Санкт-Петербургский политехнический университет Петра Великого

Институт компьютерных наук и технологий

Кафедра компьютерных систем и программных технологий

**Отчет по расчетному заданию №2**

**Дисциплина**: Теория вероятностей и математическая статистика

**Тема**: Статистическая обработка случайных последовательностей. Идентификация законов распределения

Выполнила студентка гр. 23531/3 В.В. Константинова

(подпись)

Преподаватель К.В. Никитин

(подпись)

“\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_\_ 2018 г.

Санкт-Петербург

2018

**1 Исходные данные**

В результате измерений получена выборка из генеральной совокупности с неизвестным законом распределения.

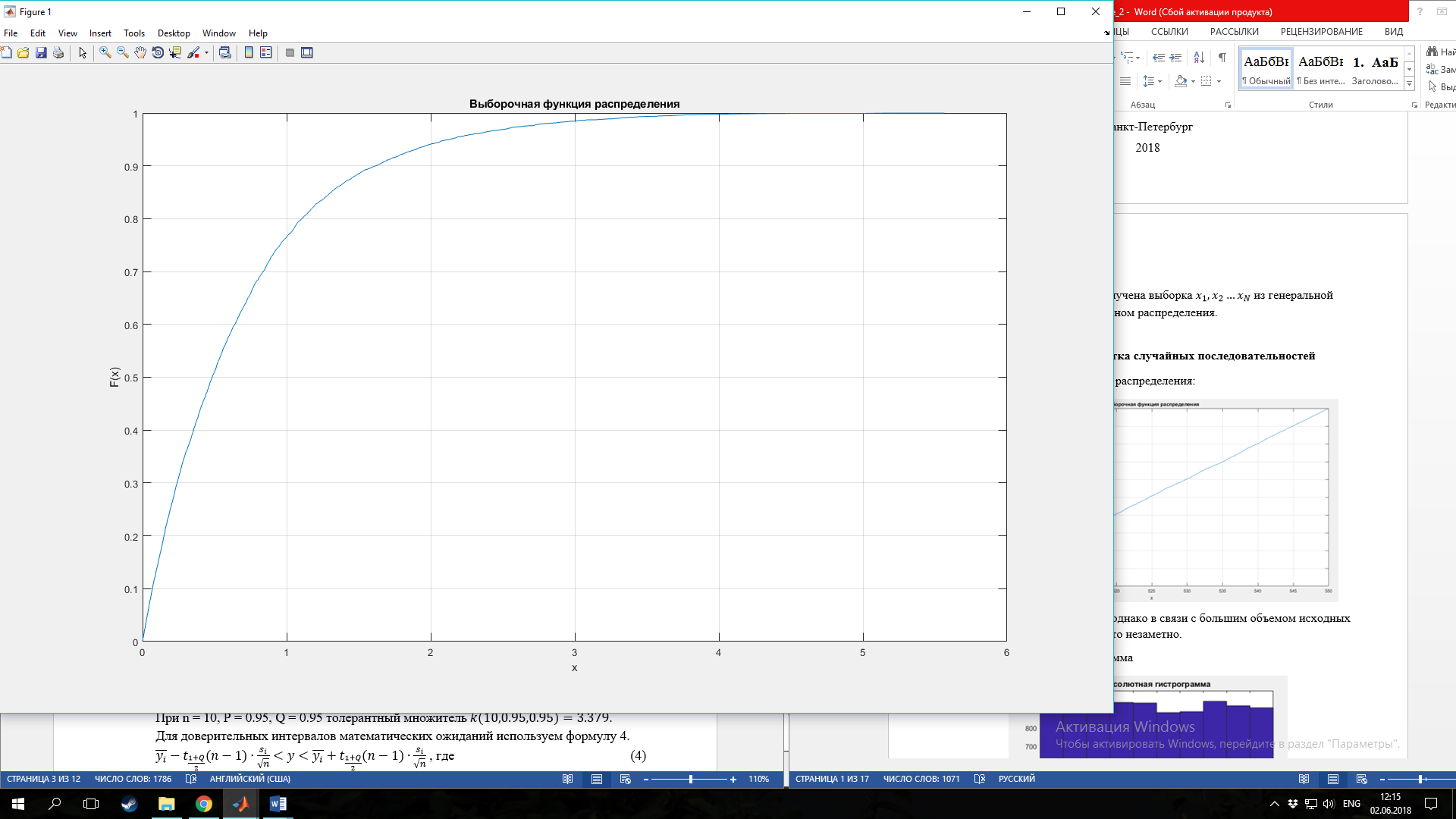
Число значений: 5800.

Значения вещественные – случай непрерывного распределения.

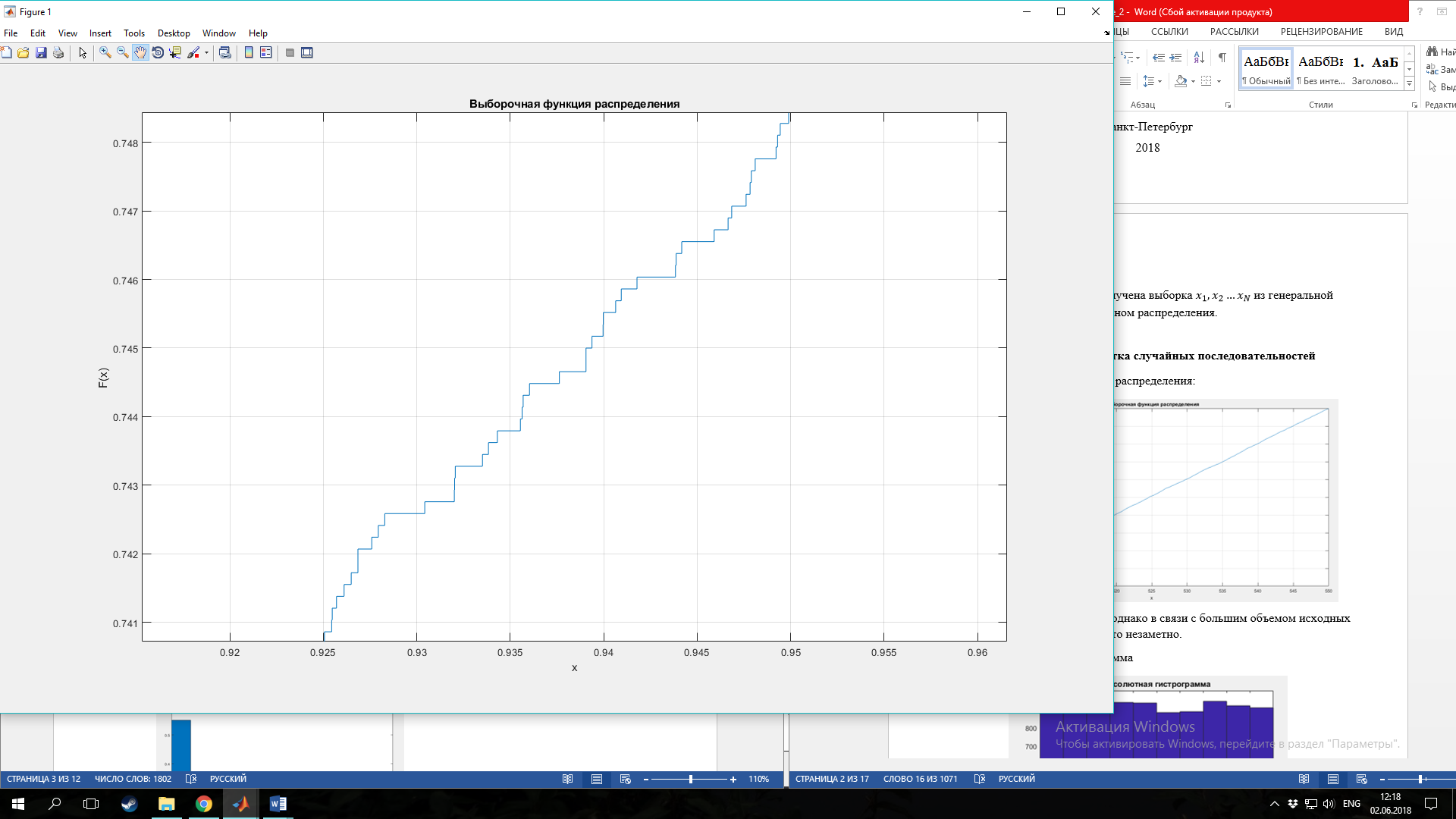
**2 Статистическая обработка случайных последовательностей**

Поcтроим по исходным данным функцию распределения, абсолютную и относительную гистограммы. Количество интервалов равно 10.

