Санкт-Петербургский государственный университет
Прикладная математика, информатика и искусственный интеллект

Отчет по учебной практике 1 (научно-исследовательская работа)

(семестр 2)

Теория вероятностей и моделирование (по книге Бродского Я.С.)

Выполнила:

Барабашева Анастасия Дмитриевна, группа 22.Б04-мм

Научный руководитель:

кандидат физико-математических наук, доцент

Голяндина Нина Эдуардовна

Кафедра статистического моделирования

Работа выполнена на хорошем уровне
и может быть зачтена с оценкой А

Санкт-Петербург
2023

I. Введение

В ходе работы я ознакомилась с некоторыми аспектами теории вероятности, решила задачи из сборника Бродского Я.С. по теории вероятности и промоделировала на языке R наиболее интересные из них. Изучила такие разделы, как «Математическое ожидание и его свойства», «Неравенство Чебышева», «Закон больших чисел», «Нормальное распределение», «Условная вероятность», «Дисперсия случайной величины», «Числовые характеристики биномиального распределения». Были решены задачи под номерами: 178, 179, 180, 181, 182, 347, 348, 349, 404, 408, 414, 415, 463, 467, 484. Из них промоделированы: 349 и 467.

II. Основная часть

1. Решение задач

Задача 1.
Условие:
Найдите значение p, при котором вероятность $P_n(m)$ того, что в n испытаниях Бернулли число «успехов» равно m, достигает наибольшего значения, и вычислите это наибольшее значение.

Решение:
Для решения этой задачи воспользуемся функцией вероятности Бернулли:
$P_n(m) = C_n^m \cdot p^m \cdot (1-p)^{n-m}$, где p - вероятность «успеха» в одном испытании, $(1-p)$ - вероятность «неудачи» в одном испытании, C_n^m - число сочетаний из n по m. Так как вероятность $P_n(m)$ зависит от p, то ее можно рассматривать как функцию вероятности:
$P_n(m) = P_n(m;p) = C_n^m \cdot p^m \cdot (1-p)^{n-m}$.

Для нахождения максимума этой функции необходимо взять ее производную и найти ее корни:

$$dP_n(m)/dp = C_n^m \cdot m \cdot p^{m-1} \cdot (1-p)^{n-m} - C_n^m \cdot (n-m) \cdot p^m \cdot (1-p)^{n-m-1} = 0.$$

Решим это уравнение:

$$C_n^m \cdot m \cdot p^{m-1} \cdot (1-p)^(n-m) = C_n^m \cdot (n-m) \cdot p^m \cdot (1-p)^{n-m-1} m \cdot (1-p) = (n-m) \cdot p$$

$p = m/n$.

Таким образом, для того чтобы вероятность $P_n(m)$ достигала наибольшего значения, необходимо, чтобы $p = m/n$. Это значит, что наиболее вероятное число «успехов» равно np, а вероятность $P_n(m)$ при этом равна:

$P_n(m) = C_n^m \cdot (m/n)^m \cdot (1 - (m/n))^{n-m}$.

Вычислим наибольшее значение этой вероятности:

$P_n(m) = C_n^m \cdot (m/n)^m \cdot (1 - (m/n))^{n-m}$.

Подставим $p = m/n$:

$P_n(m) = C_n^m \cdot (p)^m \cdot (1 - p)^{n-m}$.

Таким образом, наибольшее значение вероятности достигается при $p = m/n$, а это значение равно:

$P_n(m) = C_n^m \cdot (m/n)^m \cdot (1 - m/n)^{n-m} = C_n^m \cdot m^m \cdot (n^{n-m} - m^{n-m})/n^n$.

Задача 2.
Условие:
Устройство состоит из четырех элементов. Вероятность отказа любого элемента за время опыта равна 0,2. Найдите математическое ожидание числа опытов, в каждом из которых откажет ровно один элемент, если всего проведено 100 независимых опытов.

Решение:
Пусть событие «отказ элемента» имеет вероятность p=0.2. Число опытов, в которых откажет ровно один элемент, будет определяться распределением Бернулли. Вероятность того, что ровно один элемент откажет, равна
$P(X = 1) = C_4^1 \cdot p \cdot q^3 = 4 \cdot 0.2 \cdot 0.8^3 = 0.4096$,
где C_4^1 - число сочетаний 4 из 1, q=1-p=0.8.
Математическое ожидание числа опытов, в каждом из которых откажет ровно один элемент, равно
$E(X) = np = 100 \cdot 0.4096 = 40.96$.
Таким образом, при проведении 100 независимых опытов можно ожидать, что ровно в 41 опыте откажет ровно один элемент.

