то отношение

$$St = \frac{X_0 - m}{\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - m)^2}}$$

имеет распределение Стьюдента с n степенями свободы, которое не зависит от параметров m и σ^2 . Распределение Стьюдента для небольших n табулировано. Для достаточно больших n распределение Стьюдента близко к стандартно нормальному, поэтому можно пользоваться таблицами нормального распределения.

4. Распределение Фишера – Снедекора $F(n_1, n_2)$.

Распределение Фишера – Снедекора имеет отношение двух случайных величин, распределённых по χ^2 :

$$F(n_1, n_2) = \frac{\chi^2(n_1)}{n_1} : \frac{\chi^2(n_2)}{n_2}.$$

В математической статистике используют следующий факт: если случайные величины X_1, X_2, \dots, X_n и Y_1, Y_2, \dots, Y_k независимы и одинаково нормально распределены с параметрами m и σ^2 , то случайная величина

$$F(n, k) = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2 \bigg/ \frac{1}{k} \sum_{j=1}^k (Y_j - m)^2$$

имеет распределение Фишера – Снедекора, не зависящее от параметров m и σ^2 .

Распределение Фишера – Снедекора также табулировано.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

В практической деятельности большое значение имеют события с вероятностями, близкими к 1 или 0. Поэтому одной из задач теории вероятностей является установление закономерностей, происходящих с вероятностями, близкими к 1; при этом особую роль играют закономерности, возникающие в результате наложения большого числа случайных факторов. Группа теорем, известная под названием «закона больших чисел», является одним из таких предложений теории вероятностей.

Установленное при помощи закона больших чисел свойство случайных величин при определённых условиях вести себя как неслучайные позволяет предсказывать результаты массовых случайных явлений почти с полной определенностью.

Возможности таких предсказаний в области массовых случайных явлениях ещё больше расширяются наличием другой группы теорем, известных под названием «центральная предельная теорема». При суммировании достаточно большого числа случайных величин закон распределения суммы приближается к нормальному при соблюдении некоторых условий. Формы центральной предельной теоремы различаются между собой теми условиями на случайные величины, для которых устанавливается это свойство суммы.

Совокупность различных форм закона больших чисел вместе с различными формами центральной предельной теоремы образуют **предельные теоремы теории вероятностей.**

Для доказательства теорем, относящихся к группе «закона больших чисел», необходимо одно общее неравенство, называемое неравенством

Чебышёва. Это неравенство было открыто независимо друг от друга Пафнутием Львовичем Чебышёвым в 1867 г. и Жюлем Бьенемэ в 1853 г.

Неравенство Чебышёва. Каково бы ни было положительное число ε , вероятность того, что случайная величина X отклониться от своего математического ожидания не больше, чем на ε , ограничена сверху величиной $\frac{D(X)}{\varepsilon^2}$:

$$P\{|X - M(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2}.$$

Доказательство проведём для случая, когда случайная величина X – непрерывна. Используя свойство плотности распределения, имеем:

$$P\{|X - M(X)| \geq \varepsilon\} = \int_{|x - M(X)| > \varepsilon} f(x) dx, \quad \text{где } f(x) \text{ – плотность}$$

распределения случайной величины X . Далее:

$$\begin{aligned} D(X) &= \int_{-\infty}^{+\infty} (x - M(X))^2 f(x) dx \geq \int_{|x - M(X)| \geq \varepsilon} (x - M(X))^2 f(x) dx \geq \\ &\geq \varepsilon^2 \int_{|x - M(X)| \geq \varepsilon} f(x) dx = \varepsilon^2 P\{|X - M(X)| \geq \varepsilon\}. \quad \square \end{aligned}$$

Пример 8.1. $M(X)=m$, $D(X)=\sigma^2$. Оценим вероятность того, что случайная величина X отклониться от своего математического ожидания не больше, чем на 3σ . Для этого в неравенстве Чебышёва надо положить $\varepsilon=3\sigma$:

$$P\{|X - m| \geq 3\sigma\} \leq \frac{D(X)}{9\sigma^2} = \frac{1}{9}.$$

Заметим, что неравенство Чебышёва даёт только верхнюю границу вероятности данного отклонения. Больше этой границы вероятность быть не может ни при каком законе распределения. На практике, в большинстве случаев вероятность того, что величина X выйдет за пределы участка $m \pm 3\sigma$, значительно меньше, чем $1/9$. Например, для нормального закона эта вероятность равна $0,003$.

Закон больших чисел.

Одной из наиболее простых, но вместе с тем наиболее важных форм закона больших чисел является теорема, доказанная Чебышёвым в 1867 г.

Теорема Чебышёва. Если $X_1, X_2, \dots, X_n, \dots$ – последовательность независимых случайных величин, имеющих конечные дисперсии, то

$$P \left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{M(X_1) + M(X_2) + \dots + M(X_n)}{n} \right| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

для любого положительного ε (т.е. среднее арифметическое случайных величин становится близко к постоянной величине – среднему арифметическому их математических ожиданий).

Важным частным случаем теоремы Чебышёва является теорема Бернулли. (Яков Бернулли в конце XVII века доказал теорему, носящую теперь его имя; эта теорема впервые была опубликована в 1713 г. после смерти автора.)

