

Теория Вероятностей и Математическая  
Статистика  
Домашняя работа №1

Куприянов Кирилл  
Вариант 13

# 1 Задача

Условие: В урне один белый и пять черных шаров. Два игрока по очереди вынимают из урны шар и возвращают его обратно, после чего шары в урне перемешиваются. Выигрывает тот, кто первый извлекает белый шар. Какова вероятность того, что выиграет игрок, начинающий игру?

Решение:  $A = \{\text{Выиграл первый игрок}\}$

Исходы, при которых первый игрок выигрывает это:

(Первый игрок вытащил белый шар) ИЛИ (первый игрок вытащил черный шар И потом второй игрок вытащил черный шар И потом первый игрок вытащил белый шар) ИЛИ (и т. д.)

События эти несовместны, поэтому имеет место быть объединение (сумма) событий. Внутри скобок события независимы, поэтому вероятность произведения равна произведению вероятностей.

Пусть [Белый - Б, Черный - Ч]

Тогда наглядно выглядеть это будет так:

$$B \cup ЧЧБ \cup ЧЧЧЧБ \cup \dots \cup \underbrace{Ч \dots Ч}_{2n} Б$$

$$P(A) = \frac{1}{6} + \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6} + \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6} + \dots + \left(\frac{5}{6}\right)^{2n} \times \frac{1}{6}$$

Вынесем  $\frac{1}{6}$  за скобку.

$$P(A) = \frac{1}{6} \times \left[ 1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \dots + \left(\frac{5}{6}\right)^{2n} \right]$$

В квадратных скобках (не считая 1) получили бесконечно убывающую геометрическую прогрессию, где

$$b_1 = \left(\frac{5}{6}\right)^2, \text{ а } q = \left(\frac{5}{6}\right)^2$$

Поскольку  $|q| \leq 1$  и  $n \rightarrow \infty$ , запишем сумму  $S_n$  прогрессии как

$$S_n = \frac{b_1}{1 - q} = \frac{(5/6)^2}{1 - (5/6)^2} = \frac{25}{11}$$

Получаем  $P(A) = 1/6 \times (1 + 25/11) = 6/11 \approx 0.54$

Ответ: 0.54.

## 2 Задача

Условие: По каналу связи, подверженному воздействию помех, передается одна из команд управления в виде кодовых комбинаций 11111 или 00000, причем априорные вероятности передачи этих команд соответственно равны 0,8 и 0,2. Из-за наличия помех вероятность правильного приема каждого из символов (1 и 0) равна 0,6. Предполагается, что символы кодовых комбинаций искажаются независимо друг от друга. На выходе приемного устройства зарегистрирована комбинация 10110. Спрашивается, какая команда была передана?

Решение: Событие, которое мы рассматриваем - это  $A$ .

$$A = \{\text{Принято } 10110\}$$

Могла быть передана одна из двух комбинаций: 00000 или 11111. Поскольку других не дано и сумма их вероятностей (следует из условия) равна  $0,2 + 0,8 = 1 \Rightarrow$  они образуют полную группу, а значит, могут быть рассмотрены в качестве гипотез.

Гипотезы:

$$H_1 = \{\text{Была передана последовательность } 00000\}. P(H_1) \stackrel{\text{априор.}}{=} 0,2$$

$$H_2 = \{\text{Была передана последовательность } 11111\}. P(H_2) \stackrel{\text{априор.}}{=} 0,8$$

Условная вероятность приема 10110 вместо 00000 равна

$$P(A/H_1) = 0,4 \times 0,6 \times 0,4 \times 0,4 \times 0,6 = \frac{4^3 \times 6^2}{100000} = 0,02304$$

Условная вероятность приема 10110 вместо 11111 равна

$$P(A/H_2) = 0,6 \times 0,4 \times 0,6 \times 0,6 \times 0,4 = \frac{4^2 \times 6^3}{100000} = 0,03456$$

По формуле Байеса, найдем апостериорные вероятности:

$$P(H_1/A) = \frac{P(H_1) \times P(A/H_1)}{P(A)} = \frac{0,2 \times 0,02304}{P(A)} = \frac{0,004608}{P(A)}$$

$$P(H_2/A) = \frac{P(H_2) \times P(A/H_2)}{P(A)} = \frac{0,8 \times 0,03456}{P(A)} = \frac{0,027648}{P(A)}$$

Теперь нам надо сравнить  $P(H_1/A)$  с  $P(H_2/A)$ . Поскольку знаменатели одинаковы, можем сравнить только числители.  $0,027648 > 0,004608 \Rightarrow P(H_2/A)$  наиболее вероятно, чем  $P(H_1/A)$ . Следовательно, было передано сообщение «11111».

