СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

1. Основные понятия

Часто результатом случайного эксперимента является число. Например, можно подбросить игральную кость и получить одно из чисел: 1,2,3,4,5,6. Можно подъехать к бензоколонке и обнаружить определённое число автомашин в очереди. Можно выстрелить из пушки и измерить расстояние от места выстрела до места падения снаряда. В таких случаях будем говорить, что имеем дело со случайной величиной.

Каждому исходу случайного эксперимента поставим в соответствие единственное число x_k — значение случайной величины. Тогда естественно рассматривать случайную величину как функцию, заданную на множестве исходов случайного эксперимента.

Случайной величиной называется величина, принимающая в результате опыта одно из своих возможных значений, причем заранее неизвестно, какое именно.

Случайные величины принято обозначать заглавными латинскими буквами: *X, Y, Z*,..., а их возможные значения - соответствующими строчными буквами: *x, y, z*,...

Различают два типа случайных величин: *дискретные и непрерывные*.

Случайная величина называется ***дискретной***, если она принимает отдельные, изолированные возможные значения с определенными вероятностями.

Случайная величина называется ***непрерывной***, если множество ее возможных значений целиком заполняет некоторый конечный или бесконечный промежуток.

Примеры случайных величин:

- дискретных: число попаданий или промахов в серии выстрелов, число выпаданий герба или решки при подбрасывании монеты, в схеме Бернулли повторяющихся независимых испытаний - число появлений события при *n* испытаниях и т.п.;

- непрерывных: отклонение размера детали от номинального, ресурс (время безотказной работы) системы, физические параметры системы (температура, давление, влажность) длина тормозного пути автомобиля и т.п.

Случайная величина полностью определена с вероятностной точки зрения, если известен ее закон распределения.

Законом распределения случайной величины называется соотношение между возможными значениями этой величины и соответствующими им вероятностями.

Закон распределения может быть задан таблично, графически или аналитически.

Способы задания случайной величины

1) Табличный. Случайную величину можно задать законом распределения вероятностей.

а) Таблица, в которой перечислены возможные значения *дискретной случайной* величины и соответствующие им вероятности, называется ***рядом распределения***:

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

Ряд распределения может быть построен только для дискретных случайных величин.

б) Таблица для непрерывной случайной величины.

x	$x_1 - x_2$	$x_2 - x_3$	$x_3 - x_4$	…	…	$x_{n-1} - x_n$
p	p_1	p_2	p_3	…	…	p_n

$p_1 + p_2 + ... + p_n = 1.$

Т.к. в таблице распределения рассмотрена полная система событий, а сумма вероятностей полной системы событий равна 1.

2) *Графический. Случайная величина может быть задана полигоном и гистограммой.*

Полигоном называется многоугольник, вершинами которого являются точки (x_i, p_i), взятые из закона распределения, а сторонами – отрезки, соединяющие эти точки.

Гистограммой называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых являются интервалы изменения случайной величины, а высотами – соответствующие им вероятности, или частости, или частоты.

Замечание: непрерывную случайную величину можно свести к дискретной, заменив интервалы изменения случайной величины срединами интервалов.

2. Дискретные случайные величины

В дальнейшем для краткости будем называть величину p_i вероятностью значения x_i случайной величины. Отметим, что закон распределения содержит всю информацию о случайной величине, и задать случайную величину можно, просто представив её закон распределения.

Заметим, что событие, заключающееся в том, что случайная величина примет одно из своих возможных значений, является достоверным, поэтому

$$\sum_{i=1}^{n(\infty)} p_i = 1.$$

Пример 1. По мишени производится 4 независимых выстрела с вероятностью попадания при каждом выстреле p = 0,8. Требуется: а) найти закон распределения дискретной случайной величины x, равной числу попаданий в мишень; б) найти вероятности событий: $1 \le x \le 3$; $x > 3$; в) построить многоугольник распределения.

