Тема 7. Дисперсия и неравенство Чебышёва

Прежде, чем определить дисперсию, выведем полезную формулу для вычисления математического ожидания некоторой функции от случайной величины.

Утверждение. Пусть $Y = \varphi(X)$, где $\varphi(x)$ — заданная функция, $p_i = \mathbf{P}(X = x_i)$. Тогда верно равенство

$$\mathsf{M}Y = \mathsf{M}\varphi(X) = \sum_i \varphi(x_i) p_i. \tag{1}$$

Ради простоты выведем формулу (1) в классической вероятностной модели с равно-возможными исходами ω. Положим $\mathsf{A}_i = \{\omega\colon X(\omega) = x_i\}$. Доказательство фактически повторяет вывод формулы (1) из темы 6:

$$\mathsf{M}Y = \frac{1}{|\Omega|}\sum_{\omega\in\Omega} Y(\omega) = \frac{1}{|\Omega|}\sum_i \sum_{\omega\in\mathsf{A}_i} \varphi(x_i) = \frac{1}{|\Omega|}\sum_i \varphi(x_i)|\mathsf{A}_i| = \sum_i \varphi(x_i) p_i.$$

Пример 1. Пусть X — число очков, выпадающее на игральной кости, $\varphi(x) = x^3$. Согласно формуле (1)

$$\mathsf{M}Y = \mathsf{M}X^3 = (1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3)/6 = 441/6 = 73{,}5.$$

Рассмотрим важный частный случай формулы (1) для функции $\varphi(x) = (x - \mathsf{M}X)^2$.

Определение. *Дисперсия* $\mathsf{D}X$ случайной величины X задаётся следующей формулой:

$$\mathsf{D}X = \mathsf{M}(X - \mathsf{M}X)^2 = \sum_i (x_i - \mathsf{M}X)^2 p_i. \tag{2}$$

Дисперсия характеризует степень «разброса» распределения относительно его центра — математического ожидания $\mathsf{M}X$. Для вычисления дисперсии удобно использовать формулу

$$\mathsf{D}X = \mathsf{M}X^2 - (\mathsf{M}X)^2 = \sum_i x_i^2 p_i - (\mathsf{M}X)^2. \tag{3}$$

Пример 2. Вычислим по формуле (3) дисперсию индикатора I_A события A, где $\mathbf{P}(\mathsf{A}) = p$:

$$\mathsf{D}I_\mathsf{A} = 0^2\cdot(1-p) + 1^2\cdot p - p^2 = p - p^2 = p(1-p).$$

Простейшие свойства дисперсии

1) $\mathsf{D}c = 0,$ где c — произвольная константа.
2) $\mathsf{D}(X + c) = \mathsf{D}X.$
3) $\mathsf{D}(cX) = c^2\mathsf{D}X.$

Так же, как и для математического ожидания, в механике существует аналог дисперсии, называемый *моментом инерции относительно центра масс*.

Вопрос по механике. На тонком стержне (числовой прямой) в точках с координатами x_i находятся массы m_i (рис. 1). Где следует выбрать точку a крепления стержня к вертикальной оси, чтобы минимизировать *момент инерции* относительно неё $I(a) = \sum (x_i - a)^2 m_i$?

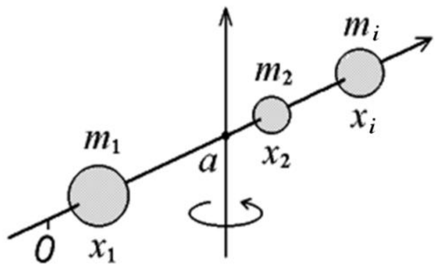

Рис. 1

Ответ. Наименьшее значение функция $I(a)=\sum (x_i-a)^2 m_i$ достигает при $a = x_{ц.м.}$ (см. задачу 7.1). В теории вероятностей аналогом $I(x_{ц.м.})$ является дисперсия случайной величины.

Неравенство Чебышёва

Выведем полезное вероятностное неравенство, впервые полученное выдающимся русским математиком Пафнутием Львовичем Чебышёвым (1821 – 1894).

Отметим, что для произвольного события A индикатор I_{A} имеет распределение Бернулли с параметром $p=\mathbf{P}(\mathrm{A})$. Поэтому $\mathrm{M} I_{\mathrm{A}}=\mathbf{P}(\mathrm{A})$.

