Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

ИРКУТСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

|  |
| --- |
| Институт информационных технологий и анализа данных |

наименование института

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ № 1

по дисциплине:

|  |
| --- |
| **Моделирование систем** |
| **«Программные средства статистического анализа»** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Выполнил | АСУб-20-2 |  |  |  | Арбакова А.В. |
|  | шифр группы |  | подпись |  | Фамилия И.О. |
| Проверил |  |  |  |  | Бучнев О.С. |
|  | должность |  | подпись |  | Фамилия И.О. |

Иркутск 2022 г.

**Содержание**

[**Введение 3**](#_Toc115379124)

[**Ход выполнения работы 4**](#_Toc115379125)

[1. Таблица и график гистограммы частот 4](#_Toc115379126)

[2. Таблица и график гистограммы накопленных частот 6](#_Toc115379127)

[3. Оценки основных числовых характеристик 8](#_Toc115379128)

[4. Доверительные интервалы для оценок математического ожидания и дисперсии 8](#_Toc115379129)

[5. Х2-критерий сравнения 9](#_Toc115379130)

[**Заключение 12**](#_Toc115379131)

[**Список литературы 13**](#_Toc115379132)

# **Введение**

**Цель лабораторной работы:** ознакомление с программными средствами статистического анализа и обработки наблюдений, а также повторение базовых понятий теории вероятностей и математической статистки.

**Язык программирования:** R

**Вариант лабораторной работы:** 3

**Выборка, соответствующая варианту:**

0.136001,0.530630,0.024833,0.901844,0.112081,1.117082,0.588929,1.417762,0.062309,0.657376,0.682872,0.292832,1.289541,0.252664,0.435376,1.508902,0.801946,0.627029,1.491898,1.320438,1.160220,0.063959,2.712940,1.114241,0.283846,1.847821,1.334505,0.154099,0.282520,0.715329,1.277721,0.823322,0.457308,0.559254,0.052565,0.353177,0.044248,0.118427,0.623034,0.902862,0.145386,0.365574,2.012537,0.805111,1.487046,1.425042,0.099711,0.371497,2.491687,0.056108,0.925483,4.106039,0.343339,0.783963,0.650423,1.415783,0.199050,0.292303,0.322167,0.106668,0.602420,0.063707,0.705936,0.553528,0.530871,0.435753,0.146514,0.949825,0.157136,0.653054,0.621129,0.326597,1.415676,0.852276,0.501352,0.485514,0.016249,0.031295,0.316009,1.540844,3.840937,1.287609,0.047888,0.570924,0.201600,0.079175,0.936836,0.716126,2.761384,2.059141,0.534188,0.270164,0.607758,1.015230,0.765101,3.399164,0.052445,1.477064,3.684560,0.243171

# **Ход выполнения работы**

1. **Таблица и график гистограммы частот**

Понятие случайной величины.

Случайная величина – это величина, которая случайно принимает какое-то значение из множества возможных значений. Случайные величины различают на дискретные и непрерывные случайные величины.

Непрерывные – случайные величины, которые могут принимать все значения из некоторого числового промежутка.

Дискретными являются случайные величины, значениями которых являются только отдельные точки числовой оси, которые могут быть конечного и бесконечного количества. Дискретную случайную величину можно характеризовать законом распределения.

Закон распределения случайно величины.

Закон распределения случайной величины – это соответствие между возможными значениями случайной величины и соответствующими им вероятностями.

Генеральная совокупность.

Генеральная совокупность – это совокупность объектов исследования, которая обладает интересующими исследователя свойствами.

Генеральная совокупность может быть конечной или бесконечной, в зависимости от конечности или бесконечности совокупности объектов. Например, производимый продукт заводом является бесконечной генеральной совокупностью, а общее число живых людей на планете – конечная генеральная совокупность.

Выборка.

Выборка – это ограниченная группа объектов генеральной совокупности для изучения ее свойств.

