



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

«МИРЭА – Российский технологический университет»

ИНСТИТУТ КИБЕРНЕТИКИ

КАФЕДРА ВЫСШЕЙ МАТЕМАТИКИ

Лабораторная работа 5

по курсу «Теория вероятностей и математическая статистика, часть 2»

Тема: Проверка статистических гипотез о равенстве математических ожиданий и дисперсий наблюдаемых нормально распределенных случайных величин

Выполнил:
Студент 3-го курса
Жолковский Д.А.

Группа: КМБО-01-16

МОСКВА 2019

Лабораторная работа по Математической статистике № 5
«Проверка статистических гипотез о равенстве математических
ожиданий и дисперсий наблюдаемых нормально распределенных
случайных величин»

Задания

Из файла DLW-5 в соответствии с номером варианта взять двумерный массив $\{u_{i,j} | 1 \leq i \leq N, 1 \leq j \leq 3\}$, столбцами которого являются выборки одинакового объёма N трёх наблюдаемых нормально распределенных случайных величин.

I. Проверить гипотезу о равенстве математических ожиданий наблюдаемых случайных величин при уровне значимости 0,06 разными способами – непосредственным расчетом по формулам из Указания и с помощью функций Python.

II. Проверить гипотезу о равенстве дисперсий наблюдаемых случайных величин при уровне значимости 0,06 разными способами – непосредственным расчетом по формулам из Указания и с помощью функций Python.

Вычисления проводить с точностью до 0,00001.

Краткие теоретические сведения

Распределение «хи-квадрат» $\chi^2(n)$, n – степень свободы.

$\chi^2(n) = \gamma\left(\frac{n}{2}; \frac{1}{2}\right)$ – частный случай гамма распределения

Плотность $f_{\xi}(x)$	$\begin{cases} 0, x \leq 0 \\ \frac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{x}{2}}, x > 0 \end{cases}$
Математическое ожидание M_{ξ}	n
Дисперсия D_{ξ}	$2n$
Характеристическая функция $g_{\xi}(t)$	$(1 - 2it)^{-\frac{n}{2}}$

При $n = 1$ получается плотность квадрата случайной величины, имеющей стандартное нормальное распределение, т.е. $\chi^2(1) = [N(0,1)]^2$

Распределение Стьюдента с n степенями свободы ($t(n)$)

Плотность $f_{\xi}(x)$	$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \cdot \Gamma(\frac{n}{2}) \cdot (1 + \frac{x^2}{n})^{\frac{n+1}{2}}}$
Математическое ожидание M_{ξ}	$\begin{cases} 0, n > 1 \\ \text{не существует, } n \leq 1 \end{cases}$
Дисперсия D_{ξ}	$\begin{cases} \frac{n}{n-2}, n > 2 \\ \text{не существует, } n \leq 2 \end{cases}$

Распределение Коши получается из распределения Стьюдента при

$$n = 1: f_{\xi}(x) = \frac{1}{\pi(1+x^2)}$$

У распределения Коши не существует математического ожидания и дисперсии.

Если $\eta \sim N(0,1)$ и $\zeta \sim \chi^2(n)$ независимы, то $\xi = \eta \sqrt{\frac{n}{\zeta}}$

Распределение Стьюдента близко к стандартному нормальному распределению при больших n .

Распределение Фишера-Снедекора $F(k_1, k_2)$ –
 F распределение с (k_1, k_2) степенями свободы

Плотность $f_{\xi}(x)$	$\begin{cases} 0, x \leq 0 \\ \frac{\Gamma(\frac{k_1 + k_2}{2}) \cdot k_1^{\frac{k_1}{2}} \cdot k_2^{\frac{k_2}{2}}}{\Gamma(\frac{k_1}{2}) \cdot \Gamma(\frac{k_2}{2})} \cdot x^{\frac{k_1}{2}-1} \cdot (k_1 x + k_2)^{-\frac{k_1+k_2}{2}}, x \geq 0 \end{cases}$
Математическое ожидание M_{ξ}	$\begin{cases} \frac{k_2}{k_2 - 2}, k_2 > 2 \\ \text{не существует}, k_2 \leq 2 \end{cases}$
Дисперсия D_{ξ}	$\begin{cases} \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)}, k_2 > 4 \\ \text{не существует}, k_2 \leq 4 \end{cases}$

Если $\eta_1 \sim \chi^2(k_1)$ и $\eta_2 \sim \chi^2(k_2)$ – независимы, то $\xi = \frac{\eta_1/k_1}{\eta_2/k_2} \sim F(k_1, k_2)$.