Пусть Y — произвольная <u>неотрицательная</u> случайная величина, константа $c>0$. Рассмотрим событие $\mathrm{A}=\{Y\ge c\}$. Применяя свойства математического ожидания из темы 6, получим:

$$\mathrm{M}Y=\mathrm{M}(Y\cdot I_{\mathrm{A}}) + \mathrm{M}(Y\cdot I_{\bar{\mathrm{A}}}) \ge \mathrm{M}(cI_{\mathrm{A}}) + \mathrm{M}(Y\cdot I_{\bar{\mathrm{A}}}) \ge c\,\mathrm{M}I_{\mathrm{A}}+0 = c\,\mathbf{P}(Y\ge c).$$

Таким образом, установлена справедливость неравенства

$$\mathbf{P}(Y\ge c) \le \frac{\mathrm{M}Y}{c}. \tag{4}$$

Пусть теперь X — произвольная случайная величина. Из неравенства (4), взяв в качестве Y неотрицательную случайную величину $(X-\mathrm{M}X)^2$, выводим ***неравенство Чебышёва***:

$$\mathbf{P}(|X-\mathrm{M}X|\ge c)=\mathbf{P}((X-\mathrm{M}X)^2\ge c^2) \le \frac{\mathrm{M}(X-\mathrm{M}X)^2}{c^2} = \frac{\mathrm{D}X}{c^2}. \tag{5}$$

Оно позволяет оценивать вероятность большого отклонения значения произвольной случайной величины X от своего математического ожидания $\mathrm{M}X$ с помощью дисперсии $\mathrm{D}X$.

Домашнее задание

Если <u>**последней**</u> буквой вашего **имени** служит:

А, Б, В, Г, Д, Е, Ё, то «своими» являются задачи 7.1 и 7.5;
Ж, З, И, Й, К, Л, М, то «своими» являются задачи 7.2 и 7.6;
Н, О, П, Р, то «своими» являются задачи 7.3 и 7.7;
С, Т, У, Ф, Х, Ц, Ч, Ш, Щ, Ъ, Ы, Ь, Э, Ю, Я, то «своими» являются задачи 7.4 и 7.8.

7.1) Для случайной величины X найти минимум по переменной a функции $g(a)=\mathrm{M}(X-a)^2$.
(*Указание*. Возведите в квадрат и используйте свойства математического ожидания из темы 6.)

7.2) Вычислить дисперсию (записать ответ в виде несократимой дроби):

а) числа очков, выпадающих при бросании двух игральных костей;

б) числа «гербов», выпадающих при одновременном бросании пяти монет.

7.3) Симметричную монету бросают 3 раза. Здесь ω = $(i,\ j,\ k)$, где компоненты вектора ω принимают значения 0 или 1; |Ω| = 2^3 = 8. Вычислить: а) $\mathrm{D}X$, где $X = i + 2jk$; б) $\mathrm{D}Y$, где $Y = 2i + jk$.

7.4) В урне лежат два шара с номерами 1 и 2. Три раза наудачу с возвращением выбирается один из шаров. Здесь ω = (i_1, i_2, i_3). Вычислить: а) $\mathrm{D}X$, где $X = \min\{i_1, i_2, i_3\}$; б) $\mathrm{D}Y$, где $Y = \max\{i_1, i_2, i_3\}$.

7.5) Найти дисперсию (запишите ответ задачи без знака суммы):

а) случайной величины S_n с биномиальным распределением: $\mathbf{P}(S_n = i) = C_n^i p^i (1-p)^{n-i},\ i = 0, 1, \dots, n$;

б) случайной величины N с пуассоновским распределением: $\mathbf{P}(N = i) = \lambda^i e^{-\lambda} / i!,\ \ i = 0, 1, 2, \dots$.

(*Указание*. Используйте тождество $i^2 = i(i-1) + i$.)

7.6) Вычислить $\mathrm{D}L$, где L имеет геометрическое распределение: $\mathbf{P}(L = i) = q^{i-1} p,\ i = 1, 2, \dots, q = 1 - p$.

7.7) Пусть ω = $(i_1,\ i_2,\ \dots,\ i_n)$ — случайная перестановка, |Ω| = $n!$. Положим $X_k(\omega) = i_k$, где k = 1, …, n. Вычислить $\mathrm{M}X_k$ и $\mathrm{D}X_k$. Записать ответ без знака суммы, приведя выражение к общему знаменателю.

7.8) По m ящикам раскладываются наудачу n занумерованных шаров. Случайная величина K — число шаров, оказавшихся в первом ящике. Вычислить $\mathrm{D}K$ (в ответе не должна присутствовать сумма).

7.9)* Пусть Z — число белых шаров при выборе без возвращения n шаров из урны с l белыми и $m - l$ чёрными шарами. Найти $\mathrm{D}Z$. Записать ответ, приведя выражение к общему знаменателю. (*Указание.* Заметим, что $Z = I_1 + \ldots + I_n$, где I_k — индикатор того, что k-й шар оказался белым.)