Решение. а) Возможные значения случайной величины x: 0, 1, 2, 3, 4. Соответствующие вероятности вычисляем по формуле Бернулли:

$$p_0 = p(x=0) = C_4^0\, 0,8^0 \cdot 0,2^4 = 0,0016.$$

$$p_1 = p(x=1) = C_4^1\, 0,8 \cdot 0,2^3 = 0,0256.$$

$$p_2 = p(x=2) = C_4^2\, 0,8^2 \cdot 0,2^2 = 0,1536.$$

$p_3 = p(x=3) = C_4^3\, 0{,}8^3 \cdot 0{,}2 \ = 0{,}4096.$

$p_4 = p(x=4) = C_4^4\, 0{,}8^4 \cdot 0{,}2^0 = 0{,}4096.$

Закон распределения *x* представится таблицей:

x_i	0	1	2	3	4
p_i	0,0016	0,0256	0,1536	0,4096	0,4096

Проверка: 0,0016 + 0,0256 + 0,1536 + 0,4096 + 0,4096 = 1.

б) Вероятность событий $1 \le x \le 3$ и $x > 3$ равны:

$p\left(1 \le x \le 3\right) = p_1 + p_2 + p_3 = 0{,}0256 + 0{,}1536 + 0{,}4096 = 0{,}5888;$

$p\left(x > 3\right) = p_4 = 0{,}4096$.

в) Полигон:

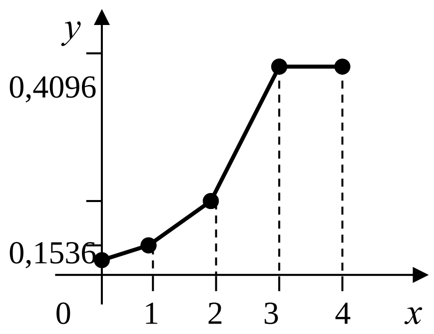

Рис. 1

3. Числовые характеристики случайных величин

Числа, позволяющие отразить наиболее существенные особенности распределения случайной величины, называются ***числовыми характеристиками случайной величины***. Важнейшими из них являются математическое ожидание и дисперсия.

3.1. Математическое ожидание

Математическим ожиданием *дискретной случайной величины* называется сумма произведений ее возможных значений на соответствующие им вероятности:

$M(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$. (3)

или

$$M(X)=\sum_{i=1}^{n} x_i p_i \qquad (4)$$

Если число возможных значений случайной величины бесконечно, то $M(X)=\sum_{i=1}^{\infty} x_i p_i$, если полученный ряд сходится абсолютно.

M(X) называют еще ***центром распределения*** или характеристикой положения случайной величины на числовой оси. Это среднее значение, вокруг которого группируются остальные возможные значения случайной величины.

Математическим ожиданием *непрерывной случайной величины X* называется среднее ожидаемое значение X, определяемое формулой:

$$M(X)=\int_{-\infty}^{+\infty} x f(x)\,dx. \qquad (5)$$

Если возможные значения непрерывной случайной величины $X\in(a,b)$, то

$$M(X)=\int_a^b xf(x)\,dx. \qquad (6)$$

Замечание 1. Математическое ожидание называют иногда ***взвешенным средним***, так как оно приближенно равно среднему арифметическому наблюдаемых значений случайной величины при большом числе опытов.

Замечание 2. Из определения математического ожидания следует, что его значение не меньше наименьшего возможного значения случайной величины и не больше наибольшего.

Замечание 3. Математическое ожидание дискретной случайной величины есть ***неслучайная*** (постоянная) величина. В дальнейшем увидим, что это же справедливо и для непрерывных случайных величин.

Пример 1. Найдем математическое ожидание случайной величины *X* – числа стандартных деталей среди трех, отобранных из партии в 10 деталей, среди которых 2 бракованных.

