

**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 №6.3

по дисциплине: «Анализ данных и основы Data science»

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Руководитель:

кандидат педагогических наук, доцент,

Светлана Викторовна Гончарова

Автор работы студент 2 курса

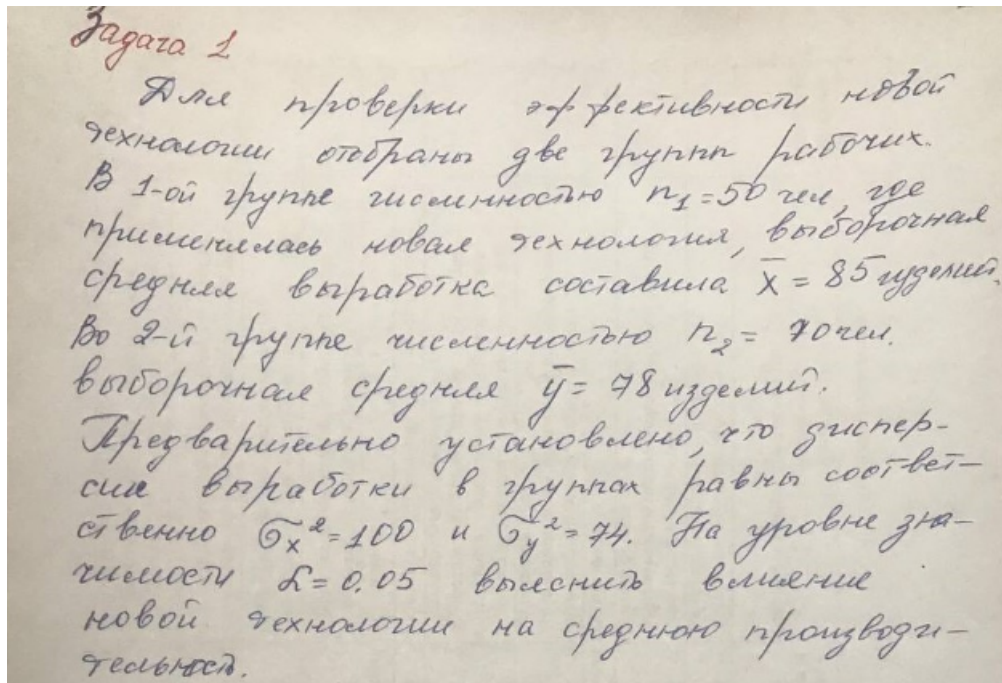
1 группы 1 подгруппы

Чирцов Тимофей Александрович

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

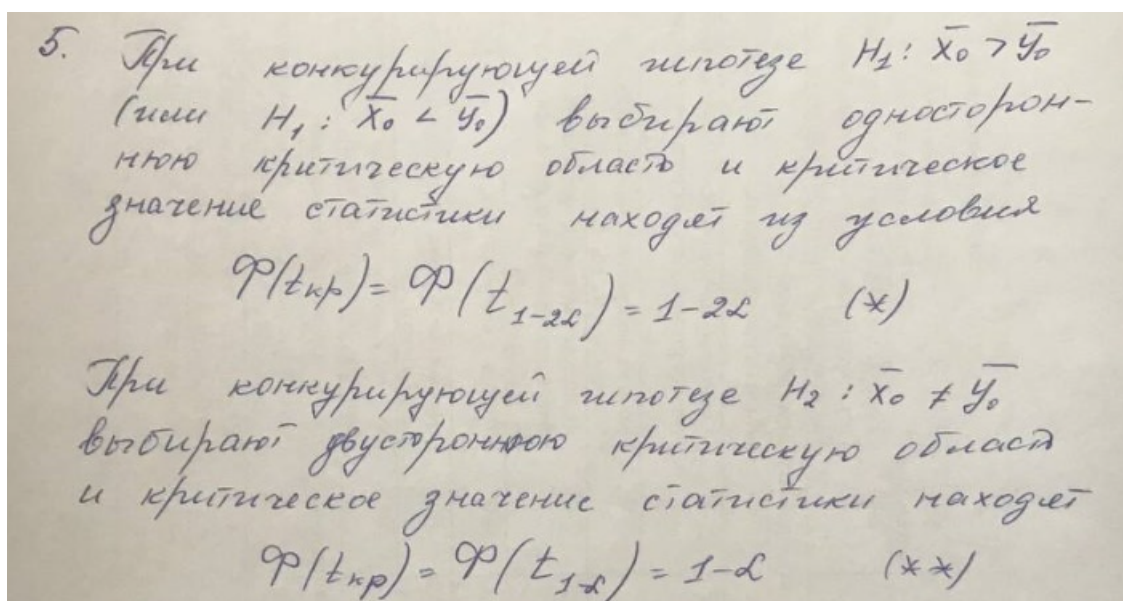
Оборудование: ПК, replit (python)