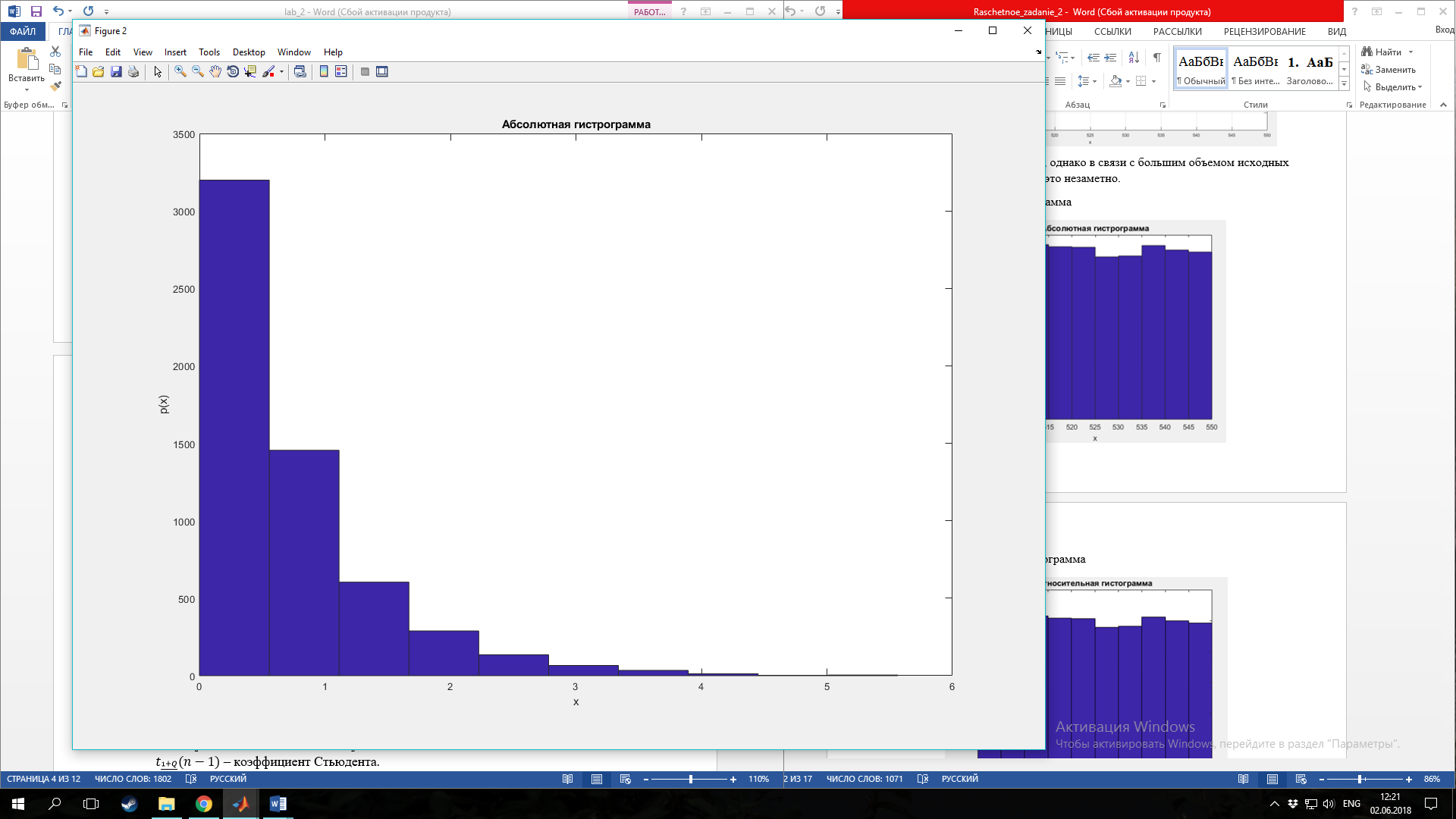
**2.1 Выборочная функция распределения**



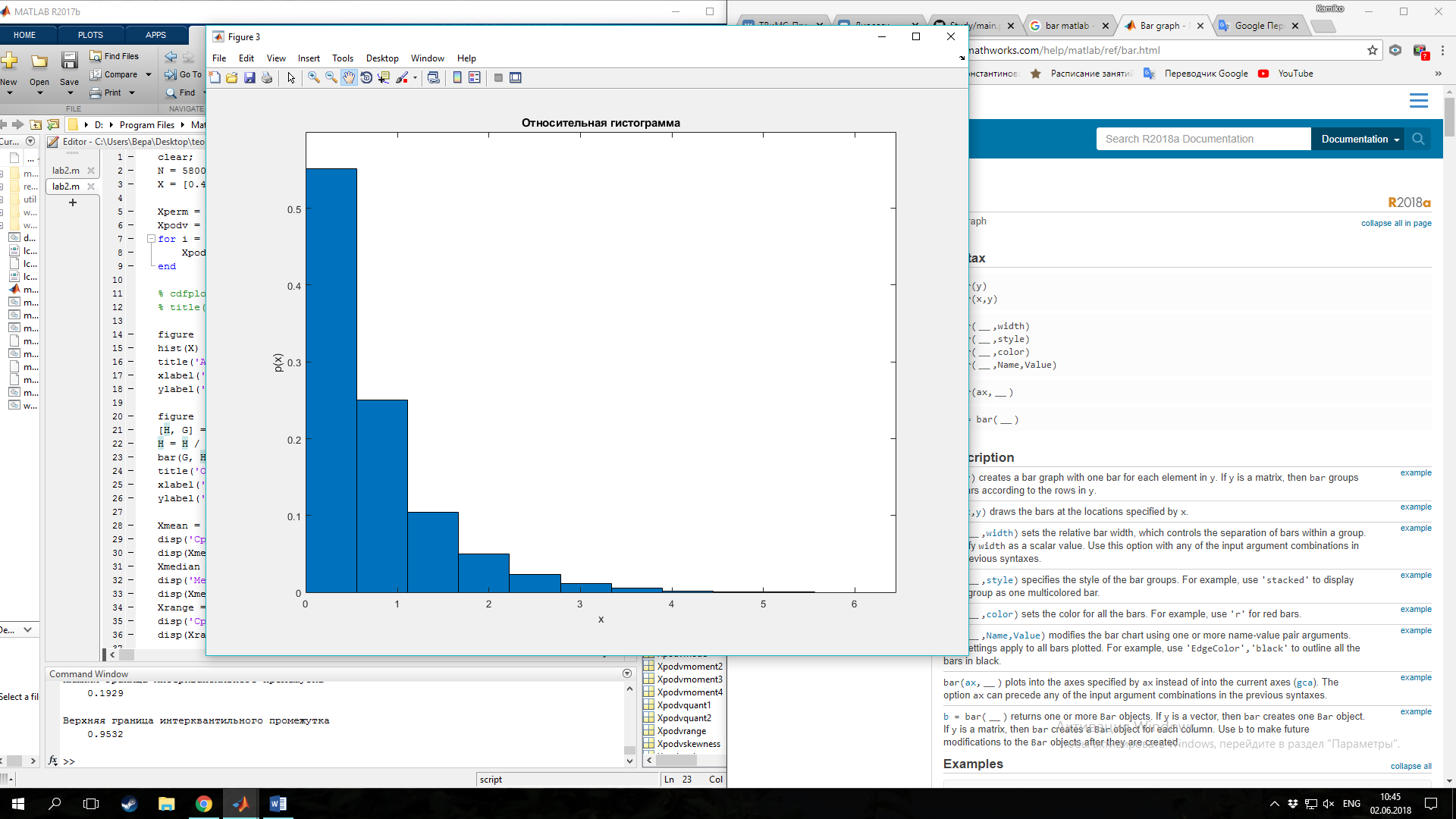
Она является ступенчатой, однако в связи с большим объемом исходных значений при данном масштабе это незаметно. Убедимся в этом использовав “Zoom in”.