Ответ: 11111.

### 3 Задача

Условие: На окружность радиуса  $R$  брошено две точки. Считая, что длина хорды – случайная величина с равномерным распределением, найти плотность распределения вероятностей длины дуги между брошенными точками.

Решение: Длина хорды  $H$  имеет равномерное распределение от 0 до  $2R$ . Плотность распределения случайной величины  $H$ :

$$f_H(x) = \begin{cases} \frac{1}{2R}, & x \in [0, 2R] \\ 0, & x \notin [0, 2R] \end{cases}$$

Найдём уравнение связи между длиной дуги  $L$  и длиной хорды  $H$ :

$$\begin{aligned} L = \alpha R &\Rightarrow \alpha = \frac{L}{R} \\ H = 2R \times \sin\left(\frac{\alpha}{2}\right) \\ H = 2R \times \sin\left(\frac{L}{2R}\right) \end{aligned}$$

Плотность её распределения определяется по формуле:

$$f_L(y) = f_H(g^{-1}(y)) \cdot |[g^{-1}(y)]'|$$

Где  $g^{-1}(y)$  – функция обратная функции, связывающей  $L$  и  $H$ .

$$\begin{aligned} g^{-1}(y) &= 2R \times \sin\left(\frac{y}{2R}\right) \\ |[g^{-1}(y)]'| &= \cos\left(\frac{y}{2R}\right) \end{aligned}$$

Длина дуги кривой может варьироваться от 0 до  $\pi R$ . В данном диапазоне  $y$  значение  $\cos\left(\frac{y}{2R}\right)$  будет всегда положительным.

$$f_L(y) = \begin{cases} 0, & y \notin [0, \pi R] \\ \frac{1}{2R} \cos\left(\frac{y}{2R}\right), & y \in [0, \pi R] \end{cases}$$

## 4 Задача

Условие: Каждая повторная передача сигнала по каналу связи увеличивает вероятность искажения сигнала на 0,1%. При передаче первого сигнала эта вероятность равна 0,05. Передано 100 сигналов. Найти границы, в которых с вероятностью 0,9 заключено число переданных без искажения сигналов.

Решение: Пусть случайная величина  $X$  - число переданных без искажения сигналов. Вероятность того, что искажён  $i$ -ый сигнал:

$$p(A_i) = 0.05 \times 1.001^{i-1}$$

Вероятность искажения полседнего сигнала:

$$p(A_{100}) = 0.05 \times 1.001^{99} \approx 0.0552$$

Поскольку вероятности отличаются незначительно, будем приближенно считать случайную величину  $X$  - количество исказившихся сигналов - распределенной по биномиальному закону. Поскольку число испытаний велико, то можно воспользоваться интегральной теоремой Лапласа, т.е. считать случайную величину распределенной нормально с математическим ожиданием  $np = 5$  и средним квадратическим отклонением  $\sqrt{npq} = \sqrt{4.75}$

Границы, в которых с вероятностью 0,9 заключено число переданных без искажения сигналов:

$$P(|X - \alpha| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sqrt{npq}}\right) = 0.9$$

$$\Phi\left(\frac{\varepsilon}{\sqrt{4.75}}\right) = 0.45$$

$$\frac{\varepsilon}{\sqrt{4.75}} = 1.645 \Rightarrow \varepsilon = 1.645 \times \sqrt{4.75} = 3.585$$

Таким образом, количество искаженных сигналов из числа всех переданных с доверительной вероятностью 0.9 лежит в интервале

$$np - 3.585 < X < np + 3.585$$

$$1.415 < X < 8.585$$

Поскольку число передаваемых сигналов может быть только целым числом, то окончательно получаем доверительный интервал для числа сигналов  $Y = 100 - X$ , переданных без искажения:

$$92 \leq Y \leq 98$$

## 5 Задача

Условие: Случайная величина  $(\xi, \eta)$  распределена по нормальному закону с математическим ожиданием  $(\mu_1, \mu_2) = (0.6; 0.3)$  и ковариационной матрицей:

$$\Sigma = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.81 \end{pmatrix}$$

Найти  $P\{\eta > 2\xi\}$ .