Задача 3:
Условие:
Вероятность получения патента равна 0.6. Если патент был получен, условная вероятность получения дохода от него составит 0.9. Однако, если патент не будет получен, условная вероятность получения дохода составит только 0.3. Найти вероятность получения дохода.

Решение:
В нашей задаче пусть событие A - получение патента, B - получение дохода.
Вероятность получения дохода при получении патента :
$P(B|A) = 0.9$
$P(B|\neg A) = 0.3$
$P(A) = 0.6$
$P(\neg A) = 1 - 0.6 = 0.4$

Формула полной вероятность для события B:
$p(B) = p(A \cap B) + (\neg A \cap B)$
Используя условную вероятность, запишем:
$p(B) = p(A \cap B) + (\neg A \cap B) = p(B|A) \cdot p(A) + p(B|\neg A) \cdot p(\neg A)$
Тогда:
$p(B) = 0.9 \cdot 0.6 + 0.3 \cdot 0.4 = 0.66.$

2. Решение задач с моделированием

Ниже я приведу решения и моделирование наиболее интересных для меня задач.

Задача 4.
Условие:
Найдите математическое ожидание и дисперсию числа делителей произвольного наугад выбранного натурального числа из совокупности 1, 2, ..., 20.

Решение:
Пусть n - произвольное натуральное число из множества 1, 2, ..., 20.
Математическое ожидание:
Математическое ожидание числа делителей можно выразить с помощью формулы:

$E(d) = (d(1) + d(2) + ... + d(20))/20$,
где d(i) - число делителей числа i. Дисперсия:
Дисперсия числа делителей можно выразить с помощью формулы:
$D(d) = E(d^2) - E(d)^2$
Теперь рассчитаем математическое ожидание и дисперсию числа делителей. Для каждого числа от 1 до 20 найдем его делители и подсчитаем их количество:
d(1) = 1,
d(2) = 2,
d(3) = 2,
d(4) = 3,
d(5) = 2,
d(6) = 4,
d(7) = 2,
d(8) = 4,
d(9) = 3,
d(10) = 4,
d(11) = 2,
d(12) = 6,
d(13) = 2,
d(14) = 4,
d(15) = 4,
d(16) = 5,
d(17) = 2,
d(18) = 6,
d(19) = 2,
d(20) = 6.

Математическое ожидание:
$E(d) = (d(1) + d(2) + ... + d(20))/20 = (1 + 2 + 2 + 3 + 2 + 4 + 2 + 4 + 3 + 4 + 2 + 6 + 2 + 4 + 4 + 5 + 2 + 6 + 2 + 6)/20 = 3.3$

$E(d^2) = (d(1)^2 + d(2)^2 + ... + d(20)^2)/20 = (1 + 4 + 4 + 9 + 4 + 16 + 4 + 16 + 9 + 16 + 4 + 36 + 4 + 16 + 16 + 25 + 4 + 36 + 4 + 36)/20 = 13.2$

Дисперсия:
$D(d) = E(d^2) - E(d)^2 = 13.2 - 3.3^2 = 2.31$

Моделирование:

```
count_divisors <- function(n) {
  count <- 0
  for (i in 1:n) {
    if (n %% i == 0) {
      count <- count + 1
    }
  }
  return(count)
}


numbers <- 1:20

num_trials <- 10000
results <- replicate(num_trials, {
  n <- sample(numbers, 1)
  count_divisors(n)
})


mean_value <- mean(results)
variance <- var(results)

print(paste("Expected value:", mean_value))
print(paste("Dispersion:", variance))
```

Сначала прописываем функцию для подсчета количества делителей числа, далее создаем множество чисел от 1 до 20. Затем моделируем выбор случайного числа и подсчитываем количество его делителей. После этого вычисляем математическое ожидание и дисперсию и выводим результат.

Посмотрим, какие значения будут у математического ожидания и дисперсии при различных количествах опытов.
M(10) = 3.8, D(100) = 2.1
M(100) = 3.41, D(100) = 2.16
M(1000) = 3.36, D(1000) = 2.42
M(10000) = 3.2844, D(10000) = 2.2867
M(100000) = 3.3061, D(100000) = 2.3172
M(1000000) = 3.3003, D(1000000) = 2.3115
Чем больше проводим опытов, тем получаем наиболее близкие к

расчетным значения. Хотя уже при 100 опытах отличие около 3%.