**2.2 Абсолютная гистограмма**



**2.3 Относительная гистограмма**



## 2.4 Определение точечных оценок

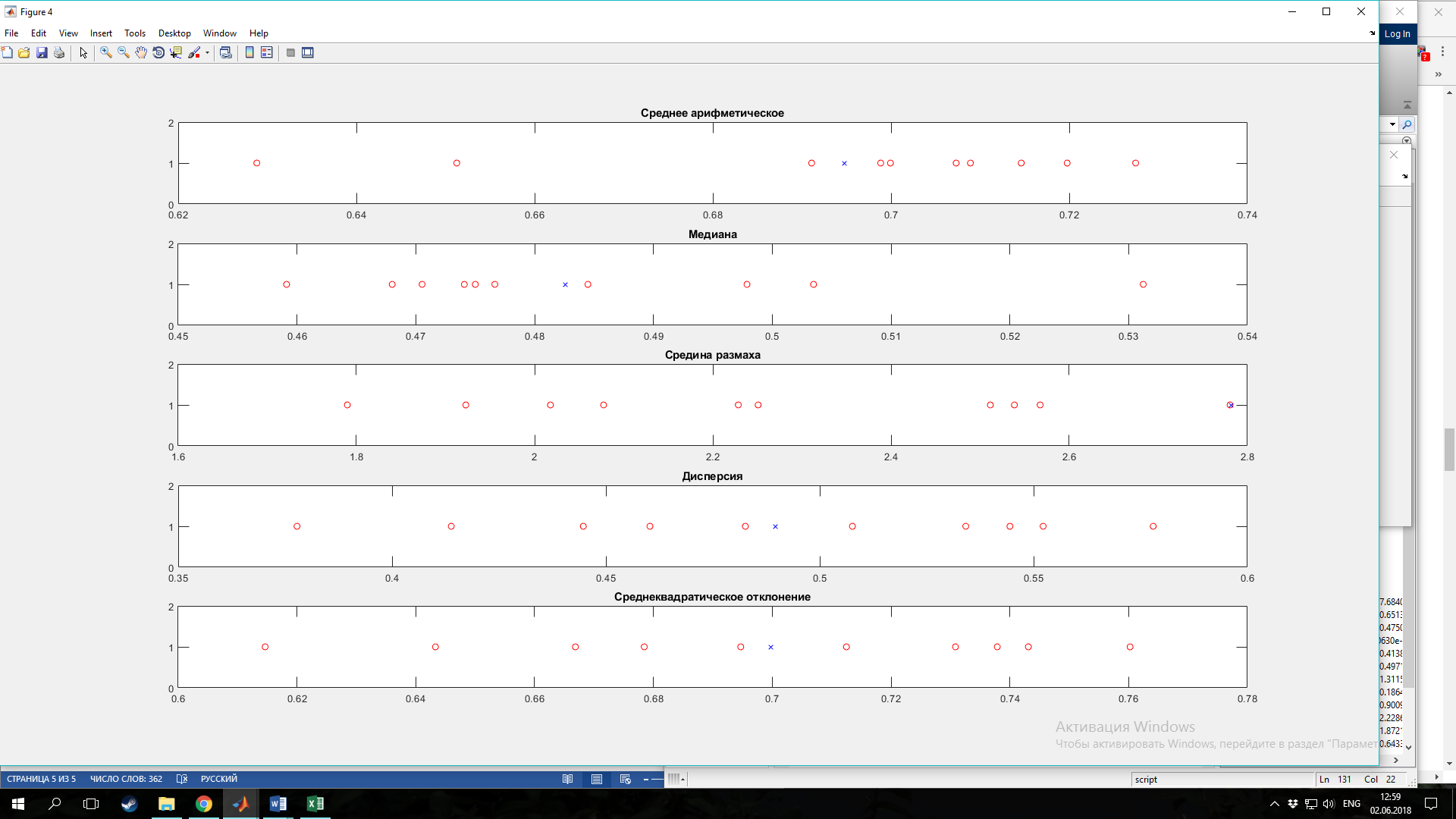
Таблица 2.4. Точечные оценки

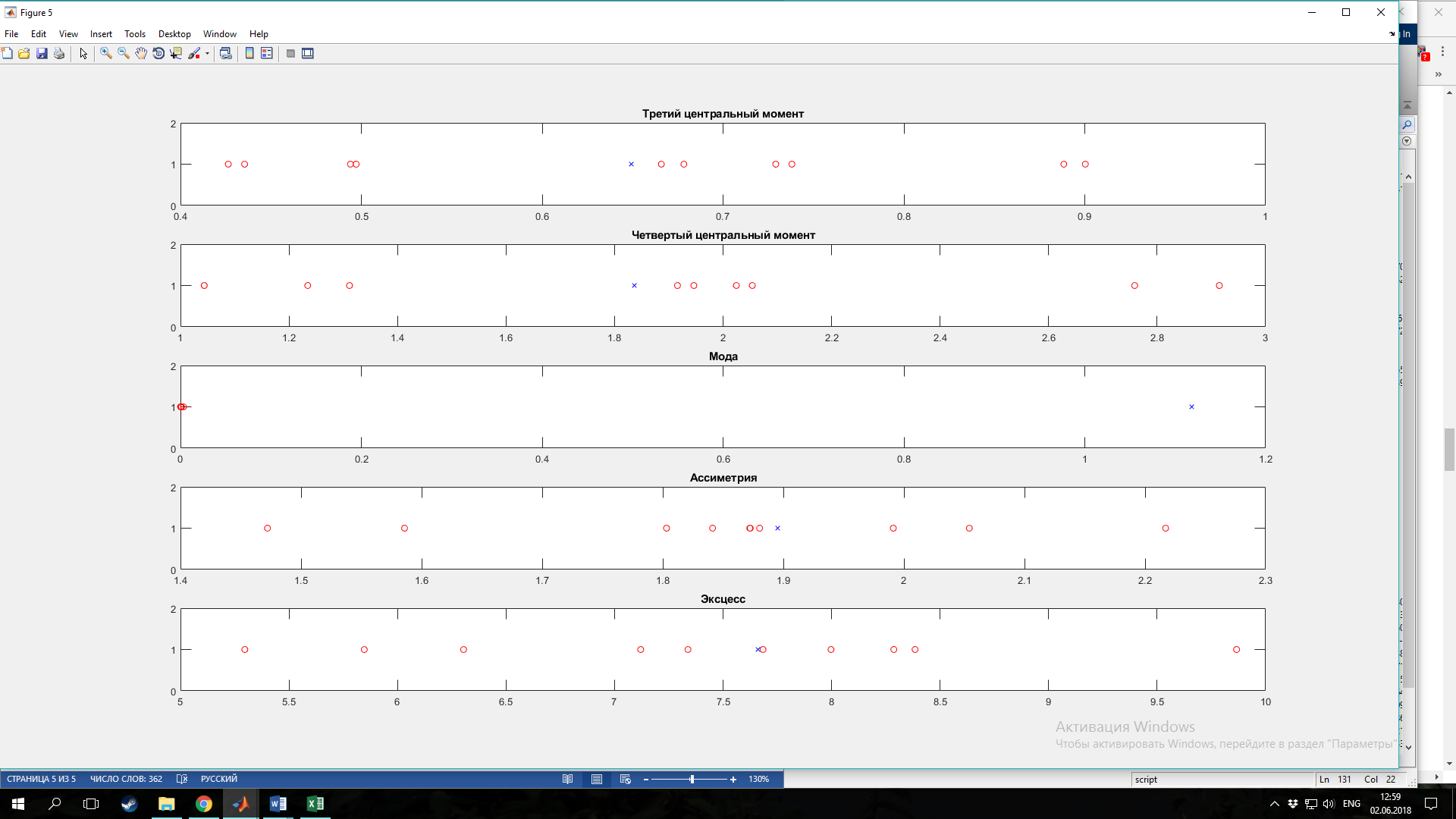
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | s |  |  | Mode | As | Ex |
| N | 0.694799 | 0,482613 | 2.782124 | 0.489683 | 0.699774 | 0.649331 | 1.836035 | 1.11831 | 1.895419 | 7.65949 |
|  | 0,688662 | 0,464723 | 2,510312 | 0,510546 | 0,714525 | 0,750077 | 2,172287 | 0,002726 | 2,061477 | 8,362711 |
|  | 0,6947 | 0,48901 | 1,966919 | 0,473698 | 0,688257 | 0,58886 | 1,556638 | 5,15E-05 | 1,810858 | 6,9612 |
|  | 0,71301 | 0,515756 | 2,567735 | 0,502766 | 0,70906 | 0,691427 | 2,078875 | 0,00036 | 1,944559 | 8,252677 |
|  | 0,682067 | 0,452293 | 2,031864 | 0,466035 | 0,682668 | 0,548633 | 1,417484 | 0,000952 | 1,728933 | 6,549074 |
|  | 0,712918 | 0,542939 | 2,249901 | 0,504389 | 0,710204 | 0,662244 | 1,809267 | 0,002239 | 1,853503 | 7,136247 |
|  | 0,696443 | 0,508521 | 2,108364 | 0,467668 | 0,683862 | 0,668649 | 1,927522 | 0,000793 | 2,096121 | 8,843481 |
|  | 0,713307 | 0,464527 | 2,537361 | 0,546316 | 0,739132 | 0,721097 | 2,094184 | 0,002238 | 1,790407 | 7,040863 |
|  | 0,700078 | 0,509496 | 2,228648 | 0,437273 | 0,661266 | 0,451533 | 1,169132 | 0,000604 | 1,565614 | 6,135599 |
|  | 0,644933 | 0,440745 | 2,165106 | 0,442725 | 0,665376 | 0,569759 | 1,479942 | 0,000528 | 1,939161 | 7,576615 |
|  | 0,701871 | 0,439145 | 2,781997 | 0,549255 | 0,741117 | 0,828626 | 2,615607 | 0,000306 | 2,0409 | 8,700086 |

Границы интерквантильного промежутка по полной выборке:

[0.1929245; 0.9532235]

## 2.5 Графическое представление результатов





## 2.6 Интервальные оценки с доверительной вероятностью Q = 0.8

Доверительный интервал для математических ожиданий:

, где

– коэффициент Стьюдента.

При Q = 0.8, n = 580 коэффициент Стьюдента .

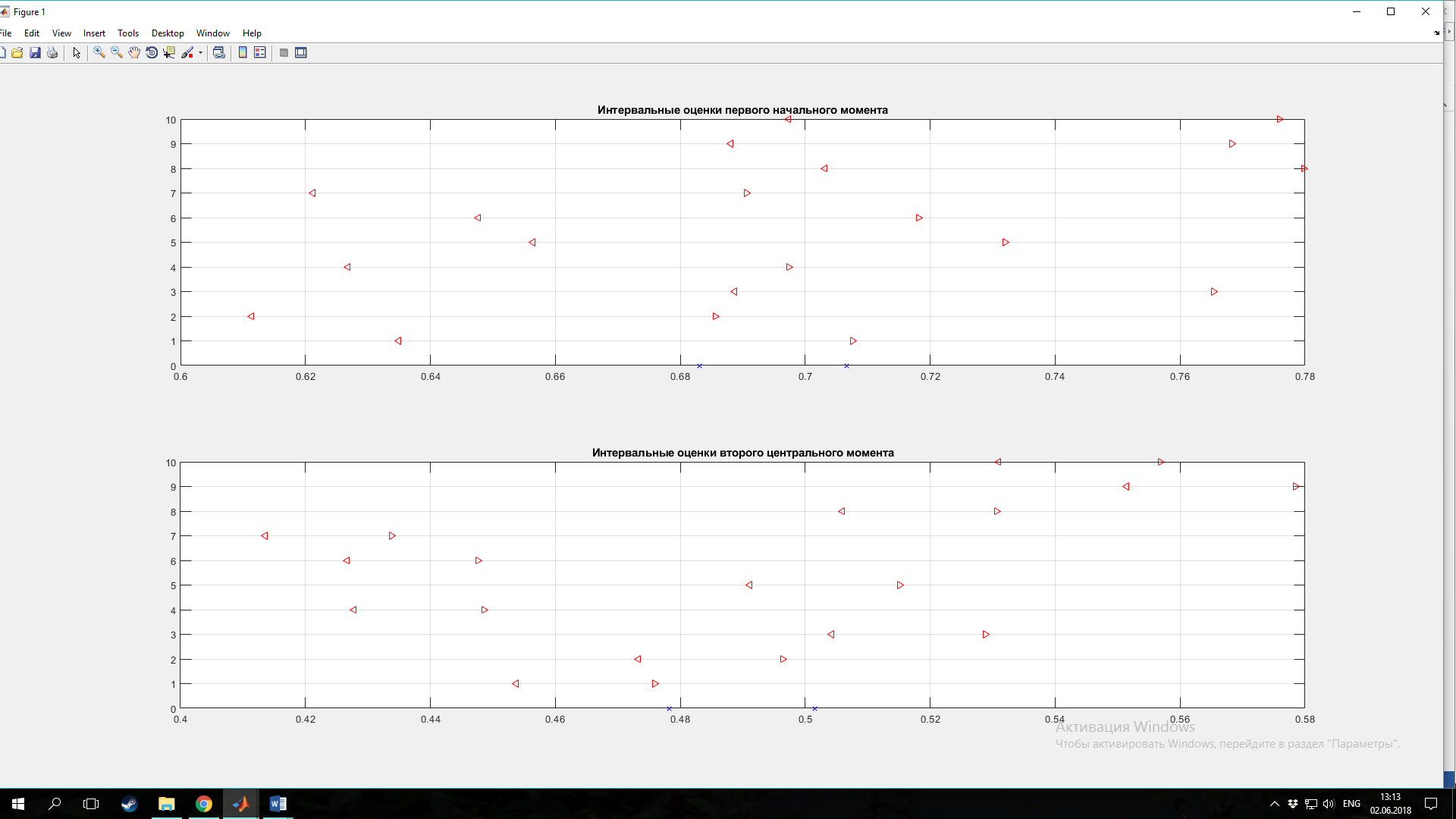
Доверительный интервал для дисперсии:

