

# ДЗ1. Вероятностные распределения

Игорь Семенов, СКБ172

1 ноября 2019 г.

## 1 Распределения и описание их основных характеристик

### 1.1 Биномиальное распределение $B(n, p)$

#### 1.1.1 Закон распределения

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k}, \quad k \in \{1, 2, \dots, n\}.$$

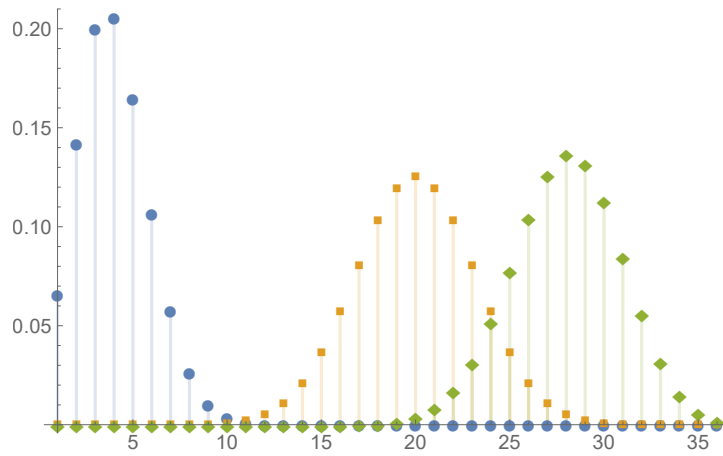


Рис. 1: Закон распределения для  $p = 0.1, 0.5, 0.7$  при  $n = 40$

#### 1.1.2 Математическое ожидание

Математическое ожидание неотрицательной дискретной случайной величины называется:

$$E\xi = \sum_i a_i p_i = \sum_i a_i P(\xi = a_i)$$

Выведем математическое ожидание для Биномиального распределения:

$$\begin{aligned} E\xi &= \sum_{k=1}^n k \cdot P(\xi = k) = \\ &= \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} = \\ &= \sum_{k=1}^n k \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} = \\ &= np \sum_{k=1}^n \frac{(n-1)!}{((n-1)-(k-1))!(k-1)!} p^{k-1} (1-p)^{(n-1)-(k-1)} = \\ &= np(p + (1-p))^{n-1} = \\ &= np \end{aligned}$$

### 1.1.3 Дисперсия

$$D\xi = E(\xi - E\xi)^2 = E\xi^2 - E^2\xi.$$

$$E\xi_i = p, \quad E\xi_i^2 = p, \quad D\xi_i = p - p^2 = p(1 - p), \quad i \in 1, \dots, n$$

$$D\xi = D\xi_1 + D\xi_2 + \dots + D\xi_n = np(p - 1)$$

### 1.1.4 Функция распределения

$$F_\xi(x) = P(\xi < x) = \sum_{k=0}^x \binom{n}{k} p^k q^{n-k}, \quad x \in R$$

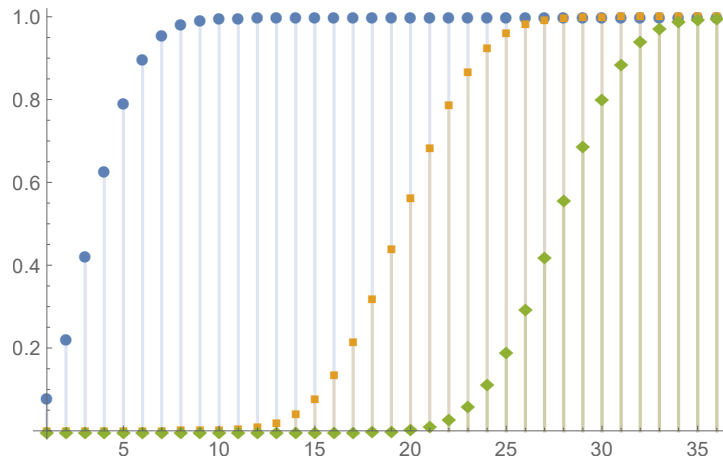


Рис. 2: Функция распределения для  $p = 0.1, 0.5, 0.7$  при  $n = 40$

### 1.1.5 Производящая функция

$$\varphi(z) = \sum_{k=0}^n z^k p_n$$

### 1.1.6 Характеристическая функция

$$g(t) = 1 - p + pe^{it}$$

## 1.2 Нормальное распределение $N(\mu, \sigma^2)$

### 1.2.1 Плотность распределения

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in R$$

### 1.2.2 Функция распределения

$$F(x) = P(\xi \leq x) = \int_{-\infty}^x f(x) dx = \left| x = \sigma t + \mu \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

### 1.2.3 Математическое ожидание и Дисперсия

$\xi \in N(\mu, \sigma^2)$ , тогда  $\eta = (\xi - \mu)/\sigma \in N(0, 1)$ , где  $E\eta = 0$  и  $D\eta = 1$