Решение. Составим ряд распределения для X. Из условия задачи следует, что X может принимать значения 1, 2, 3.

$$p(1)=\frac{C_8^1\cdot C_2^2}{C_{10}^3}=\frac{1}{15},\ \ p(2)=\frac{C_8^2\cdot C_2^1}{C_{10}^3}=\frac{7}{15},\ \ p(3)=\frac{C_8^3}{C_{10}^3}=\frac{7}{15}.$$

X	1	2	3
p	1/15	7/15	7/15

Тогда, $M(X)=1\cdot\frac{1}{15}+2\cdot\frac{7}{15}+3\cdot\frac{7}{15}=2{,}4.$

Свойства математического ожидания

1) Математическое ожидание постоянной равно самой постоянной:

$$M(C) = C. \tag{7}$$

2) Постоянный множитель можно выносит за знак математического ожидания:

$$M(CX) = C\,M(X). \tag{8}$$

Две случайные величины называются ***независимыми***, если закон распределения одной из них не зависит от того, какие значения приняла другая. В противном случае случайные величины ***зависимы***.

Назовем ***произведением независимых случайных величин X и Y*** случайную величину XY, возможные значения которой равны произведениям всех возможных значений X на все возможные значения Y, а соответствующие им вероятности равны произведениям вероятностей сомножителей.

3) Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(XY) = M(X)M(Y). \tag{9}$$

Замечание 1. Аналогично можно доказать это свойство для большего количества возможных значений сомножителей.

Замечание 2. Свойство 3 справедливо для произведения любого числа независимых случайных величин, что доказывается методом математической индукции.

Определим ***сумму случайных величин X и Y*** как случайную величину $X + Y$, возможные значения которой равны суммам каждого возможного значения X с каждым возможным значением Y; вероятности таких сумм равны произведениям вероятностей слагаемых (для зависимых случайных величин – произведениям вероятности одного слагаемого на условную вероятность второго).

4) Математическое ожидание суммы двух случайных величин (зависимых или независимых) равно сумме математических ожиданий слагаемых:

$$M(X + Y) = M(X) + M(Y). \qquad (10)$$

Замечание. Из свойства 4 следует, что сумма любого числа случайных величин равна сумме математических ожиданий слагаемых.

Кроме математического ожидания в качестве характеристик положения случайной величины применяется иногда *медиана* и *мода.*

Модой $\boldsymbol{M_o}$ дискретной случайной величины называется ее наиболее вероятное значение, ***модой M*** непрерывной случайной величины – значение, в котором плотность вероятности максимальна.

Если максимум один, то распределение называется ***одномодальным.***

Замечание 1. Если кривая распределения имеет больше одного максимума, распределение называется *полимодальным (многомодальным*), если эта кривая не имеет максимума, но имеет минимум – ***анти-модальным***.

Замечание 2. В общем случае мода и математическое ожидание не совпадают. Но, если распределение является симметричным и модальным (то есть кривая распределения симметрична относительно прямой $x = M$) и имеет математическое ожидание, оно совпадает с модой.

Медианой Me $\left(\mu(X)\right)$ непрерывной случайной величины называют такое ее значение, для которого

$$p\left(X < Me\right) = p\left(X > Me\right). \qquad (11)$$

Графически прямая $x = Me$ делит площадь фигуры, ограниченной кривой распределения, на две равные части.

Замечание. Для симметричного модального распределения медиана совпадает с математическим ожиданием и модой.

Для непрерывной случайной величины x это равенство принимает вид

$$\int_{-\infty}^{Me} p(x)dx = \int_{Me}^{\infty} p(x)dx = \frac{1}{2}.$$

Для дискретных случайных величин медиана определяется неоднозначно и практически не употребляется. Геометрически медиана - это абсцисса точки, в которой площадь, ограниченная графиком плотности распределения, делится пополам.

При описании непрерывного распределения используют иногда *квантили*. Для случайной величины X с функцией распределения $F(X)$ ***квантилью порядка p*** $(0 < p < 1)$ называется число K_p такое, что $F(K_p) \le p$, $F(K_p + 0) \ge p$. В частности, если $F(X)$ строго монотонна, K_p: $F(K_p) = p$.