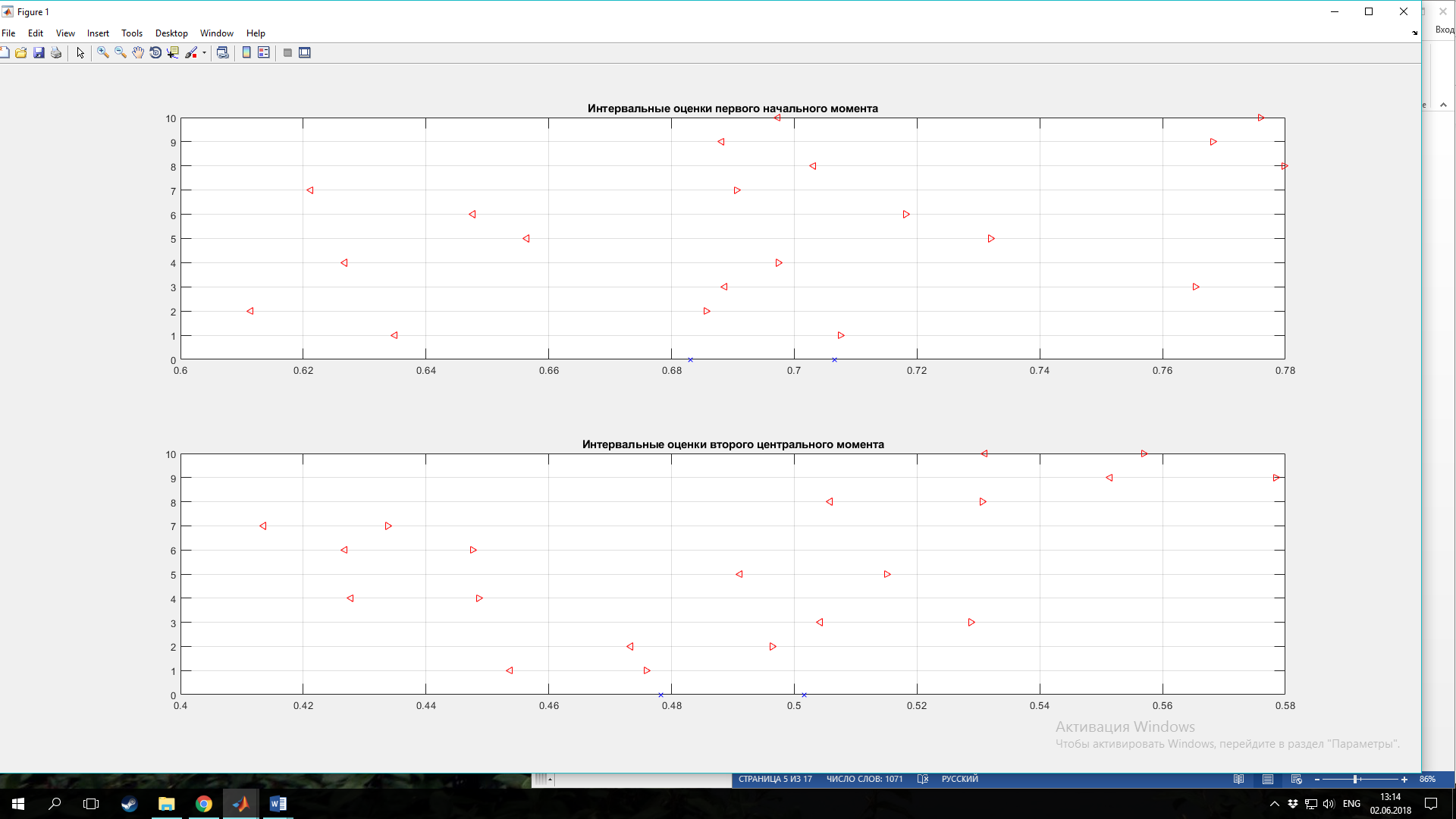
, где

и – квантили плотности распределения “хи-квадрат”.

Таблица 2.6.1. Интервальные оценки при Q = 0.8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Выборка | Мат. ожидание | | Дисперсия | |
| Слева | Справа | Слева | Справа |
| Полная | 0.683022115891545 | 0.706575824180282 | 0.478266088433898 | 0.501583458354177 |
| 1 | 0.642096979449845 | 0.718847408705327 | 0.506780622721782 | 0.531488189355093 |
| 2 | 0.668065523233978 | 0.743456948569471 | 0.488992595311486 | 0.512832925012660 |
| 3 | 0.710899889424238 | 0.788288293265418 | 0.515240690063368 | 0.540360718555306 |
| 4 | 0.597121579842160 | 0.668926776399221 | 0.443578129605464 | 0.465204323865703 |
| 5 | 0.625516472238349 | 0.694118651854756 | 0.404887345258666 | 0.424627210228710 |
| 6 | 0.670618541649472 | 0.745065303695357 | 0.476815101321098 | 0.500061729861051 |
| 7 | 0.696556973779322 | 0.776014182452749 | 0.543156502463994 | 0.569637537600794 |
| 8 | 0.662381156721941 | 0.735395955322887 | 0.458648690468322 | 0.481009634381618 |
| 9 | 0.629736800284149 | 0.700546210691713 | 0.431360457504287 | 0.452390991760948 |
| 10 | 0.672686390277470 | 0.749639362860460 | 0.509458926258463 | 0.534297070818651 |



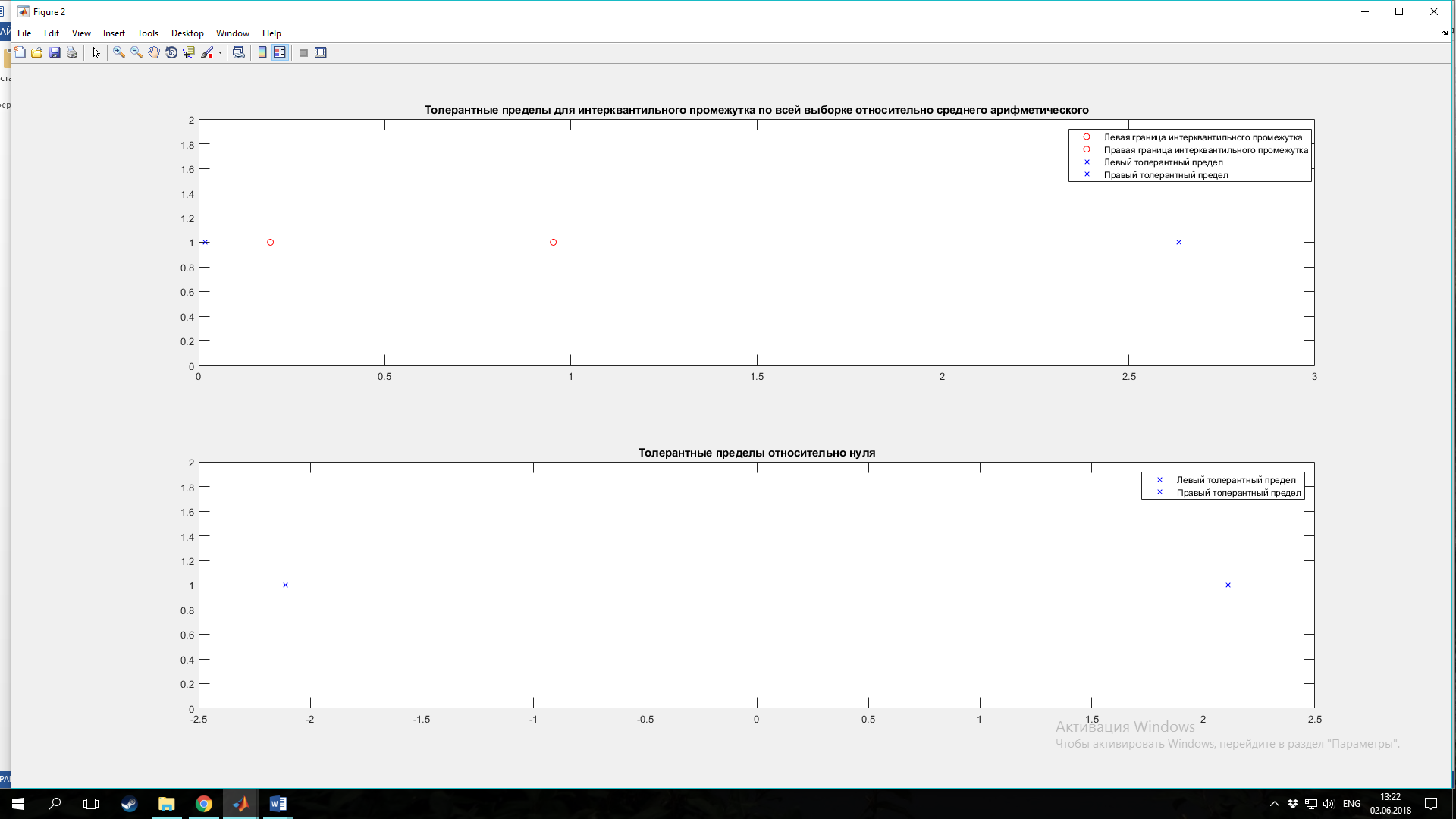


## 2.7 Интервальные оценки интерквантильного промежутка для P = 0.95

Непараметрические толерантные пределы для всей выборки симметричные относительно среднего значения. Кол-во отбрасываемых точек было найдено с помощью биномиального распределения.

