

Теория вероятностей и математическая статистика
Характеристики случайных величин. Биномиальное распределение.

Глеб Карпов

ФКН ВШЭ

Свойства математического ожидания

1. **Линейность:** Для любых констант a и b :

$$E[aX + b] = aE[X] + b$$

2. **Константа:** Если c - константа, то:

$$E[c] = c$$

3. **Ожидание функции:** Для функции $Y = g(X)$:

$$E[g(X)] = \sum_{x \in \Omega_X} g(x)P(X = x)$$

Доказательство:

$$\begin{aligned} E(Y) &= \sum_{y \in \Omega_Y} (y P(Y = y)) = \\ &= \sum_{y \in \Omega_Y} \left(y \sum_{x \in g^{-1}(y)} P(X = x) \right) = \sum_{x \in \Omega_X} g(x)P(X = x) \end{aligned}$$

Математическое ожидание

Пример

Посмотрим на примеры вычисления математического ожидания на основе уже посчитанной ранее задачи.

Пусть распределение случайной величины задано таблицей

x	0	-1	-2	1	2	3
$p_X(x)$	0.1	0.1	0.1	0.2	0.2	0.3

Постройте распределение (функцию вероятности) случайной величины $Y = X^2 + 10$.

Наша задача: восстановить новое пространство элементарных исходов Ω_Y для величины Y и выразить новую функцию вероятности $P_Y(y)$ через старую известную $P_X(x)$.

Математическое ожидание

Пример

x	0	-1	-2	1	2	3
$p_X(x)$	0.1	0.1	0.1	0.2	0.2	0.3

y	10	11	14	19
$p_Y(y)$	0.1	0.3	0.3	0.3

- $E[X] = \sum_{i=1}^6 x_i P_X(X = x_i) = -0.1 - 0.2 + 0.2 + 0.4 + 0.9 = 1.2$
- Первый подход к вычислению $E[Y]$ - по определению, зная таблицу функции вероятности для Y :

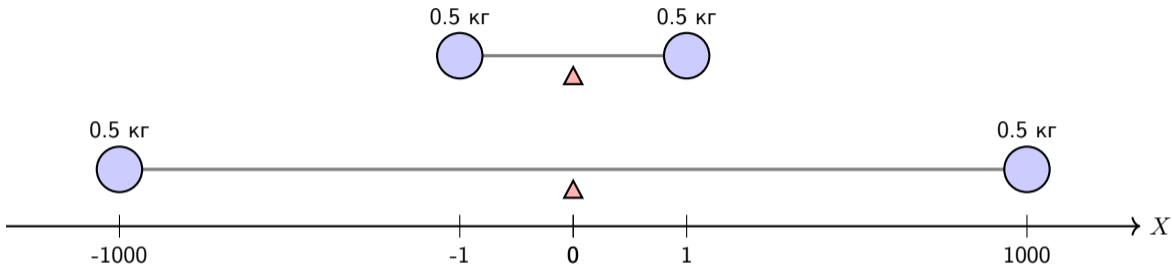
$$E[Y] = \sum_{i=1}^4 y_i P_Y(Y = y_i) = 10 \cdot 0.1 + 11 \cdot 0.3 + 14 \cdot 0.3 + 19 \cdot 0.3 = 14.2$$

- Второй подход к вычислению $E[Y]$ - по свойству **3** с предыдущего слайда. Главная идея: мы можем миновать шаг построения таблицы вероятности для Y и сразу предсказать, каким будет $E[Y]$, по таблице вероятностей для X и зная вид функции $Y = g(X)$, в нашем примере $Y = X^2 + 10$:

$$\begin{aligned} E[Y] &= \sum_{i=1}^6 g(x_i) P_X(X = x_i) = \sum_{i=1}^6 (x_i^2 + 10) P_X(X = x_i) = \\ &= 10 \cdot 0.1 + 11 \cdot 0.1 + 14 \cdot 0.1 + 11 \cdot 0.2 + 14 \cdot 0.2 + 19 \cdot 0.3 = 14.2 \end{aligned}$$