При проверке гипотез о равенстве математических ожиданий и дисперсий используются следующие формулы расчета характеристик выборок $\{x_1, x_2, x_3, \dots, x_N\}$ и $\{y_1, y_2, y_3, \dots, y_M\}$:

Выборочное среднее:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{M} \sum_{i=1}^M y_i$$

Выборочная несмещённая дисперсия:

$$\overline{x^2} = \frac{1}{N} \sum_{i=1}^N x_i^2, \quad \overline{y^2} = \frac{1}{M} \sum_{i=1}^M y_i^2$$

$$S_x^2 = \frac{N}{N-1} (\overline{x^2} - \bar{x}^2), \quad S_y^2 = \frac{M}{M-1} (\overline{y^2} - \bar{y}^2)$$

Проверка гипотезы о равенстве математических ожиданий наблюдаемых случайных величин.

1. Проверка гипотезы о равенстве математических ожиданий двух случайных величин по выборкам $\{x_1, x_2, x_3, \dots, x_N\}$ и $\{y_1, y_2, y_3, \dots, y_M\}$ с использованием распределения Стьюдента с числом степеней свободы $N + M - 2$.

Расчет значения критерия $T_{N,M}$:

$$S_x^2(N-1) = N(\overline{x^2} - \bar{x}^2), \quad S_y^2(M-1) = M(\overline{y^2} - \bar{y}^2)$$

$$T_{N,M} = \frac{\bar{x} - \bar{y}}{\sqrt{S_x^2(N-1) + S_y^2(M-1)}} \sqrt{\frac{MN(N+M-2)}{N+M}}$$

Если гипотеза о равенстве математических ожиданий нормально распределенных случайных величин X и Y верна, то $T_{N,M}$ имеет распределение $t(N+M-2)$ – распределение Стьюдента с числом степеней свободы $N + M - 2$.

В данной лабораторной работе $N = M$ и рассматривается распределение $t(2N-2)$.

2. Проверка гипотезы о равенстве математических ожиданий трёх случайных величин по выборкам, являющимся столбцами массива $U = \{u_{i,j} | 1 \leq i \leq N, 1 \leq j \leq m = 3\}$, с использованием однофакторного дисперсионного анализа.

Расчёт общего среднего значения и групповых средних:

$$\bar{u} = \frac{1}{Nm} \sum_{j=1}^m \sum_{i=1}^N u_{ij}, \quad \bar{u}_{.j} = \frac{1}{N} \sum_{i=1}^N u_{ij}, \quad j = 1, \dots, m$$

Расчёт общей суммы квадратов отклонений $S_{\text{общ}} = \sum_{j=1}^m \sum_{i=1}^N (u_{ij} - \bar{u})^2$.

Расчёт общей суммы квадратов отклонений $S_{\text{факт}} = N \sum_{j=1}^m (u_{.j} - \bar{u})^2$.

Расчёт общей суммы квадратов отклонений $S_{\text{ост}} = \sum_{j=1}^m \sum_{i=1}^N (u_{ij} - u_{.j})^2 = S_{\text{общ}} - S_{\text{факт}}$.

Расчёт значения критерия $F_{N,m}$: $F_{N,m} = \frac{s_{\text{факт}}^2}{s_{\text{ост}}^2}$, где $s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{m-1}$, $s_{\text{ост}}^2 = \frac{S_{\text{ост}}}{m(N-1)}$

Если гипотеза о равенстве математических ожиданий m нормально распределенных случайных величин верна, то $F_{N,m}$ имеет распределение Фишера-Снедекора с числом степеней свободы (k_1, k_2) , $k_1 = m - 1$, $k_2 = m(N - 1)$.

Проверка гипотезы о равенстве дисперсий наблюдаемых случайных величин.

Проверка гипотезы о равенстве дисперсий двух случайных величин по выборкам 1 $\{x_1, x_2, x_3, \dots, x_N\}$ и $\{y_1, y_2, y_3, \dots, y_M\}$ с использованием распределения Фишера-Снедекора.

Расчёт значения критерия $F_{N,M}$:

$$F_{N,M} = \frac{S_1^2}{S_2^2}, \text{ где } S_1^2 = \max(S_x^2, S_y^2), S_2^2 = \min(S_x^2, S_y^2)$$

Если гипотеза о равенстве дисперсий двух нормально распределенных случайных величин верна, то $F_{N,M}$, имеет распределение Фишера-Снедекора с числом степеней свободы (k_1, k_2) , где

$$k_1 = \begin{cases} N - 1, S_1^2 = S_x^2 \\ M - 1, S_1^2 = S_y^2 \end{cases} \quad k_2 = \begin{cases} N - 1, S_2^2 = S_x^2 \\ M - 1, S_2^2 = S_y^2 \end{cases}$$

В данной лабораторной работе $N = M$, $k_1 = k_2 = N - 1$ и рассматривается распределение Фишера-Снедекора с числом степеней свободы $(N - 1, N - 1)$.