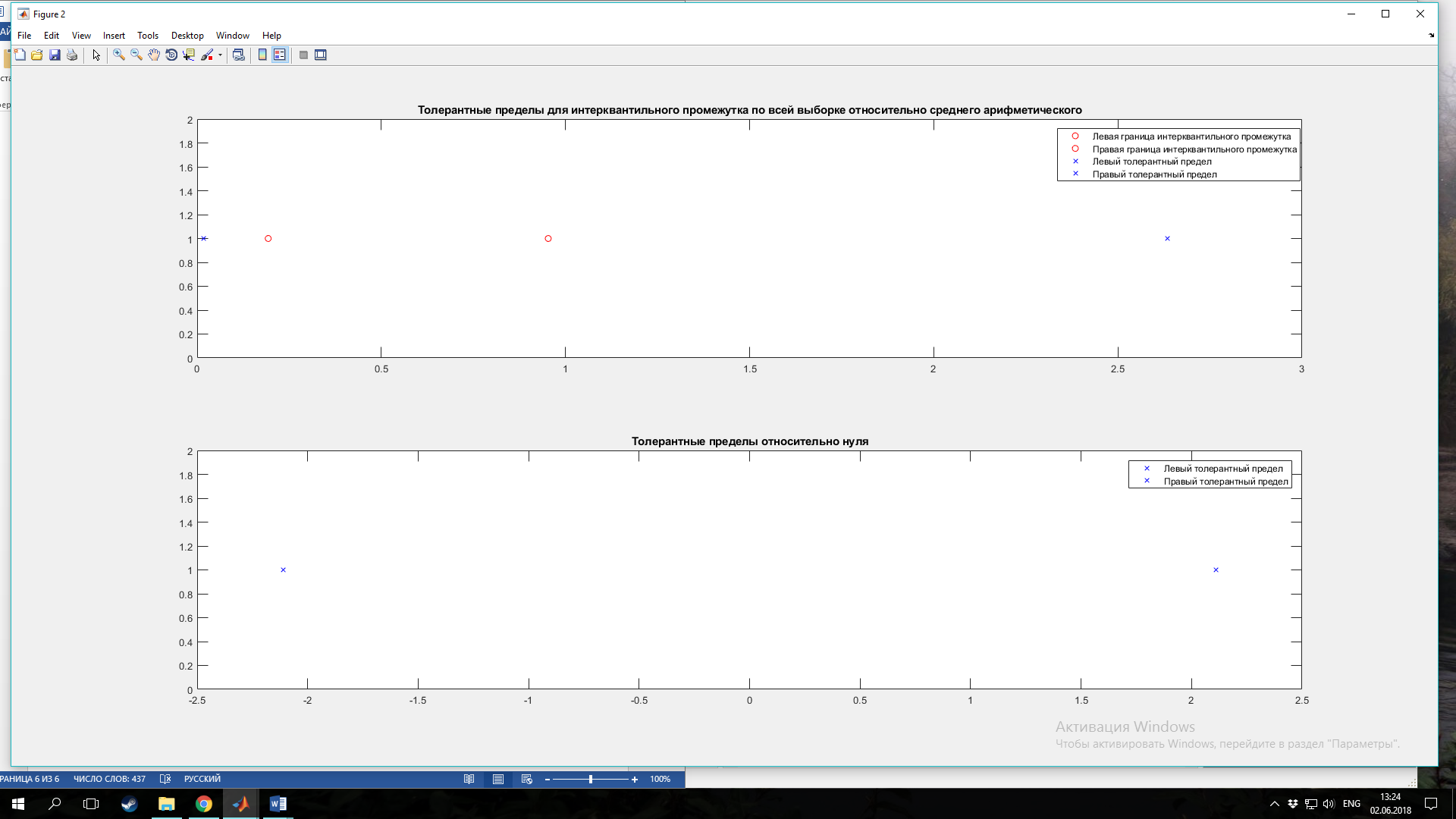
Толерантные пределы всей выборки симметричные относительно среднего значения:

[0.016793, 2.63384]



Толерантные пределы всей выборки симметричные относительно нуля:

[-2.10983, 2.10983]

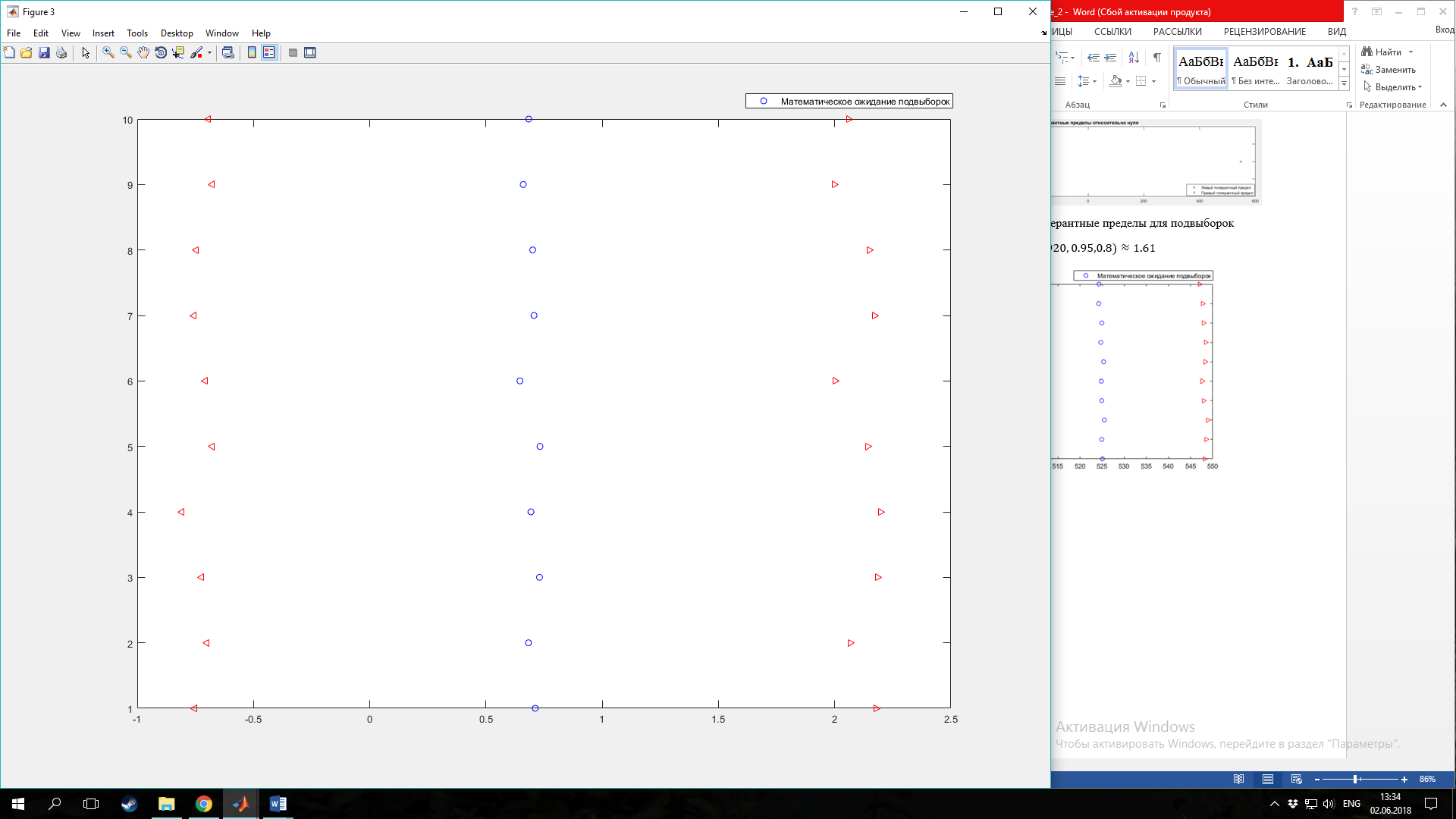


## 2.8 Параметрические толерантные пределы для подвыборок

Параметр

Таблица 2.8.1. Толерантные пределы подвыборок

|  |  |  |
| --- | --- | --- |
| Подвыборка | Левый | Правый |
| 1 | -0.781802540522055 | 2.142746928677227 |
| 2 | -0.730621303491977 | 2.142143775295425 |
| 3 | -0.724835547494858 | 2.224023730184514 |
| 4 | -0.735032336440767 | 2.001080692682147 |
| 5 | -0.647214007218220 | 1.966849131311325 |
| 6 | -0.710542577172869 | 2.126226422517698 |
| 7 | -0.777559619260024 | 2.250130775492094 |
| 8 | -0.692213699453464 | 2.089990811498293 |
| 9 | -0.683942959775577 | 2.014225970751439 |
| 10 | -0.754970780949391 | 2.177296534087322 |



1. **Идентификация закона и параметров распределения**
   1. **Начальный выбор распределения**

В качестве распределений-гипотез были выбраны следующие: гамма, экспоненциальное, Рэлея.

