

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического образования

Кафедра компьютерные технологии и электронного обучения

Основная профессиональная образовательная программа

Направление подготовки 09.03.01 Информатика и вычислительная техника

Направленность (профиль) «Технологии разработки программного
обеспечения»

форма обучения – очная

ЛАБОРАТОРНАЯ РАБОТА №6.2

по дисциплине: «Анализ данных и основы Data science»

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Руководитель:

кандидат педагогических наук, доцент,

Светлана Викторовна Гончарова

Автор работы студент 2 курса

1 группы 1 подгруппы

Чирцов Тимофей Александрович

Цель работы: проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

Оборудование: ПК, replit (python, C++)

Задание 1

По результатам $n=9$ замеров установлено, что выборочное среднее время 1 секунду подготовки детали $\bar{x} = 48$. Предполагая, что время приготовления нормально распределенная случайная величина с дисперсией $\sigma^2 = 9$, рассмотрим на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы: <https://replit.com/@mshestak2016/zadan1>

```
import math
l=float(0.95);
x=48;
n=9;
o=9;
a=int(input('Введите чему равно a = '));
t=((x-a)*math.sqrt(n))/math.sqrt(o);
print(t)
tkr=1.96
if abs(t)<tkr:
    print('гипотеза принимается')
else:
    print('гипотеза не принимается')
```

Результат программы:

```
Введите чему равно a = 49
-1.0
гипотеза принимается
```

В данной формуле мы заметили неточность, так как $\sigma^2 = 9$, то $\sigma = \pm 3$.

1) $a = 49$
 $\sigma = 0.95$
 $\bar{x} = 48$
 $n = 9$
 $\sigma^2 = 9$

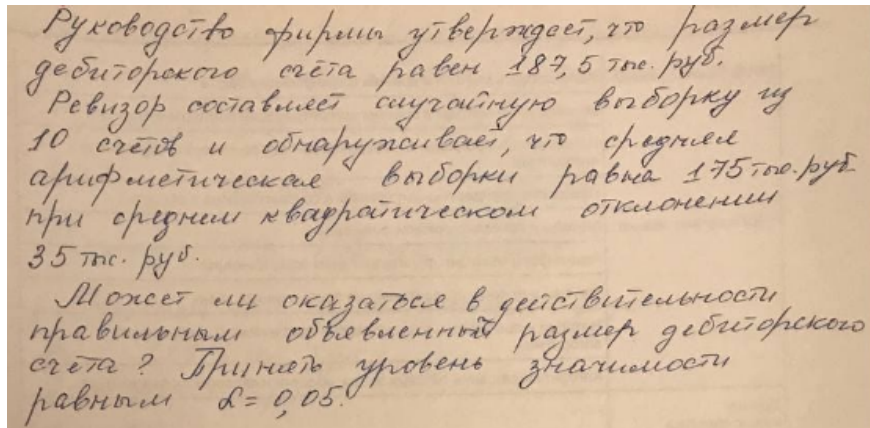
$$t = \frac{(48 - 49) \sqrt{9}}{3} = -1$$

Работаем с таблицей «Значения функции Лапласа» $\gamma = 0,95$ – ищем в таблице соответствует $t_{кр} = 1,96$.

Анализируем и делаем вывод

$|t| < t_{кр}$ ($|-1| < 1,96$) \Rightarrow гипотеза принимается.

Задание 2



Код программы: <https://replit.com/@mshestak2016/zadan2>

```
import math
l=0.05;
ll=0.95;
x=175;
a=187.5;
n=10;
tkr=2.26;
S=35;
t=((x-a)*math.sqrt(n))/S
if abs(t)<tkr:
    print('гипотеза H0 о среднем размере дебиторского счёта принимается на
    уровне доверия y=0.95')
else:
    print('гипотеза H0 о среднем размере дебиторского счёта не принимается на
    уровне доверия y=0.95')
print(t)
```

