****МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ

(национальный исследовательский университет)»

Институт № 3 «Системы управления, информатика и электроэнергетика»

Кафедра 311 «Прикладные программные средства и математические методы»

**КУРСОВАЯ РАБОТА**

**Дисциплина: «Алгоритмы и обработка данных»**

Выполнила:

Студентка гр. М3О-216Б-22

Хутиева Эрика Арсеновна

Научный руководитель:

Проф. Агамиров Л.В.

Оценка КР(КП): **\_\_\_\_\_\_\_\_\_\_\_\_\_**

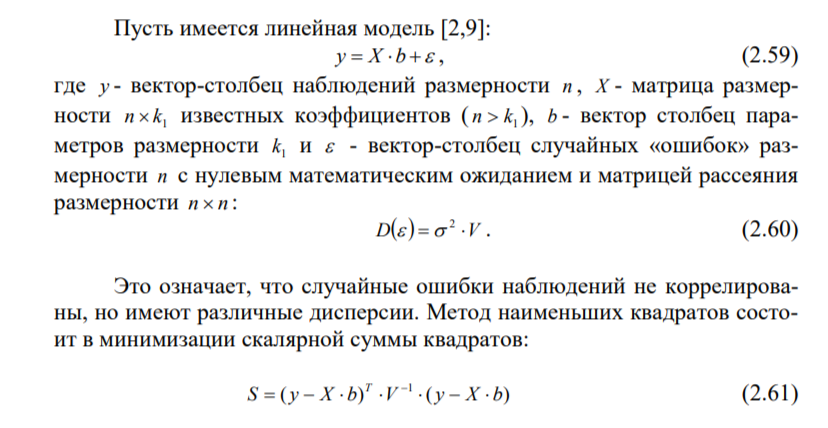
Москва 2023г.

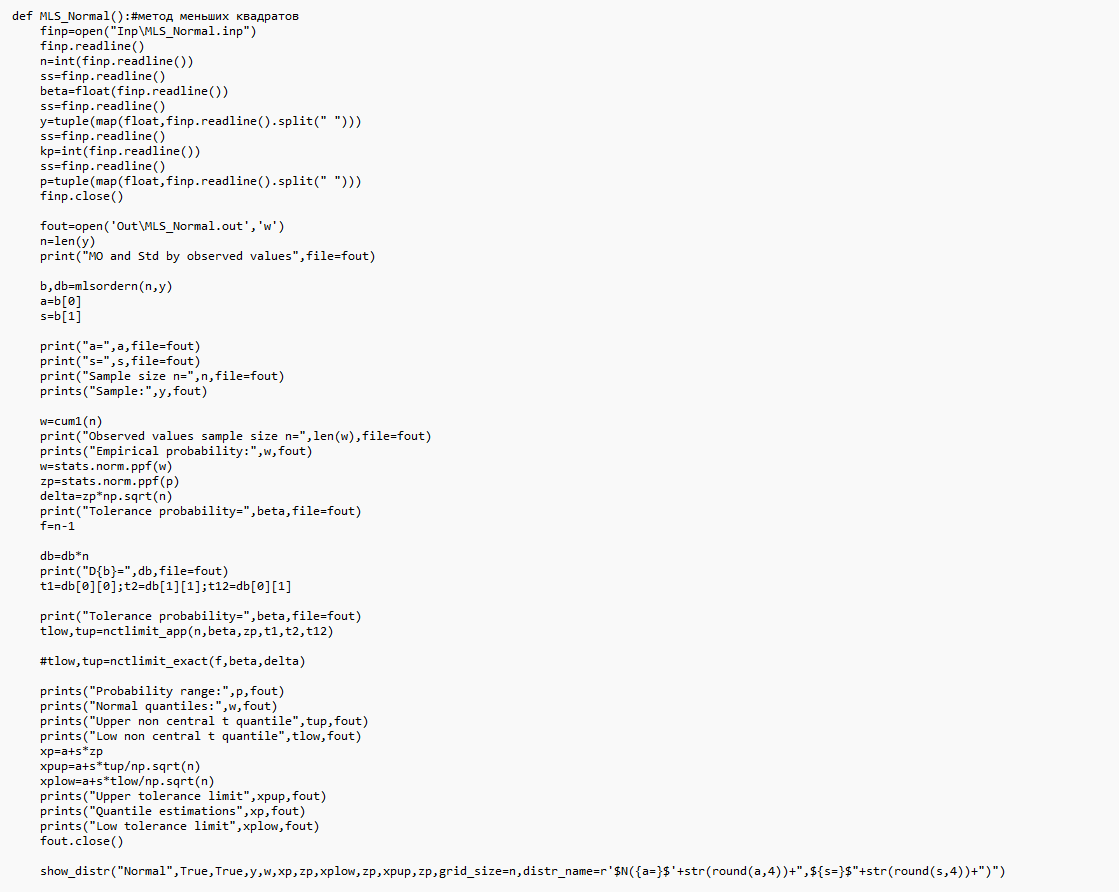
Оглавление

1. [Метод наименьших квадратов 3](#_Toc154063713)
2. [Определение квантиля и дисперсии 5](#_Toc154063714)
3. [Планирование дисперсии 6](#_Toc154063715)
4. [Метод максимального правдоподобия 8](#_Toc154063716)
5. [Распределение Вейбула-Гнеденко 10](#_Toc154063717)
6. [Критерий Граббса. 11](#_Toc154063718)
7. [Критерий Фишера (F-критерий) – об однородности двух дисперсий 13](#_Toc154063719)
8. [Критерий Стьюдента (t-критерий) 15](#_Toc154063720)
9. [Критерий Бартлета 17](#_Toc154063721)
10. [Однофакторный дисперсионный анализ 18](#_Toc154063722)
11. [Критерий Шапиро-Уилка 20](#_Toc154063723)
12. [Критерий Смирнова 22](#_Toc154063724)
13. [Критерий Андерсона-Дарлинга 23](#_Toc154063725)
14. [Критерий 24](#_Toc154063726)
15. [Критерий знаков для медианы 25](#_Toc154063727)
16. [Двухвыборочный критерий Уилкоксона 26](#_Toc154063728)
17. [Критерий Колмогорова-Смирнова 27](#_Toc154063729)
18. [Критерий Краскела-Уоллиса 28](#_Toc154063730)

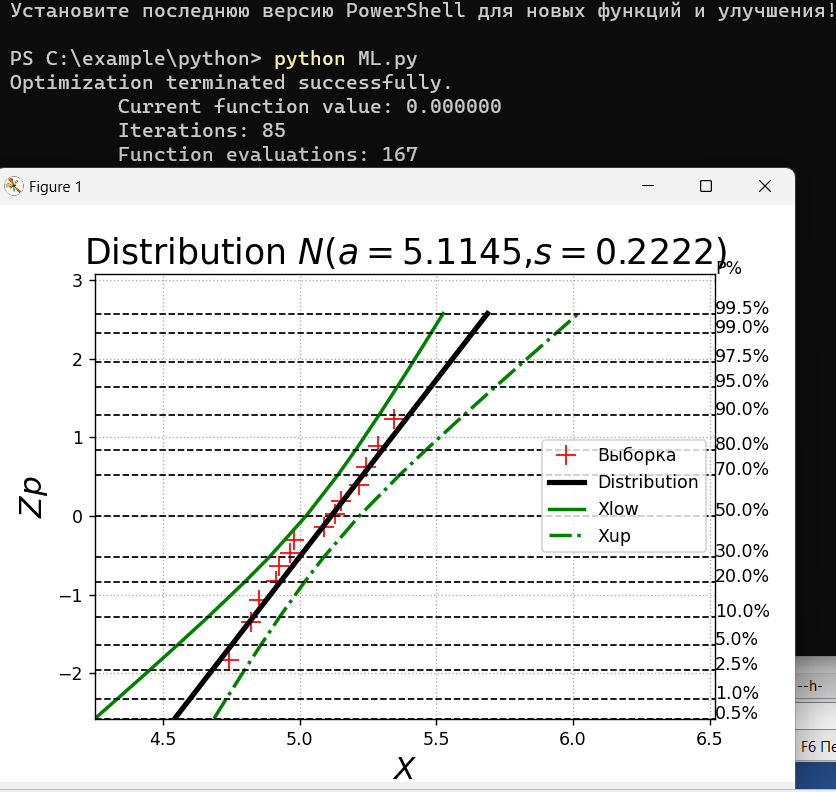
**Все программы вы можете найти по ссылке https://github.com/erikahutieva/Python/tree/main/matstat**