* 1. **Определение параметров теоретических распределений**
* **Метод моментов**

***Гамма-распределение***

***Экспоненциальное распределение***

***Распределение Рэлея***

* **Метод максимального правдоподобия**

Пусть дана некоторая функция. Возьмем от нее логарифм.

***Гамма-распределение***

Получение дальнейшего результата аналитическим методом затруднено, поэтому воспользуемся готовым численным методом – функцией :

***Экспоненциальное распределение***

Выражение, полученное данным методом для экспоненциального распределения вычислено:

Оно может быть получено функцией в Matlab. Значения совпали с точностью до 15 знака после запятой.

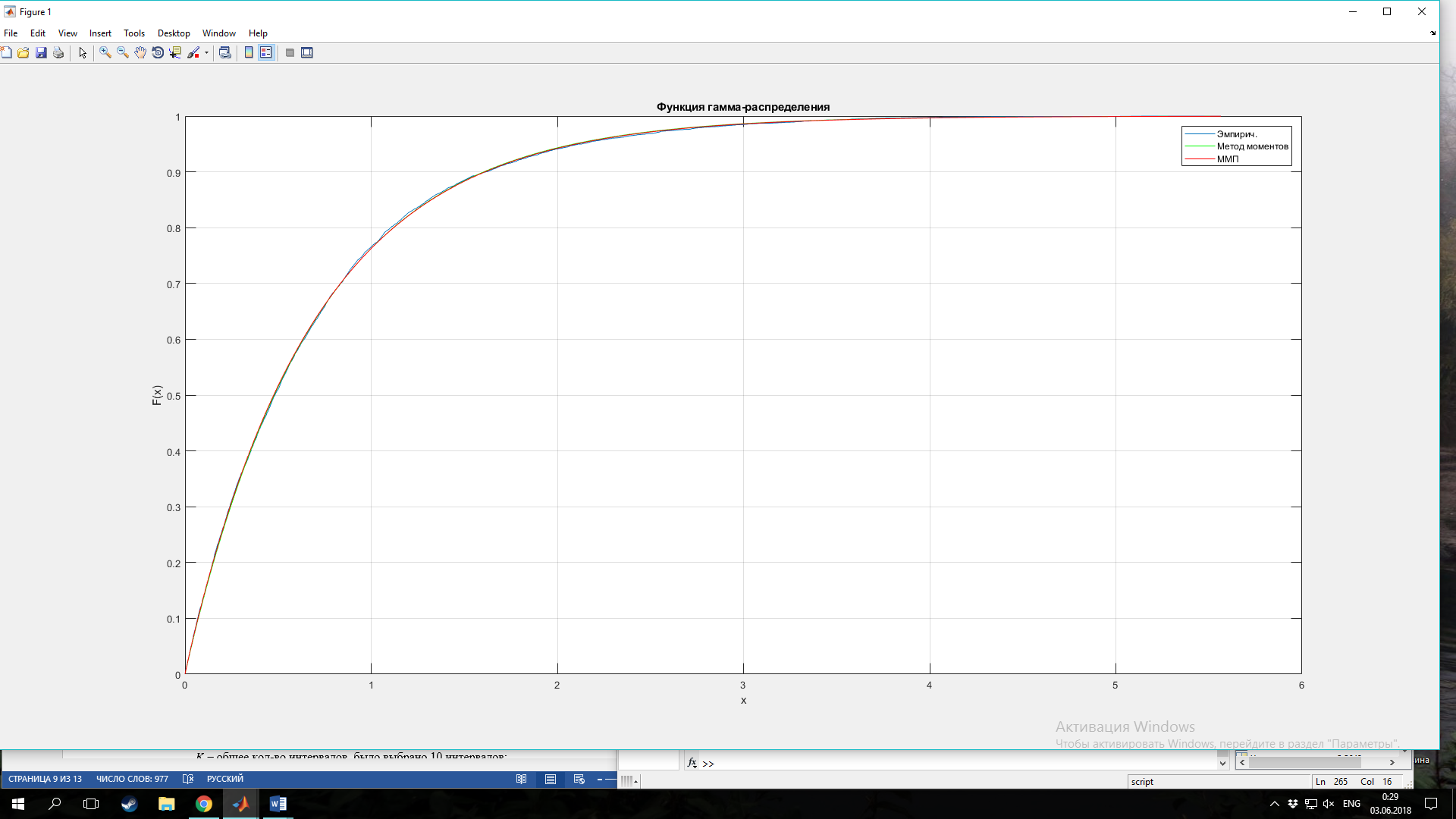
***Распределение Рэлея***

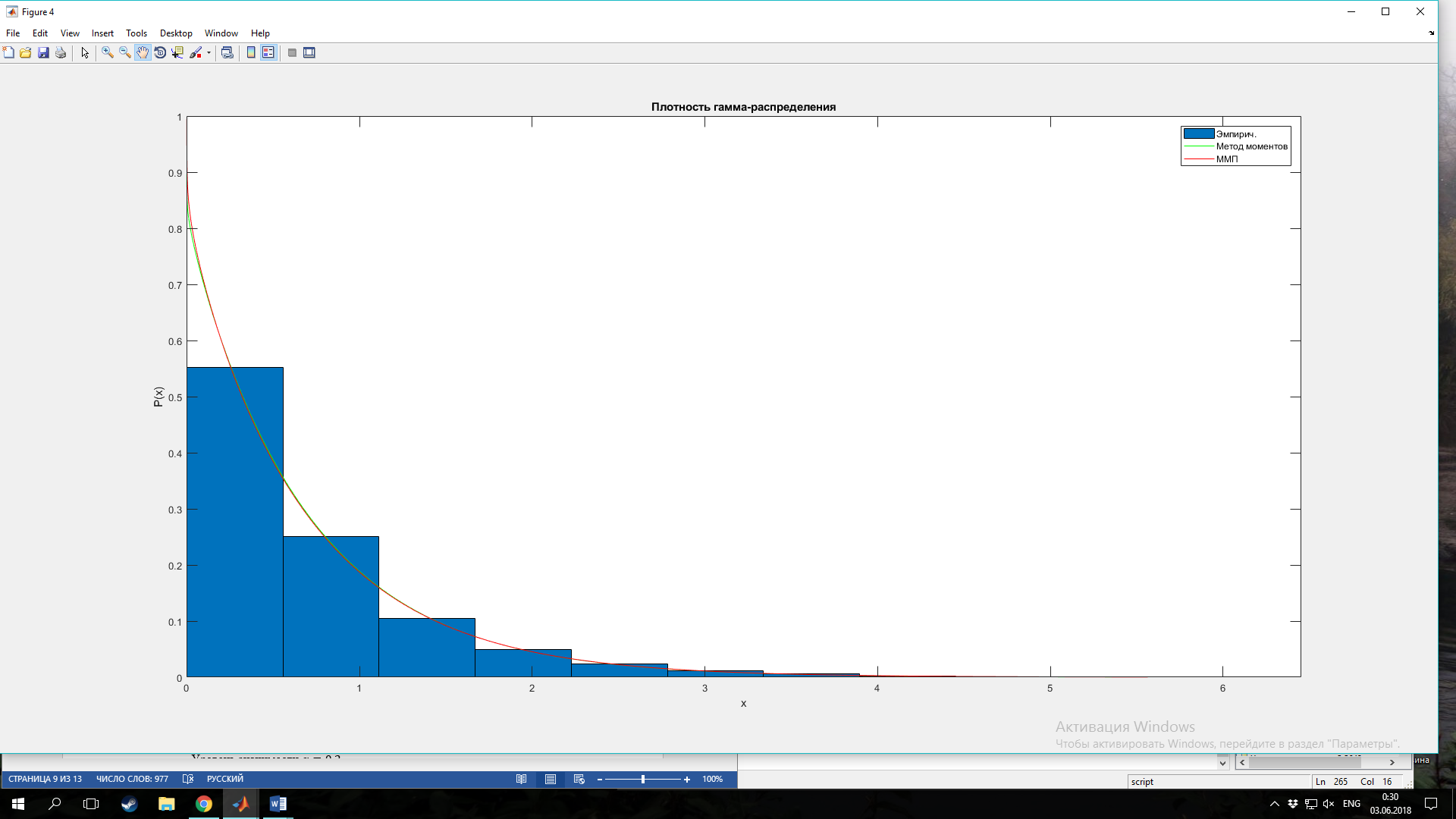
Продифференцируем по и найдем его значение:

Функция дала такой же результат.