Решение: Каждая составляющая вектора  $(\xi, \eta)$  распределена по нормальному закону:

$$\xi \sim N(0.6, 0.25), \eta \sim N(0.3, 0.81)$$

Рассмотрим случайную величину  $\zeta = 2\xi - \eta$ . Данная случайная величина имеет нормальное распределение:

$$\zeta \sim N(2m_\xi - m_\eta, 4\sigma_\xi^2 + \sigma_\eta^2 + 2cov(2\xi - \eta))$$

$$m_\zeta = 2m_\xi - m_\eta = 2 \times 0.6 - 0.3 = 0.9$$

$$4\sigma_\xi^2 + \sigma_\eta^2 + 2cov(2\xi, -\eta) = 4 \times 0.25 + 0.81 - 4cov(\xi, \eta) = 1.81 - 4 \times 0.25 = 0.81$$

$$\zeta \sim N(0.9, 0.81), \sigma_\zeta = 0.9$$

Необходимо найти:

$$P\{\eta > 2\xi\} = P\{2\xi - \eta < 0\} = P\{\zeta < 0\} = 0.5 + \Phi\left(\frac{0 - m_\zeta}{\sigma_\zeta}\right),$$

где  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-s^2/2} ds$  - интегральная функция Лапласа.

$$P\{\eta > 2\xi\} = 0.5 + \Phi\left(\frac{0 - 0.9}{0.9}\right) = 0.5 + \Phi(-1) = 0.5 - \Phi(1) \approx 0.5 - 0.3413$$

$$P\{\eta > 2\xi\} = 0.1587$$

## 6 Задача

Условие: Для заданной выборки:

1. постройте вариационный ряд выборки
2. пользуясь формулой Стерджесса, определите количество интервалов разбиения выборки
3. постройте таблицу статистического ряда, в первой строке которой указаны интервалы разбиения, а во второй-частоты попадания элементов выборки в соответствующие интервалы
4. постройте гистограмму
5. найдите реализации точечных оценок математического ожидания и дисперсии
6. на основе анализа результатов наблюдений выдвинете гипотезу о виде закона распределения наблюдаемой случайной величины

Масса одного колоса пшеницы сорта Sonnora (Япония) при плотности посева  $15 \times 2.5$  см,г.

1,80	1,40	1,12	2,30	2,70	3,30	1,30	1,13	1,70	1,40
1,25	1,90	1,64	1,47	1,65	1,50	1,85	1,68	1,51	1,48
1,95	0,80	2,80	2,40	2,95	2,50	2,30	2,90	1,84	2,20
1,68	2,50	2,52	1,29	3,30	1,85	2,10	3,60	2,40	2,55
1,50	1,29	1,85	1,58	1,31	1,69	1,28	1,90	1,87	1,70
1,49	2,10	1,90	1,49	1,80	2,45	2,30	3,00	3,10	3,10
1,60	1,88	2,20	1,63	0,80	1,63	1,45	1,29	1,47	2,55
1,49	2,40	2,55	1,26	0,80	1,25	2,10	0,70	2,00	1,85
0,90	1,90	2,10	2,55	2,55	2,40	0,60	2,10	0,40	2,50
1,50	1,69	2,70	1,48	1,50	1,69	1,46	1,48	1,52	1,30

Решение:

- 1) Построим вариационный ряд, т.е. упорядочим выборку по возрастанию (столбцами по 10 значений).

