Математическое ожидание, дисперсия, ковариация

Определение математического ожидания

- значение случайной величины нам заранее не известно, тем не менее нам интересно понять, как она устроена “в среднем”;
- **определение:** математическим ожиданием случайной величины $X(\omega)$ назовем
$$\mathbb{E}\,X = \sum_{k=1}^{\infty} X(\omega_k)\,\mathbb{P}(\omega_k),$$
где ряд предполагается сходящимся абсолютно (иначе мы не можем менять порядок слагаемых);
- **пример:** математическое ожидание числа, выпавшего при подбрасывании кубика;
- когда речь идет о случайной величине и ее численных характеристиках, нам достаточно знать лишь ее распределение (нам не важно как устроено вероятностное пространство, на котором случайная величина определена);
- как же переписать математическое ожидание, пользуясь лишь ее распределением?
$$\mathbb{E}\,X = \sum_{k} x_k\,\mathbb{P}_X(k),$$
где $\mathbb{P}_X(\cdot)$ – функция масс случайной величины X (переформуллировка справедлива, если ряд в правой части сходится абсолютно);
- переформуллировка математического ожидания в терминах X^+ и X^-.

Математическое ожидание функции от случайной величины

- пусть имеем набор случайных величин $X_1, X_2, \ldots, X_m$, принимающих значения $x_{1,1}, \ldots, x_{1,n_1};\ \ldots;\ x_{m,1}, \ldots, x_{m,n_m}$, соответственно; пусть также $g : \mathbb{R}^m \to \mathbb{R}$ – измеримое отображение, тогда $g(X_1, \ldots, X_m)$ – случайная величина;
- как искать ее математическое ожидание?
- **первый способ:** найти распределение этой случайной величины (с какой вероятностью она принимает конкретные значения) – может занять много усилий;
- **второй способ:** воспользоваться формулой
$$\mathbb{E}\,g(X_1, \ldots, X_m) = \sum_{i_1,\ldots,i_m} g(x_{1,i_1}, \ldots, x_{m,i_m})\,\mathbb{P}_{\overrightarrow{X}}(x_{1,i_1}, \ldots, x_{m,i_m}),$$
если ряд сходится абсолютно (здесь $\mathbb{P}_{\overrightarrow{X}}(\cdot, \ldots, \cdot)$ – функция масс случайного вектора $\overrightarrow{X} = (X_1, \ldots, X_m)$);
- **третий способ:** пусть X – целочисленная неотрицательная случайная величина, тогда ее **производящей функцией** назовем
$$\phi_X(s) = \sum_{k=0}^{\infty} \mathbb{P}_X(k) s^k,$$
где $\mathbb{P}_X(\cdot)$ – функция масс;

- заметим, что **степенной** ряд в определении $\phi_X(s)$ сходится в точке $s = 1$, а значит сходится и в интервале $(-1, 1)$;

- более того, внутри этого интервала сумма ряда $\phi_X(s)$ дифференцируема и ее производная равна
$$\phi'_X(s) = \sum_{k=1}^{\infty} k\,\mathbb{P}_X(k)s^{k-1};$$

- если ряд в правой части сходится в точке $s = 1$ (то есть математическое ожидание существует), то у $\phi_X(s)$ в точке $s = 1$ существует производная слева, и она равна
$$\phi'_X(1) = \sum_{k=1}^{\infty} k\,\mathbb{P}_X(k) = \mathbb{E}\,X;$$

- таким образом, если математическое ожидание существует, то его можно найти по алгоритму:
  - найти производящую функцию $\phi_X(s)$;
  - найти ее производную;
  - подставить в результат точку $s = 1$.

Свойства математического ожидания

- математическое ожидание константы: $\mathbb{E}\,c = c$, где c – число;

- математическое ожидание индикатора события: $\mathbb{E}\,\mathbb{I}_A = \mathbb{P}(A)$, где A – некоторое событие;

- линейность: для любых чисел a и b и случайных величин X, Y с конечным математическим ожиданием
$$\mathbb{E}\,(aX + bY) = a\,\mathbb{E}\,X + b\,\mathbb{E}\,Y;$$

- мультипликативность: если случайные величины X и Y независимы и имеют конечное математическое ожидание, то
$$\mathbb{E}\,XY = \mathbb{E}\,X\,\mathbb{E}\,Y.$$