Генеральная совокупность

Выборка

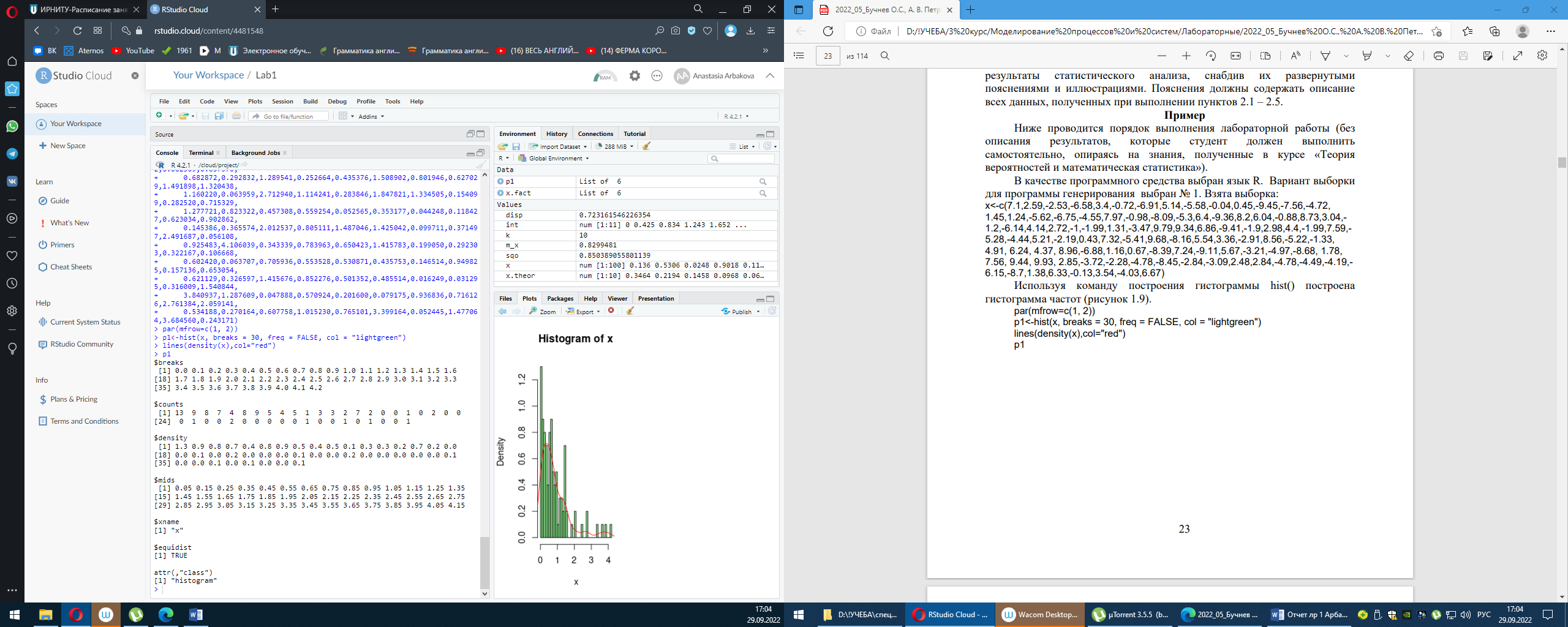


Рисунок 1 – Таблица частот

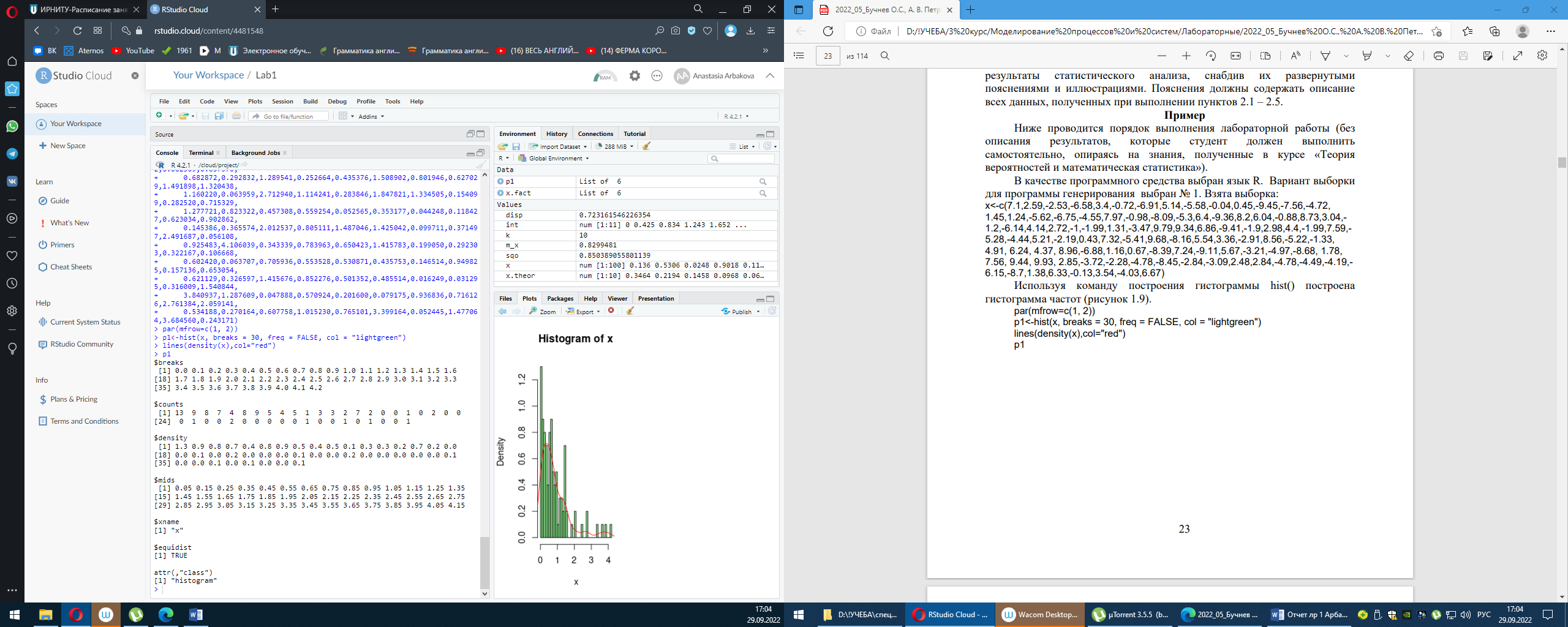


Рисунок 2 – График гистограммы частот

1. **Таблица и график гистограммы накопленных частот**

Гистограмма частот.

Гистограмма частот – это ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы.

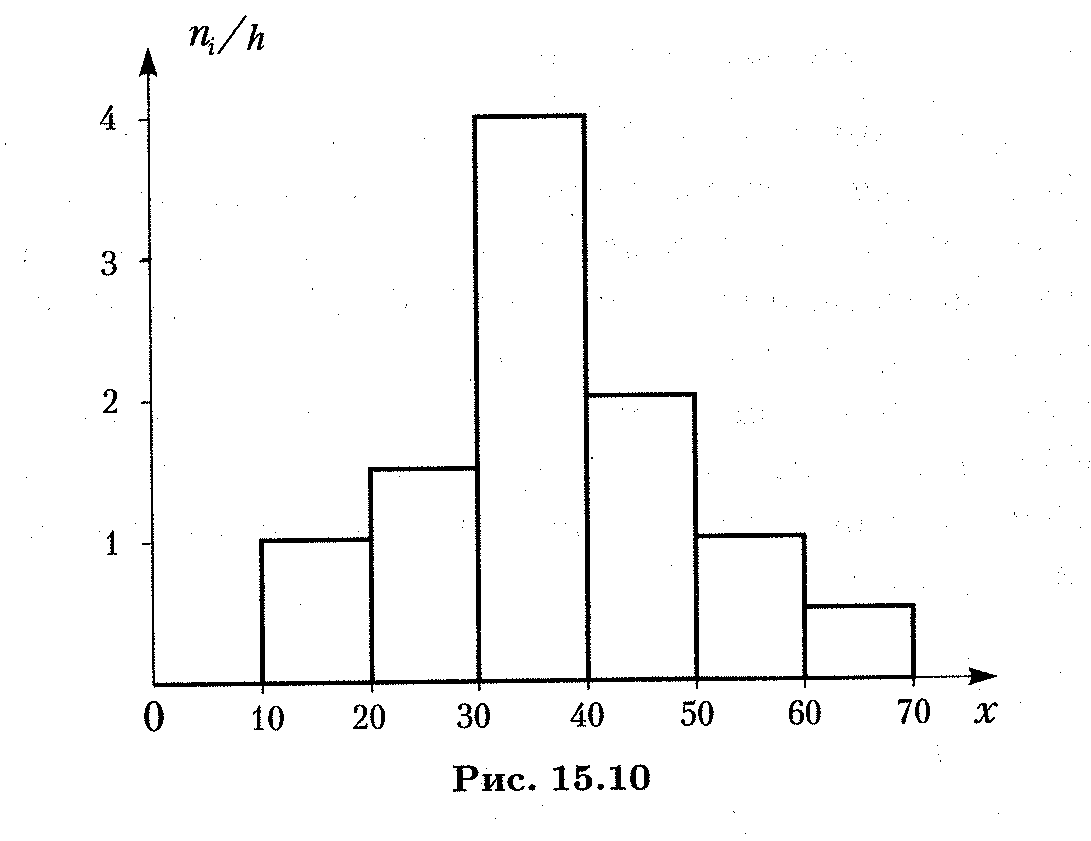


Рисунок 3 – Пример гистограммы частот

Частотой случайного события в серии испытаний называется отношение числа испытаний, в которых это событие наступило, к числу всех испытаний.