## Метод наименьших квадратов

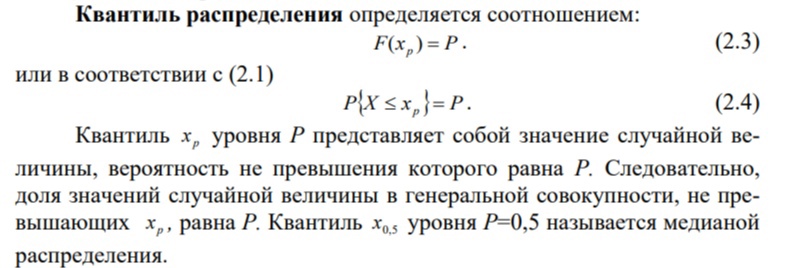
****

****

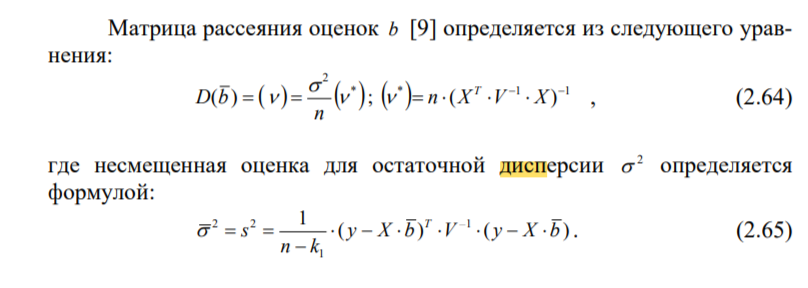
Функция не принимает ничего на вход, но читает данные или файла INP, а выводит верхние и нижние пределы, а также оценки квантиля графическое представление выборочных данных и полученного нормального распределения.