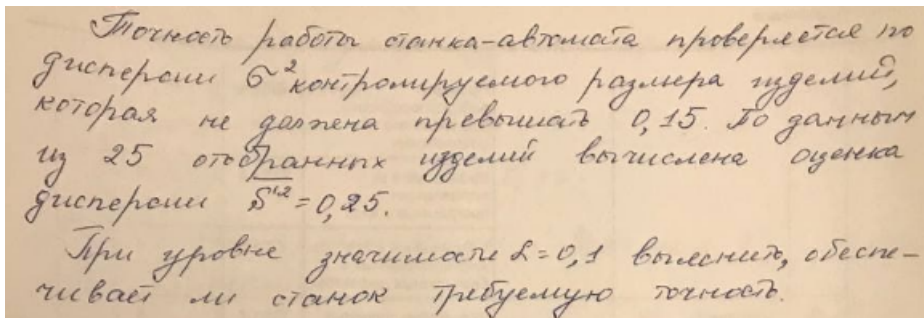
Результат программы:

```
гипотеза H0 о среднем размере дебиторского счёта принимается на уровне доверия y=0.
-1.1293848786315641
```

Работаем с таблицей «Значения t – критерия Стьюдента»). По таблице при $\gamma = 0,95$ находим $t_{кр}$ для числа степеней свободы $(n-1) = 9$. $t_{кр9,095} = 2,26$.

Анализируем и принимаем решение. Так как $|t| = 1,129 < t_{9, кр}$, то гипотеза H_0 о среднем размере дебиторского счета принимается на уровне доверия $\gamma = 0,95$

Задание 3



Код программы: <https://replit.com/@mshestak2016/zadan3>

```
o=0.15
n=24
S=0.25
tkr=15.7
t=((n-1)*S)/o;
print(t)
if t>tkr:
    print('гипотеза о достижении требуемой точности отклоняется')
else:
    print('гипотеза о достижении требуемой точности принимается ')
```

Результат программы:

```
38.333333333333336
гипотеза о достижении требуемой точности отклоняется
>
```

Анализ и принятие решения. Так как $t = 38.333 > t_{кр} = 15,7$, то гипотеза о достижении требуемой точности отклоняется.

Задание 4

Расход сырья x_i и y_i на единицу продукции по старой и новой технологиям приведен в таблице 1:

	По старой технологии				По новой технологии				
Расход сырья	x_i	304	307	308	y_j	303	304	306	308
Число изделий	n_i	1	4	4	n_j	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними μ_1 и μ_2 .

Требуется проверить гипотезу $H_0: \mu_1 = \mu_2$ против гипотезы $H_1: \mu_1 \neq \mu_2$ на уровне значимости $\alpha = 0.1$.

Код программы: <https://replit.com/@mshestak2016/zadan4>

```
#include <iostream>
#include <math.h>
using namespace std;

int main()
{
    setlocale(LC_ALL, "Russian");
    float t, sum2y = 0, sum2x = 0, s2y, s2x, x[3] = { 304, 307, 308 }, y[4] =
    {303,304,306,308}, ni[3] = {1, 4, 4}, nj[4] = {2,6,4,1}, nx = 9, ny = 13, sumx = 0,
    sumy = 0, x_op, y_op;
    int i, j;
    for (i = 0; i < 3; i++) {
        sumx += x[i] * ni[i];
    }

    x_op = 1/nx * sumx;
    cout << x_op;
    cout << "\n";
    for (i = 0; i < 3; i++) {
        sum2x += (x[i]-x_op)*(x[i]-x_op);
    }
    s2x = (1 / (nx-1))* sum2x;
    cout << s2x;
    cout << "\n";
    for (j = 0; j < 4; j++) {
        sumy += y[j]*nj[j];
    }
    y_op = 1 / ny * sumy;
    cout << y_op;
```

```

cout << "\n";
for (j = 0; j < 4; j++) {
    sum2y += (y[j] - y_op)*(y[j] - y_op);
}
s2y = (1 / (ny-1)) * sum2y;
cout << s2y;
cout << "\n";
t = (x_op - y_op) / sqrt((nx * s2x + ny * s2y) / (nx + ny - 2) * (1 / nx + 1 / ny));
cout << t;
cout << "\n";
}

```

Результат программы:

```

307.111
1.31019
304.769
1.3062
4.50266

```

Анализируем и принимаем решение. Так как $|t| > t_{20, \text{кр}}$, то гипотезу H_0 отвергаем. Это значит, что при переходе на новую технологию происходит изменение среднего расхода сырья.

Вывод по лабораторной работе: с помощью программного кода нам удалось проверить статистическую гипотезу о нормальном законе распределения данных.