Средства языка программирования Python, которые использованы в программе расчета

- `max(arr)` – выбор максимального значения в массиве `arr`;
- `min(arr)` – выбор минимального значения в массиве `arr`;
- `sum(x)` и `x.sum()` возвращают сумму элементов массива `x`.
- `len(x)` и `x.size` возвращают количество элементов в массиве `x`.
- `scipy.stats.f_oneway(X,Y,Z)` - проверка гипотезы о равенстве математических ожиданий трёх случайных величин
- `scipy.stats.ttest_ind(X,Y)` - проверка гипотезы о равенстве математических • ожиданий пары случайных величин;
- `scipy.stats.ttest_ind(X,Y,equal_var=False)` - проверка гипотезы о равенстве математических ожиданий пары случайных величин;
- `scipy.stats.bartlett (X,Y,Z)` -проверка гипотезы о равенстве дисперсий трёх случайных величин `X, Y, Z`;
- `scipy.stats.barlett(X,Y)` - проверка гипотез о равенстве дисперсий пары случайных величин;
- `scipy.stats.t.ppf(x,2N-2)` – определение критического значения $t_{кр,\alpha}(2N - 2)$, $x=1-\frac{\alpha}{2}$
- `scipy.stats.f.ppf(1- α , k_1 , k_2)` – нахождение критического значения $F(\alpha, k_1, k_2)$

Извлечение данных из PDF-файла было реализовано с помощью библиотеки `pdfminer`. А построение таблиц было реализовано средствами библиотеки `docx`.

Результаты расчетов с комментариями

1.23711	0.85209	-0.64836
0.89850	2.69360	-1.31322
-0.62379	-3.10639	-1.79283
-1.99397	3.14631	-0.47805
2.31921	-3.63701	-2.49926
-1.13083	1.08540	-0.82844
-4.48306	-2.62187	-3.46558
-1.88237	-0.95353	1.21530
1.23293	2.31578	3.29139
0.69600	-0.28418	-2.80116
-5.98482	-1.49992	3.94803
-1.84592	-0.01161	-0.36222
-2.20889	-3.69538	4.76533
-1.89087	0.07593	2.09255
-2.12074	1.07141	0.28011
-1.73964	-3.80871	2.80244
-0.17203	-1.47559	1.43270
1.58969	2.34929	0.56925
4.38888	-0.26896	-2.25504
-2.68677	0.74149	-0.00000

Задание 1) гипотеза о равенстве математических ожиданий

Пункт 1)

Столбцы	\bar{x}	\bar{y}	$\overline{x^2}$	$\overline{y^2}$	S_x^2	S_y^2	$T_{N,N}$
(1,2)	-0.82007	-0.35159	6.16238	4.75239	5.77881	4.87239	-0.64195
(1,3)	-0.82007	0.19765	6.16238	5.14433	5.77881	5.37396	-1.36286
(2,3)	-0.35159	0.19765	4.75239	5.14433	4.87239	5.37396	-0.76735

Пункт 2)

$$S_{\text{общ}} = 314.85725$$

$$S_{\text{факт}} = 10.37920$$

$$S_{\text{ост}} = 304.47805$$

$$s_{\text{факт}}^2 = 5.18960$$

$$s_{\text{ост}}^2 = 7.61195$$

$$k_1 = 3$$

$$k_2 = 57$$

$$F_{N,m} = 0.68177$$

Пункт 3) $pval = anova(U) = 0.38469$

Пункт 4)

Столбцы	pval [t_test_2]
(1,2)	0.52476
(1,3)	0.18095
(2,3)	0.44762

Пункт 5)

Столбцы	pval [welch_test]
(1,2)	0.52479
(1,3)	0.18096
(2,3)	0.44763

Задание 2) гипотеза о равенстве дисперсий

Пункт 1)

Столбцы	S_1^2	S_2^2	k_1	k_2	$F_{N,N}$
(1,2)	5.77881	4.87239	19	19	1.18603
(1,3)	5.77881	5.37396	19	19	1.07534
(2,3)	5.37396	4.87239	19	19	1.10294

Пункт 2)

$pval = bartlett_test(X, Y, Z) = 0.93452$

Пункт 3)

Столбцы	pval [var_test]
(1,2)	0.71375
(1,3)	0.87584
(2,3)	0.83309

Анализ результатов и выводы:

Задание 1

Пункт 1)

столбцы	$ T_{NN} $	$t_{кр,\alpha}(2N - 2)$	ВЫВОД
(1; 2)	0.64195	1.93863	верно
(1; 3)	1.36286	1.93863	верно
(2; 3)	0.76735	1.93863	верно