Неравенства для математических ожиданий (дополнительно)

- неравенство Йенсена: пусть f – выпуклая функция, тогда $\mathbb{E}\,f(X) \geq f(\mathbb{E}\,X)$, если математические ожидания существуют;

- неравенство Ляпунова: пусть $p > q > 0$, $\mathbb{E}|X|^p < \infty$, тогда
$$(\mathbb{E}|X|^q)^{\frac{1}{q}} \leq (\mathbb{E}|X|^p)^{\frac{1}{p}};$$

- неравенство Гельдера: пусть $p, q > 1$, $1/p + 1/q = 1$ и математические ожидания $\mathbb{E}|X|^q < \infty$, $\mathbb{E}|X|^p < \infty$, тогда
$$\mathbb{E}\,|XY| \leq (\mathbb{E}|X|^q)^{\frac{1}{q}}\,(\mathbb{E}|X|^p)^{\frac{1}{p}};$$

- неравенство Коши-Буняковского: частный случай неравенства Гельдера для $p = q = 2$.

Дисперсия и ковариация

- математическое ожидание описывает среднее; логично рассмотреть также дисперсию – величину, характеризующую разброс вокруг среднего;
- **определение:** дисперсией случайной величины X называется число
$$\mathbb{D}\,X = \mathbb{E}\,(X - \mathbb{E}\,X)^2 \geq 0;$$
- более удобная формула:
$$\mathbb{D}\,X = \mathbb{E}\,X^2 - (\mathbb{E}\,X)^2\,;$$
- еще один способ подсчета: если известна производящая функция $\phi_X(s)$ и дисперсия существует, то
$$\mathbb{D}\,X = \phi''_X(1) + \phi'_X(1) - \left(\phi'_X(1)\right)^2.$$
- **определение:** ковариацией случайных величин X, Y называют число
$$\operatorname{cov}(X,Y) = \mathbb{E}\,(X - \mathbb{E}\,X)\,(Y - \mathbb{E}\,Y)\,.$$

Свойства дисперсии и ковариации

- $\mathbb{D}\,X = \operatorname{cov}(X,X)\,;$
- $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)\,;$
- $\operatorname{cov}(X,Y+c) = \operatorname{cov}(X,Y)\,,\ \mathbb{D}\,(X+c) = \mathbb{D}\,X;$
- если X, Y независимы, то $\operatorname{cov}(X,Y) = 0;$
- $\mathbb{D}\,(aX) = a^2\,\mathbb{D}\,X;$
- $\mathbb{D}\,(X_1 + \ldots + X_n) = \sum_{i=1}^n \mathbb{D}\,X_i + 2\sum_{i<j}\operatorname{cov}(X_i, X_j)\,.$

Задачи

Задача 1. *При подбрасывании коробка со спичками вероятность того, что он выпадет этикеткой кверху (этикетка наклеена только на одну грань коробка) равна p (назовем такое событие "успехом"). Коробок подбросили n раз. Случайная величина X – число "успехов". Найдите $\mathbb{E}\,X$ и $\mathbb{D}\,X$.*

Задача 2. *Студент выучил m вопросов из n (здесь $m < n$). Экзаменатор спросил k вопросов (предположим, что $k \leq m$ и $k \leq n - m$). Случайная величина X – число известных студенту вопросов, среди заданных. Найдите $\mathbb{E}\,X$ и $\mathbb{D}\,X$.*

Задача 3. *Первый стрелок попадает в мишень с вероятностью* 0.4, *а второй с вероятностью* 0.8. *Они сделали по n выстрелов каждый. Случайная величина X – разность между числом попаданий второго и первого. Найдите $\mathbb{E}\,X$ и $\mathbb{D}\,X$.*

Задача 4. $X \sim Poiss(\lambda)$, *найдите* $\mathbb{E}\,X,\ \mathbb{D}\,X$.

Задача 5. *Совместное распределение пары (X,Y) задано таблицей*

X Y	0	1	2
-1	0.2	0.05	0.15
0	0.05	0.1	0.05
1	0.05	0.15	0.2

Найдите $\mathbb{E}\,X$, $\mathbb{E}\,Y$, $\mathbb{E}\,XY$, $\mathbb{D}\,X$, $\mathbb{D}\,Y$, $\operatorname{cov}(X,Y)$.