Задача 5.
Условие:
Правильную монету подбрасывают до тех пор, пока она не выпадет цифрой вверх, либо до трех последовательных выпадений герба. Найдите математическое ожидание числа подбрасываний при одном выполнении этого эксперимента.

Решение:
После первого броска возможны исходы: Г (герб) или Ц (цифра), вероятность того, что подбрасывание будет одна - 0.5. При второй подбрасывании возможные исходы ГГ(тогда подбрасываний будет больше двух), ГЦ (тогда подбрасываний ровно 2), вероятность того, что подбрасываний будет два равна $0.5 \cdot 0.5$. При третьем подбрасывании возможные исходы ГГЦ или ГГГ - оба подходят, тогда вероятность того, что подбрасываний будет 3, равна $0.5 \cdot 0.5 \cdot 0.5 \cdot 2$ Нетрудно доказать, что у нас не может быть больше трех подбрасываний. Для доказательства рассмотрим всевозможные исходы для $n > 3$. Во всех исходах, в которых присутствует Ц на позициях, меньших 3 - число подбрасываний меньше 3 (мы кидаем до первого появления цифры, то есть такие исходы в рамках нашей задачи невозможны). А в остальных исходах на первых трех позиция Г, тогда мы кидали ровно 3 раза, что мы и учли. Таким образом, достаточно кинуть 3 раза, чтобы точно достичь успеха.
Теперь посчитаем математическое ожидание.
$M(X) = 1 \cdot 0.5 + 2 \cdot 0.5 \cdot 0.5 + 3 \cdot (0.5 \cdot 0.5 \cdot 0.5 \cdot 2) = 1.75.$

Смоделируем задачу на языке R.

```
p <- 0.5
q <- 1 - p

coin_toss <- function() {
  count <- 0

  flag <- FALSE

  while (!flag) {
    x <- runif(1)

    if (x < p) {
```

```
        count <- count + 1
        break
      }

      if (x >= p) {
        count <- count + 1

        if (count >= 3) {
          flag <- TRUE
        }
      }
    }

    return(count)
  }


coin_toss_mean <- function(n) {

  counts <- numeric(n)

  for (i in seq_len(n)) {
    counts[i] <- coin_toss()
  }

  return(mean(counts))
}

cat("Number of flips in one experiment:", coin_toss(), "\n")

cat("Mathematical expectation of the number of tosses:",

coin_toss_mean(100000), "\n")
```

При моделировании сначала задаем вероятность выпадения орла и герба, затем прописываем функцию для моделирования одного эксперимента: в ней счетчик числа подбрасываний, флаг для проверки трех последовательных выпадений герба. Далее подбрасываем монету до тех пор, пока не выпадет орел или не будет три последовательных выпадений герба: генерируем случайное число от 0 до 1; если выпал орел, то эксперимент заканчивается; если выпал герб, то увеличиваем

счетчик и проверяем флаг; возвращаем число подбрасываний в этом эксперименте. Прописываем функцию для моделирования нескольких экспериментов и вычисления математического ожидания: создаем вектор для хранения числа подбрасываний в каждом эксперименте; моделируем n экспериментов и сохраняем результаты в вектор; вычисляем математическое ожидание числа подбрасываний. Моделируем один эксперимент и выводим результат на экран. Моделируем один эксперимент и выводим результат на экран. Получаем примерно 1.75, что согласуется с расчетами.
Посчитаем математическое ожидание при различных количествах экспериментов.
M(10)= 1.7
M(100)= 1.79
M(1000)= 1.743
M(10000)= 1.7471
M(100000)= 1.7527
Уже начиная с 10000 математическое ожидание принимает значение, очень близкое к расчетному. Чем больше n, тем ближе будут эти значения. Уже при n = 10 отличие от расчетов меньше 3%.

III. Заключение

В ходе выполнения работы я укрепила знания в области теории вероятностей, решила ряд задач и исследовала некоторые из них моделированием на языке R. Результаты моделирования совпали с вычисленными значениями в пределах 3%. Моделирование позволило посмотреть на величины (дисперсию и математическое ожидание) при различном количестве проведенных экспериментов.

IV. Список литературы

1. Бродский Я.С. «Статистика. Вероятность. Комбинаторика.» - М.:ООО «Издательство Оникс»:ООО «Издательство «Мир и Образование», 2008. - 544 с.: ил. - (Школьный курс математики)