Гипотеза о равенстве математических ожиданий не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 2)

$$z_{\alpha} = 2.95696$$

$$F_{NM} = 0.68177$$

Гипотеза о равенстве математических ожиданий трёх случайных величин, выборки которых находятся в столбцах массива $U = \{u_{i,j} | 1 \leq i \leq N, 1 \leq j \leq 3\}$, не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 3)

$$pval = \text{anova}(U) = 0.38469 > \alpha = 0.06$$

Гипотеза о равенстве математических ожиданий трёх случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 4)

столбцы	pval [t_test_2]	α	ВЫВОД
(1,2)	0.52476	0.06	верно
(1,3)	0.18095	0.06	верно
(2,3)	0.44762	0.06	верно

Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 5)

столбцы	pval [welch_test]	α	ВЫВОД
(1,2)	0.52479	0.06	верно
(1,3)	0.18096	0.06	верно
(2,3)	0.44763	0.06	верно

Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Задание 2

Пункт 1)

столбцы	F_NM	$F(\alpha/2, k_1, k_2)$	ВЫВОД
(1; 2)	1.18603	2.43072	верно
(1; 3)	1.07534	2.43072	верно
(2; 3)	1.10294	2.43072	верно

Гипотеза о равенстве дисперсий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 2)

$pval = \text{bartlett_test}(X, Y, Z) = 0.93452 > 0.06 = \alpha$

Гипотеза о равенстве дисперсий трёх случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Пункт 3)

строка	$pval[\text{bartlett_test}]$	α	вывод
(1; 2)	0.71375	0.06	верно
(1; 3)	0.87584	0.06	верно
(2; 3)	0.83309	0.06	верно

Гипотеза о равенстве дисперсий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.

Список использованной литературы

1. Лобузов А.А. Математическая статистика [Электронный ресурс]: Методические указания по выполнению лабораторных работ / под ред. Ю. И. Худака. Москва: Московский технологический университет (МИРЭА), 2017. 36 с.
2. Чернова Н. И. Математическая статистика: Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2007. 148 с

Приложение (Листинг программы)

```

import io

from pdfminer.converter import TextConverter
from pdfminer.pdfinterp import PDFPageInterpreter
from pdfminer.pdfinterp import PDFResourceManager
from pdfminer.pdfpage import PDFPage

import numpy as np
import scipy.stats as stats

from docx import Document

def extract_text_from_pdf(pdf_path):
    resource_manager = PDFResourceManager()
    fake_file_handle = io.StringIO()
    converter = TextConverter(resource_manager, fake_file_handle)
    page_interpreter = PDFPageInterpreter(resource_manager,
converter)

    with open(pdf_path, 'rb') as fh:
        for page in PDFPage.get_pages(fh,
                                     caching=True,
                                     check_extractable=True):
            page_interpreter.process_page(page)

            text = fake_file_handle.getvalue()

        # close open handles
        converter.close()
        fake_file_handle.close()

    if text:
        return text

def var_in_pdf_length(filename = 'DLW-5_2019_var_1-20.pdf', v =
10):
    s = extract_text_from_pdf(filename)
    n0 = s.find('Вариант ' + str(v))
    n1 = s.find('Вариант ' + str(v+1))
    if n0 == -1:

```

```

    n0 = 0
    if n1 == -1:
        n1 = len(s)
    space0 = 'Вариант '+str(v)+' '
    space1 = ' '
    #print(len(space0),len(space1))
    w = s[n0+len(space0):n1-len(space1)]
    w = w.replace(',', '.')
    w = w.split(' ')
    l = [[],[],[]]
    for i in range(len(w)):
        if i%3 == 0:
            l[0].append(float(w[i]))
        if (i-1)%3 == 0:
            l[1].append(float(w[i]))
        if (i-2)%3 == 0:
            l[2].append(float(w[i]))
    return l

```

```

ql = var_in_pdf_length()
ql = [np.array(i) for i in ql]
for i in ql:
    print(i)

```

```

print()
N = ql[0].size
m = len(ql)
print(N,m)

```

```

def mean1(x):
    return sum(x)/x.size
def mean2(x):
    return sum(x**2)/x.size

```

```

for i in range(m):
    print('mean1('+str(i+1)+') =',mean1(ql[i]),'      mean2('+str(i+1)+')
    =',mean2(ql[i]))
    print()

```

```

def S2(x):

```

```

    return (mean2(x)-mean1(x)**2)*x.size/(x.size-1)

for i in range(m):
    print('S2('+str(i+1)+') =',S2(ql[i]))
print()

def TNM(x,y):
    o1 = mean1(x)-mean1(y)
    o2 = S2(x)*(x.size-1) + S2(y)*(y.size-1)
    o3 = x.size*y.size*(x.size+y.size-2)/(x.size+y.size)
    return o1*((o3/o2)**0.5)

for k in range(m):
    for i in range(k):
        print('TNM('+str(i+1)+';'+str(k+1)+') =',TNM(ql[i],ql[k]))