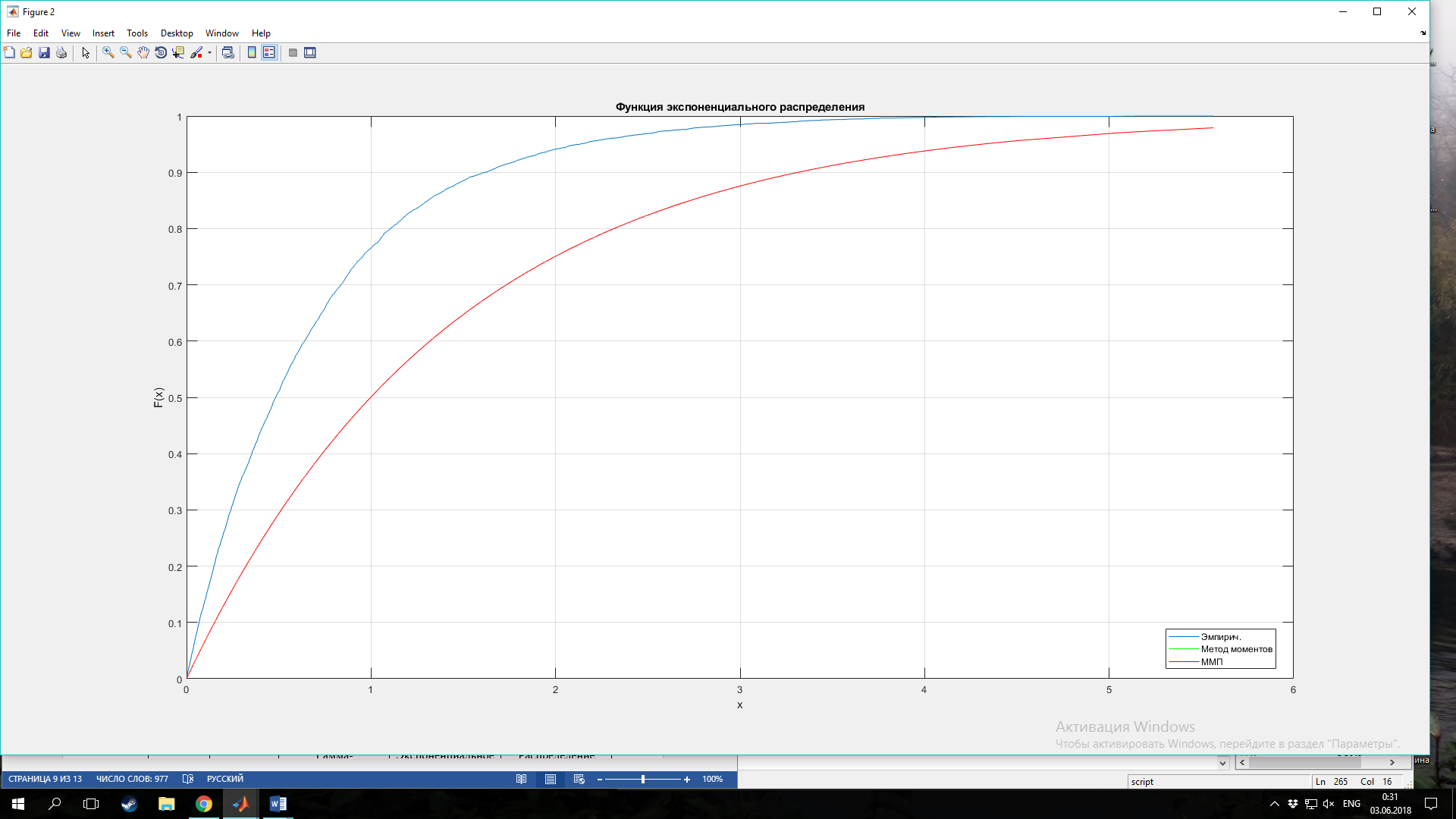
* **Графики сравнения результатов**

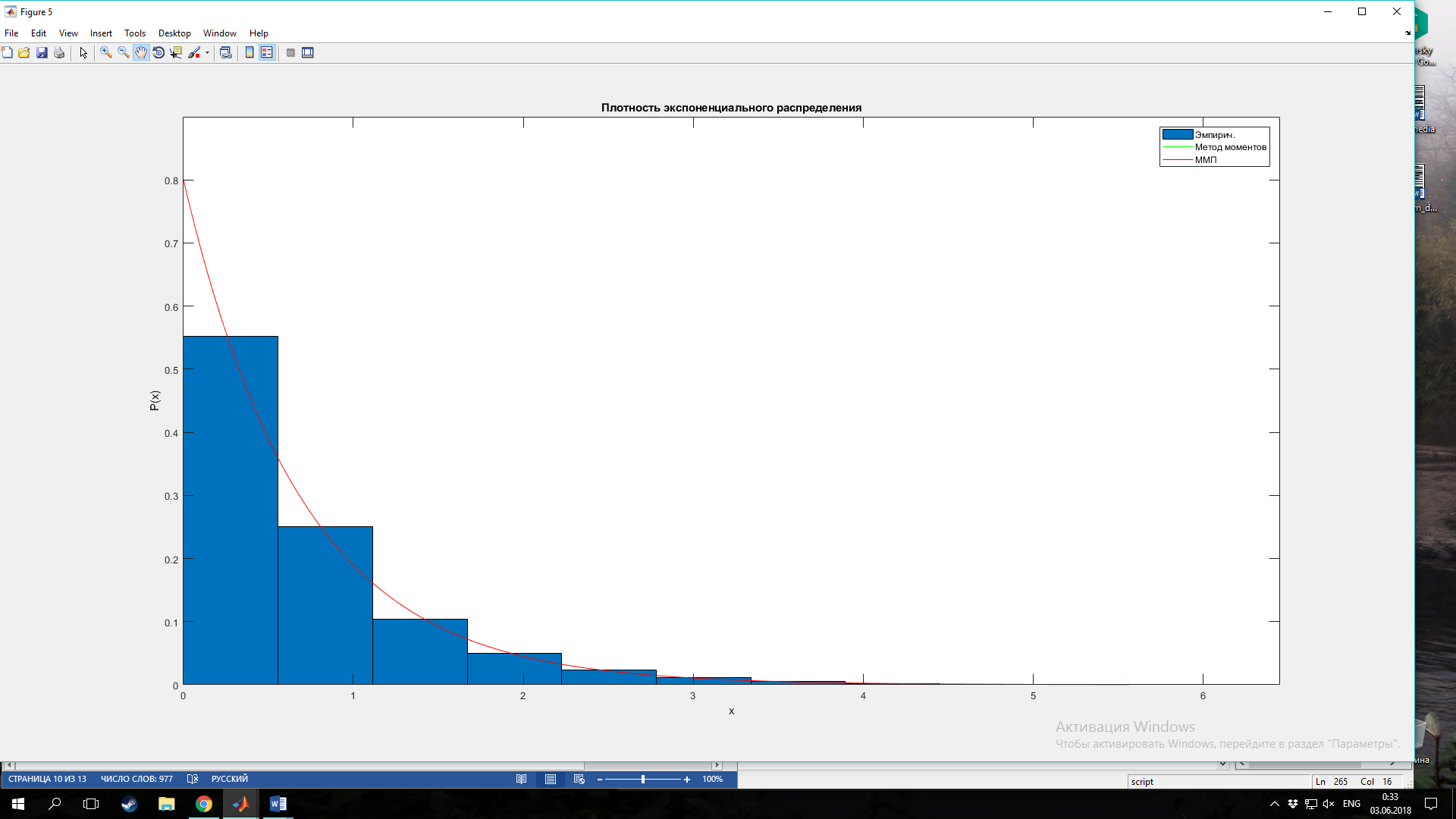
***Гамма-распределение***



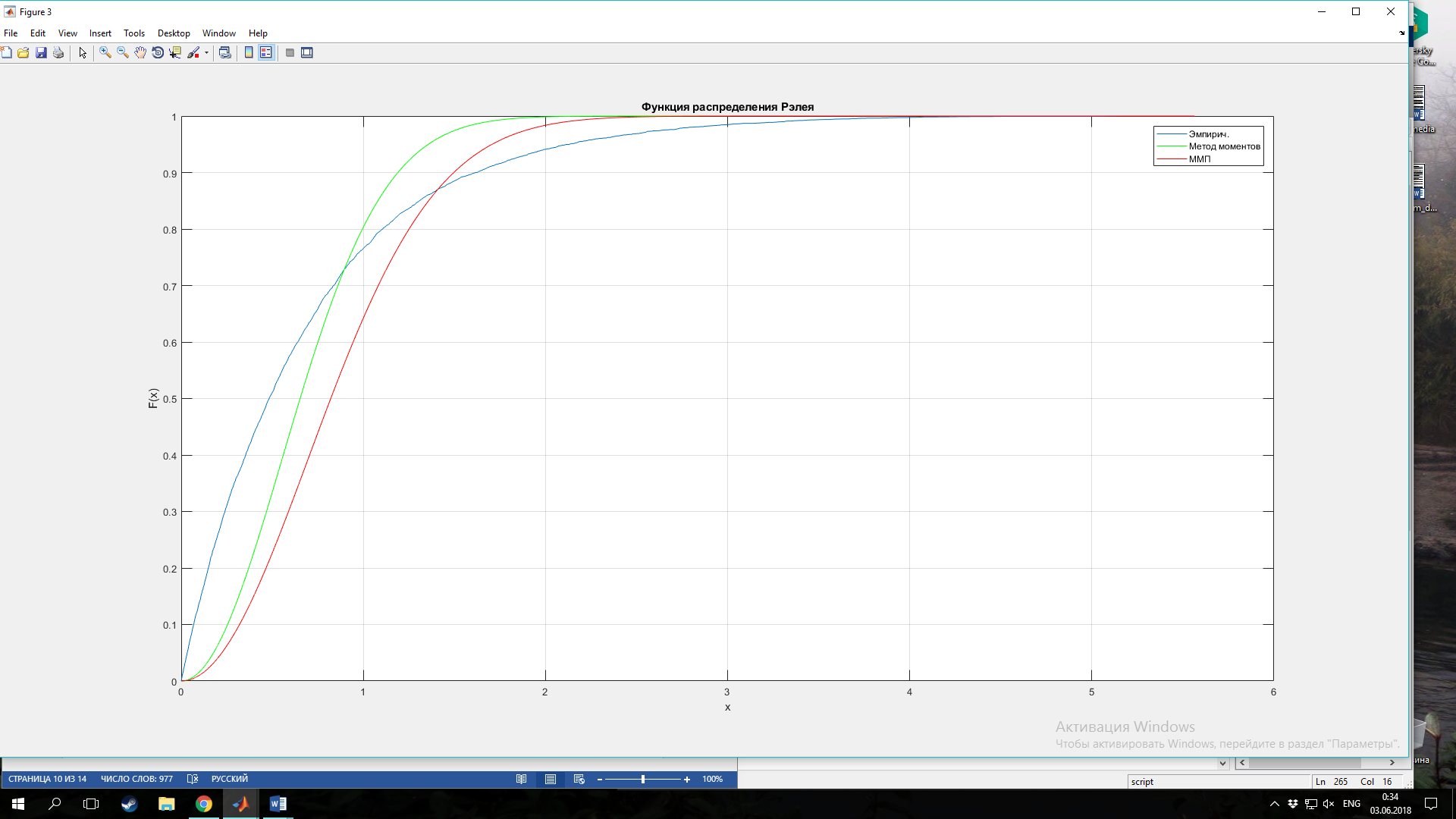


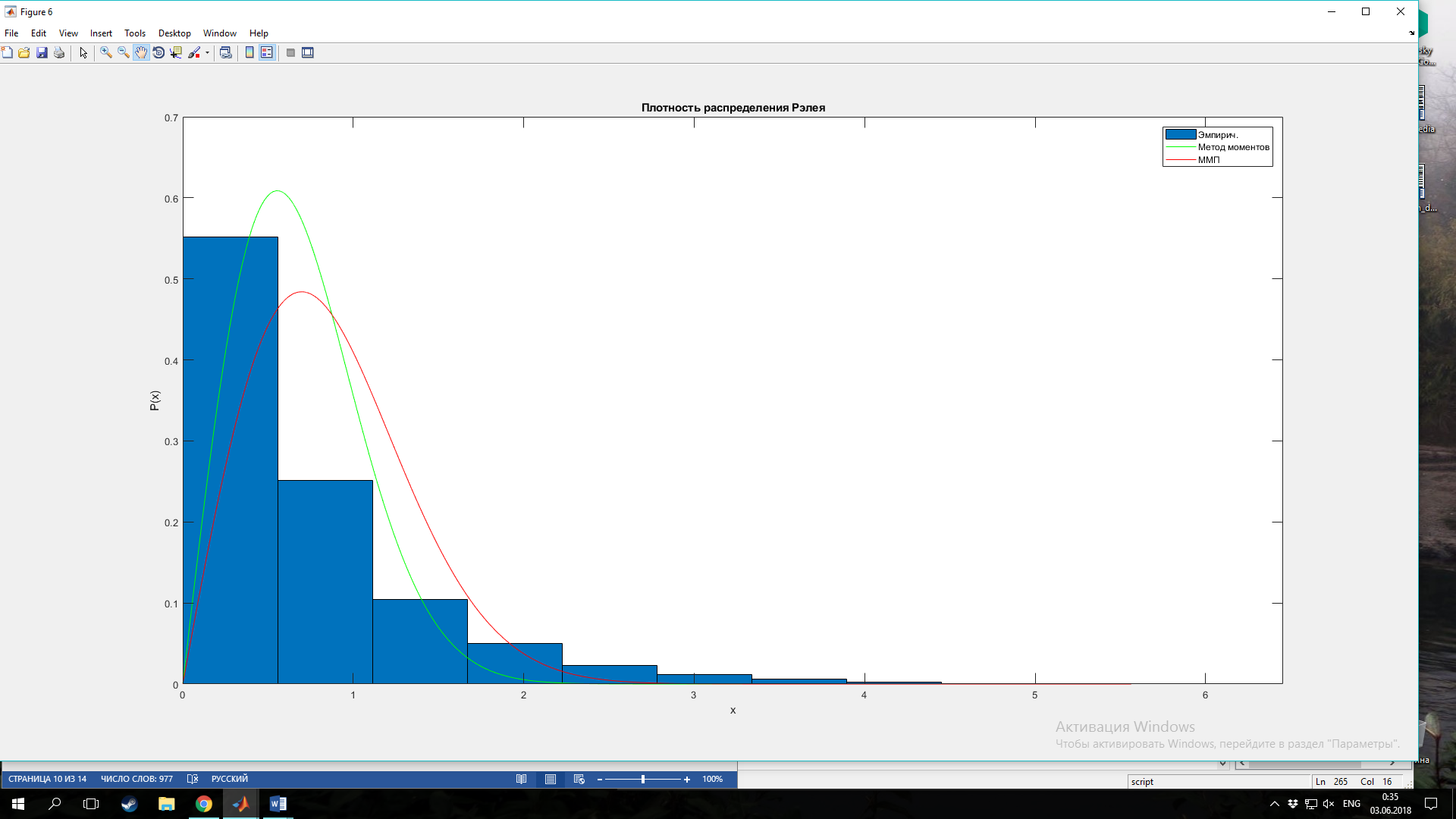
***Экспоненциальное распределение***





***Распределение Рэлея***





На некоторых графиках не видно кривой полученной методом моментов. Это связано с тем, что метод моментов и ММП дали очень близкие значения параметров, и одна кривая находится под другой. Для гамма-распределения и экспоненциального распределения оценки параметров, полученные методом моментов и методом максимального правдоподобия очень близки. Для распределения Рэлея оценки параметров отличаются.

Таблица 3.2.1. Оценки параметров распределения

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Параметры | Гамма-распределение | Экспоненциальное распределение | Распределение Рэлея |
| Метод моментов |  | 0.98583 | 1.43927 | - |
|  | 0.70478 | - | 0.55437 |
| ММП |  | 0.97257 | 1.43927 | - |
|  | 0.71439 | - | 0.69726 |

* 1. **Проверка гипотез**

Уровень значимости

* **Критерий «Хи-квадрат»**

Критическое значение

Статистика

, где

– объем выборки;

– кол-во выборочных значений, попавших в интервал ;

– общее кол-во интервалов, было выбрано 10 интервалов;

– кол-во оцениваемых параметров (1 или 2);

– вероятность попадания в k-ый интервал.

Для гамма-распределения распределения: . Найдем его с помощью функции chi2inv(0.8, 8) в MATLAB.

Для экспоненциального распределения и Рэлея: .

Таблица 3.3.1. Оценки параметров по критерию «Хи-квадрат»

|  |  |  |  |
| --- | --- | --- | --- |
|  | Гамма - распределение | Экспоненциальное распределение | Распределение Рэлея |
| Метод моментов | 2.50326 | 4.87015 | Inf |
| ММП | 2.63769 | 5.00457 | 5811.58916 |

Гипотезы о том, что плотность распределения есть плотность гамма- распределения или экспоненциального распределения с параметрами, полученными методом моментов и ММП подтверждаются, а гипотеза о распределении Рэлея отклоняется.

* **Критерий Колмогорова-Смирнова**

Критическое значение при больших вычисляется по формуле

При получаем

Статистика Колмогорова-Смирнова вычисляется по формуле

Таблица 3.3.2. Оценки параметров по критерию Колмогорова-Смирнова

|  |  |  |  |
| --- | --- | --- | --- |
|  | Гамма - распределение | Экспоненциальное распределение | Распределение Рэлея |
| Метод моментов | 0.00718 | 0.01032 | 0.22214 |
| ММП | 0.00702 | 0.01032 | 0.28866 |

По данному критерию подтверждается гипотеза о гамма-распределении с параметрами, полученными ММП и методом моментов. Гипотезы о распределении Рэлея и экспоненциальном распределении отвергаются.