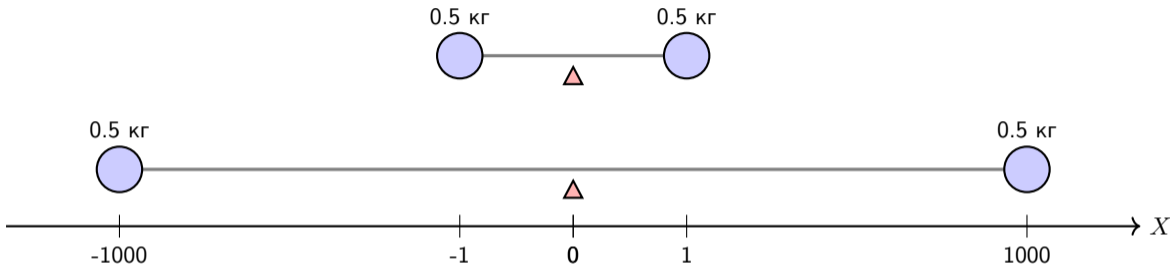
Слайд для записей

Физическое Лирическое отступление 2



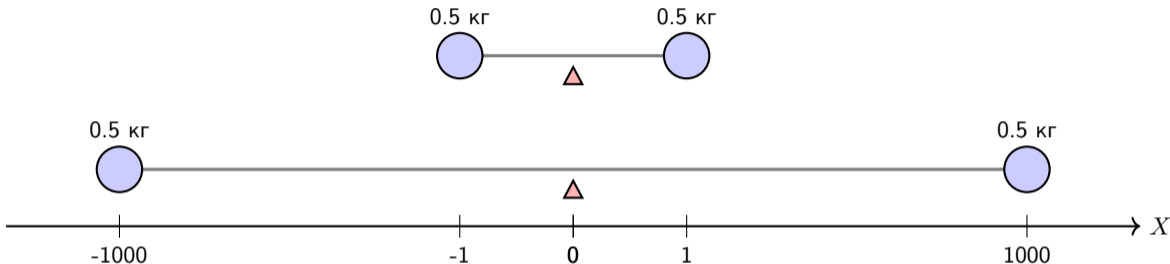
- В вероятностном мире интерпретация картинки такая, что может быть две случайные величины с одинаковым математическим ожиданием, но с кардинально разными значениями.

Физическое Лирическое отступление 2



- В вероятностном мире интерпретация картинки такая, что может быть две случайные величины с одинаковым математическим ожиданием, но с кардинально разными значениями.
- Представьте, что в одном случае вы кидаете монетку с привязанными численными значениями $+1$ и -1 (например, выигрыш - проигрыш), а в другом ту же монетку, но с привязанными значениями уже $+1000$ и -1000 . Математическое ожидание в обоих случаях будет 0, но сравните магнитуды возможных значений!

Физическое Лирическое отступление 2



- В вероятностном мире интерпретация картинки такая, что может быть две случайные величины с одинаковым математическим ожиданием, но с кардинально разными значениями.
- Представьте, что в одном случае вы кидаете монетку с привязанными численными значениями $+1$ и -1 (например, выигрыш - проигрыш), а в другом ту же монетку, но с привязанными значениями уже $+1000$ и -1000 . Математическое ожидание в обоих случаях будет 0 , но сравните магнитуды возможных значений!
- Это помогает нам понять мотивацию введения новой характеристики для случайной величины - **дисперсии** - которая сможет охарактеризовать степень разброса значений случайной величины.

Характеристики случайной величины

Дисперсия

- Дисперсия случайной величины - это число, **константа**, которое помогает нам понять степень разброса потенциальных значений случайной величины относительно среднего.

Характеристики случайной величины

Дисперсия

- Дисперсия случайной величины - это число, **константа**, которое помогает нам понять степень разброса потенциальных значений случайной величины относительно среднего.

i Definition

Для случайной величины X её дисперсия обозначается $Var[X]$ (Variance) и определяется как:

$$Var[X] = E [(X - E[X])^2]$$

Характеристики случайной величины

Дисперсия

- Дисперсия случайной величины - это число, **константа**, которое помогает нам понять степень разброса потенциальных значений случайной величины относительно среднего.

i Definition

Для случайной величины X её дисперсия обозначается $Var[X]$ (Variance) и определяется как:

$$Var[X] = E [(X - E[X])^2]$$

- Формально говоря, дисперсия это математическое ожидание новой случайной величины:

$$Y = g(X) = (X - E[X])^2,$$

описывающей квадрат расстояния между значением случайной величины X и её математическим ожиданием.

Характеристики случайной величины

Дисперсия

- Дисперсия случайной величины - это число, **константа**, которое помогает нам понять степень разброса потенциальных значений случайной величины относительно среднего.

i Definition

Для случайной величины X её дисперсия обозначается $Var[X]$ (Variance) и определяется как:

$$Var[X] = E [(X - E[X])^2]$$

- Формально говоря, дисперсия это математическое ожидание новой случайной величины:

$$Y = g(X) = (X - E[X])^2,$$

описывающей квадрат расстояния между значением случайной величины X и её математическим ожиданием.

- Также вводится понятие стандартного отклонения (standard deviation, std) величины X , которое является квадратным корнем из дисперсии: $std[X] = \sqrt{Var[X]}$.

Характеристики случайной величины

Дисперсия

Вычисление дисперсии

Существует упрощенная формула:

$$\begin{aligned} E[(X - E[X])^2] &= E[X^2 - 2XE[X] + (E[X])^2] = \\ &= E[X^2] - E[2E[X]X] + E[(E[X])^2] = E[X^2] - 2E[X]E[X] + (E[X])^2 = \\ & \quad \boxed{E[X^2] - (E[X])^2} \end{aligned}$$

В процессе мы используем свойство линейности математического ожидания, а также знание того, что $E[X]$ - это константа, при этом $E[c] = c$, то есть $E[E[X]] = E[X]$.

Дисперсия

Свойства

1. **Неотрицательность:** Для любой случайной величины X :

$$\text{Var}[X] \geq 0$$

2. **Дисперсия константы:** Если c - константа, то:

$$\text{Var}[c] = 0$$

3. **Масштабирование:** Для любой константы c :

$$\text{Var}[cX] = c^2 \text{Var}[X]$$

Дисперсия

Пример из двух частей: часть 1

Посчитаем дисперсию на примере нашей задачи. Считаем уже известным, что $E[X] = 1.2$.

x	0	-1	-2	1	2	3
$p_X(x)$	0.1	0.1	0.1	0.2	0.2	0.3

- Нам нужно вычислить $E[Y]$, где $Y = (X - E[X])^2$. Сразу воспользуемся подходом, позволяющим миновать шаг построения таблицы вероятности для Y :

$$\begin{aligned} E[Y] &= \sum_{i=1}^6 g(x_i) P_X(X = x_i) = \sum_{i=1}^6 (x_i - 1.2)^2 P_X(X = x_i) = \\ &= 1.44 \cdot 0.1 + 4.84 \cdot 0.1 + 10.24 \cdot 0.1 + 0.04 \cdot 0.2 + 0.64 \cdot 0.2 + 3.24 \cdot 0.3 = 2.76 \end{aligned}$$

- Но для дисперсии у нас есть ещё более умный подход, описанный на одном из слайдов выше:

$$\begin{aligned} \text{Var}[X] &= E[X^2] - (E[X])^2 = 4.2 - (1.2)^2 = 2.76 \\ E[X^2] &= \sum_{i=1}^6 (x_i)^2 P_X(X = x_i) = 0.1 + 4 \cdot 0.1 + 1 \cdot 0.2 + 4 \cdot 0.2 + 9 \cdot 0.3 = 4.2 \end{aligned}$$

- Voilà!

Дисперсия

Пример из двух частей: часть 2

Предположим, у нас есть другая случайная величина W и ее функция вероятности задана:

w	-6	0	1	15
$p_W(x)$	0.1	0.5	0.3	0.1

- Посчитаем математическое ожидание:

$$E[W] = -6 \cdot 0.1 + 0 + 1 \cdot 0.3 + 15 \cdot 0.1 = 1.2 = E[X] \text{ из предыдущего примера}$$

Дисперсия

Пример из двух частей: часть 2

Предположим, у нас есть другая случайная величина W и ее функция вероятности задана:

w	-6	0	1	15
$p_W(x)$	0.1	0.5	0.3	0.1

- Посчитаем математическое ожидание:

$$E[W] = -6 \cdot 0.1 + 0 + 1 \cdot 0.3 + 15 \cdot 0.1 = 1.2 = E[X] \text{ из предыдущего примера}$$

- Посчитаем дисперсию для W :

$$\text{Var}[W] = E[W^2] - (E[W])^2 = 26.4 - (1.2)^2 = 24.96$$

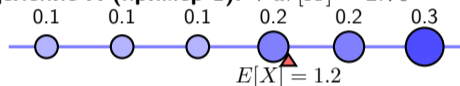
$$E[W^2] = \sum_{i=1}^4 (w_i)^2 \cdot P_W(W = w_i) = 36 \cdot 0.1 + 0 + 1 \cdot 0.3 + 225 \cdot 0.1 = 26.4$$

Дисперсия

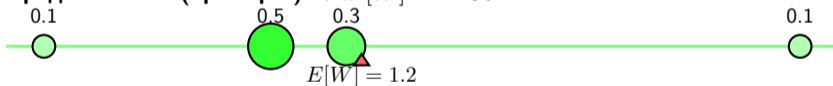
Снова о физической интерпретации

Дисперсия показывает, насколько “разбросаны” значения относительно среднего. Представим случайные величины как массы на стержне:

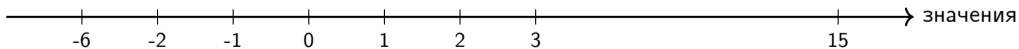
Распределение X (пример 1): $Var[X] = 2.76$



Распределение W (пример 2): $Var[W] = 24.96$



Масштаб не соблюден - для наглядности



Распределение Бернулли

Начнем с простейшего случая - **одного случайного эксперимента** с двумя возможными исходами.

i Definition

Распределение Бернулли - это распределение случайной величины X , которая принимает только два значения:

$$X = \begin{cases} 1 & \text{с вероятностью } p \text{ ("успех")} \\ 0 & \text{с вероятностью } q = 1 - p \text{ ("неудача")} \end{cases}$$

Обозначается как $X \sim \text{Bern}(p)$, где $p \in [0, 1]$ - параметр распределения.

Характеристики распределения Бернулли

💡 Основные характеристики

Для случайной величины $X \sim \text{Bern}(p)$:

- Математическое ожидание:

$$E[X] = 0 \cdot (1 - p) + 1 \cdot p = p$$

- Дисперсия:

$$\text{Var}[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

- Стандартное отклонение:

$$\sigma = \sqrt{p(1 - p)}$$

Примеры распределения Бернулли

1. **Подбрасывание монеты:** "Орел" = 1, "Решка" = 0, $p = 0.5$ (для честной монеты)
2. **Производственный контроль:** "Дефектное изделие" = 1, "Качественное" = 0
3. **Медицинский тест:** "Положительный результат" = 1, "Отрицательный" = 0
4. **Маркетинг:** "Покупка товара" = 1, "Отказ от покупки" = 0

От одного испытания к процессу

- Что если мы хотим провести несколько независимых испытаний Бернулли?
- Например: подбросить монету 10 раз, проверить 100 изделий на дефекты, провести 50 медицинских тестов.
- Это приводит нас к понятию процесса Бернулли - последовательности независимых испытаний Бернулли.

Процесс Бернулли

i Definition

Процесс Бернулли - это последовательность независимых одинаково распределенных случайных величин X_1, X_2, \dots, X_n , где каждая $X_i \sim \text{Bern}(p)$:

$$X_i = \begin{cases} 1 & \text{с вероятностью } p \text{ ("успех")} \\ 0 & \text{с вероятностью } 1 - p \text{ ("неудача")} \end{cases}$$

Ключевое отличие от одного испытания: теперь у нас есть **последовательность** результатов.

Свойства процесса Бернулли

1. **Независимость:** Результат каждого испытания X_i не зависит от результатов других испытаний
2. **Два исхода:** Каждое испытание может закончиться только успехом (1) или неудачей (0)
3. **Постоянная вероятность:** Вероятность успеха p не изменяется от испытания к испытанию

Примеры процесса Бернулли

1. **Серия подбрасываний монеты:** Последовательность $(X_1, X_2, \dots, X_{10})$ для 10 подбрасываний
2. **Контроль качества:** Проверка партии из n изделий, где каждое может быть дефектным независимо от других
3. **Клинические испытания:** Тестирование лекарства на n пациентах, где каждый может показать положительную реакцию

Биномиальное распределение

- Теперь, когда у нас есть процесс Бернулли, возникает естественный вопрос: **“Сколько успехов мы получим в n испытаниях?”**

i Definition

Биномиальное распределение описывает количество успехов в n независимых испытаниях Бернулли с вероятностью успеха p . Если X_1, X_2, \dots, X_n - процесс Бернулли с параметром p , то случайная величина:

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

имеет биномиальное распределение и обозначается $Y \sim Bin(n, p)$. Y может принимать значения $0, 1, 2, \dots, n$ (от 0 до n успехов).

Функция вероятности биномиального распределения

Вероятность получить ровно k успехов в n испытаниях:

$$P(Y = k) = C_n^k p^k (1 - p)^{n-k}$$

Интуиция формулы

- p^k - вероятность k успехов
- $(1 - p)^{n-k}$ - вероятность $(n - k)$ неудач
- C_n^k - количество способов выбрать k позиций из n для размещения успехов

Характеристики биномиального распределения

💡 Основные характеристики

Для случайной величины $Y \sim \text{Bin}(n, p)$:

Математическое ожидание:

$$E[Y] = np$$

Дисперсия:

$$\text{Var}[Y] = np(1 - p)$$

Стандартное отклонение:

$$\sigma = \sqrt{np(1 - p)}$$

Вывод математического ожидания:

$$E[Y] = E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] + \dots + E[X_n] = p + p + \dots + p = np$$

Визуализация биномиального распределения

Форма биномиального распределения зависит от параметров n и p . Рассмотрим, как влияет значение p на распределение при фиксированном $n = 20$:

Сравнение различных значений p

Функция вероятности биномиального распределения
 $n = 20$ испытаний

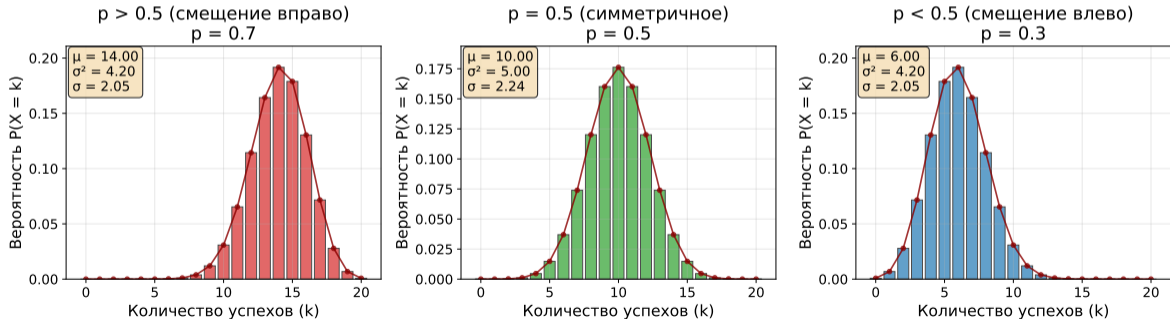


Рис. 1: Сравнение биномиального распределения для разных значений p

Визуализация биномиального распределения

i Наблюдения

$p < 0.5$ (например, $p = 0.3$): Распределение смещено влево, большинство значений сосредоточено в области малых k

$p = 0.5$: Распределение симметрично относительно среднего значения

$p > 0.5$ (например, $p = 0.7$): Распределение смещено вправо, большинство значений сосредоточено в области больших k

Пример: подбрасывание монеты

Подбросим честную монету 10 раз. Какова вероятность получить ровно 6 орлов?

- $n = 10, p = 0.5, k = 6, Y \sim Bin(10, 0.5)$

$$P(Y = 6) = C_{10}^6 (0.5)^6 (0.5)^4 = \frac{10!}{6! \cdot 4!} (0.5)^{10}$$

$$= \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} \cdot \frac{1}{1024} = 210 \cdot \frac{1}{1024} \approx 0.205$$

Ожидаемое количество орлов: $E[Y] = 10 \cdot 0.5 = 5$

Процесс Бернулли, биномиальное распределение

Пример 1

Инвестор владеет акциями 7 предприятий одной отрасли. Известно, что вероятность роста цены акций по каждому из предприятий равна 0.4, вероятность падения равна 0.3. (будем считать, что акции ведут себя независимо)

1. Найти вероятность того, что изменится цена акций шести предприятий.
2. Найти вероятность того, цена акций вырастет более чем у двух предприятий.

Слайд для записей

Процесс Бернулли, биномиальное распределение

Пример 2

В команде 10 хороших стрелков, попадающих в цель при одном выстреле с вероятностью 0.8, и 3 плохих, попадающих с вероятностью 0.5. Один стрелок производит 5 выстрелов. Чему равна вероятность того, что это хороший стрелок, если он попал более двух раз?

Слайд для записей