0,4	1,25	1,4	1,49	1,63	1,8	1,9	2,2	2,5	2,7
0,6	1,26	1,45	1,5	1,64	1,8	1,9	2,2	2,5	2,8
0,7	1,28	1,46	1,5	1,65	1,84	1,9	2,3	2,5	2,9
0,8	1,29	1,47	1,5	1,68	1,85	1,95	2,3	2,52	2,95
0,8	1,29	1,47	1,5	1,68	1,85	2	2,3	2,55	3
0,8	1,29	1,48	1,51	1,69	1,85	2,1	2,4	2,55	3,1
0,9	1,3	1,48	1,52	1,69	1,85	2,1	2,4	2,55	3,1
1,12	1,3	1,48	1,58	1,69	1,87	2,1	2,4	2,55	3,3
1,13	1,31	1,49	1,6	1,7	1,88	2,1	2,4	2,55	3,3
1,25	1,4	1,49	1,63	1,7	1,9	2,1	2,45	2,7	3,6

2) Найдем наибольшие и наименьшие значения выборки:  $x_{max} = 3.6$ ,  $x_{min} = 0.4$ . Размах колебаний выборки:

$$R = x_{max} - x_{min} = 3.6 - 0.4 = 3.2$$

Число интервалов разбиения определяем по формуле Стерджеса (объём выборки  $n = 100$ )

$$k = 1 + 3.322 \lg 100 = 7$$

Длина интервала:

$$h = \frac{R}{7} \approx 0.46$$

За начало первого интервала примем

$$x_0 = 0.39$$

3) Границы интервалов  $x_i$  покажем в таблице:

$i$	0	1	2	3	4	5	6	7
$x_i$	0,39	0,85	1,31	1,77	2,23	2,69	3,15	3,61

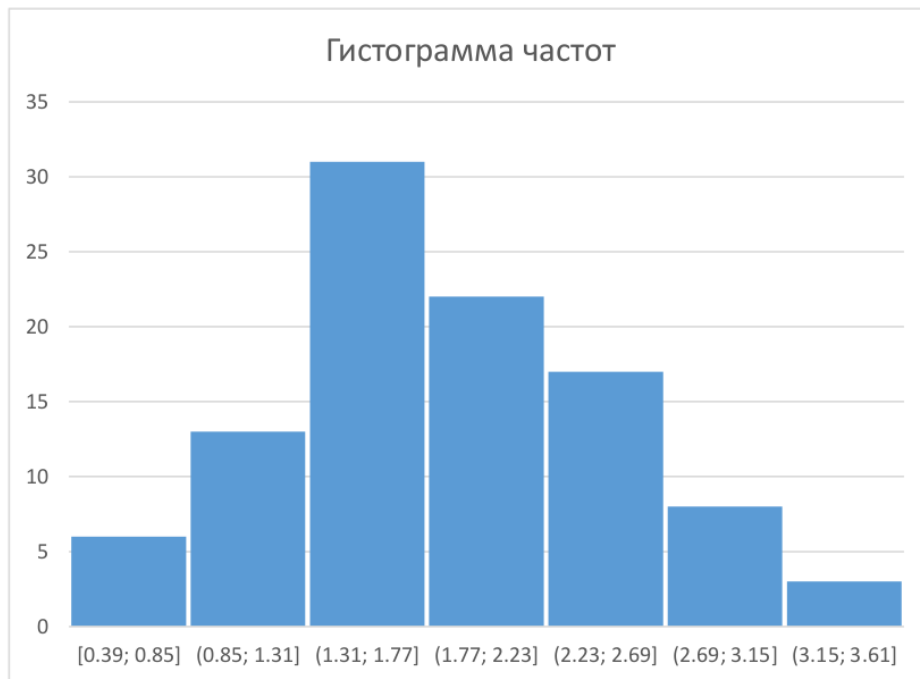
В результате получим интервальный статистический ряд:

Интервал ( $x_i, x_{i+1}$ ]	[0.39, 0.85]	(0.85, 1.31]	(1.31, 1.77]	(1.77, 2.23]	(2.23, 2.69]	(2.69, 3.15]	(3.15, 3.61]
Частота $n_i$	6	13	31	22	17	8	3

Частота - это количество значений признака, встречающееся в данном интервале. Например, в интервал  $[0.85, 1.31]$  попадает 13 значений.

4) Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которой служат частичные интервалы длиной  $h$ , а высоты равны  $n_i$  (иногда принимают  $\frac{n_i}{h}$ ).





5) Построим дискретный вариационный ряд. Для этого интервалы заменим их серединами  $\tilde{x}_i$ , причем частоты остаются прежними. Рассчитаем точечные оценки для математического ожидания и дисперсии.

