

**Школа лингвистики, 2023-24 уч. год****Теория вероятностей****Дискретные случайные величины: биномиальное распределение, геометрическое распределение, распределение Пуассона (30.01.2024)***И. В. Щуров, Д. А. Филимонов, Р. Я. Будылин*

**Задача 1.** Монетка имеет смещённый центр тяжести и выпадает орлом вверх с вероятностью  $p \in (0, 1)$ . Монетку подкинули  $n$  раз. Случайная величина  $\xi$  — число выпавших орлов. Как называется такое распределение?

- (a) Найти ряд распределения  $\xi$  для  $n = 4$ ,  $n = 5$ ,  $n = 6$  для честной монеты;
- (b) Написать ряд распределения  $Bin(3, \frac{1}{3})$ , т.е. для  $\xi$  при  $n = 3$  и  $p = \frac{1}{3}$
- (c) Найти  $E\xi$ ,  $D\xi$  и  $\sigma(\xi)$  для  $n = 3$ ,  $n = 4$ ,  $n = 5$  для честной монеты.
- (d) Найти  $E\xi$ ,  $D\xi$  и  $\sigma(\xi)$  для  $n = 100$  и  $p = \frac{1}{4}$ .

**Задача 2.** Пусть мы бросаем игральный кубик 100 раз и  $\xi$  - число выпавших шестёрок. Найти математическое ожидание и дисперсию  $\xi$ .

**Задача 3.** У меня в кармане две монетки — одна нормальная, а другая с двумя орлами. Я достаю случайную монетку и подбрасываю её. Если выпала решка, я прекращаю подбрасывание, а если орёл, возвращаю монетку в карман и продолжаю процесс. Случайная величина  $\xi$  — число выпавших орлов до первого выпадения решки.

- (a) Как называется такое распределение? Найти значения параметров.
- (b) Какие значения может принимать  $\xi$ ?
- (c) Чему равняется  $P(\xi = 0)$ ?  $P(\xi = 1)$ ?  $P(\xi = k)$  для произвольного  $k$ ?
- (d) Найти математическое ожидание этой случайной величины.
- (e) Чему равняется  $P(\xi < 4)$ ?
- (f) Чему равняется  $P(\xi > 3)$ ?
- (g) Что поменяется, если мы наоборот, будем подбрасывать монетку до выпадения первого орла?
- (h) (\*) Чему равняется  $P(\xi > 5 | \xi > 2)$  (то есть вероятность того, что  $\xi$  оказалось больше 5, при условии, что  $\xi > 2$ )?

**Задача 4.** Частота русской буквы "о" примерно 0.1

- (a) Сколько в среднем букв "о" встретится в случайно взятом тексте на русском языке длиной 10000 знаков?
- (b) Найти дисперсию числа букв "о".
- (c) Каково стандартное отклонение числа букв "о" в таком тексте?

**Задача 5.** Найдите среднее число бросаний игрального кубика, если мы останавливаемся, получив шестёрку.

**Задача 6.** Пусть вероятность аварии на участке дороги длиной в 1 км в среднем равна 0.001. Сколько в среднем можно проехать по такой дороге без аварий?

**Задача 7.** Найдите среднее число бросаний монеты, если мы останавливаемся, получив двух орлов (не обязательно подряд). А до 10 орлов?

**Задача 8.** Найдите дисперсию числа бросаний монеты до 2 орлов (не обязательно подряд). А для 10 орлов?

**Задача 9.** В ходе исследования вопроса об авторстве «Федералиста» (сборник статей в поддержку ратификации Конституции США) использовались статистические методы [Mosteller and Wallace (1984, 2.3)]: в частности, подсчёт частотности вхождения тех или иных слов. Для сравнения был взят длинный текст, принадлежащий перу Александра Гамильтона. Его разбили на большое количество фрагментов примерно одинаковой длины. Выбрали случайный фрагмент и посчитали, сколько раз в нём встречается слово «апу». Оказалось, что полученная случайная величина распределена по Пуассону<sup>1</sup> с параметром  $\lambda = 0.67$ .

- (a) С какой вероятностью в выбранном фрагменте слово «апу» встречалось три раза?
- (b) Сколько в среднем раз встречалось слово «апу» в фрагменте?
- (c) С какой вероятностью в фрагменте встретилось меньше трёх слов «апу»?

**Задача 10.** На заводе для приготовления булочек с изюмом на 10 тонн теста добавляют миллион изюминок. Какова вероятность, что в одной булочке весом 100 грамм нет ни одной изюминки?

**Задача 11.** Примерно четвертая часть носителей русского языка считает, что слово «виски» среднего рода. Вы решили проверить это и сделали форму, где респонденту предлагается вписать окончание прилагательного. Далее программа выбирает случайного человека из 1000 и предлагает ему ответить на вопрос. Программа собрала 10 ответов, однако так получилось, что один и тот же человек мог более одного раза ответить на вопрос. Предположим, что гипотеза верна, какое распределение имеет количество людей, считающих, что слово «виски» среднего рода? А какое у этой случайной величины математическое ожидание и стандартное отклонение?

**Задача 12.** В условиях предыдущей задачи исправим программу так, чтобы каждый человек мог ответить только один раз. Как изменится распределение количества людей, считающих, что слово «виски» среднего рода? А среднее?

## Дополнительные задачи

**Задача 13.** В условиях предыдущей задачи найти стандартное отклонение. Записать его в общем виде для вероятности успеха  $p$ , количества человек  $N$  и количества ответов  $n$ .

**Задача 14.** За час рыбак ловит в среднем 6 рыб. Какова вероятность, что он за 10 минут поймает больше 2 рыб?

---

<sup>1</sup>Вообще говоря, частотность слов далеко не всегда распределена по Пуассону, но в данном случае это оказалось так.