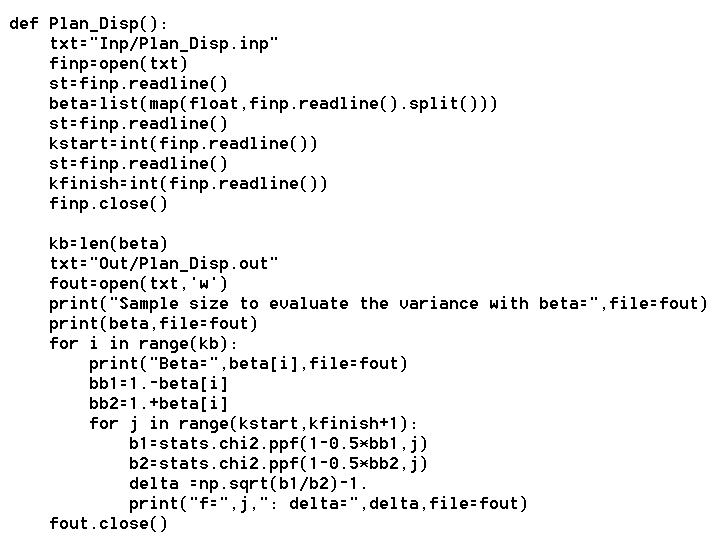


## Определение квантиля и дисперсии

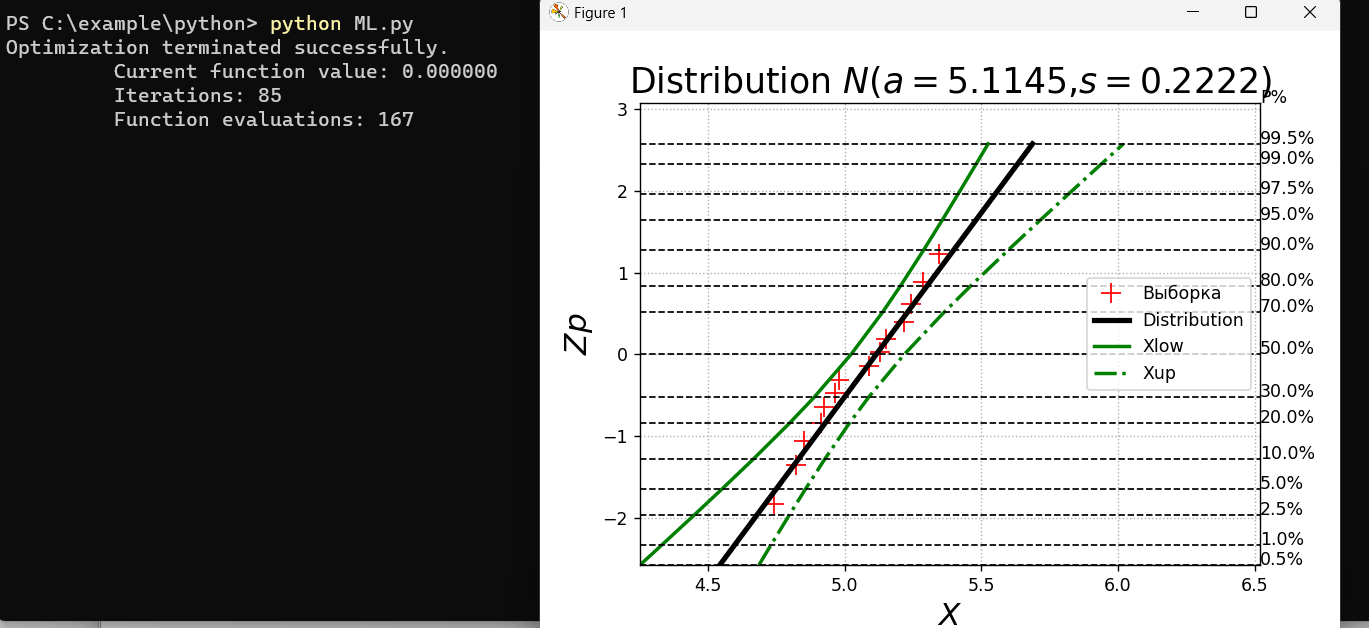


## Планирование дисперсии

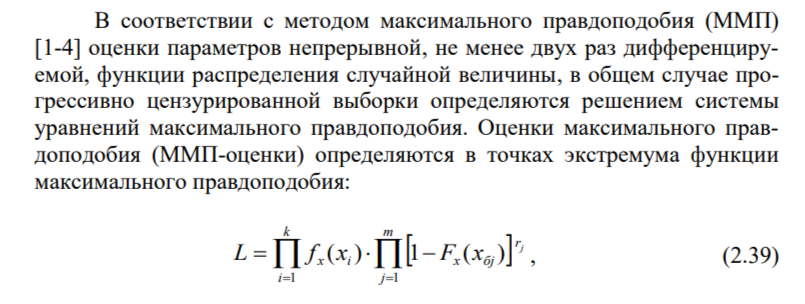
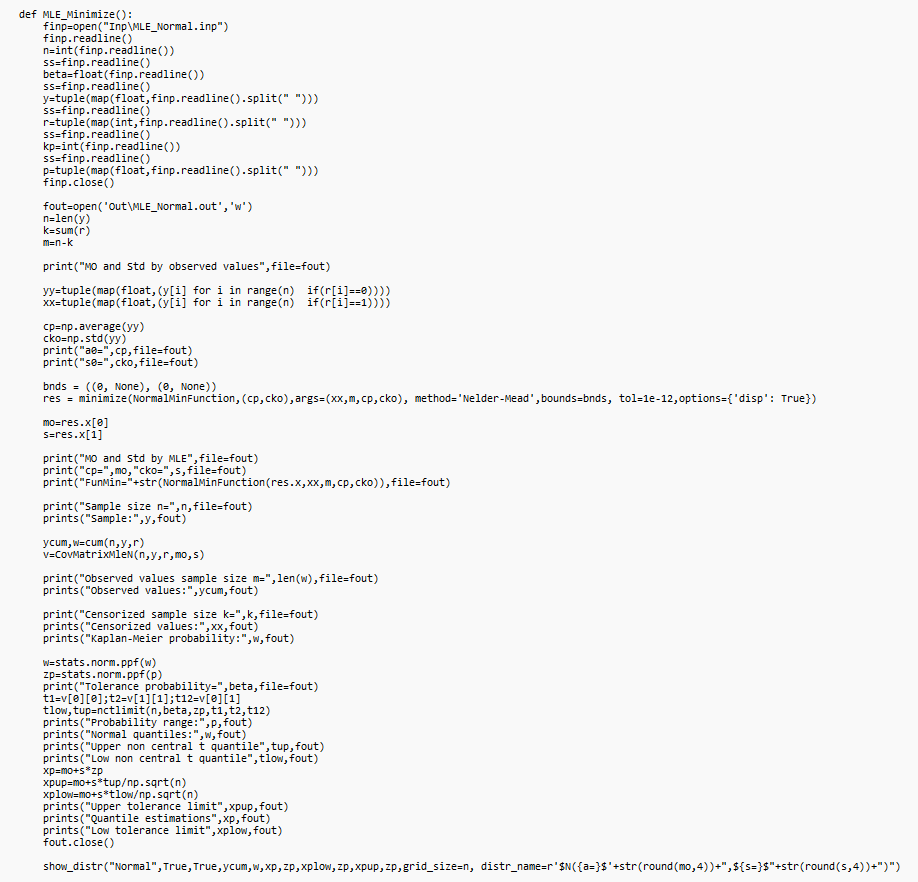
****

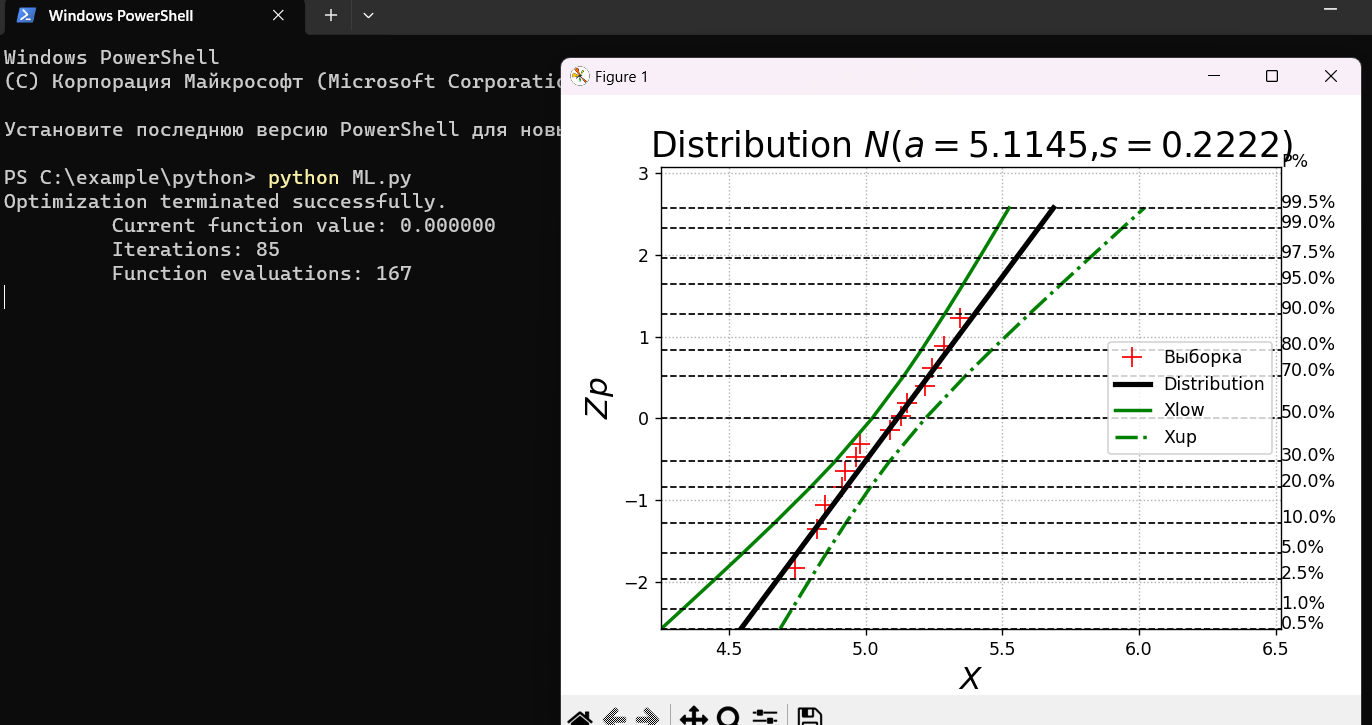
****

Уравнения (2.63), (2.64) позволяют оценивать параметры расположения (сдвига) и масштаба, исходя из порядковых статистик, то есть выборочных наблюдений, упорядоченных по величине.