```

```

def u_all(ql):
    s = 0
    n = 0
    for i in ql:
        n += i.size
        s += i.sum()
    return s/n

def u_one(x):
    return x.sum()/x.size

```

```

def s_all(ql):
    s = 0
    u = u_all(ql)
    for i in ql:
        for k in i:
            s += (u - k)**2
    return s

```

```

def s_fakt(ql):
    s = 0
    u = u_all(ql)
    for i in ql:
        s += i.size*(u-u_one(i))**2
    return s

```

```

def s_ost(ql):
    return s_all(ql) - s_fakt(ql)

def FNM_1(ql):
    sf2 = s_fakt(ql)/(len(ql)-1)
    so2 = s_ost(ql)/((len(ql)-1)*ql[0].size)
    return sf2/so2

print('u_all = ', u_all(ql))
for i in range(len(ql)):
    print('u_'+str(i+1)+' = ', u_one(ql[i]))
print('s_all = ', s_all(ql))
print('s_fakt = ', s_fakt(ql))
print('s_ost = ', s_ost(ql))
print('FNM_1 = ', FNM_1(ql))


def Sm1(x,y):
    return max(S2(x),S2(y))

def Sm2(x,y):
    return min(S2(x),S2(y))

def FNM_2(x,y):
    return Sm1(x,y)/Sm2(x,y)

def k1(x,y):
    if Sm1(x,y) == S2(x):
        return x.size-1
    if Sm1(x,y) == S2(y):
        return y.size-1

def k2(x,y):
    if Sm2(x,y) == S2(x):
        return x.size-1
    if Sm2(x,y) == S2(y):
        return y.size-1

for k in range(3):
    for i in range(k):
        print('Sm1('+str(i+1)+';'+str(k+1)+') = ', Sm1(ql[i],ql[k]))
        print('Sm2('+str(i+1)+';'+str(k+1)+') = ', Sm2(ql[i],ql[k]))
        print('FNM_2('+str(i+1)+';'+str(k+1)+') = ', FNM_2(ql[i],ql[k]))

```



```

        print('k1('+str(i+1)+';'+str(k+1)+')'          =      ',k1(ql[i],ql[k]),'
k2('+str(i+1)+';'+str(k+1)+') = ',k2(ql[i],ql[k]))
        print()

```

```

print('N1')
print('3'),'ANOVA(1,2,3) = ', stats.f_oneway(ql[0],ql[1],ql[2]).pvalue)
for k in range(3):
    for i in range(k):
        print('4'),'t_test_2(' + str(i+1) + ';' + str(k+1)+') = ',
stats.ttest_ind(ql[i],ql[k]).pvalue)
    for k in range(3):
        for i in range(k):
            print('5'),'welch_test(' + str(i+1) + ';' + str(k+1)+') = ',
stats.ttest_ind(ql[i],ql[k], equal_var = False).pvalue)
    print()
print('N2')
print('2'),'bartlett_test(1,2,3) = ',stats.bartlett(ql[0],ql[1],ql[2]).pvalue)
for k in range(3):
    for i in range(k):
        print('3'),'var_test(' + str(i+1) + ';' + str(k+1)+') = ',
stats.bartlett(ql[i],ql[k]).pvalue)

```

```

head_1_1_str = ['столбцы', 'x', 'y', 'x^2', 'y^2', 'S_x^2',
'S_y^2','T_NM']
def out_1_1(ql):
    out = []
    for i in range(len(ql)):
        for k in range(i+1,len(ql)):
            l = [
                ('+str(i+1)+';'+str(k+1)+')',
                mean1(ql[i]),
                mean1(ql[k]),
                mean2(ql[i]),
                mean2(ql[k]),
                S2(ql[i]),
                S2(ql[k]),
                TNM(ql[i],ql[k])
            ]
            out.append(l)
    return out

```

```

head_1_2_str = ['S_общ', 'S_факт', 'S_ост', 'S_факт^2', 'S_ост^2',
'k_1', 'k_2', 'F_NM']
def out_1_2(ql):
    sf2 = s_fakt(ql)/(len(ql)-1)
    so2 = s_ost(ql)/((len(ql)-1)*ql[0].size)
    k1 = len(ql)
    k2 = 0
    for i in ql:
        k2 += (i.size-1)
    return [
        s_all(ql), s_fakt(ql), s_ost(ql), sf2, so2, k1, k2, FNM_1(ql)
    ]