Гистограмма накопленных частот.

Гистограмма накопленных частот отличается от гистограммы распределения тем, что высота каждого столбца пропорциональна частоте, накопленной к интервалу.

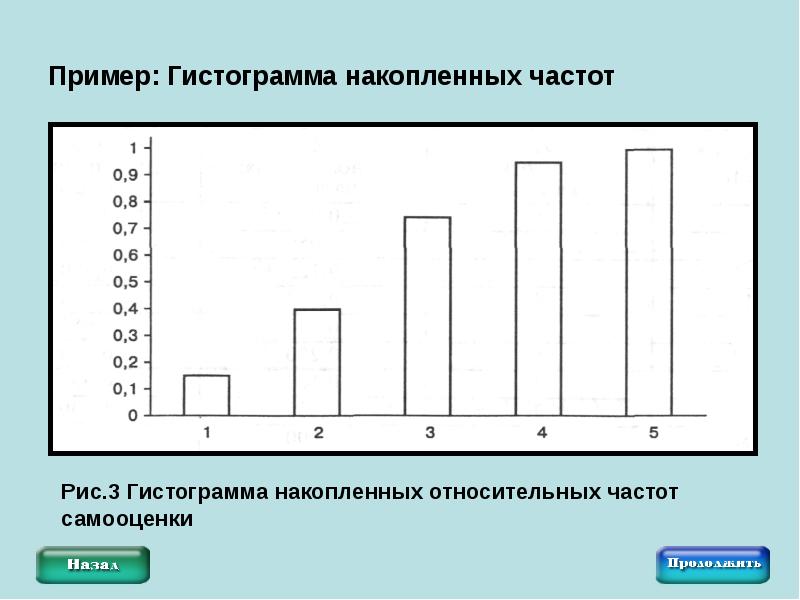


Рисунок 4 – Пример гистограммы накопленных частот

Накопленные частоты показывают, сколько единиц совокупности не превышают заданного значения x.

Функция распределения.

Функция F(x) распределения равна вероятности того, что случайная величина X примет значение меньше заданного числа x.

F(x)=P(X<x)

Плотность распределения.

Плотность распределения показывает, как часто значение случайной величины X появляется в окрестности точки x при повторении опытов.