* **Критерий Мизеса**

Критическое значение при равно

Статистика критерия вычисляется по следующей формуле:

Таблица 3.3.3. Оценки параметров по критерию Мизеса

|  |  |  |  |
| --- | --- | --- | --- |
|  | Гамма - распределение | Экспоненциальное распределение | Распределение Рэлея |
| Метод моментов | 0.06018 | 0.11259 | 107.77521 |
| ММП | 0.05155 | 0.11259 | 214.71919 |

По критерию Мизеса прошли проверку только гипотезы о гамма- распределении и экспоненциальном распределении с параметрами, полученными ММП и методом моментов.

Таблица 3. Итоговые результаты

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Название | Гамма | | | Экспоненциальное | | | Рэлея | | |
| Фор-ла плотности |  | | |  | | |  | | |
|  | Метод моментов | ММП - аналитич | ММП – числ. | Метод моментов | ММП - аналитич | ММП – числ. | Метод моментов | ММП - аналитич | ММП – числ. |
|  | 0.9858 |  |  |  |  |  |  |  |  |
|  | 0.7048 |  |  |  |  |  |  |  |  |
| Хи-квадрат статистика | 2.5033 |  | 2.6377 | 4.8701 | 5.0046 | 5.0046 | Inf | 5811.5892 | 5811.5892 |
| Хи-квадрат критич. знач. |  | | |  | | | | | |
| Хи-квадрат вывод |  |  |  |  | + |  | X | X | X |
| Колм. - Смирн. статистика | 0.0072 |  | 0.00702 | 0.0103 | 0.0103 | 0.0103 | 0.2221 | 0.2887 | 0.2887 |
| Колм. - Смирн. критич. знач. | 0,00732 | | | | | | | | |
| Колм. - Смирн. вывод |  |  |  | X | X | X | X | X | X |
| Мизеса - статистика | 0.0602 |  | 0.0516 | 0.1126 | 0.1126 | 0.1126 | 107.7752 | 214.7192 | 214.7192 |
| Мизеса критич. знач. | 0,2415 | | | | | | | | |
| Мизеса вывод |  |  |  |  | + |  | X | X | X |

1. **Вывод**

В данной лабораторной работе была проанализирована некоторая выборка случайных величин. Были получены точечные и интервальные оценки. С помощью точечных оценок и гистограммы были выдвинуты гипотезы о соответствии выборки различным распределениям. В качестве распределений-гипотез рассматривались следующие три: гамма-распределение, экспоненциальное и распределение Рэлея. Далее гипотезы были проверены с помощью различных критериев.

Во время проверки гипотез на основе критериев «хи-квадрат», Колмогорова-Смирнова и Мизеса статистические значения распределения Рэлея с полученными параметрами во всех случаях не удовлетворяли критическим. Таким образом, распределение Рэлея выбыло из рассмотрения. Экспоненциальное распределение было отвергнуто на этапе проверки критерием Колмогорова-Смирнова. Гамма-распределение всюду удовлетворяло критическим значениям, а, следовательно, подошло для данной выборки.

Таким образом, исследуемый набор случайных величин имеет гамма-распределение с параметрами: