МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования «Самарский национальный исследовательский университет   
имени академика С.П. Королева»

(Самарский университет)

Институт информатики, математики и электроники

Факультет информатики

Кафедра программных систем

Дисциплина методы планирования эксперимента и статистической обработки информации

**ОТЧЁТ**

по лабораторной работе №2 по дисциплине «Методы планирования эксперимента и статистической обработки информации»

по теме «Расчет доверительной оценки»

Студент группы № В.А. Артамонов

Проверил В. В. Любимов

Дата сдачи:

Оценка:

Самара  
2022

Постановка задачи

Задание к лабораторной работе:

Провести расчёт доверительной оценки опытных данных с помощью программы MicrosoftOfficeExcel.

Теоретические сведения

Интервальная оценка, доверительный уровень которой равен 95%, интерпретируется следующим образом: если из генеральной совокупности извлечь все выборки, имеющие объем n, и вычислить их выборочные средние, то 95% доверительных интервалов, построенных на их основе, будут содержать математическое ожидание генеральной совокупности, а 5% − нет. На практике, как правило, из генеральной совокупности извлекается только одна выборка, а математическое ожидание генеральной совокупности μ не известно. По этой причине невозможно гарантировать, что некий конкретный доверительный интервал содержит величину μ. Можно лишь утверждать, что вероятность этого события равна 95%.

Как правило, доверительный уровень обозначают следующим образом: , где величина представляет собой площадь, ограниченную хвостом распределения, выходящим за пределы доверительного интервала. Величину α называют уровнем значимости доверительного интервала. Уровень значимости − это вероятность отклонить нулевую гипотезу, когда на самом деле она верна. Кроме того, в качестве синонима для доверительного уровня иногда употребляется выражение «доверительная вероятность». Площади, ограниченные как левым, так и правым хвостами распределения, выходящими за пределы доверительного интервала, равны α/2.



На практике как математическое значение генеральной совокупности, так и его дисперсия часто бывают неизвестными. Следовательно, необходимо построить доверительный интервал, содержащий математическое значение генеральной совокупности, используя лишь выборочные статистики и S.



* 1. Доверительный интервал для математического ожидания нормальной выборки

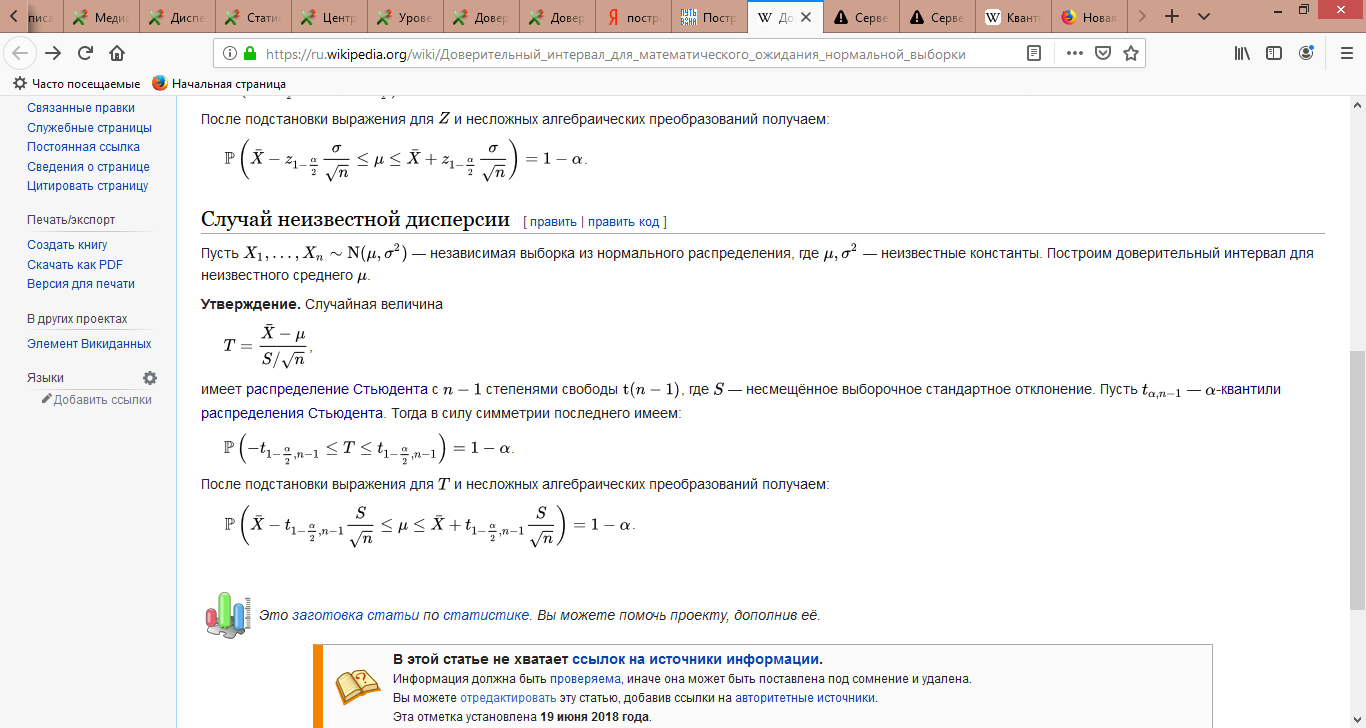
Пусть − независимая выборка из нормального распределения, где −неизвестные константы. Построим доверительный интервал для неизвестного среднего .



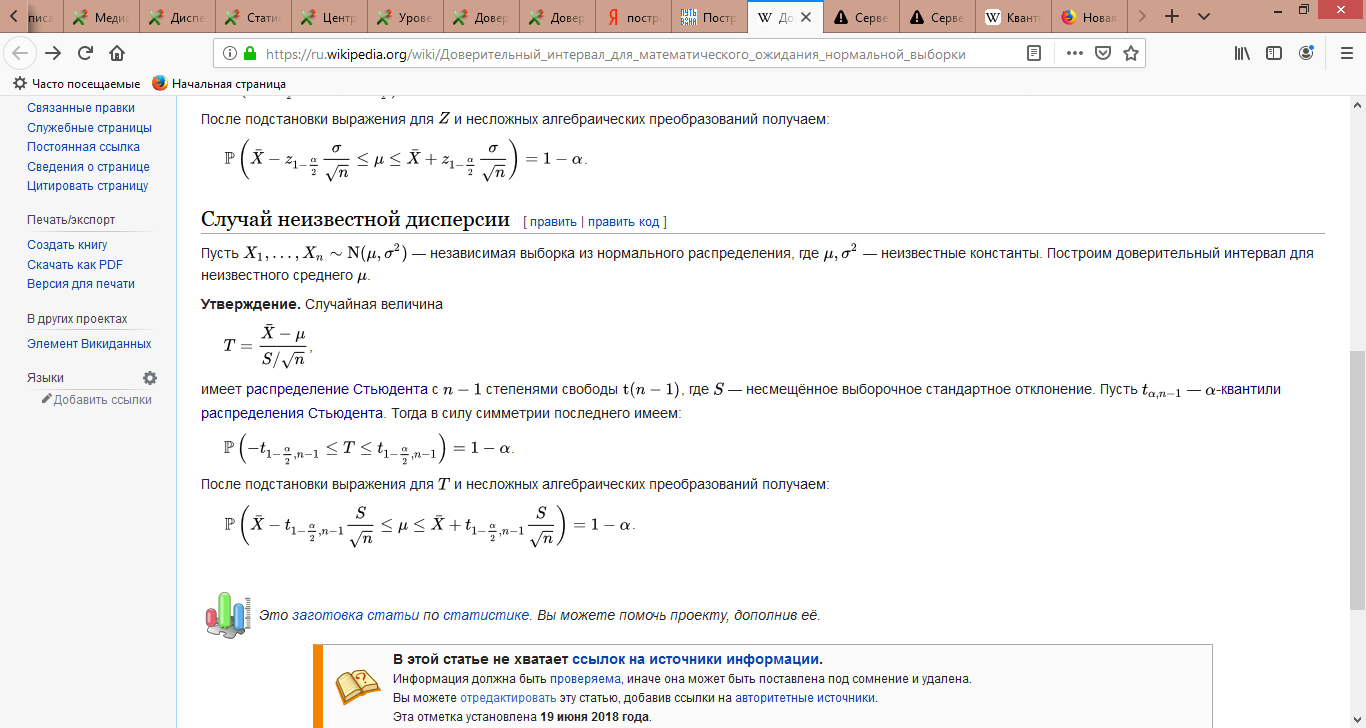
Утверждение: случайная величина



имеет распределение Стьюдента с n-1 степенями свободы, где S−несмещенное выборочное стандартное отклонение. Пусть −α-квантили распределения Стьюдента. Тогда в силу симметрии последнего имеем:



После подстановки выражения для Tполучаем:



* 1. Нормальное распределение

Нормальное распределение, также называемое распределением Гаусса или Гаусса − Лапласа − распределение вероятностей, которое в одномерном случае задаётся функцией плотности вероятности, совпадающей с функцией Гаусса:

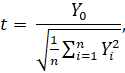


где параметр μ − математическое ожидание (среднее значение), медиана и мода распределения, а параметр σ − среднеквадратическое отклонение распределения.

Стандартным нормальным распределением называется нормальное распределение с математическим ожиданием μ = 0 и стандартным отклонением σ = 1.

* 1. Распределение Стьюдента

Пусть − независимые стандартные нормальные случайные величины, такие что . Тогда распределение случайной величины где



называется распределением Стьюдента с степенями свободы.



Квантиль в математической статистике − значение, которое заданная случайная величина не превышает с фиксированной вероятностью. Если вероятность задана в процентах, то квантиль называется процентилем или перцентилем.

Этапы выполнения лабораторной работы

1. Сгенерировать таблицу случайных чисел размерности 1000 разными законами распределения случайных величин.:

* Нормальное распределение между 0 и 1;

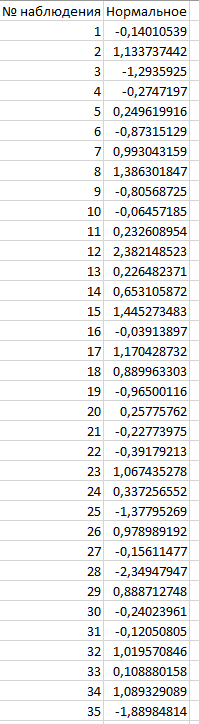


Рисунок 1 – Исходные данные лабораторной работы

На рисунке 1 показана часть данных выборок, которые использовались для выполнения лабораторной работы.

1. Произвести описательную статистику сгенирированного закона распределения.

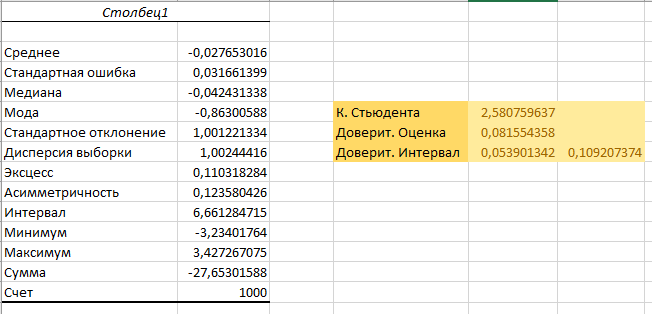


Рисунок 2 - Однофакторный дисперсионный анализ сгенерированного закона распределения

1. Сделать выводы и написать отчёт.

Специальное название стандартной ошибки обусловлено стремлением подчеркнуть, что она показывает величину неопределенности выборочного среднего. Стандартная ошибка оценивает насколько выборочное среднее отличается от истинного среднего значения исходного распределения. А термин стандартное отклонение используют для обозначения величины изменчивости отдельных элементов выборки от среднего.

Медиана выборки − число, характеризующее выборку (например, набор чисел). Если все элементы выборки различны, то медиана − это такое число выборки, что ровно половина из элементов выборки больше него, а другая половина меньше него.

Мода − значение во множестве наблюдений, которое встречается наиболее часто.

Коэффициент Стьюдента (или квантиль α/2) вычисляется с помощью статистической функции «СТЬЮДЕНТ.ОБР.2Х» с параметрами (уровень значимости, объем выборки − 1).

Доверительная оценка вычисляется с помощью статистической функции «ДОВЕРИТ.НОРМ» с параметрами (уровень значимости, стандартное отклонение, объем выборки).