```

```

def out_1_3(ql):
    return stats.f_oneway(ql[0],ql[1],ql[2]).pvalue

```

```

head_1_4_str = ['столбцы','pval[t_test_2]']
def out_1_4(ql):
    l = []
    for i in range(3):
        for k in range(i):

```

```

l.append(['('+str(k+1)+'+',str(i+1)+'+)',stats.ttest_ind(ql[k],ql[i]).pvalue])
    return l

```

```

head_1_5_str = ['столбцы','pval[welch_test]']
def out_1_5(ql):
    l = []
    for i in range(3):
        for k in range(i):

```

```

l.append(['('+str(k+1)+'+',str(i+1)+'+)',stats.ttest_ind(ql[k],ql[i],equal_var
False).pvalue])
    return l

```

```

head_2_1_str = ['столбцы', 'S_1^2','S_2^2','k_1','k_2','F_NM']
def out_2_1(ql):
    out = []
    for i in range(3):
        for k in range(i):
            x = k
            y = i

```

```

        l = [
            '+' + str(x+1) + '+'; '+' + str(y+1) + '+',
            Sm1(ql[x],ql[y]),
            Sm2(ql[x],ql[y]),
            k1(ql[x],ql[y]),
            k2(ql[x],ql[y]),
            FNM_2(ql[x],ql[y])
        ]
        out.append(l)
    return out

```

```

def out_2_2(ql):
    return stats.bartlett(ql[0],ql[1],ql[2]).pvalue

```

```

head_2_3_str = ['столбцы', 'pval[var_test]']
def out_2_3(ql):
    out = []
    for i in range(3):
        for k in range(i):
            x = k
            y = i
            l = [
                '+' + str(x+1) + '+'; '+' + str(y+1) + '+',
                stats.bartlett(ql[x],ql[y]).pvalue
            ]
            out.append(l)
    return out

```

```

expert = {
    True : 'верно',
    False : 'не верно'
}

```

```

anal_head_1_1_str = ['столбцы', 'T_NN|', 't_кр,alpha(2N-2)', 'вывод']
def anal_1_1(ql):
    ao = []
    N = ql[0].size
    for k in range(3):
        for i in range(k):
            l = [
                '+' + str(i+1) + '+'; '+' + str(k+1) + '+',

```

```

        abs(TNM(ql[i],ql[k])),
        stats.t.ppf(0.97,2*N-2),
        expert[abs(TNM(ql[i],ql[k])) <= stats.t.ppf(0.97,2*N-2)]
    ]
    ao.append(l)
return ao

anal_head_1_2_str = ['F_NM', 'alpha', 'z_alpha', 'ВЫВОД']
def anal_1_2(ql):
    m = len(ql)-1
    n = len(ql)*(ql[0].size-1)
    f = FNM_1(ql)
    z = stats.f.ppf(0.94,m,n)
    ao = [f, '0.06', z, expert[f <= z]]
    return [ao]

anal_head_1_3_str = ['pval[anova]', 'alpha', 'ВЫВОД']
def anal_1_3(ql):
    anova = stats.f_oneway(ql[0],ql[1],ql[2]).pvalue
    return [[anova, '0.06', expert[anova >= 0.06]]]

anal_head_1_4_str = ['столбцы', 'pval [t_test_2]', 'alpha', 'ВЫВОД']
def anal_1_4(ql):
    o_1_4 = out_1_4(ql)
    ao = []
    for i in o_1_4:
        l = [i[0],i[1], '0.06', expert[i[1] >= 0.06]]
        ao.append(l)
    return ao

anal_head_1_5_str = ['столбцы', 'pval [welch_test]', 'alpha', 'ВЫВОД']
def anal_1_5(ql):
    o_1_5 = out_1_5(ql)
    ao = []
    for i in o_1_5:
        l = [i[0],i[1], '0.06', expert[i[1] >= 0.06]]
        ao.append(l)
    return ao

anal_head_2_1_str = ['столбцы', 'F_NM', 'F(alpha/2,k_1,k_2)', 'ВЫВОД']
def anal_2_1(ql):
    z = stats.f.ppf(0.97,N-1,N-1)
    o_2_1 = out_2_1(ql)

```

```

ao = []
for i in o_2_1:
    l = [i[0], i[5], z, expert[i[5] <= z]]
    ao.append(l)
return ao

```

```

anal_head_2_2_str = ['pval[bartlett_test]', 'alpha', 'вывод']
def anal_2_2(ql, alpha = 0.06):
    pv = out_2_2(ql)
    return [[pv, '0.06', expert[pv >= 0.06]]]

