ОТЧЕТ

по лабораторной работе № 3

«Оценка числовых характеристик случайных величин»

по дисциплине

«ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА»

Выполнил студент группы ИС/б-22о

Горбенко К.Н.

Проверил:

Кузнецов С.А.

* 1. ЦЕЛЬ РАБОТЫ

Изучить методы нахождения числовых характеристик случайных величин (с.в.).

Произвести экспериментальные исследования зависимости точности оценок числовых характеристик от объема выборки случайной величины.

* 1. КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

*Неслучайные параметры*, выражающие в сжатой форме наиболее существенные особенности распределения *случайной величины*, называются ее *числовыми характеристиками.* Эти числовые характеристики находятся, как правило, путем осреднения по всему числу испытаний некоторых неслучайных функций исследуемой с.в.

Допустим, что с.в.  в *j*-м испытании приняла конкретное значение,  и полное число этих испытаний есть *N*. Тогда среднее арифметическое величины , обозначаемое как, есть

. (2.1)

Эта величина случайна, однако при  она в силу статистической устойчивости стремится к некоторому пределу, носящему название *математического ожидания* (МО) величины . Оно обозначается как . Для дискретной с.в. оно выражается формулой

, (2.2)

а для непрерывной – формулой

. (2.3)

В формуле (2.2) *n* величин  представляют собой полную совокупность значений, которые может принимать дискретная с.в. , а - вероятности этих значений. В формуле (2.3)  есть плотность вероятности непрерывной с.в..

Строго говоря, не совпадает с , и это совпадение достигается только при . Следовательно, точное значение МО может быть найдено по формулам (2.2) и (2.3) при точном знании или , которые не всегда известны. В то же время экспериментально-расчетным путем по формуле (2.1) может быть найдено только его приближенное значение , которое в связи с этим называется *оценкой* математического ожидания.

Итак, в силу данных выше определений является числовой характеристикой с.в., а - ее приближенной оценкой. Величина определяет некоторую среднюю величину , вокруг которого группируются ее все возможные значения.

Другие числовые характеристики с.в.  находятся путем осреднения некоторых детерминированных функций случайного аргумента . Если число испытаний, конечно, то по аналогии с формулой (2.1) получим оценки таких характеристик в виде

. (2.4)

При  они переходят в МО этих функций:

 (2.5)

Для дискретной с.в.  и

 (2.6)

для непрерывной.

На практике наибольшую применимость имеют *центральные моменты* различных порядков, обозначаемые как . Для них , где порядок *k* – целые неотрицательные числа. Величина , получаемая из каждого значения *исходной* с.в.  вычитанием ее МО, называется *центрированной*, а сама процедура этого вычитания – *центрированием*. Итак, имеем оценку момента *k* – го порядка

 (2.6)

При  отсюда получим

 (2.8)

для дискретной с.в. и

 (2.9)

для непрерывной с.в.

Физическая размерность  и  есть

. (2.10)

Центральный момент *второго порядка*

 (2.11)

называется *дисперсией* с.в. , а квадратный корень из нее *среднеквадратическим отклонением*  с.в. . Величина

 (2.12)

есть *оценка* этой дисперсии, а  - оценка среднеквадратического значения с.в. . Величина характеризует полуширину распределения вероятности или плотности распределения вероятности.

Нетрудно показать, что центральный момент третьего порядка  равен нулю, если распределение симметрично относительно своего МО, и отличен от нуля в противном случае. Однако применять его непосредственно для оценки степени асимметрии распределения неудобно, так как он имеет размерность . Для этого применяют *безразмерную* величину

, (2.13)

называемую *коэффициентом асимметрии* величины . Этот коэффициент характеризует *скошенность* распределения или плотности распределения вероятности. Одновершинное распределение с  имеет левостороннюю (отрицательную) асимметрию, т.е. распределение имеет слева «хвост». Если , оно имеет «хвост» справа. Для симметричного распределения .

Безразмерная величина

. (2.14)

называется *коэффициентом эксцесса* распределения и характеризует степень его *островершинности* в сравнении с нормальным (гауссовским) распределением. Для гауссовского распределения эта величина равна нулю. Для более островершинного распределения . Для менее островершинного . При этом сравнении необходимо считать, что у всех рассматриваемых распределений величина одинакова.

Cтатистический пакет Statistics&Machine Learning Toolbox системы MATLAB поддерживает 20 видов распределений вероятности: 14 непрерывных и 6 дискретных (таблица 2.1).

В этой таблице: A, B, MU, NU, NU1, NU2, V1, V2, DELTA, LAMBDA, NN, M, K, P, RR – параметры, описывающие распределения; R – матрица размером m×n, состоящая из случайных величин , имеющих указанное распределение; M –математическое ожидание  и V – дисперсия с.в. . Команда из столбца IV дает возможность вычислить теоретическое значение МО , обозначаемое здесь как M, и теоретическую дисперсию , обозначаемую как V.

Входящий в MATLAB пакет Statistics Toolbox имеет в своем составе демонстрационные программы, создающие интерактивную среду для генерации случайных чисел, изучения их различных распределений вероятностей и других целей.

* 1. ЭКСПЕРИМЕНТАЛЬНЫЙ РАСЧЕТ ОЦЕНОК ЧИСЛОВЫХ ХАРАКТЕРИСТИК

Для варианта № 4 задано экспоненциальное распределение, MU = 3.

Для оценки числовых характеристик был разработан MATLAB скрипт, использующий библиотеку StatisticsToolbox:

m = 1;

n = 1000;

MU = 3;

ME = zeros(1, n);

CM1 = zeros(1, n);

CM2 = zeros(1, n);

CM3 = zeros(1, n);

CM4 = zeros(1, n);

D = zeros(1, n);

AK = zeros(1, n);

EK = zeros(1, n);

R = exprnd(3, m, n);

for i=1:n

ME(i) = MathematicalExpectation(R, i);

CM1(i) = CentralMoment(R, ME(i), i, 1);

CM2(i) = CentralMoment(R, ME(i), i, 2);

CM3(i) = CentralMoment(R, ME(i), i, 3);

CM4(i) = CentralMoment(R, ME(i), i, 4);

D(i) = sqrt(CM2(i));

AK(i) = AssymetryKoefficient(CM3(i), D(i));

EK(i) = ExcessKoefficient(CM4(i), D(i));

end

figure(1);

subplot(8, 1, 1);

plot(ME(1:n));

xlabel('N');

ylabel('ME');

subplot(8, 1, 2);

plot(CM1(1:n));

xlabel('N');

ylabel('CM1');

subplot(8, 1, 3);

plot(CM2(1:n));

xlabel('N');

ylabel('CM2');

subplot(8, 1, 4);

plot(CM3(1:n));

xlabel('N');

ylabel('CM3');

subplot(8, 1, 5);

plot(CM4(1:n));

xlabel('N');

ylabel('CM4');

subplot(8, 1, 6);

plot(D(1:n));

xlabel('N');

ylabel('D');

subplot(8, 1, 7);

plot(AK(1:n));

xlabel('N');

ylabel('AK');

subplot(8, 1, 8);

plot(EK(1:n));

xlabel('N');

ylabel('EK');

figure(2);

subplot(8, 1, 1);

semilogx(ME(1:n));

xlabel('N');

ylabel('ME');

subplot(8, 1, 2);

semilogx(CM1(1:n));

xlabel('N');

ylabel('CM1');

subplot(8, 1, 3);

semilogx(CM2(1:n));

xlabel('N');

ylabel('CM2');

subplot(8, 1, 4);

semilogx(CM3(1:n));

xlabel('N');

ylabel('CM3');

subplot(8, 1, 5);

semilogx(CM4(1:n));

xlabel('N');

ylabel('CM4');

subplot(8, 1, 6);

plot(CM4(1:n));

xlabel('N');

ylabel('D');

subplot(8, 1, 7);

semilogx(AK(1:n));

xlabel('N');

ylabel('AK');

subplot(8, 1, 8);

semilogx(EK(1:n));

xlabel('N');

ylabel('EK');

Функции, использующиеся в скрипте:

function y = MathematicalExpectation (R, n)

y = sum(R(1:n)) / n;

end

function y = CentralMoment(R, M, n, k)

temp = (R(1:n) - M).^k;

y = sum(temp) / n;

end

function y = AssymetryKoefficient(CM3, D)

y = CM3 / D^3;

end

function y = ExcessKoefficient (CM4, D)

y = CM4 / D^4 - 3;

end

Графики зависимости числовых характеристик от количества испытаний *N* изображены на рисунках № 1 и № 2:

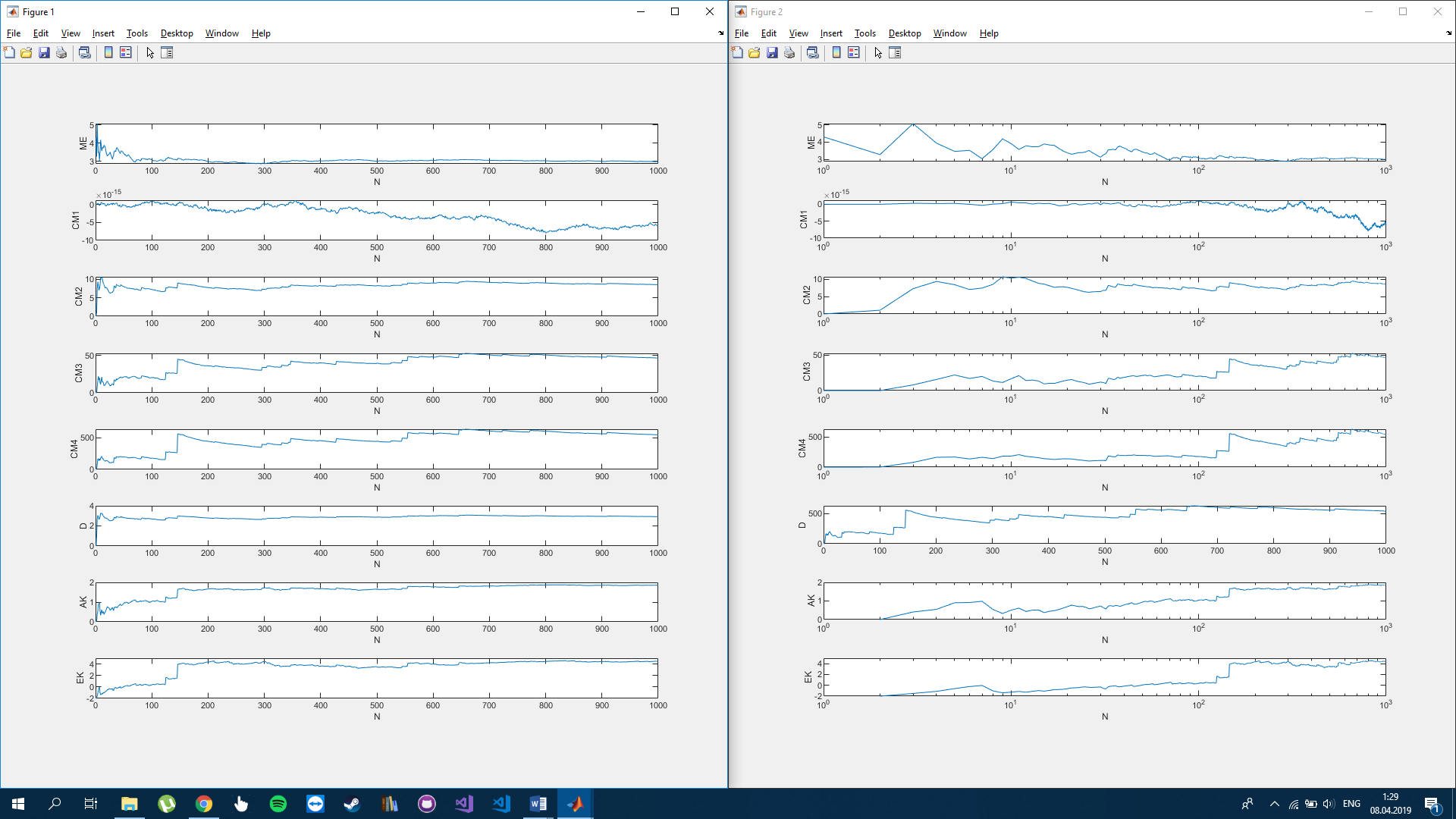


Рисунок № 1 – Графики зависимости числовых характеристик от количества испытаний N,

где ME – математическое ожидание;

СМ1 – центральный момент первого порядка;

СМ2 – центральный момент первого порядка (дисперсия);

СМ3 – центральный момент третьего порядка;

СМ4 – центральный момент четвертого порядка;

D – среднеквадратичное значение дисперсии;

АК – коэффициент ассиметрии;

ЕК – коэффициент эксцесса.

График зависимости числовых характеристик от количества испытаний *N* в полулогарифмическом масштабе изображены на рисунке № 2:

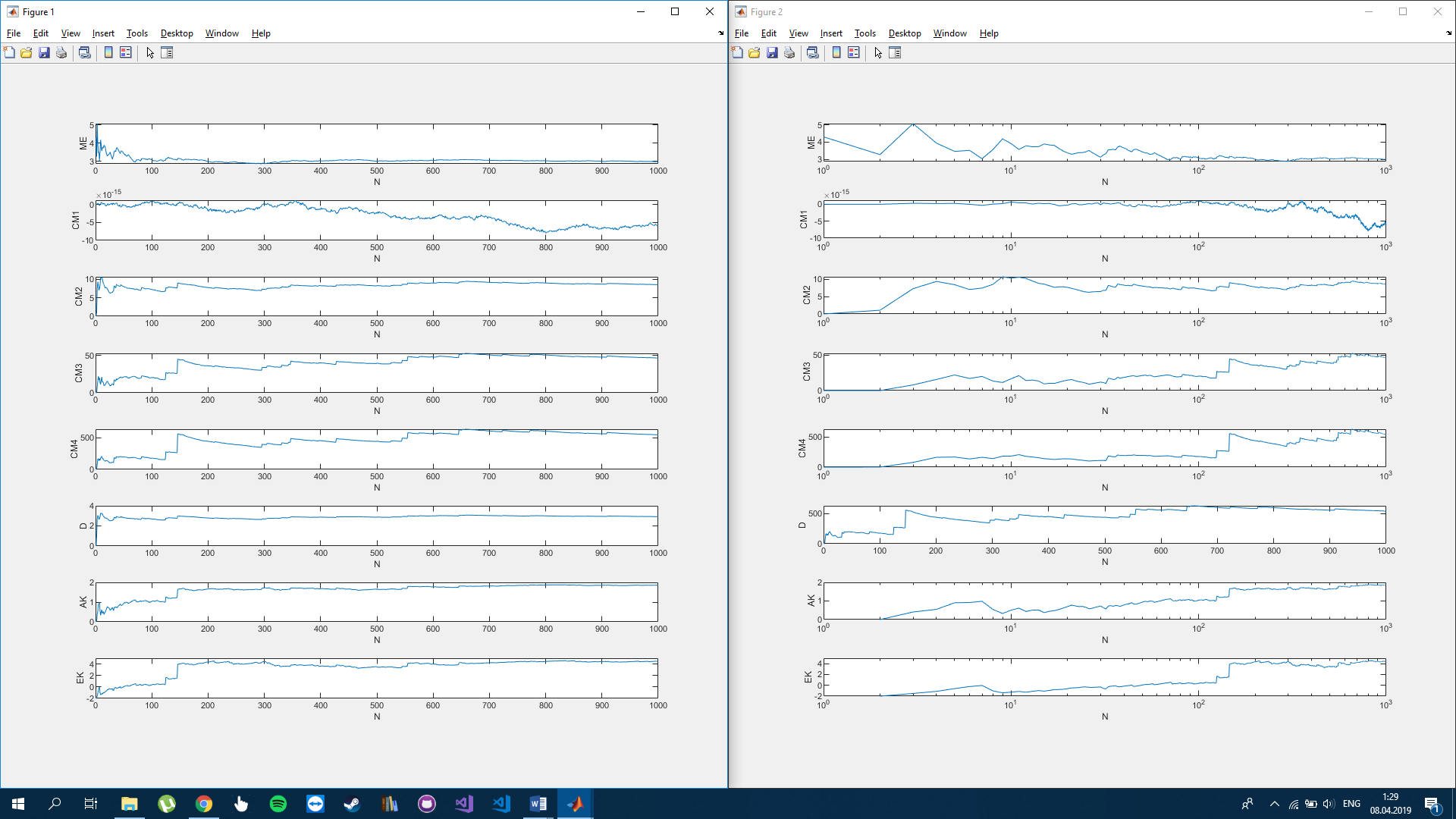


Рисунок № 1 – Графики зависимости числовых характеристик от количества испытаний N в полулогарифмическом масштабе,

где ME – математическое ожидание;

СМ1 – центральный момент первого порядка;

СМ2 – центральный момент первого порядка (дисперсия);

СМ3 – центральный момент третьего порядка;

СМ4 – центральный момент четвертого порядка;

D – среднеквадратичное значение дисперсии;

АК – коэффициент ассиметрии;

ЕК – коэффициент эксцесса.

Используя оператор disttool определим вид теоретической кривой для заданного распределения:

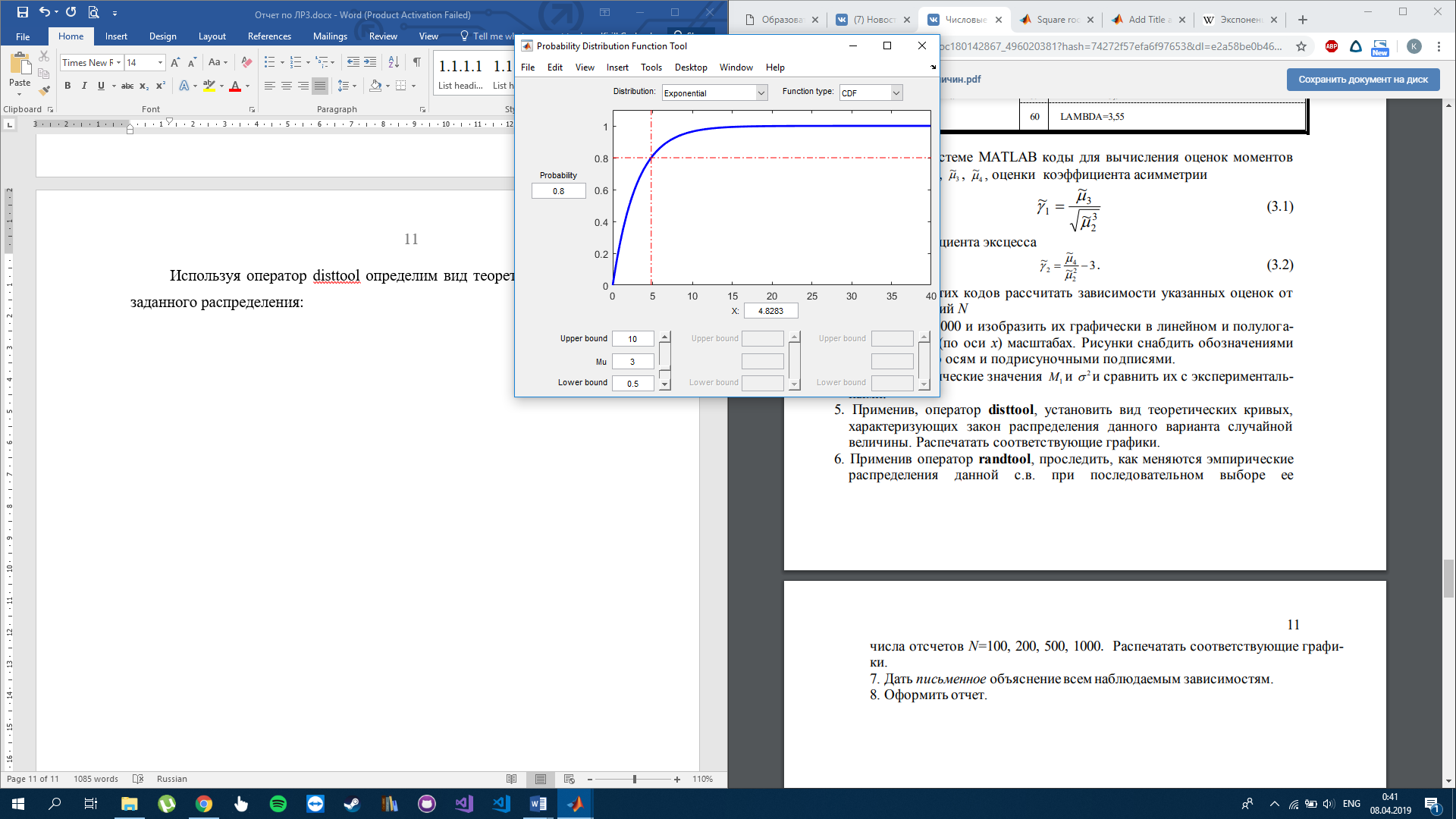


Рисунок № 3 – Теоретическая кривая, характеризующая экспоненциальное распределение

Используя оператор randtool построим экспоненциальные распределения для различного числа испытаний:

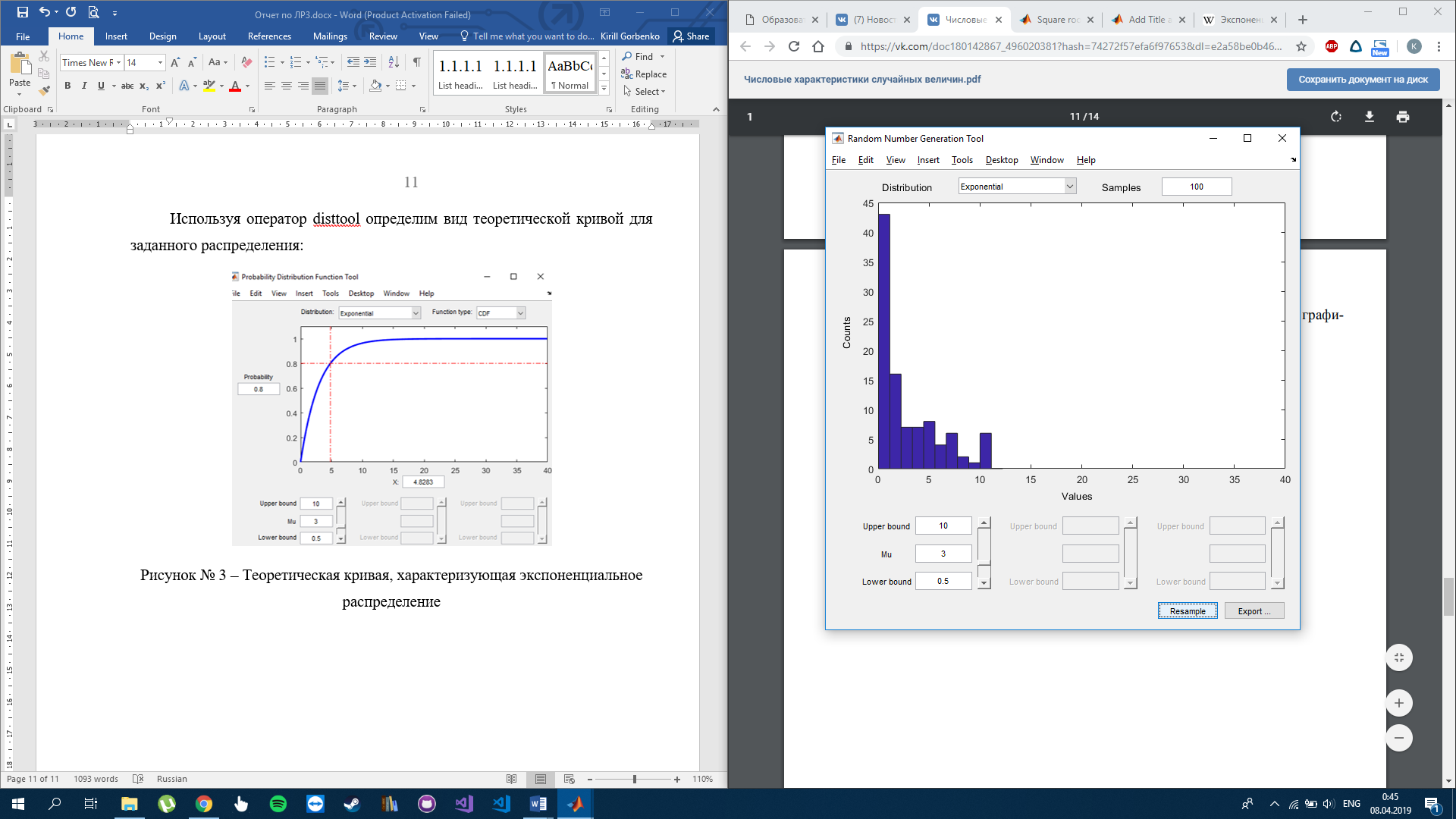


Рисунок № 4 – Эмпирическое распределение с.в. при N = 100

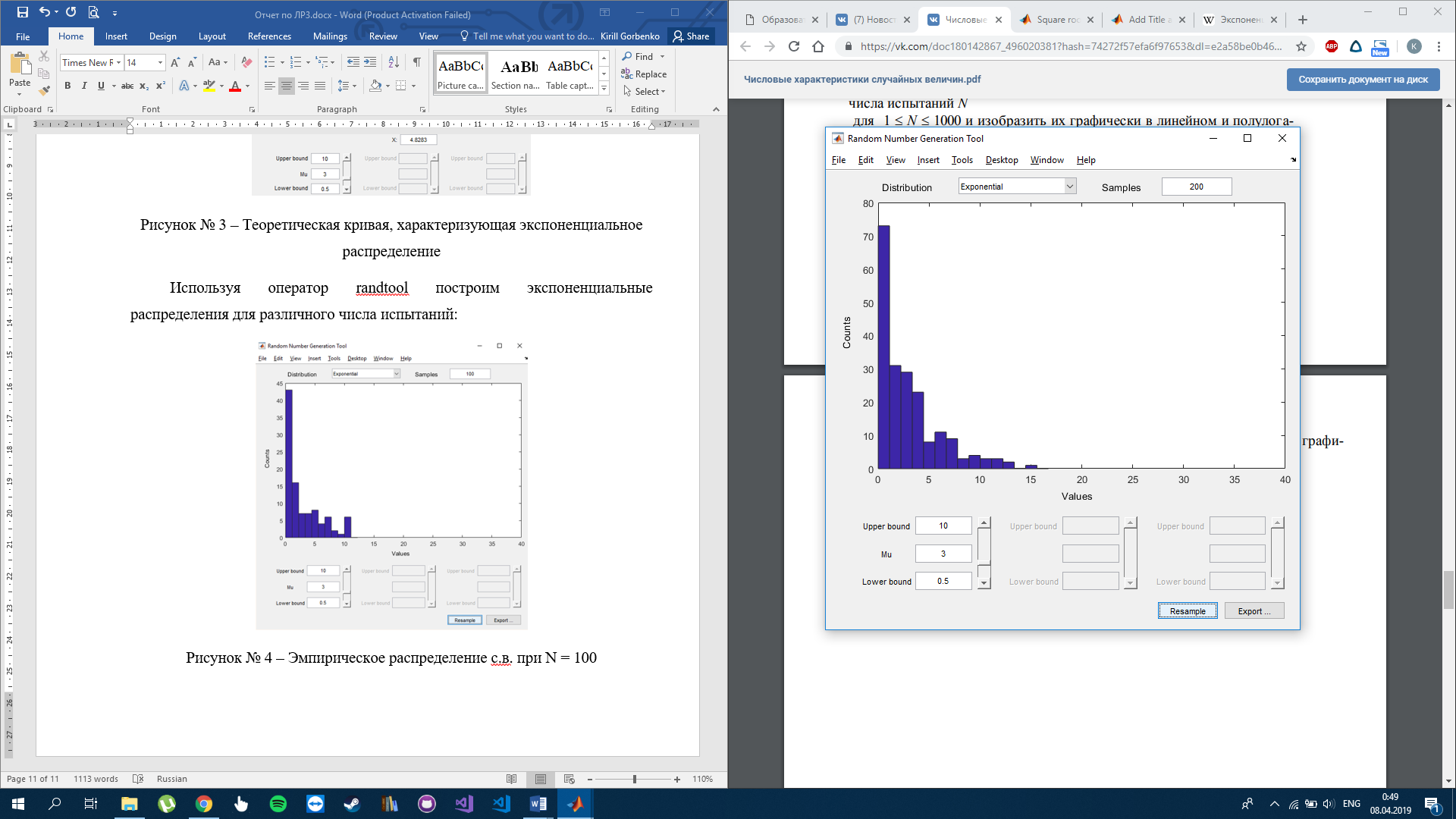


Рисунок № 5 – Эмпирическое распределение с.в. при N = 200

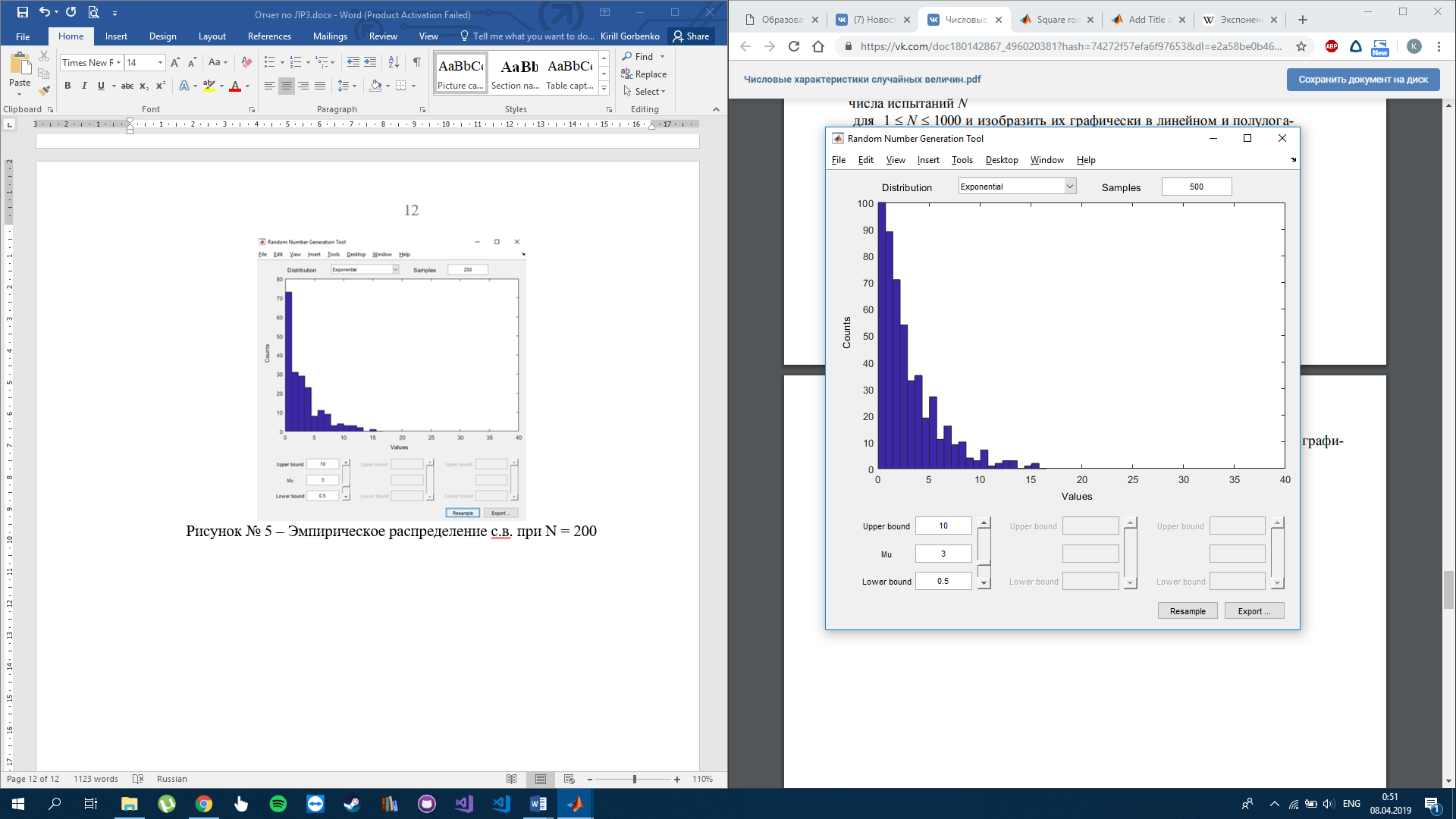


Рисунок № 6 – Эмпирическое распределение с.в. при N = 500

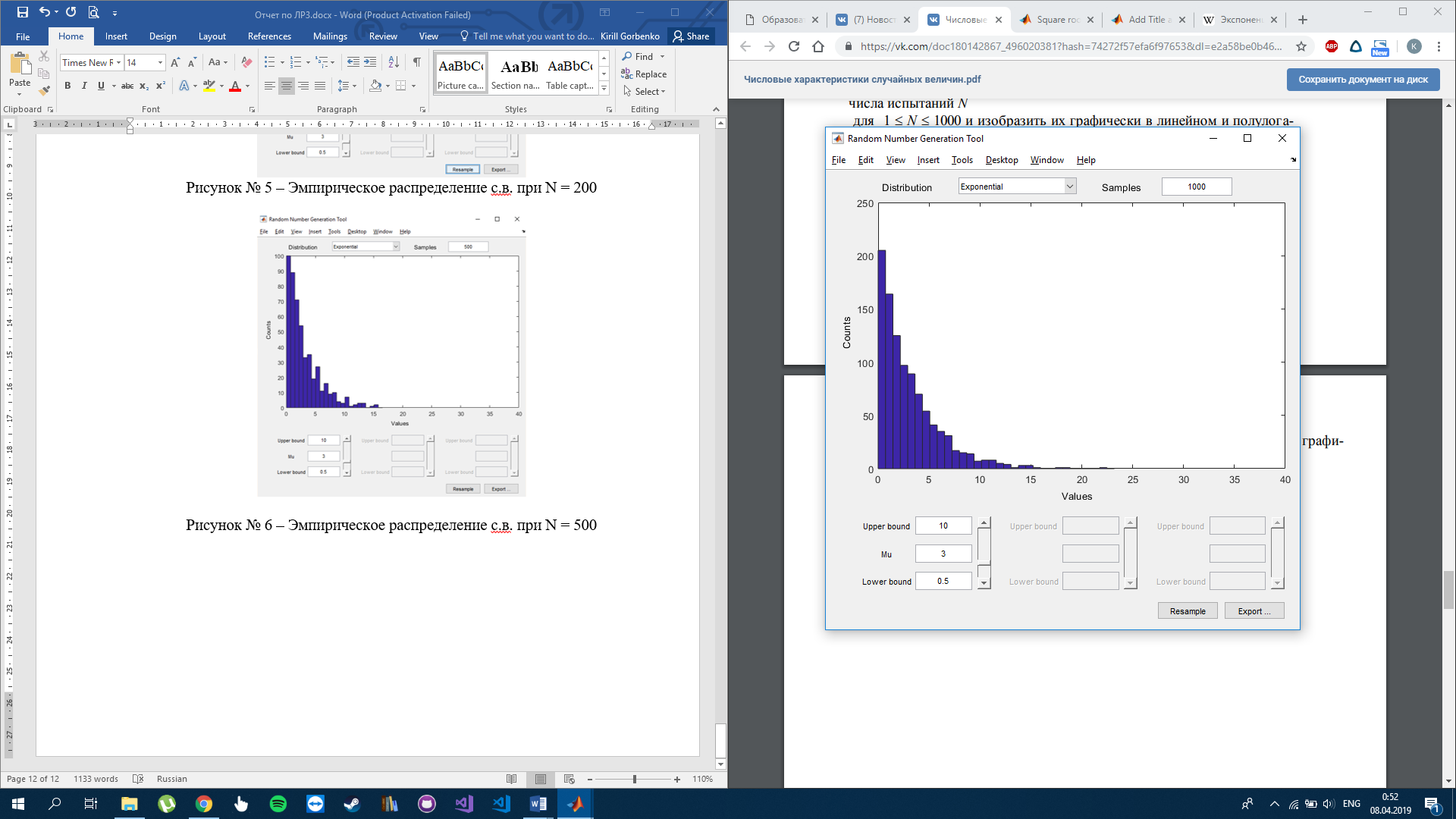


Рисунок № 7 – Эмпирическое распределение с.в. при N = 1000

* 1. ВЫВОД

В ходе лабораторной работы были оценены следующие числовые характеристики с.в., заданной экспоненциальным распределением: математическое ожидание, центральные момент до 4 порядка, коэффициенты асимметрии и эксцесса. Все перечисленные оценки характеристик стремятся к соответствующему действительному значению числовой характеристики при увеличении числа испытаний, что позволяет использовать эти оценки вместо теоретических значений при достаточном количестве испытаний.

Судя по высокому значению дисперсии, экспоненциальное распределение характеризуется большим разбросом значений случайной величины. Так как коэффициент асимметрии > 0, то распределение асимметрично (смещено вправо). Также, поскольку коэффициент эксцесса > 0, распределение более островершинно, чем распределение Гаусса.