Из св-ва Математического ожидания:

$$E\xi = E(\sigma\eta + \mu) = \sigma E\eta + \mu = \mu$$

Из св-ва Дисперсии:

$$D\xi = D(\sigma\eta + \mu) = \sigma^2 D\eta = \sigma^2$$

### 1.2.4 Производящая и характеристическая функции

$$\varphi(z) = \exp(\mu z + \frac{\sigma^2 z^2}{2})$$

$$g(t) = \exp(i\mu t + \frac{\sigma^2 t^2}{2}).$$

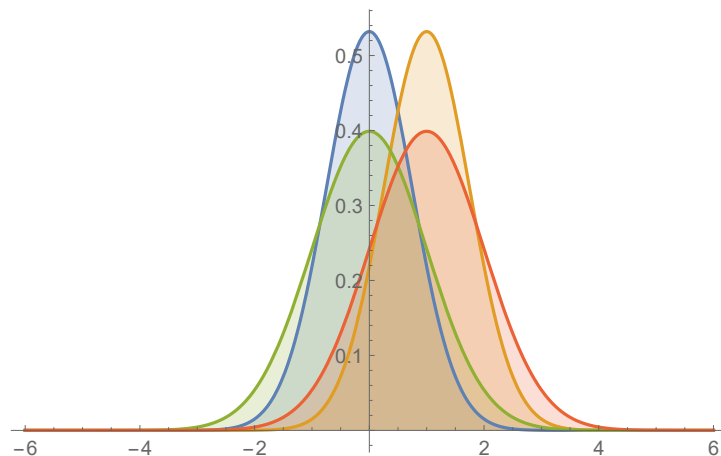


Рис. 3: Плотности распределения для  $\mu = 0$ ;  $\sigma = 0.75$  — синий,  $\mu = 1$ ;  $\sigma = 0.75$  — желтый,  $\mu = 0$ ;  $\sigma = 1$  — зеленый,  $\mu = 1$ ;  $\sigma = 1$  — красный

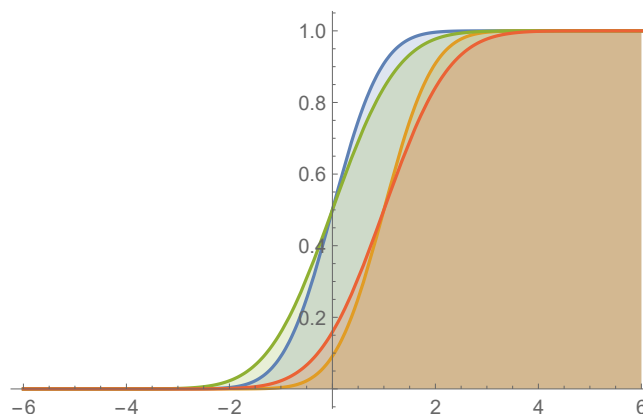


Рис. 4: Функции распределения для  $\mu = 0$ ;  $\sigma = 0.75$  — синий,  $\mu = 1$ ;  $\sigma = 0.75$  — желтый,  $\mu = 0$ ;  $\sigma = 1$  — зеленый,  $\mu = 1$ ;  $\sigma = 1$  — красный

## 2 Поиск примеров событий

### 2.1 Биномиальное распределение

#### 2.1.1 Типичная интерпретация

Биномиальное распределение определяет количество «успехов» в последовательности из  $n$  независимых случайных экспериментов, таких, что вероятность «успеха» в каждом из них постоянна и равна  $p$ .

Самый распространенный пример использования биномиального распределения — это подсчет количества выпадения орлов у монетки за  $n$  бросков.

Пусть вероятность выпадения орла постоянна и равна  $0.3 - p = 0.3$ , посчитаем вероятность выпадения 5 орлов за 6 бросков:

$$P(\xi = 5) = \binom{6}{5} (0.3)^5 (0.7) = 0,010206$$

### 2.1.2 Нетипичная интерпретация

С помощью биномиального распределения можно считать не только успех броска монетки. Представим что мы хотим улучшить доходность нашего колл центра, где сотрудники за счет звонков продают наш товар[1]

Нам известна следующая информация:

- В среднем один работник делает 50 звонков в день
- Вероятность успешного звонка 4%
- Доход с одного успешного звонка 100 \$
- В нашем колл центре работает 100 работников
- Каждый из них получает 200\$ в день

Создадим симуляцию одного дня нашего колл центра в Python

```
# Симуляция КоллЦентра
import numpy as np
import matplotlib.pyplot as plt

# Количество работников в центре
employees = 100

# Зарплата работника
wage = 200

# Количество звонков от одного работника
n = 50

# Вероятность успеха каждого звонка
p = 0.04

# Доход за звонок
revenue = 100

# Биномиальные случайные величины каждого сотрудника
conversions = np.random.binomial(n, p, size=employees)

print('Среднее значение случайной величины: ' + str(round(np.mean(conversions), 2)))
print('Стандартное отклонение: ' + str(round(np.std(conversions), 2)))
print('Всего звонков: ' + str(np.sum(conversions)))
print('Общий доход: ' + str(np.sum(conversions)*revenue))
print('Общий расход: ' + str(employees*wage))
print('Итоговая прибыль: ' + str(np.sum(conversions)*revenue - employees*wage))
```

```
-----

Среднее значение случайной величины для одного работника: 2.09
Стандартное отклонение: 1.4
Всего звонков: 209
Общий доход: 20900
Общий расход: 20000
Итоговая прибыль: 900
```

Проведем симуляцию целого года:

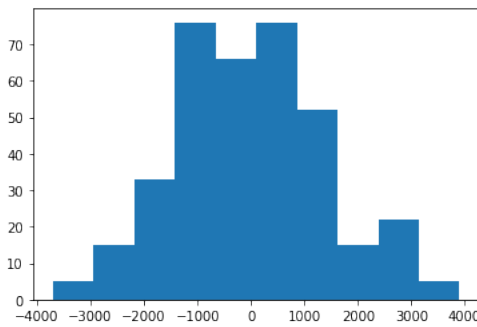


Рис. 5: Статистика доходов колл центра за год

На гистограмме видно, что существует очень высокая вероятность потери прибыли. Давайте попробуем увеличить прибыль нашего колл центра не уменьшая зарплату работников. Например мы разработали алгоритм поиска потенциальных клиентов, благодаря этому увеличивается вероятность успешного звонка, а также, ориентируясь на конкретного клиента, мы уменьшаем время самого звонка.

В конечном итоге:

- $n = 53$  звонков в день
- $p = 4.5\%$

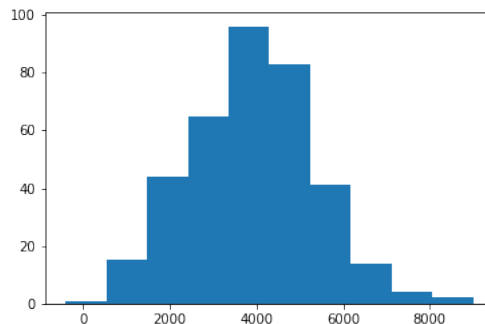


Рис. 6: Статистика доходов колл центра за год после внедрения алгоритма поиска потенциальных клиентов

Мы увеличили прибыль нашего колл центра! :)

## 2.2 Нормальное распределение

### 2.2.1 Типичные интерпретации

Если некая величина образуется в результате сложения многих случайных слабо взаимосвязанных величин, каждая из которых вносит малый вклад относительно общей суммы, то центрированное и нормированное распределение такой величины при увеличении числа наблюдений стремится к нормальному распределению.

Это вытекает из центральной предельной теоремы теории вероятностей. Таких величин в окружающем нас мире очень много, поэтому такое распределение величин и названо нормальным.

Примеры:

- Рост и вес человека;
- Ошибки наблюдений;
- Скорость молекул;

- Гауссовский процесс;
- Отклонение при стрельбе.

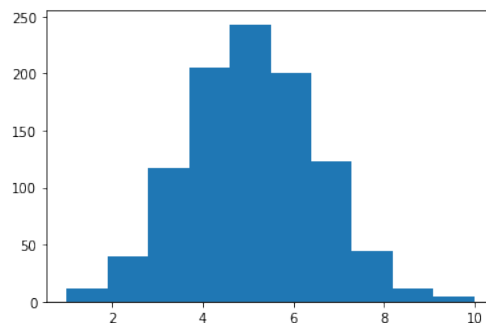
### 3 Описание способа моделирования выбранных случайных величин

Случайные величины для Биномиального и Нормального распределения можно смоделировать при помощи библиотеки NumPy из Python.

`np.random.binomial(n, p, size)` — функция для моделирования случайных величин Биномиального распределения, где  $n$  — это количество независимых случайных испытаний,  $p$  — шанс успеха одного испытания, `size` — объем выборки.

Пример использования: Посчитаем количество выпадения орлов для 10 бросков с объемом выборки `size = 1000`

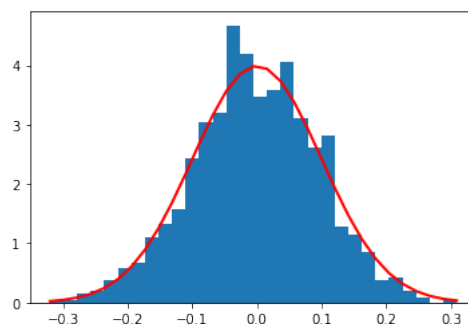
```
np.random.binomial(n = 10, p = 0.5, size = 1000)
```



`np.random.normal( $\mu$ ,  $\sigma$ , size)` — функция для моделирования случайных величин Нормального распределения, где  $\mu$  и  $\sigma$  — это параметры нормального распределения математическое ожидание и среднеквадратичное отклонение соответственно, `size` — объем выборки.

Пример использования:

```
np.random.normal(mu = 0, sigma = 0.1, size = 1000)
```



### Список литературы

- [1] Tony Yiu. Fun with the binomial distribution, 2019. URL <https://towardsdatascience.com/fun-with-the-binomial-distribution-96a5ecabf65b>.