Средины интервалов $\tilde{x}_i$	0,62	1,08	1,54	2	2,46	2,92	3,38
Частота $n_i$	6	13	31	22	17	8	3

Точечной оценкой математического ожидания является выборочная средняя  $\bar{x}$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^m \tilde{x}_i n_i$$

$$\begin{aligned} \bar{x} &= \frac{0.62 \times 6 + 1.08 \times 13 + 1.54 \times 31 + 2 \times 22 + 2.46 \times 17 + 2.92 \times 8 + 3.38 \times 3}{100} = \\ &= \frac{184.82}{100} = 1.8482 \end{aligned}$$

Итак,  $\bar{x} = 1.8482$

Точечной оценкой дисперсии является несмещённая выборочная дисперсия  $s^2$ :

$$s^2 = \frac{\sum_{i=1}^m (\tilde{x}_i - \bar{x})^2 n_i}{n - 1}$$

$$s^2 = \frac{(0.62 - 1.8482)^2 \times 6 + (1.08 - 1.8482)^2 \times 13 + (1.54 - 1.8482)^2 \times 31}{99} +$$

$$+ \frac{(2 - 1.8482)^2 \times 22 + (2.46 - 1.8482)^2 \times 17 + (2.92 - 1.8482)^2 \times 8}{99} +$$

$$+ \frac{(3.38 - 1.8482)^2 \times 3}{99} = \frac{42.766}{99} = 0.4320$$

Несмещённое среднее квадратическое отклонение:

$$s = \sqrt{s^2} \approx 0.6573$$

6) Анализируя результаты исследования, гистограмму распределения, можно предположить, что данная выборка (масса одного колоса пшеницы сорта *Sonnora*) имеет нормальное распределение со средней массой одного колоса примерно 1.85г и средним квадратическим отклонением 0.6573г.

## 7 Задача

Условие: Построить 90%-ный доверительный интервал для вероятности попадания снаряда в цель, если после 220-ти выстрелов в цель попало 75 снарядов (предполагается, что случайная величина - попадания снаряда в цель - имеет нормальное распределение).

Решение: Точечная оценка для вероятности попадания снаряда в цель

$$w = \frac{m}{n} = \frac{75}{220} = \frac{15}{44} \approx 0.3409$$

Чтобы найти интервальную оценку для генеральной доли  $p$ , зная доверительную вероятность  $\gamma$ , необходимо воспользоваться формулами:

$$w - u_\gamma \sigma_w \leq p \leq w + u_\gamma \sigma_w,$$
$$\sigma_w = \sqrt{\frac{w(1-w)}{n}}$$

Здесь  $u_\gamma$  - квантиль нормального распределения:  $2\Phi(u_\gamma)$ . По таблицам значений интегральной функции Лапласа находим:

$$2\Phi(u_{0.9}) = 0.9 \Rightarrow \Phi(u_{0.9}) = 0.45 \Rightarrow u_{0.9} \approx 1.645$$

Вычислим среднюю квадратическую ошибку  $\sigma_w$  для доли повторной выборки:

$$\sigma_w = \sqrt{\frac{\frac{15}{44} \left(1 - \frac{15}{44}\right)}{220}} \approx 0.032$$

Найдем доверительный интервал:

$$0.3409 - 1.645 \times 0.032 \leq p \leq 0.3409 + 1.645 \times 0.032$$

$$0.3409 - 0.0526 \leq p \leq 0.3409 + 0.0526$$

$$0.2883 \leq p \leq 0.3935$$

## 8 Задача

Условие: Будем считать, что наблюдаемая в задаче №6 СВ имеет гауссовское распределение.

а) Постройте двусторонние доверительные интервалы уровня надёжности 0.99 для математического ожидания и дисперсии наблюдаемой случайной величины.

б) Проверьте на уровне значимости 0.05 гипотезу о том, что математическое ожидание наблюдаемой СВ равно 2, а дисперсия равна 1.

Решение: Выборочная средняя  $\bar{x}$  :

$$\bar{x} = 1.8482$$

Несмещённое среднее квадратическое отклонение:

$$s = 0.6573$$

Число наблюдений  $n = 100$ .