```

```

anal_head_2_3_str = ['строка', 'pval[bartlett_test]', 'alpha', 'вывод']
def anal_2_3(ql, alpha = 0.06):
    o_2_3 = out_2_3(ql)
    ao = []
    for i in o_2_3:
        l = [i[0], i[1], '0.06', expert[i[1] >= 0.06]]
        ao.append(l)
    return ao
print('N1')
print(anal_1_1(ql))
print(anal_1_2(ql))
print(anal_1_3(ql))
print(anal_1_4(ql))
print(anal_1_5(ql))
print('N2')
print(anal_2_1(ql))
print(anal_2_2(ql))
print(anal_2_3(ql))

```

```

def stick(anal):#Sick of Truth
    g = True
    for i in anal:
        g = g and (i[len(i)-1] == 'верно')
    return g

```

```

def anal_1_1_str(ql):
    s = {

```

True : 'Гипотеза о равенстве математических ожиданий не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве математических ожиданий противоречит экспериментальным данным (не может быть принята) при уровне значимости $\square \square 0,06$.'

```

    }
    g = True
    N = len(ql[0])
    for i in range(3):
        for k in range(i):
            g = g and (abs(TNM(ql[k],ql[i])) <= stats.t.ppf(0.97,2*(N-1)))
    return [s[g]]

```

```

def anal_1_2_str(ql):
    m = len(ql)-1
    n = len(ql)*(ql[0].size-1)
    z = stats.f.ppf(0.94,m,n)
    f = FNM_1(ql)
    s1 = 'z_alpha = ' + str('%0.5f' % z)
    s2 = 'F_NM = ' + str('%0.5f' % f)
    s3 = {

```

True : 'Гипотеза о равенстве математических ожиданий трёх случайных величин, выборки которых находятся в столбцах массива $U=\{u_{i,j}|1 \leq i \leq N, 1 \leq j \leq 3\}$, не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве математических ожиданий трёх случайных величин, выборки которых находятся в столбцах массива $U=\{u_{i,j}|1 \leq i \leq N, 1 \leq j \leq 3\}$, противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```

    }
    return [s1,s2,s3[f<=z]]

```

```

def anal_1_3_str(ql):
    ao = anal_1_3(ql)
    s1 = 'pval = anova(U) = ' + str('%0.5f' % ao[0][0])
    if ao[0][0] > 0.06:
        s1 = s1 + '>'
    if ao[0][0] == 0.06:
        s1 = s1 + '='
    if ao[0][0] < 0.06:
        s1 = s1 + '<'
    s1 = s1 + 'alpha = 0.06'
    s2 = {

```

True : 'Гипотеза о равенстве математических ожиданий трёх случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве математических ожиданий трёх случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```
}
return [s1, s2[ao[0][0] > 0.06]]
```

```
def anal_1_4_str(ql):
```

```
    s = {
```

True : 'Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False: 'Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```
    }
    g = True
    for i in out_1_4(ql):
        g = g and (i[1] >= 0.06)
    return [s[g]]
```

```
def anal_1_5_str(ql):
```

```
    s = {
```

True : 'Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False: 'Гипотеза о равенстве математических ожиданий соответствующих пар случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```
    }
    g = True
    for i in out_1_5(ql):
        g = g and (i[1] >= 0.06)
    return [s[g]]
```

```
def anal_2_1_str(ql):
```

```
    s = {
```

True : 'Гипотеза о равенстве дисперсий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве дисперсий соответствующих пар случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```

}
g = True
for i in anal_2_1(ql):
    g = g and (i[3] == 'верно')
return [s[g]]

```

```

def anal_2_2_str(ql):
    ao = anal_2_2(ql)
    s1 = 'pval = bartlett_test (X, Y, Z) = ' + str('%0.5f' % ao[0][0])
    if ao[0][0] > 0.06:
        s1 = s1 + '>'
    if ao[0][0] == 0.06:
        s1 = s1 + '='
    if ao[0][0] < 0.06:
        s1 = s1 + '<'
    s1 = s1 + '0.06'
    s2 = {

```

True : 'Гипотеза о равенстве дисперсий трёх случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве дисперсий трёх случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```

}
return [s1,s2[ao[0][2]=='верно']]

```

```

def anal_2_3_str(ql):

```

```

    s = {

```

True : 'Гипотеза о равенстве дисперсий соответствующих пар случайных величин не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,06$.'