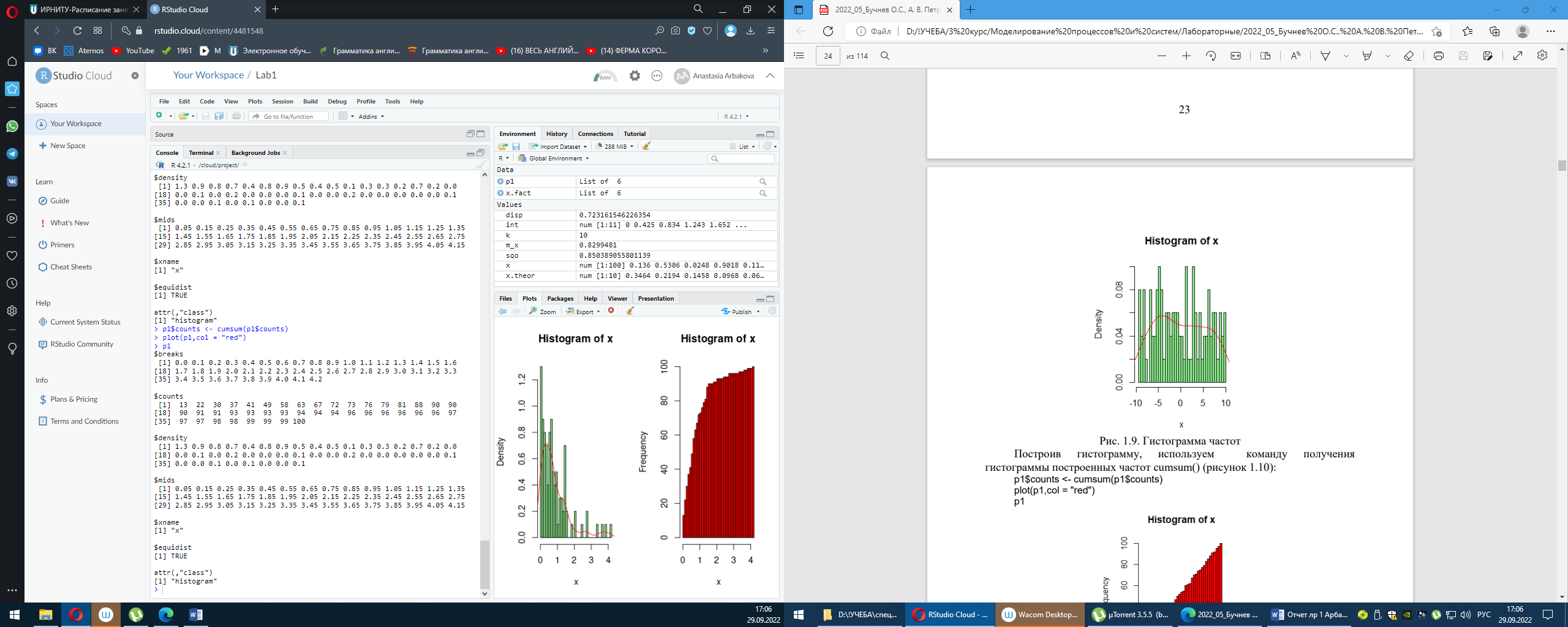


Рисунок 5 – Таблица накопленных частот

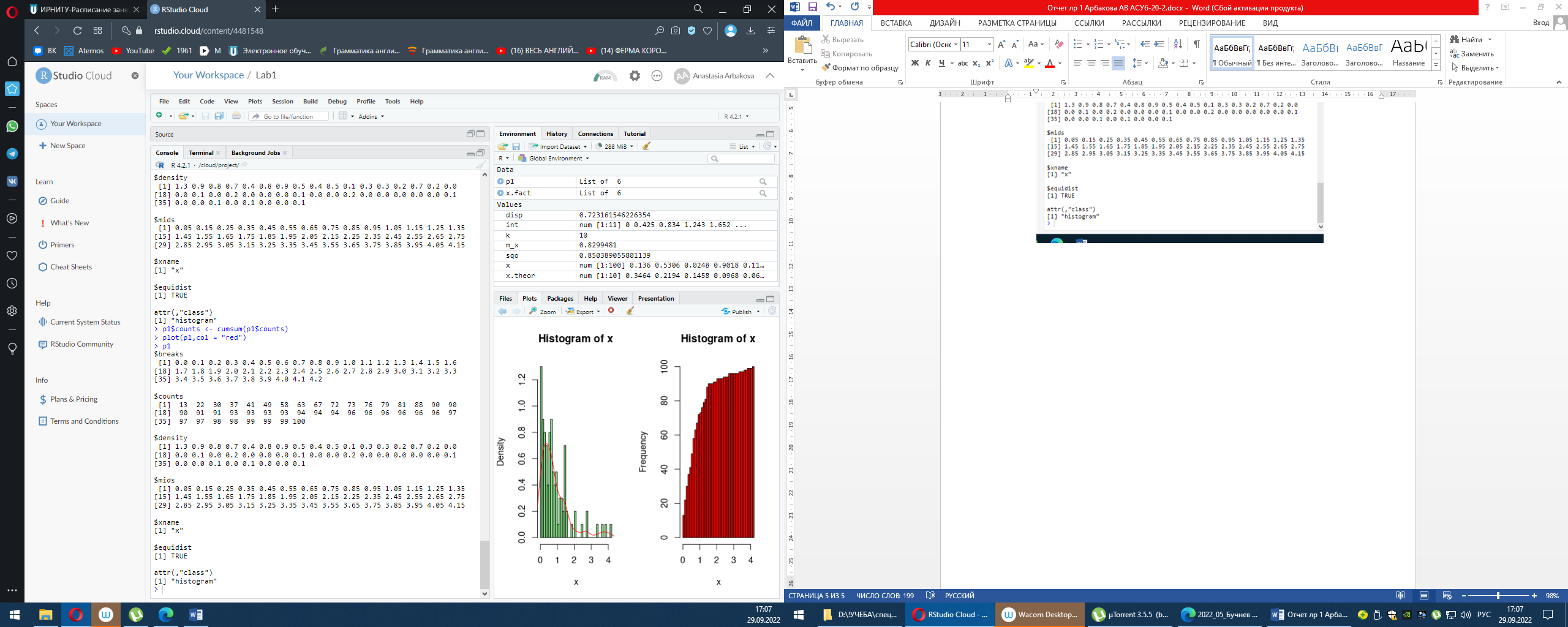


Рисунок 6 – График гистограммы накопленных частот

1. **Оценки основных числовых характеристик**

Оценка среднего и стандартного отклонения.

Среднее отклонение – это среднее из абсолютных значений отклонений отдельных вариантов от их средней величины.

Стандартное отклонение – оценка среднеквадратичного отклонения случайной величины x относительно её математического ожидания на основе несмещенной оценки её дисперсии.

Дисперсия.

Дисперсия – это средний квадрат отклонений индивидуальных значений признаков от их средней величины.

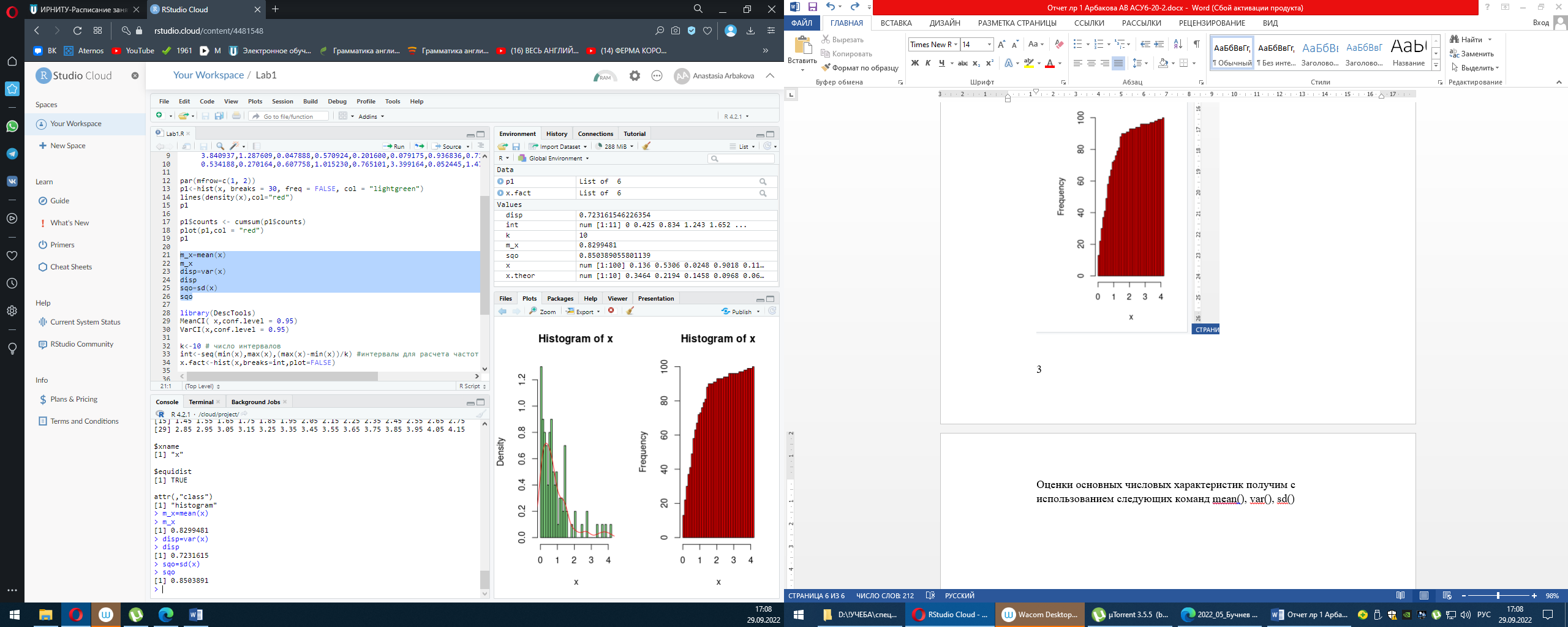


Рисунок 7 – Оценки основных числовых характеристик

Из результатов можно сказать, что математическое ожидание (m\_x) равно 0,8299481, дисперсия (disp) 0,7231615 и среднеквадратическое отклонение (sqo) 0,8503891.

1. **Доверительные интервалы для оценок математического ожидания и дисперсии**

Доверительный интервал для среднего.

Наряду с выборочным средним, которое является точечной оценкой параметра, часто приводит интервальные оценки.