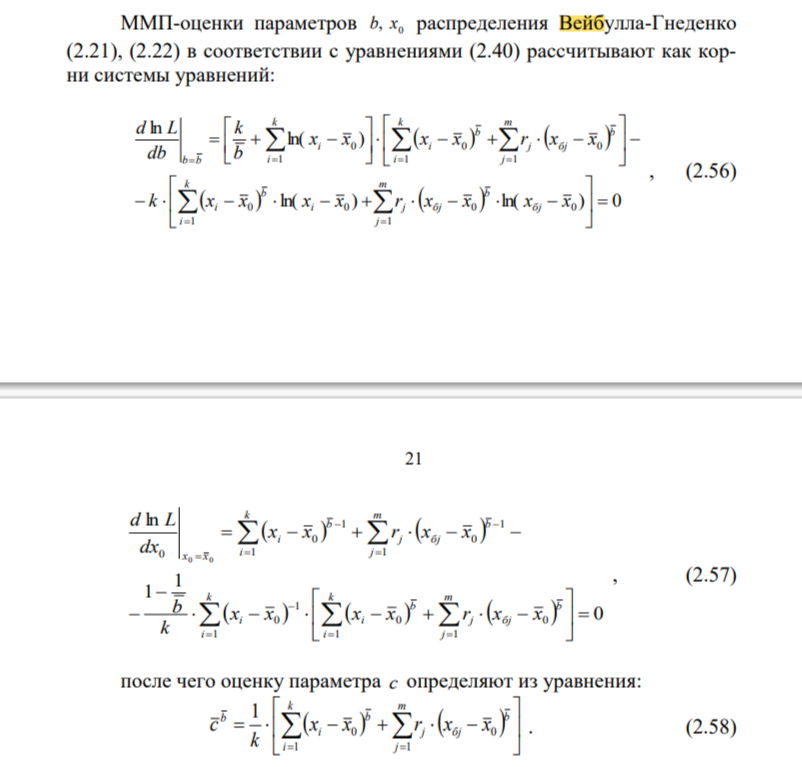


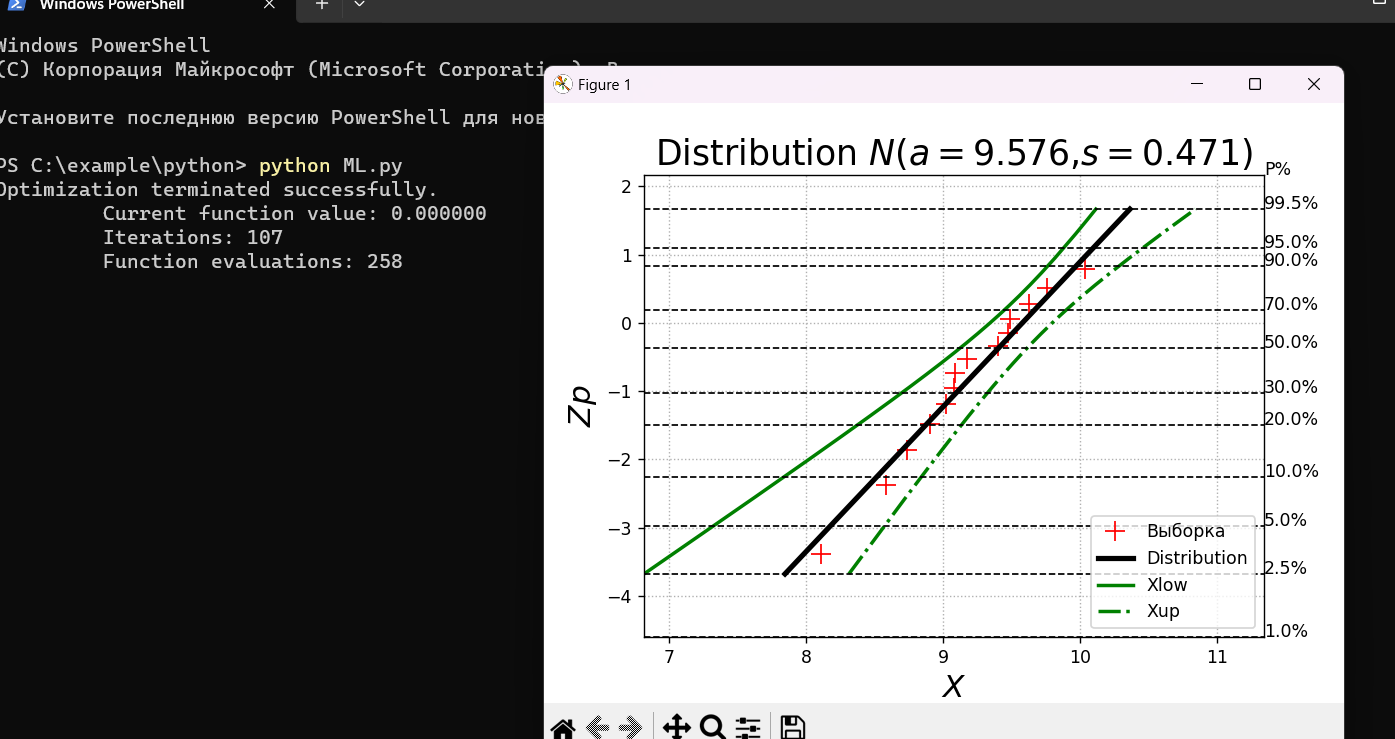
## Метод максимального правдоподобия



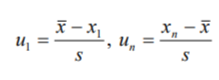
## Распределение Вейбула-Гнеденко

****

****

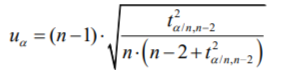
## Критерий Граббса

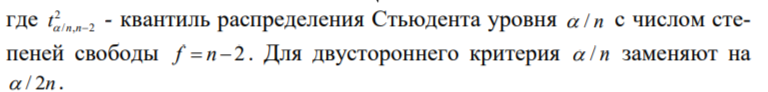
Критерий Граббса применяют в тех случаях, когда имеются статистические данные по рассматриваемой выборке. Для этого рассчитывают статистики:

****

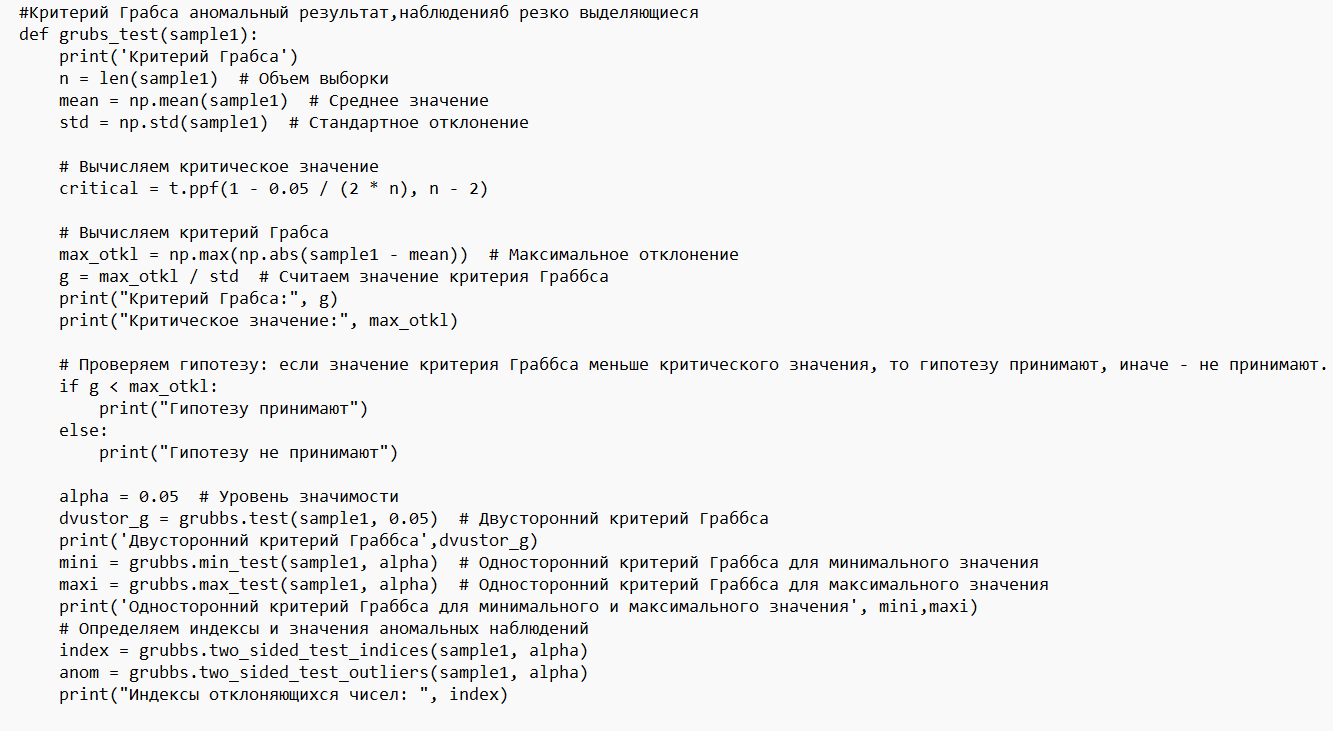
где x, s - выборочное среднее и среднее квадратичное отклонение, n x , x 1 - крайние члены вариационного ряда.

Рассчитанное значение u сопоставляют с критическим uα для заданного уровня значимости α и объема выборки п. Критические значения определяются из уравнения:

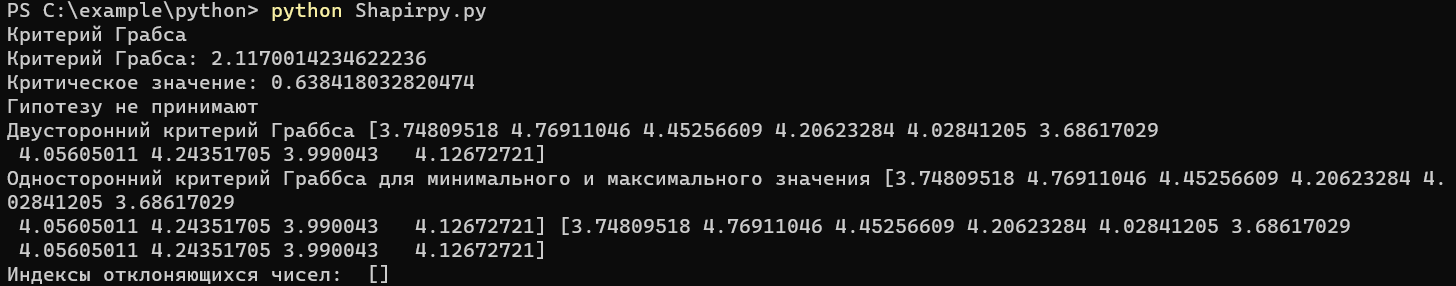
****

****

Нулевую гипотезу принимают, если u ≤ uα и отвергают в противном случае.

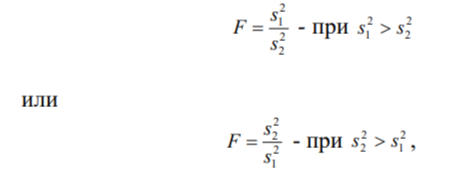


Функция принимает на вход выборку, а выводит статистику критерий Граббса, критические значения, сообщение о том, принимают ли гипотезу, двусторонний критерий Граббса, односторонний критерий Граббса для максимального и минимального значения, индексы отклоняющихся чисел.

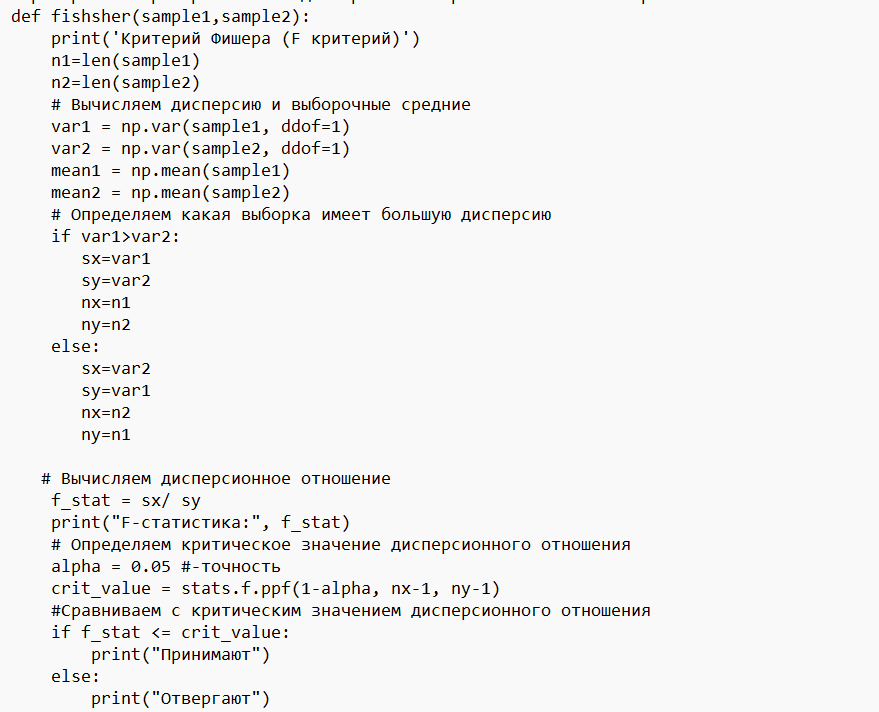


## Критерий Фишера (F-критерий) – об однородности двух дисперсий

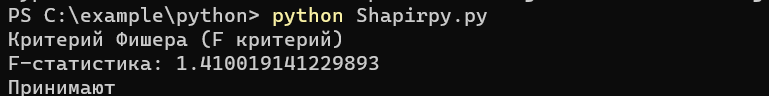
Дисперсии двух совокупностей объемами n1 и n2 , подчиняющихся нормальному (логарифмически нормальному) закону распределения, сравнивают с помощью двустороннего критерия F. Для этого рассчитывают дисперсионное отношение F по формуле:

Дисперсионное отношение F сопоставляют с критическим значением Fα для заданного уровня значимости α и чисел степеней свободы f1 = n1 −1, f2 = n2 −1 , где f1- число степеней свободы для большей дисперсии. В случае соблюдения условия F ≤ Fα , принимают нулевую гипотезу о равенстве генеральных дисперсий. В противном случае нулевая гипотеза отвергается.

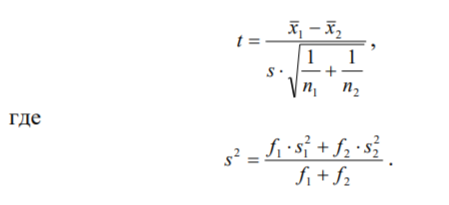


Функция принимает 2 выборки, а выводит F-статистику (дисперсионное отношение), критическое значение и сообщение о том, принимают ли нулевую гипотезу.

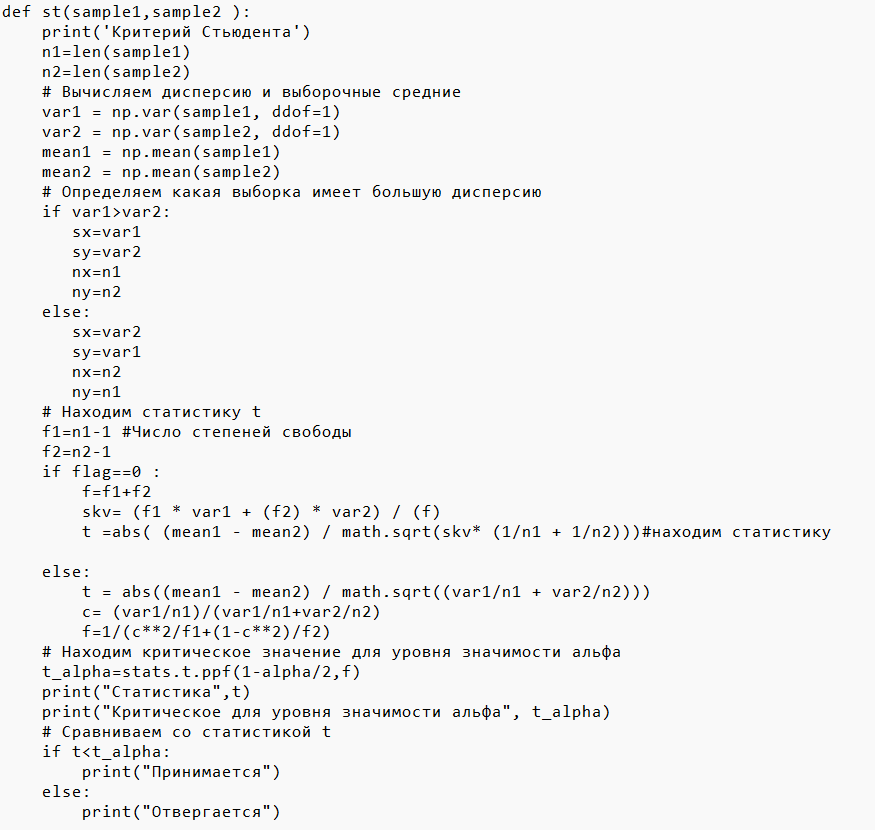


## Критерий Стьюдента (t-критерий)

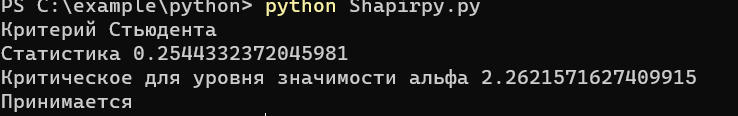
Критерий Стьюдента применяют для сравнения средних значений двух нормально распределенных совокупностей при неизвестных, но равных дисперсиях (σ1)^2 = (σ2)^2 . Нулевая гипотеза заключается в предположении о равенстве средних H0 : a = a . Для проверки этой гипотезы рассчитывают статистику t:

****

Полученное значение t критерия сравнивают с критическим для уровня значимости α и числа степеней свободы f = f1 + f2 . Если t ≤ t(α/2) , то нулевую гипотезу о равенстве средних принимают. В противном случае a1 ≠ a2 .

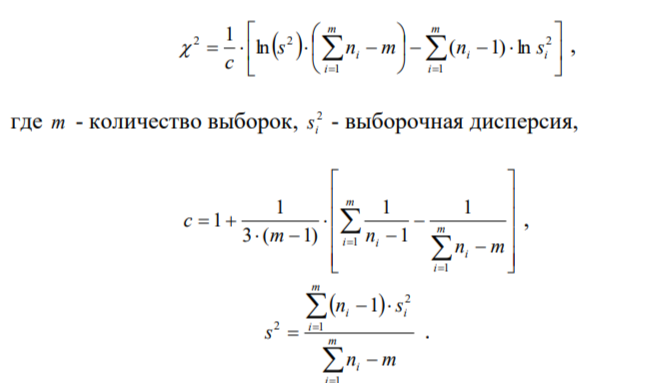
****

Функция принимает на вход две выборки, а выводит статистику критерия t, его критическое значение и сообщение о том, принимают ли гипотезу о равенстве средних.

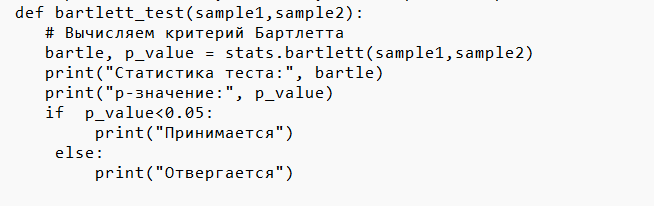


## Критерий Бартлета

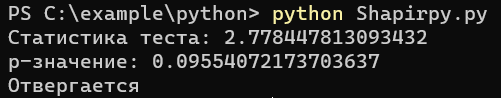
Однородность (равенство) дисперсий ряда выборок из нормально распределенных совокупностей оценивают с помощью критерия Бартлета. Для этого рассчитывают статистику критерия по формуле:

****

Если **** , то нулевая гипотеза об однородности ряда дисперсий подтверждается. В противном случае принимается альтернативная гипотеза о неравенстве дисперсий.

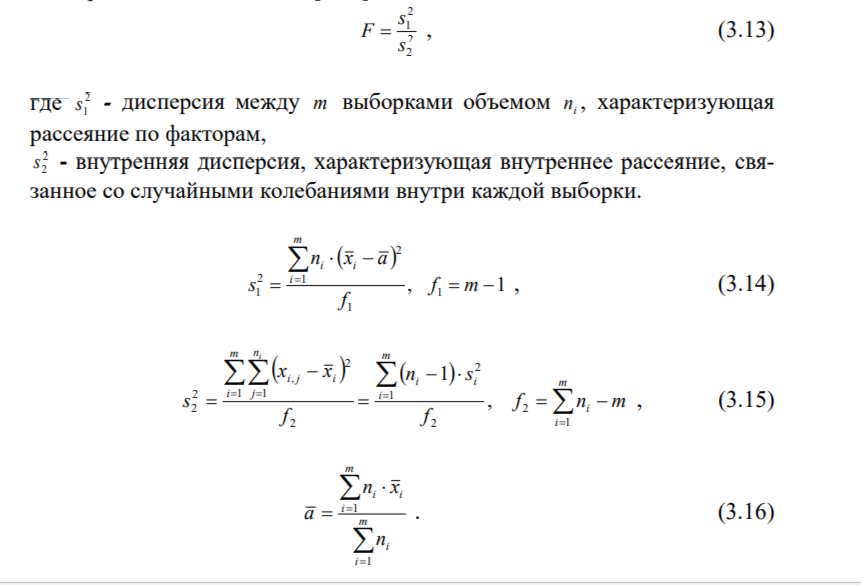


Функция принимает 2 выборки, а возвращает статистику критерия Бартлета, критическое значение, сообщение о том, принимается ли нулевая гипотеза.

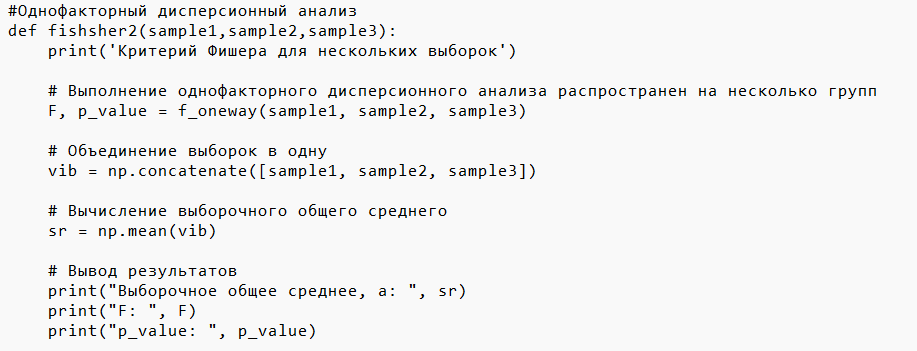


## Однофакторный дисперсионный анализ

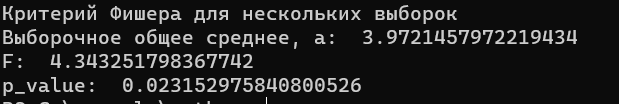
Равенство (однородность) ряда средних значений оценивают с помощью однофакторного дисперсионного анализа. В основе его лежит предположение о нормальности закона распределения случайной величины в каждой выборке и однородности ряда дисперсий. Проверка нулевой гипотезы о равенстве всех средних производят с помощью F- критерия дисперсионного отношения:



Если дисперсионное отношение (3.13) окажется меньше критического значения Fα критерия Фишера, найденного для уровня значимости α и чисел степеней свободы f2 , f1 , то нулевая гипотеза о равенстве средних a1 = a2 = .... = am = a подтверждается

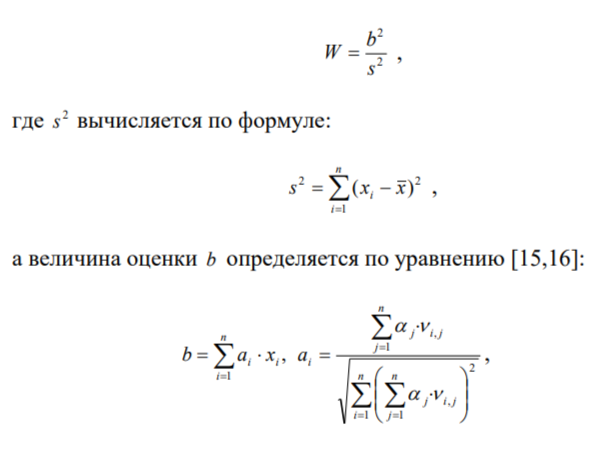


Функция принимает 3 выборки, а выводит выборочное общее среднее, F – статистику и p\_value.

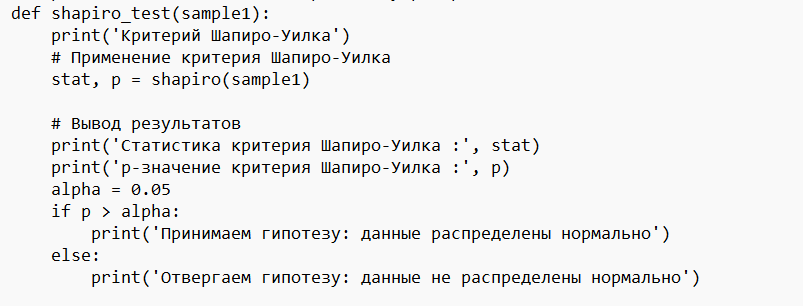


## Критерий Шапиро-Уилка

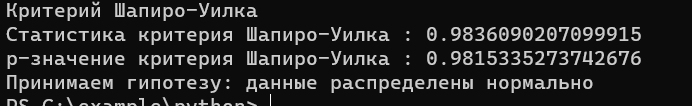
Критерий Шапиро-Уилка (W-критерий) предназначен для проверки гипотезы о нормальном (логарифмически нормальном) распределении. Результаты испытаний располагают в вариационный ряд и вычисляют статистику критерия:



Если W больше критического значения Wα для объема выборки n, то нулевая гипотеза принимается. В противном случае принимается альтернативная гипотеза.

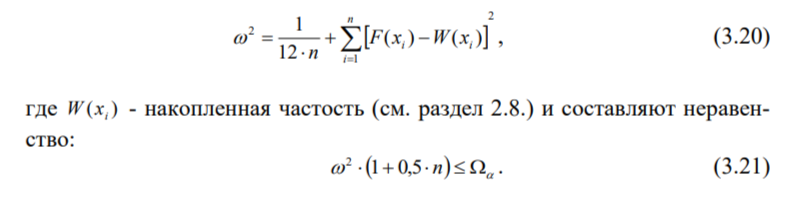


Функция принимает выборку, а выводит статистику критерия Шапиро-Уилка, p-значение и сообщение о том, нормальная ли выборка.

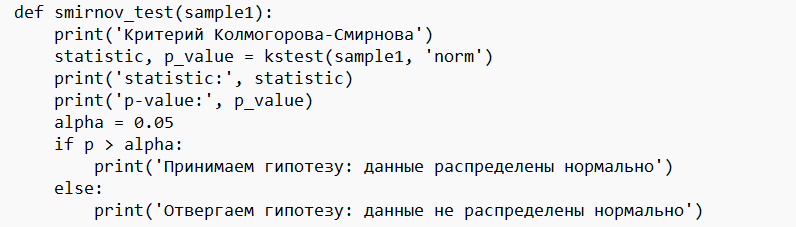


## Критерий Смирнова

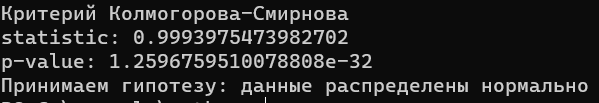
Критерий Смирнова рекомендуется использовать для проверки гипотезы о соответствии эмпирического распределения нормальному закону. Для этого вычисляют статистику:



Если неравенство (3.21) выполняется, то нулевая гипотеза принимается, в противном случае нулевая гипотеза отвергается.

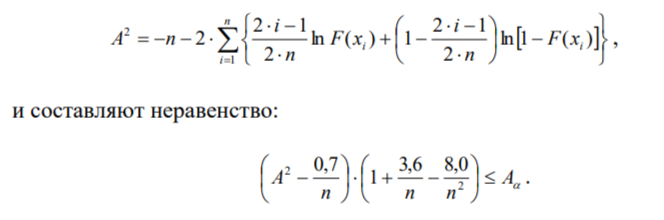


Функция принимает выборку, а выводит статистику критерия Смирнова, p-значение и сообщение о том, нормальная ли выборка

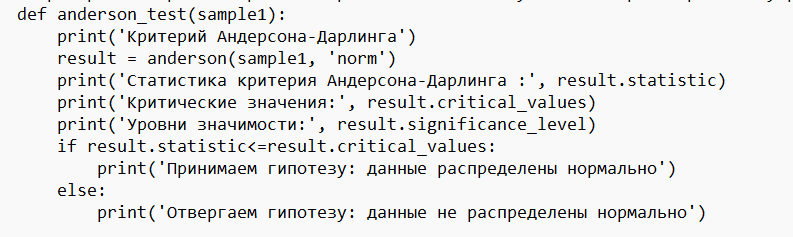


## Критерий Андерсона-Дарлинга

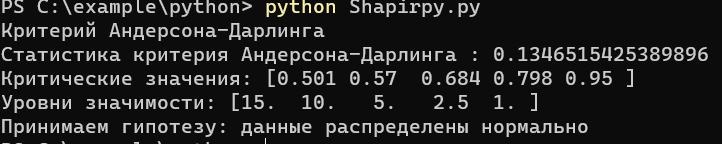
Критерий Андерсона-Дарлинга используют для проверки нормальности в тех случаях, когда больший интерес представляет соответствие эмпирической функции распределения теоретической в области крайних значений случайной величины. С этой целью вычисляют статистику:

****

Если неравенство выполняется, то нулевая гипотеза принимается, в противном случае нулевая гипотеза отвергается.

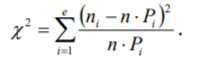
****

Функция принимает выборку, а выводит статистику критерия Андерсона-Дарлинга, p-значение, уровни значимости и сообщение о том, нормальная ли выборка

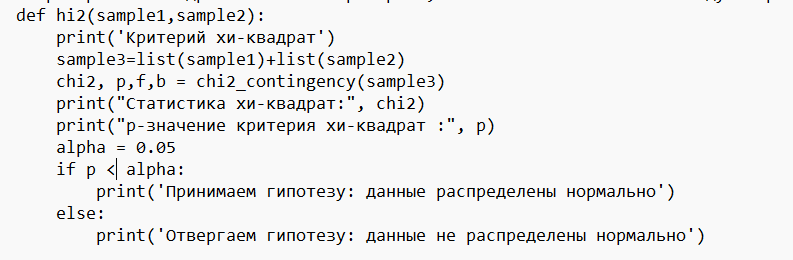


## Критерий

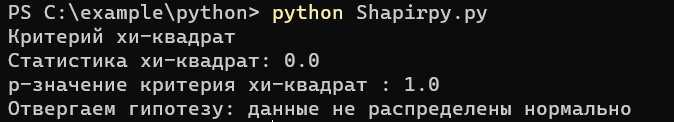
Критерий согласия применяется для проверки гипотезы о соответствии эмпирического распределения произвольному теоретическому распределению, параметры которого оцениваются по выборке. С этой целью рассчитывают статистику:

****

Если значение статистики меньше критического, то нулевая гипотеза о соответствии опытных данных выбранному гипотетическому распределению принимается, в противном случае нулевая гипотеза отвергается.

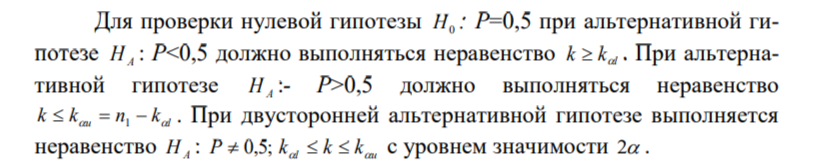
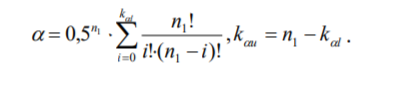
****

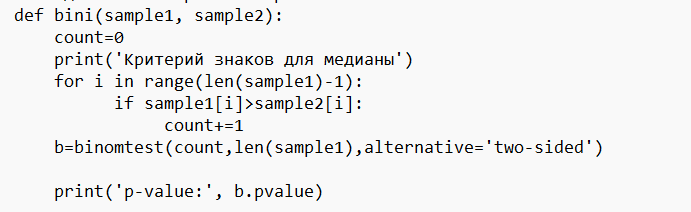
Функция принимает 2 выборки, а выводит статистику критерия согласия, p-значение и сообщение о том, нормальная ли выборка

****

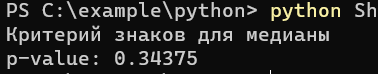
## Критерий знаков для медианы

Пусть в n пар испытаний получены k положительных разностей, т - отрицательных и l нулевых; n1=n- l. Нулевую гипотезу о равенстве медиан двух совокупностей не отвергают, если число k попадает в область допустимых значений с уровнем значимости α. Границы допустимых значений рассчитывают по формулам :



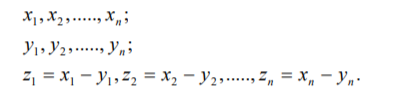


Функция принимает 2 выборки, а выводит р-значение.

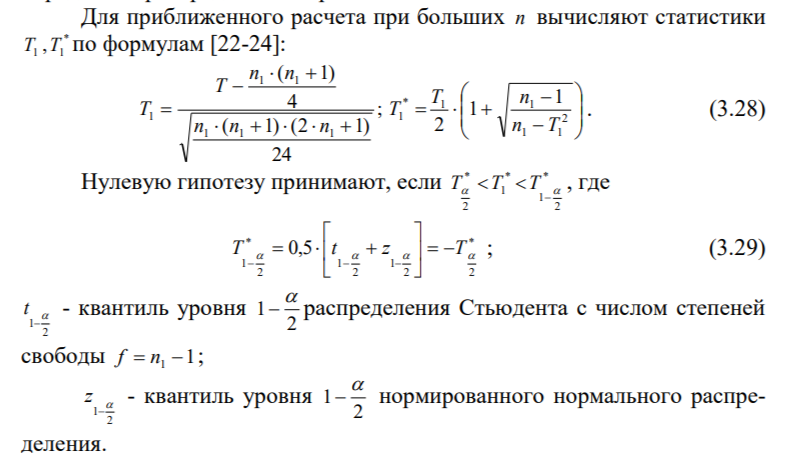


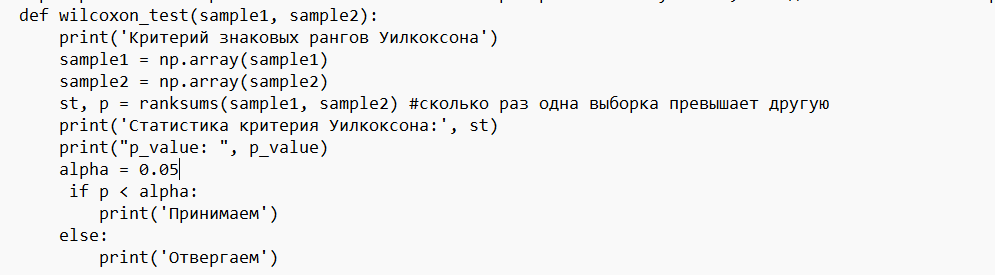
## Двухвыборочный критерий Уилкоксона

Критерий знаковых рангов Уилкоксона учитывает расстояние наблюдений относительно нуля посредством рангов.

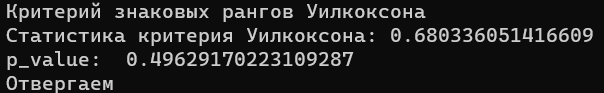
****

Абсолютные значения разностей z располагают в порядке возрастания (ранжируют) и подсчитывают сумму рангов Т (порядковых номеров) положительных значений z в этом ряду.

****

****

Функция принимает 2 выборки, а выводит статистику критерия Уилкоксона , p-значение и сообщение о том, принимаем ли гипотезу

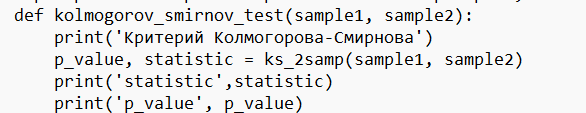


## Критерий Колмогорова-Смирнова

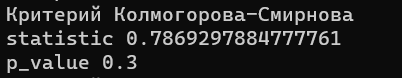
Критерий предназначен для проверки гипотезы о принадлежности двух независимых выборок одной и той же генеральной совокупности. При произвольном распределении в качестве статистики критерия Колмогорова-Смирнова служит наибольшая разность между накопленными частостями, которые рассчитывают для каждого значения случайной величины Х обеих выборочных совокупностей объемом n1 и n2 :



Рассчитанное значение k сравнивают с критическим kα . Если k ≤ kα , гипотеза о принадлежности двух независимых выборок одной генеральной совокупности подтверждается, в противном случае нулевая гипотеза отвергается.

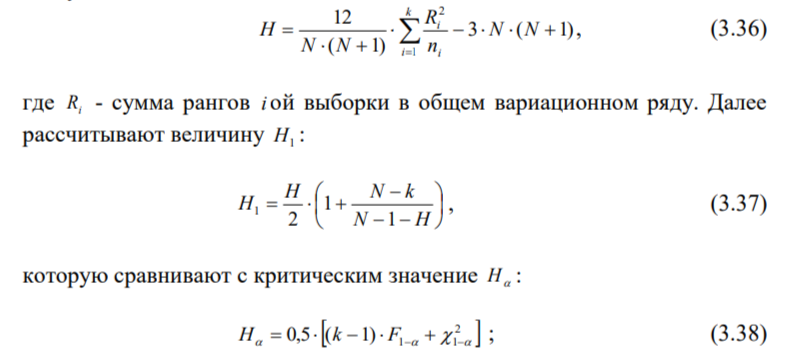
****

Функция принимает 2 выборки, а выводит статистику критерия Колмогорова-Смирнова, p-значение.

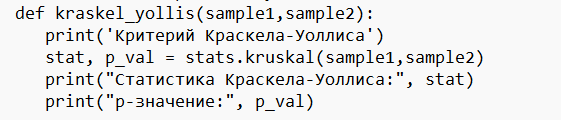
****

## Критерий Краскела-Уоллиса

Критерий Краскела-Уоллиса обобщает задачу о двух выборках на случай k выборок. Нулевая гипотеза утверждает, что k выборок из произвольных совокупностей можно рассматривать как одну (объединенную) выборку из общей совокупности, то есть утверждается равенство параметров сдвига , когда не задано значение общего параметра масштаба против альтернативы не все равны. Для проверки нулевой гипотезы строят общий вариационный ряд наблюдений и рассчитывают статистику:



Нулевую гипотезу принимают, если H1 ≤ Hα с уровнем значимости α . В противном случае принимают альтернативную гипотезу.



Функция принимает 2 выборки, а выводит статистику критерия Колмогорова-Смирнова, p-значение.

