Высшая школа экономики

Факультет прикладной политологии

Математика для политологов

И.А.Хованская, И.В.Щуров, К.И.Сонин (РЭШ), А.В.Михайлович

15 октября 2012 г.

Лекция 7. Элементы теории вероятностей, часть 5

Характеристики случайных величин

1 Математическое ожидание случайной величины

Допустим, случайная величина принимает два значения — 0 и 5000. Какое у неё будет *среднее значение*? На первый взгляд кажется разумным взять среднее арифметическое:

$$\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 5000 = 2500. \tag{1}$$

Однако, рассмотрим такой пример:

Пример 1.1. *Вероятность выиграть в мгновенную лотерею равна* 1/100. *Сумма выигрыша — 5000 руб. Если рассмотреть выигрыш как случайную величину, каково её среднее?*

Распределение такой случайной величины (мы будем обозначать её через z) приведено в табл. 1. Сейчас

значение z	0	5000
вероятность	1/100	99/100

Таблица 1: Распределение выигрыша в лотерею y

кажется очевидным, что «средний выигрыш» в такой игре не будет равен 2500 руб, а будет существенно меньше. Чему же он тогда равен? Допустим, мы играем в лотерею $100'000$ раз. Из них мы примерно $100'000 \cdot \frac{99}{100} = 99'000$ раз проиграем, и примерно $100'000 \cdot \frac{1}{100} = 1000$ раз выиграем. За каждый выигрыш мы получим 5000 руб., а за проигрыш не получим ничего. То есть общий выигрыш составит примерно $99'000 \cdot 0 + 1000 \cdot 5000 = 5'000'000$ руб. В среднем на одну игру выигрыш составит

$$\frac{99'000 \cdot 0 + 1000 \cdot 5000}{100'000} = \frac{99}{100} \cdot 0 + \frac{1}{100} \cdot 5000 = 50 \text{ руб.} \tag{2}$$

Это и есть «настоящее» среднее описанной случайной величины. (Или, что то же самое, «честная» стоимость лотерейного билетика. Если билетик стоит дороже, излишек идет в карман организатору лотереи, если дешевле — организатор быстро разорится.) Нетрудно видеть, что в отличие от среднего арфиметического, каждое слагаемое здесь берется с множителем, равным вероятности его реализации. (Ср. формулы 1 и 2.) Аналогичным образом — просуммировав значения, которые принимает случайная величина, умноженные на соответствующие вероятности — можно найти среднее (оно называется *математическим ожиданием*) любой дискретной случайной величины.

Рассмотрим еще один пример.

Пример 1.2. *Мы подкидываем монетку три раза. Случайная величина (обозначим её за x) — число выпавших орлов.*

Каково «среднее» число орлов, выпадающих в трёх бросаниях? Поскольку орлы и решки равновероятны, представляется логичным, что среднее число орлов равно среднему числу решек. С другой стороны, в каждом испытании сумма числа выпавших орлов и числа выпаших решек равна 3 (количеству подкидываний монетки). Значит, в «среднем» должно выпадать 1,5 орла и 1,5 решки. Давайте проверим, так ли это. Для этого используем таблицу распределения рассматриваемой случайной величины:

значение x	исходы	вероятность
0	PPP	1/8
1	PPO, POP, OPP	3/8
2	OOP, OPO, POO	3/8
3	OOO	1/8

Таблица 2: Распределение значений случайной величины x

Вычислим математическое ожидание случайной величины x (оно обозначается Ex, от *Expected value*):

$$Ex = \frac{1}{8} \cdot 0 + \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 = \frac{12}{8} = \frac{3}{2} = 1{,}5 \tag{3}$$

Интуиция нас не обманула: среднее действительно равно 1,5.

2 Дисперсия случайной величины

Математическое ожидание — важная характеристика случайной величины. Но зачастую знания только мат.ожидания недостаточно, чтобы делать какие-либо выводы. Давайте представим себе «честное казино». Вы можете в нём поставить произвольную сумму, например, на «красное» (или на «нечет»), и с вероятностью 1/2 выиграть и удвоить свою ставку (то есть выиграть столько же, сколько поставили), а с вероятностью 1/2 — проиграть и потерять её. Допустим, один человек ставит 10 рублей, а другой — миллиард рублей. Можно рассмотреть две случайные величины — выигрыш первого игрока (x_1) и выигрыш второго игрока (x_2). (Если игрок проигрывает свою ставку, то его выигрыш отрицателен.) Распределения этих случайных величин приведены в таблице 3. Нетрудно видеть, что в обоих случаях математическое ожидание равно

вероятность	1/2	1/2
значение x_1	-10	10
значение x_2	-1000'000'000	1000'000'000

Таблица 3: Распределение выигрыша в «честном казино»

нулю (поэтому мы и назвали такое казино «честным» — на практике, конечно, так не бывает, и математическое ожидание выигрыша в казино всегда отрицательно). Тем не менее, очевидно, что второй игрок рискует гораздо большими суммами (хотя может и выиграть гораздо больше), чем первый. На практике, это явно различные ситуации — вы, вероятно, были бы готовы рискнуть десятью рублями (имея некоторую толику азартности), но вряд ли согласитесь поставить на кон большую сумму, даже если знаете, что «в среднем» в обоих случаях ничего не потеряете и не приобретете.

Случайные величины x_1 и x_2 отличаются тем, насколько далеко они отклоняются от своего среднего значения (в данном случае, нуля). Если мы попытаемся просто посчитать «среднее отклонение от среднего значения», то есть $E(x - Ex)$, то получим всегда нуль (потому что отклонение в положительную сторону и в отрицательную берутся с разными знаками и компенсируют друг друга при сложении). Чтобы избежать этой проблемы рассматривают *средний квадрат отклонения от среднего значения*, то есть $E((x-Ex)^2)$. Эта величина и называется *дисперсией* случайной величины.

Пример 2.1. *Вычислим дисперсию средней величины y — «число очков, выпавших на кубике».*

Математическое ожидание Ey равно 3,5 (проверьте!). Чтобы вычислить дисперсию, вместе с величиной y нужно рассмотреть еще две случайные величины: $y - Ey = y - 3.5$ (например, если y принимает значение 1, случайная величина $y - 3.5$ принимает значение $1 - 3.5 = -0.5$) и $(y - Ey)^2$ (если $y - Ey$ принимает значение 1.5, то $(y - Ey)^2$ принимает значение $1.5^2 = 2.25$). Математическое ожидание последней и будет дисперсией y. Рассмотрим соответствующие таблицы распределений (см. табл. 4).

Таким образом,

$$Dy = E((y - Ey)^2) = \frac{1}{6} \cdot 6.25 + \frac{1}{6} \cdot 2.25 + \frac{1}{6} \cdot 0.5 + \frac{1}{6} \cdot + \frac{1}{6} \cdot 2.25 + \frac{1}{6} \cdot 6.25 = 15/6 \tag{4}$$

Это и есть дисперсия величины y.

вероятность	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$
значение y	1	2	3	4	5	6
значение $y - Ey$	-2.5	-1.5	-0.5	0.5	1.5	2.5
значение $(y - Ey)^2$	6.25	2.25	0.25	0.25	2.25	6.25

Таблица 4: Вычисление дисперсии y