а) Найдем интервальные оценки математического ожидания  $a$  и дисперсии  $\sigma^2$  гауссовского распределения. Из таблиц найдем квантиль распределения Стьюдента для заданной доверительной вероятности 0.99 (уровень значимости  $\alpha = 0.01$ ) и числу степеней свободы  $v = n - 1 = 100 - 1 = 99$  :  $t_{\alpha/2;v} = t_{0.005;99}$ . Предельная погрешность интервальной оценки математического ожидания:

$$\varepsilon = t_{\alpha/2;n-1} \frac{s}{\sqrt{n}} = 2.626 \frac{0.6573}{\sqrt{100}} \approx 0.1726$$

Искомый доверительный интервал, накрывающий среднее значение случайной величины с заданной доверительной вероятностью  $P = 0.99$ , равен

$$\begin{aligned}\bar{x} - \varepsilon &< a < \bar{x} + \varepsilon \\ 1.8482 - 0.1726 &< a < 1.8482 + 0.1726 \\ 1.6756 &< a < 2.028\end{aligned}$$

Для определения доверительного интервала, накрывающего неизвестное среднее квадратическое отклонение  $\sigma$  с доверительной вероятностью 0.99 (уровень значимости  $\alpha = 0.01$ ) найдем квантили  $\chi^2$  - распределения  $\chi_{\alpha/2;n-1}^2$  и  $\chi_{1-\alpha/2;n-1}^2$  с числом степеней свободы  $v = n - 1$ :

$$\chi_{0.005;99}^2 = 138.99, \chi_{0.995;99}^2 = 66.51$$

Доверительный интервал для дисперсии  $\sigma^2$  определяется следующим образом:

$$s^2 \frac{n-1}{\chi_{\alpha/2;n-1}^2} < \sigma^2 < s^2 \frac{n-1}{\chi_{1-\alpha/2;n-1}^2}$$

Доверительный интервал для дисперсии равен

$$0.432 \times \frac{99}{138.99} < \sigma^2 < 0.432 \times \frac{99}{66.51} \\ 0.3077 < \sigma^2 < 0.6431$$

б) Проверим гипотезу о равенстве математического ожидания  $H_0 : a = 2$  (конкурирующая гипотеза:  $H_1 : a \neq 2$ ). В качестве критерия проверки нулевой гипотезы примем величину

$$T_{\text{набл}} = \frac{(\bar{x} - a)\sqrt{n}}{s} = \frac{(1.8482 - 2)\sqrt{100}}{0.6573} = -2.3095$$

Квантиль распределения Стьюдента для заданной доверительной вероятности 0.95 (уровень значимости  $\alpha = 0.05$ ) и числу степеней свободы  $v = n - 1 = 100 - 1 = 99$  :  $t_{\alpha/2;v} = t_{0.025;99} = 1.984$ . Область принятия нулевой гипотезы описывается следующим неравенством:

$$|T_{\text{набл}}| < t_{\alpha/2;v}$$

В данном случае  $2.3095 > 1.984$ . Таким образом, нулевая гипотез отклоняется, принимается альтернативная гипотеза  $H_1$ : математическое ожидание не равно 2. Проверим гипотезу о равенстве дисперсии  $H_0 : \sigma^2 = \sigma_0^2 = 1$  (конкурирующая гипотеза:  $H_1 : \sigma^2 \neq 1$ ).

$$\chi_{\text{набл}}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{99 \times 0.432}{1} = 42.768$$

Данная величина имеет  $\chi^2$  распределение со степенью свободы  $n-1 = 99$ . Находим квантили для уровня значимости 0.025 и 0.0975:

$$\chi_{0.025;99}^2 = 128.42; \chi_{0.975;99}^2 = 73.36$$

Область принятия нулевой гипотезы описывается следующим неравенством:

$$\chi_{0.975;99}^2 < \chi_{\text{набл}}^2 < \chi_{0.025;99}^2$$

Поскольку  $42.768 \notin (73.36; 128.42)$ , гипотеза о равенстве дисперсии  $\sigma^2 = 1$  отвергается на уровне значимости 0.05. Принимаем  $H_1 : \sigma^2 \neq 1$ .