Теорема Бернулли. Пусть μ_n – число наступлений события A в n независимых опытах (испытаниях Бернулли) и p есть вероятность наступления события A в каждом из испытаний. Тогда

$$P \left\{ \left| \frac{\mu_n}{n} - p \right| > \varepsilon \right\} \xrightarrow{n \rightarrow \infty} 0$$

при любом $\varepsilon > 0$ (т.е. с ростом числа испытаний частота события приближается к его вероятности).

Доказательство. Рассмотрим случайные величины X_k – индикаторы появления события A в k -ом опыте, тогда

$$\mu_n = X_1 + X_2 + \dots + X_n$$

А так как $M(X_k) = p$; $D(X_k) = p(1-p)$, отсюда следует, что теорема Бернулли является простейшим частным случаем теоремы Чебышёва.

Так как на практике часто неизвестные вероятности приходится определять из опыта приближенно, то для проверки соответствия теоремы Бернулли было проведено большое число опытов. При этом рассматривались события, вероятности которых можно считать по тем или иным соображениям известными, относительно которых легко проводить испытания и обеспечить независимость этих испытаний, а также постоянство вероятностей в каждом из испытаний. Все такие опыты дали прекрасное совпадение с теорией. Приведём некоторые из них. Французский естествоиспытатель Жорж Луи Бюффон бросал монету 4040 раз, герб выпал при этом 2048 раз. Частота появления герба в опыте Бюффона равна 0,507. Английский статистик Эгон Шарп Пирсон бросал монету 12000 раз и при этом наблюдал 6019 выпадений герба. Частота выпадения герба в этом опыте

Пирсона равна 0,5016. В другой раз он бросил монету 24000 раз, и герб выпал 12012 раз; частота оказалась равна 0,5005. Во всех приведенных опытах частоты лишь немного уклонялись от вероятности 0,5.

Закон больших чисел дает теоретическое основание правилу среднего арифметического, постоянно употребляющемуся в теории измерений. Предположим, что проводится измерение некоторой физической величины a . Повторив измерение n раз в одинаковых условиях, наблюдатель часто получает не вполне совпадающие результаты: x_1, x_2, \dots, x_n . В качестве приближенного значения a принято брать среднее арифметическое из результатов наблюдений:

$$a \approx \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Если измерения лишены систематической ошибки (т.е. если математические ожидания величин x_1, x_2, \dots, x_n равны нулю), то согласно закону больших чисел при достаточно больших n с вероятностью, близкой к 1, указанным путем можно получить значение, достаточно близкое к a .

Центральная предельная теорема.

Все формы центральной предельной теоремы посвящены установлению условий, при которых возникает нормальный закон распределения. Так как на практике эти условия часто выполняются, нормальный закон является самым распространённым. Он возникает во всех случаях, когда исследуемая случайная величина может быть представлена в виде суммы достаточно большого числа независимых или слабо зависимых слагаемых. Например, при измерении физических величин возникающие ошибки складываются из многочисленных независимых элементарных ошибок, порождаемых различными причинами. Поэтому ошибки измерения в большинстве случаев распределены по нормальному закону. Некоторые показатели в экономике также представляют собой сумму большого числа слагаемых, каждое из которых имеет малый вклад в суммарный показатель. В этом случае исследуемый показатель имеет приближенно нормальное распределение.

Различные формы центральной предельной теоремы отличаются условиями, накладываемыми на распределение слагаемых, образующих сумму.

Одна из наиболее простых форм центральной предельной теоремы – теорема, относящаяся к случаю одинакового распределения слагаемых.

Центральная предельная теорема для одинаково распределенных слагаемых.

Если $X_1, X_2, \dots, X_n, \dots$ – последовательность независимых одинаково распределённых случайных величин, с математическим ожиданием m и дисперсией σ^2 , то при неограниченном увеличении n закон распределения нормированной суммы неограниченно приближается к стандартно нормальному распределению:

$$P \left\{ \alpha < \frac{\sum_{k=1}^n X_k - n \cdot m}{\sqrt{n} \cdot \sigma} < \beta \right\} \xrightarrow{n \rightarrow \infty} \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} dx.$$

Частным случаем последней теоремы является теорема Муавра – Лапласа, которая в общем виде была доказана французским математиком и физиком Пьером Симоном Лапласом в 1812 году, а один частный случай был известен ещё английскому математику Абрахаму де Муавру (1730 г.).

Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p , то справедливо соотношение

$$P \left\{ \alpha < \frac{Z_n - n \cdot p}{\sqrt{n \cdot p \cdot q}} < \beta \right\} \approx \int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} dx,$$

где Z_n – число появлений события A в n опытах; $q=1-p$; n – достаточно велико.

Доказательство. Представим случайную величину Z_n – общее число появлений события A в n опытах в виде суммы

$$Z_n = \sum_{k=1}^n X_k,$$

Где X_k – индикатор (число появлений) события A в k опыте.

Случайная величина X_k имеет биномиальное распределение с параметром p , поэтому

$$M(X_k) = p; D(X_k) = p \cdot q.$$

В силу центральной предельной теоремы для одинаково распределенных слагаемых отсюда сразу следует доказываемое утверждение.