Доверительный интервал – это диапазон значений, внутри которого находится средняя популяция с вероятностью 95%, 99%.

Сравнение эмпирического среднего с генеральным средним.

Эмпирическая функция распределения выборки служит для оценки теоретической функции распределения генеральной совокупности.

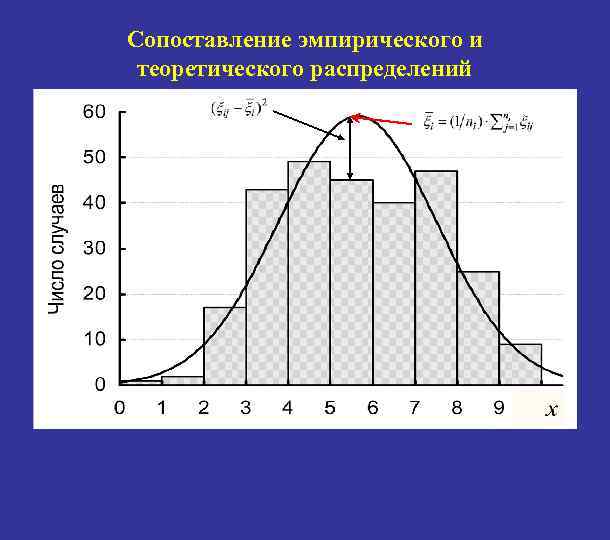


Рисунок 8 – Сравнение эмпирического среднего с генеральным средним

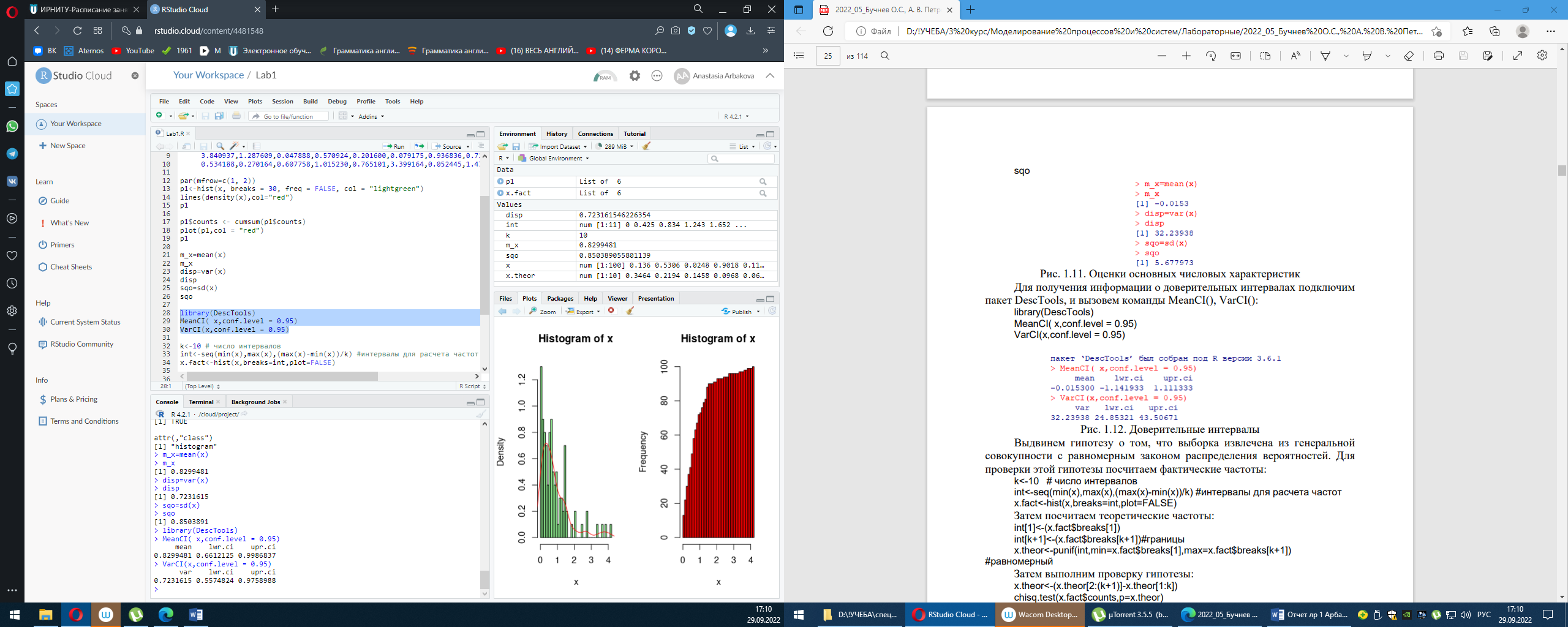


Рисунок 9 – Доверительные интервалы

1. **Х2-критерий сравнения**

Нулевая гипотеза.

Нулевая гипотеза – это гипотеза об отсутствии отличий. Это то, что мы хотим опровергнуть, если перед нами стоит задача доказать значимость различий. Гипотеза называется нулевой, потому что содержит 0.

X1 – X2 = 0

X1 и X2 – сопоставляемые значения признаков.

Ошибки 1-го и 2-го рода.

Ошибка 1 рода происходит, когда мы отвергаем нулевую гипотезу, когда она верна.

Ошибка 2 рода происходит, когда мы принимаем нулевую гипотезу, когда она неверна.

Уровень значимости.

Уровень значимости – вероятность отклонения нулевой гипотезы при условии, что она истинная. Уровень значимости – вероятность ошибки 1 рода.

Х2-критерий сравнения.

Критерий согласия Пирсона X2 – один из основных, который можно представить, как сумму отношений квадратов расхождений между теоретическими (fT) и эмпирическими (f) частотами к теоретическим частотам.

Результаты неподходящих выборке законов распределения вероятностей:

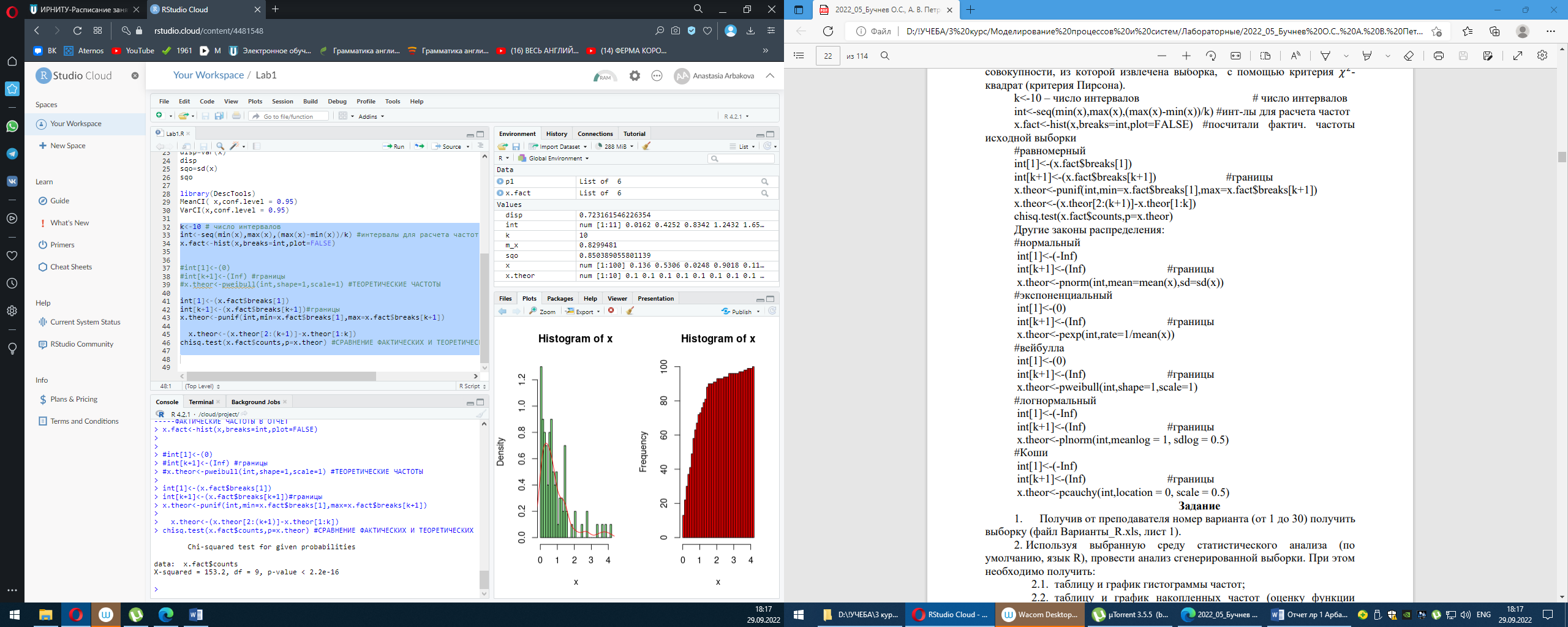


Рисунок 10 – Равномерный закон распределения вероятностей

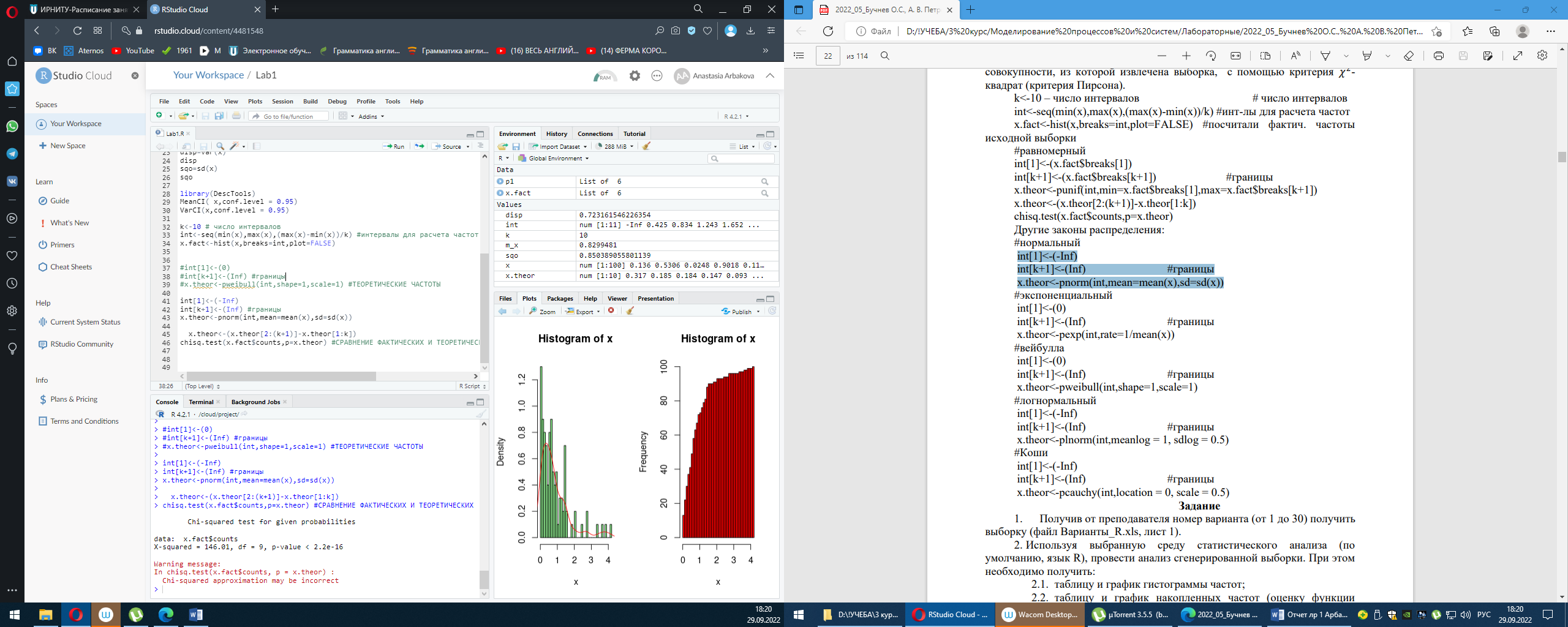


Рисунок 11 – Нормальный закон распределения вероятностей

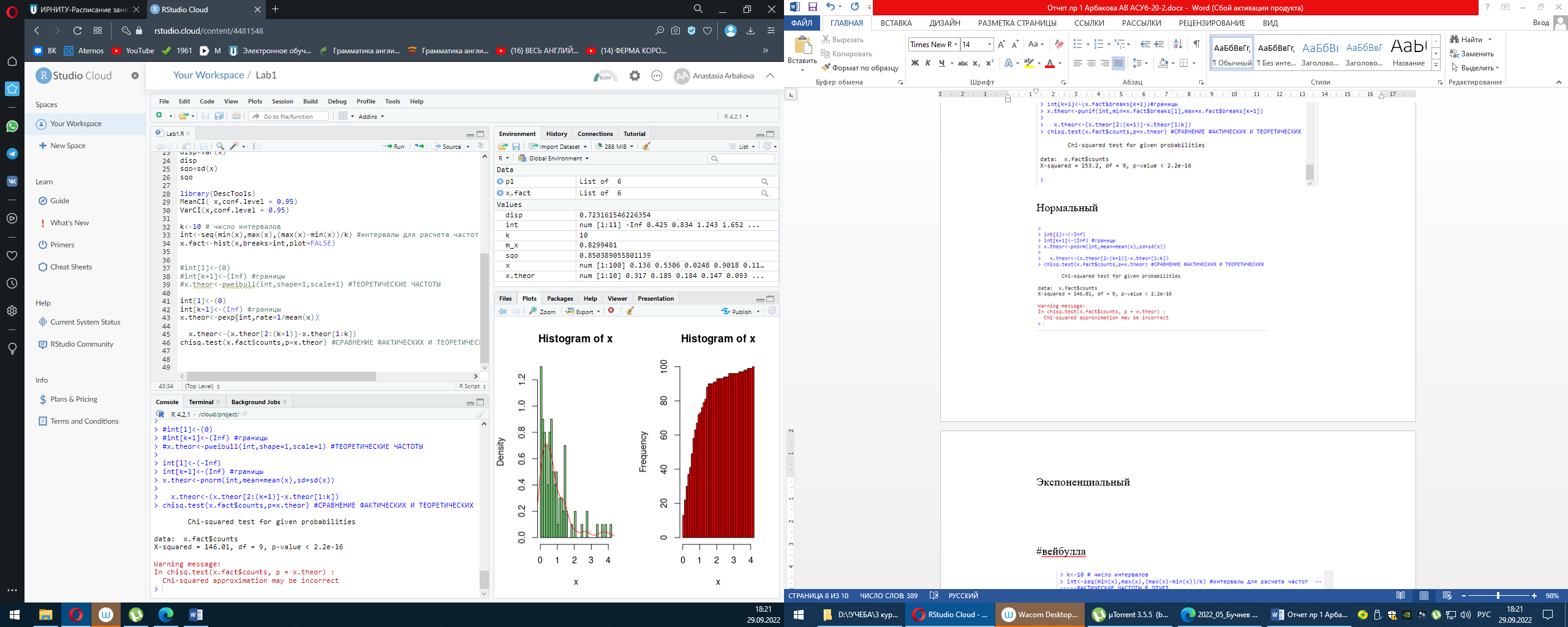


Рисунок 12 – Экспоненциальный закон распределения вероятностей

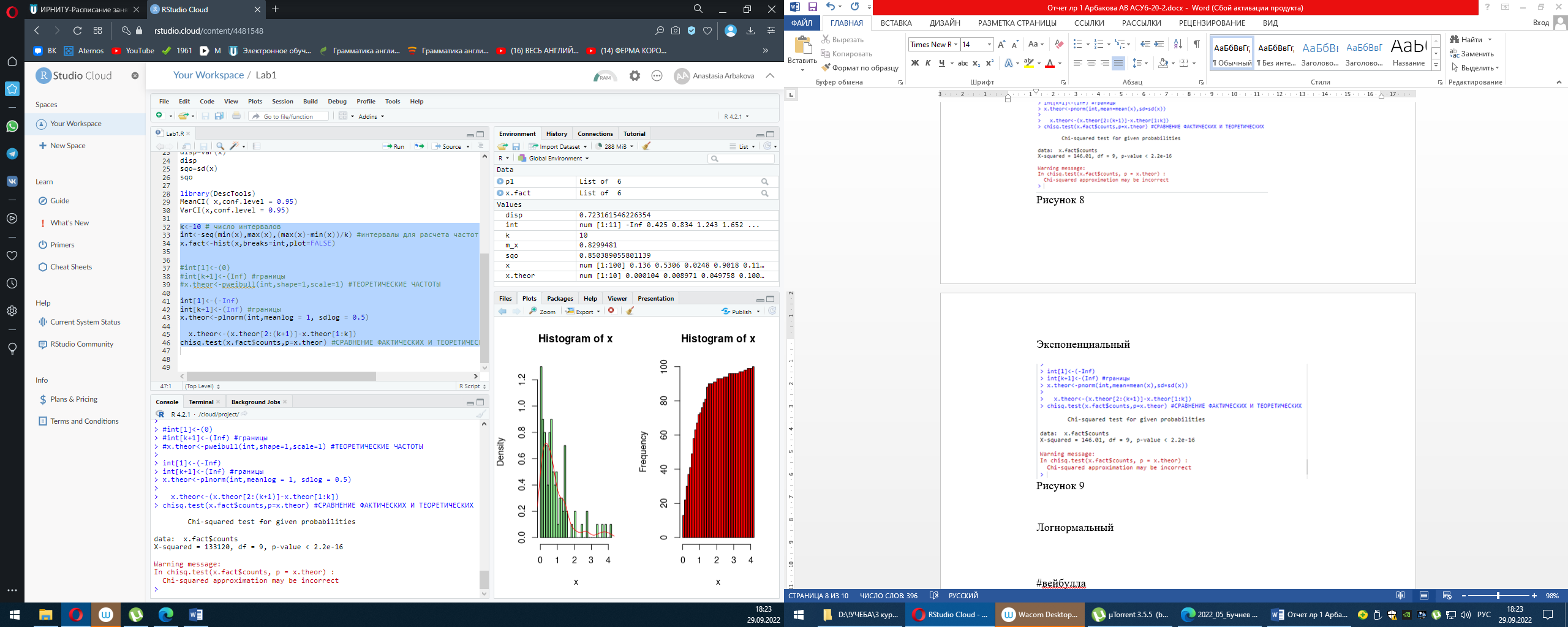


Рисунок 13 – Логнормальный закон распределения вероятностей

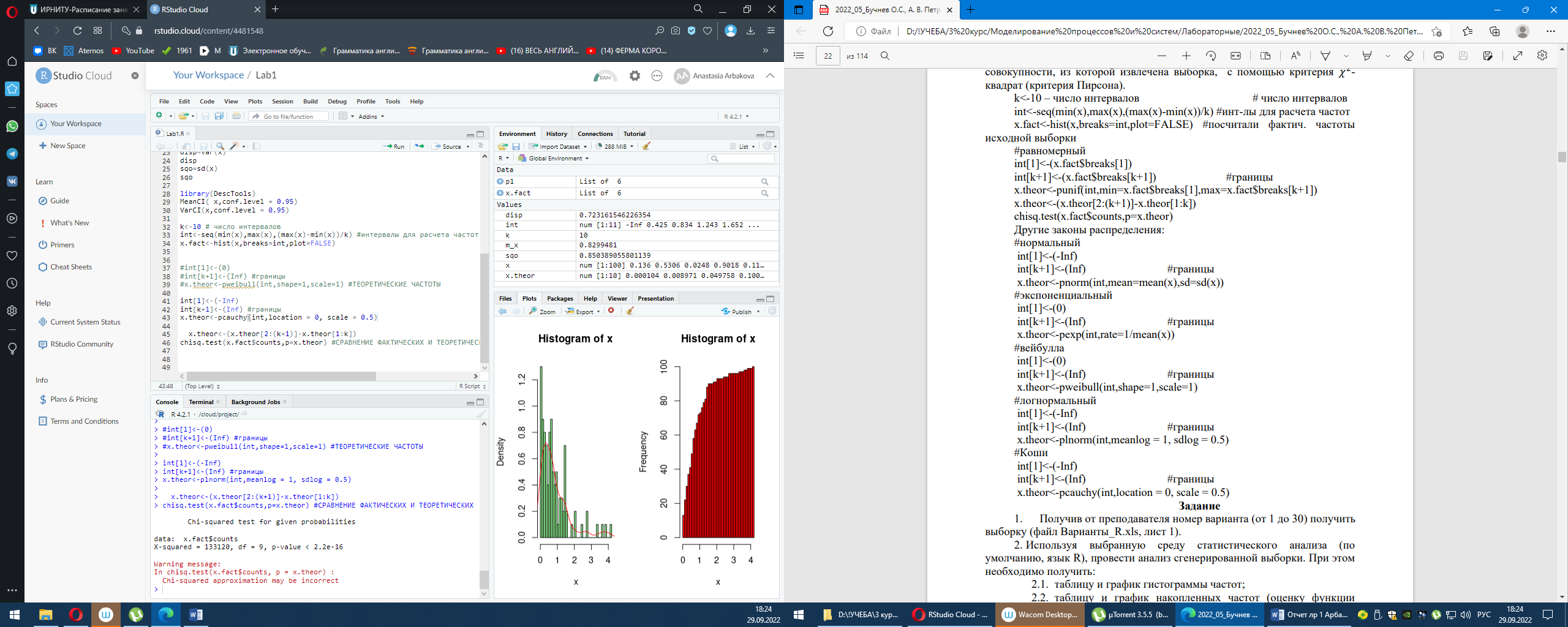


Рисунок 14 – Закон распределения вероятностей Коши

Подходящий закон распределения вероятностей для данной выборки – Вейбулла.

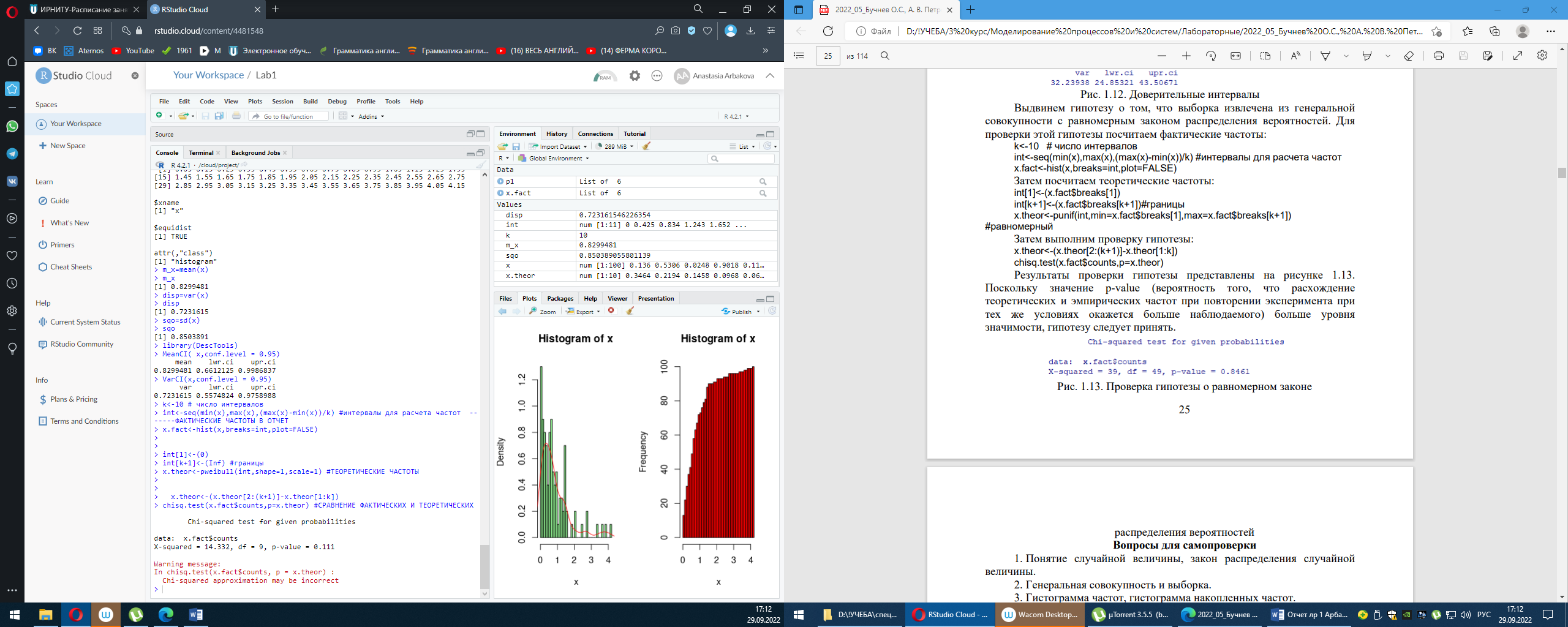


Рисунок 15 – Х2 критерий сравнения с законом распределения Вейбулла

# **Заключение**

В ходе выполнения лабораторной работы по теме «Программные средства статистического анализа» была выполнена цель работы по ознакомлению с программными средствами статистического анализа и обработке наблюдений, а также по повторению базовых понятий теории вероятностей и математической статистки.

Было изучены такие темы как: гистограммы частот, гистограммы накопленных частот, оценки основных числовых характеристик, доверительные интервалы для оценок математического ожидания и дисперсии, и X2-критерий сравнения.

Знания были закреплены во время использования среды разработки программного обеспечения RStudio и использования языка программирования R, в котором были построены гистограммы и вычислены требуемые заданием значения.

# **Список литературы**

1. Петров А. В., Бучнев О. С. Моделирование процессов и систем: лабораторный практикум – Иркутск, 2022. – 114 с
2. Советов Б. Я., Яковлев С. А. Моделирование систем: Учеб. для вузов — 3-е юд., перераб. и доп. — М.: Высш. шк., 2001. — 343 с: ил.
3. RStudio https://rstudio.cloud