4. 2. Дисперсия

Для того чтобы иметь представление о поведении случайной величины, недостаточно знать только ее математическое ожидание. Рассмотрим две случайные величины: X и Y, заданные рядами распределения вида

X	49	50	51
p	0,1	0,8	0,1

Y	0	100
p	0,5	0,5

Найдем $M(X) = 49 \cdot 0,1 + 50 \cdot 0,8 + 51 \cdot 0,1 = 50$, $M(Y) = 0 \cdot 0,5 + 100 \cdot 0,5 = 50$. Как видно, математические ожидания обеих величин равны, но если для X $M(X)$ хорошо описывает поведение случайной величины, являясь ее наиболее вероятным возможным значением (причем остальные значения ненамного отличаются от 50), то значения Y существенно отстоят от $M(Y)$. Следовательно, наряду с математическим ожиданием желательно знать, насколько значения случайной величины отклоняются от него. Для характеристики этого показателя служит дисперсия.

Дисперсией (рассеянием) случайной величины называется математическое ожидание квадрата ее отклонения от ее математического ожидания:

$$D(X) = M\left[X - M(X)\right]^2. \qquad (12)$$

Для вычисления дисперсии удобно пользоваться формулой

$$D\left(X\right) = M\left(X^2\right) - M^2\left(X\right). \qquad (13)$$

Для дискретных случайных величин $D(X) = \sum_{i=1}^{n}[x_1 - M(X)]^2 p_1$,

для непрерывных случайных величин

$$D(X) = \int_{-\infty}^{\infty}[x - M(X)]^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - M^2(X). \qquad (14)$$

Если возможные значения X принадлежат интервалу (a,b), то

$$D(X) = \int_{a}^{b}[x - M(X)]^2 f(x)dx = \int_{a}^{b} x^2 f(x)dx - M^2(X). \qquad (15)$$

Замечание 1. В определении дисперсии оценивается не само отклонение от среднего, а его квадрат. Это сделано для того, чтобы отклонения разных знаков не компенсировали друг друга.

Замечание 2. Из определения дисперсии следует, что эта величина принимает только неотрицательные значения.

Пример 1. Вычислим дисперсии случайных величин X и Y, рассмотренных в начале этого раздела.

Решение. $M(X) = (49^2 \cdot 0{,}1 + 50^2 \cdot 0{,}8 + 51^2 \cdot 0{,}1) - 50^2 = 2500{,}2 - 2500 = 0{,}2$. $M(Y) = (0^2 \cdot 0{,}5 + 100^2 \cdot 0{,}5) - 50^2 = 5000 - 2500 = 2500$. Итак, дисперсия второй случайной величины в несколько тысяч раз больше дисперсии первой. Таким образом, даже не зная законов распределения этих величин, по известным значениям дисперсии мы можем утверждать, что X мало отклоняется от своего математического ожидания, в то время как для Y это отклонение весьма существенно.

Свойства дисперсии

1) Дисперсия постоянной величины C равна нулю:

$$D(C) = 0. \qquad (16)$$

2) Постоянный множитель можно выносить за знак дисперсии, возведя его в квадрат:

$$D(CX) = C^2 D(X). \qquad (17)$$

3) Дисперсия суммы двух независимых случайных величин равна сумме их дисперсий:

$$D(X + Y) = D(X) + D(Y). \qquad (18)$$

Следствие 1. Дисперсия суммы нескольких взаимно независимых случайных величин равна сумме их дисперсий.

Следствие 2. Дисперсия суммы постоянной и случайной величин равна дисперсии случайной величины.

4) Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(X - Y) = D(X) + D(Y). \qquad (19)$$

Дисперсия дает среднее значение квадрата отклонения случайной величины от среднего; для оценки самого отклонения служит величина, называемая средним квадратическим отклонением.

Средним квадратическим отклонением σ случайной величины X называется квадратный корень из дисперсии:

$$\sigma = \sqrt{D(X)}. \qquad (20)$$