False : 'Гипотеза о равенстве дисперсий соответствующих пар случайных величин противоречит экспериментальным данным (не может быть принята) при уровне значимости $\alpha = 0,06$.'

```

}
g = True
for i in anal_2_3(ql):
    g = g and (i[3]=='верно')
return [s[g]]

```

```

print(anal_1_1_str(ql))
print(anal_1_2_str(ql))
print(anal_1_3_str(ql))
print(anal_1_4_str(ql))
print(anal_1_5_str(ql))
print(anal_2_1_str(ql))

```



```

print(anal_2_2_str(ql))
print(anal_2_3_str(ql))
print()

```

```

def out_1_2_str(ql):
    o = out_1_2(ql)
    l = []
    for i in range(len(o)):
        s = head_1_2_str[i] + ' = '
        if type(o[i]) != int:
            s = s + str('%0.5f' % o[i])
        else:
            s = s + str(o[i])
        l.append(s)
    return l

```

```

for i in out_1_2_str(ql):
    print(i)

```

```

def strm(a):
    n = len(a)
    b = a[0]
    m = len(b)
    for i in range(n):
        for k in range(m):
            if type(a[i][k]) == int:
                a[i][k] = str(a[i][k])
            elif type(a[i][k]) != str:
                a[i][k] = str('%0.5f' % a[i][k])

```

```

def tabler(document, out, head = None):
    tl = []
    if head != None:
        tl.append(head)
    for i in out:
        tl.append(i)
    strm(tl)

```

```

table = document.add_table(rows = len(tl), cols = len(tl[0]))
for i in range(len(tl)):
    hdr_cells = table.rows[i].cells
    for k in range(len(tl[0])):
        if type(tl[i][k]) == str:
            hdr_cells[k].text = tl[i][k]

```

```

else:
    pass #LaTeX

def stringer(document, strs):
    for i in strs:
        document.add_paragraph(i)

def docker(ql, filename):
    document = Document()
    document.add_paragraph('Результаты расчетов')

    document.add_paragraph('Данный двумерный массив:')
    qh = [[ql[i][k] for i in range(len(ql))] for k in range(len(ql[0]))]
    tabler(document, qh)
    document.add_paragraph("")

    document.add_paragraph('Задание 1: гипотеза о равенстве
математических ожиданий')
    document.add_paragraph('Пункт 1')
    tabler(document, out_1_1(ql), head_1_1_str)
    document.add_paragraph("")
    document.add_paragraph('Пункт 2')
    #tabler(document, [out_1_2(ql)], head_1_2_str)
    stringer(document, out_1_2_str(ql))
    document.add_paragraph("")
    document.add_paragraph('Пункт 3')
    document.add_paragraph('pval = anova(U) = ' + str("%.5f"%
out_1_3(ql)))
    document.add_paragraph("")
    document.add_paragraph('Пункт 4')
    tabler(document, out_1_4(ql), head_1_4_str)
    document.add_paragraph("")
    document.add_paragraph('Пункт 5')
    tabler(document, out_1_5(ql), head_1_5_str)
    document.add_paragraph("")
    document.add_paragraph("")
    document.add_paragraph("")

    document.add_paragraph('Задание 2: гипотеза о равенстве
дисперсий')
    document.add_paragraph('Пункт 1')
    tabler(document, out_2_1(ql), head_2_1_str)
    document.add_paragraph("")
    document.add_paragraph('Пункт 2')

```

```

        document.add_paragraph('pval = bartlett_test (X, Y, Z ) =' +
str('%0.5f% out_2_2(ql)))
        document.add_paragraph("")
        document.add_paragraph('Пункт 3)')
        tabler(document, out_2_3(ql), head_2_3_str)
        document.add_paragraph("")
        document.add_paragraph("")
        document.add_paragraph("")
        document.add_paragraph("")
        document.add_paragraph("")

document.add_paragraph('Анализ результатов и выводы:')
document.add_paragraph('Задание 1')
document.add_paragraph('Пункт 1)')
tabler(document, anal_1_1(ql), anal_head_1_1_str)
stringer(document, anal_1_1_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 2)')
tabler(document, anal_1_2(ql), anal_head_1_2_str)####
stringer(document, anal_1_2_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 3)')
tabler(document, anal_1_3(ql), anal_head_1_3_str)####
stringer(document, anal_1_3_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 4)')
tabler(document, anal_1_4(ql), anal_head_1_4_str)
stringer(document, anal_1_4_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 5)')
tabler(document, anal_1_5(ql), anal_head_1_5_str)
stringer(document, anal_1_5_str(ql))
document.add_paragraph("")
document.add_paragraph("")
document.add_paragraph("")

document.add_paragraph('Задание 2')
document.add_paragraph('Пункт 1)')
tabler(document, anal_2_1(ql), anal_head_2_1_str)
stringer(document, anal_2_1_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 2)')
tabler(document, anal_2_2(ql), anal_head_2_2_str)####
stringer(document, anal_2_2_str(ql))
document.add_paragraph("")
document.add_paragraph('Пункт 3)')

```

```
tabler(document,anal_2_3(ql),anal_head_2_3_str)  
stringer(document,anal_2_3_str(ql))  
document.add_paragraph("")
```

```
document.save(filename + '.docx')
```

```
print('Done!')
```

```
docker(ql, 'result')
```