Задание 1



Математическая модель:

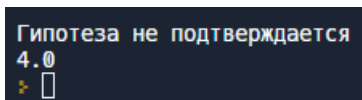
$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$



Код программы: <https://replit.com/@mshestak2016/lab63zadan1#main.py>

```
import math
x=85
y=78
o1=100
o2=74
n1=50
n2=70
l=float(0.05)
t=float(round((x-y)/(math.sqrt((o1/n1)+(o2/n2)))))
f1=1-2*1
t1=t/100
tkr1=t1+1.6
f2=1-1
tkr2=1.96
if t<tkr1 or t<tkr2:
    print('Гипотеза подтверждается')
else:
    print('Гипотеза не подтверждается')
print(t)
```

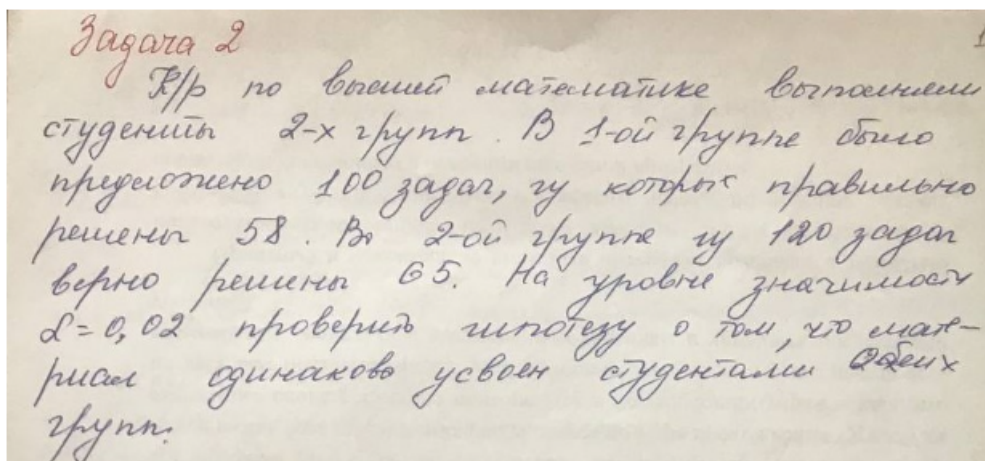
Результат программы:



```
Гипотеза не подтверждается
4.0
```

Вывод: так как фактически наблюдаемое значение $t = 4.0 > t_{кр}$ (при любой из взятых конкурирующих гипотез, то гипотеза H_0 отвергается, то есть на 5%-ном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

Задание 2



Задача 2

Р/р по внешней математике составили студентов 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решено 58. В 2-ой группе из 120 задач верно решено 65. На уровне значимости $\alpha = 0,02$ проверить гипотезу о том, что математическое ожидание усвоения студентами обеих групп.

Математическая модель:

$$w_1 = \frac{m_1}{n_1}$$
$$w_2 = \frac{m_2}{n_2}$$
$$p = \frac{m_1 + m_2}{n_1 + n_2}$$
$$t = \frac{w_1 - w_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Код программы: <https://replit.com/@mshestak2016/lab63zadan2#main.py>

```
import math
m1=58
n1=100
m2=65
n2=120

w1=round(m1/n1, 3)
w2=round(m2/n2, 3)

p=round((m1+m2)/(n1+n2), 3)

t=round((w1-w2)/(math.sqrt(p*(1-p)*(1/n1+1/n2))), 3)

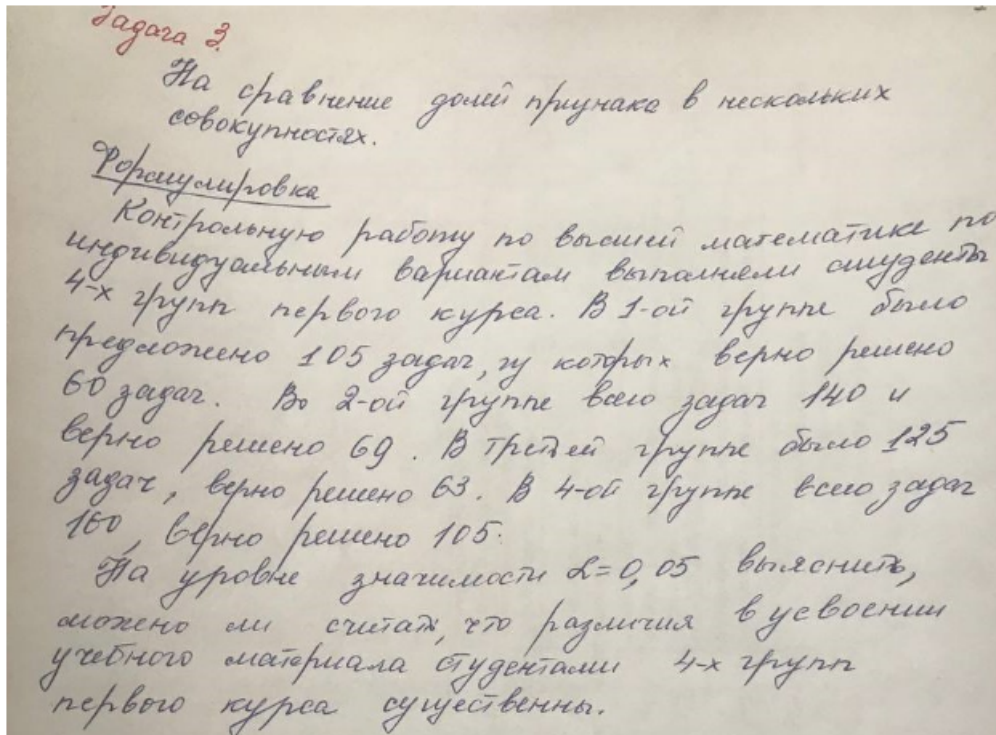
tkr=2.3
if abs(t)<tkr:
    print(t, 'Гипотезу H0 об одинаковой усваиваемости материала принимается ')
else:
    print('Гипотезу H0 об одинаковой усваиваемости материала не принимается ')
    print('')
```

Результат программы:

```
Гипотезу H0 об одинаковой усваиваемости материала принимается
✦ □
```

Вывод: $t = 0.565$. Для уровня значимости $\alpha = 0,02$ ($\gamma = 0,98$) по таблицам находим $t_{кр} = 2,33$. Так как $|t| < t_{кр}$, то гипотезу H_0 об одинаковой усваиваемости материала принимаем.

Задание 3



Математическая модель:

$$w_i = \frac{m_i}{n_i}, \quad w_1 = \frac{m_1}{n_1}, \dots, w_l = \frac{m_l}{n_l},$$

где m_i - число элементов i -ой выборки ($i=1, 2, \dots, l$), обладающих данными признаками

$$\chi^2 = \frac{1}{p(1-p)} \sum_{i=1}^l n_i (w_i - p)^2$$

$$p = \frac{\sum_{i=1}^l m_i}{\sum_{i=1}^l n_i}$$

Код программы: <https://replit.com/@mshestak2016/lab63zadan3#main.py>

```
import math
```

```
m1=60
```

```
m2=69
```

```
m3=63
```

```
m4=105
```

```

m=[m1,m2,m3,m4]

n1=105
n2=140
n3=125
n4=160
n=[n1,n2,n3,n4]

p=round(sum(m)/sum(n),2)

w1=round(m1/n1, 2)
w2=round(m2/n2, 2)
w3=round(m3/n3, 2)
w4=round(m4/n4, 2)
w=[w1,w2,w3,w4]

y=0;

for i in range(0,4):

    y=round(n[i]*(w[i]-p)**2,2)+y

x=round(1/(p*(1-p))*y, 2)

x005=7.82

if x>x005:
    print('гипотеза Н0 отвергается')
else:
    print('гипотеза Н0 принимается')

```

Результат программы:

```

гипотеза Н0 отвергается
>

```

Вывод: по таблице «Значения критерия Пирсона» $X^2_{0.05; 2} = 7.82$. Так как $X^2 > X^2_{0.05; 3}$, то гипотеза H_0 отвергается, то есть различие в усвоении уч. Материала студентами 4-х групп значимо или существенно на уровне $\alpha = 0,05$

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