# 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

Задага 1

Вле проверки эфрекливность новой чехнологии отобрано две группп робочих. В 1-ой группе гистиностью 
$$n_1=50$$
 чем гре приминением выработка составшия  $\chi=85$  премий средний выработка составшия  $\chi=85$  премий выборогной гредний. Во 2-й группе мистемностью  $r_2=70$  чем. выборогной средний  $y=78$  издений. Предваритемно установыемо, го дисперсие выработки в группом равны соответся выно  $G_{\chi}^2=100$  и  $G_{\chi}^2=74$ . На уровне значими  $\chi=100$  и  $\chi=10$ 

### Математическая модель:

5. При конкурирующей пилотеде 
$$H_2$$
:  $X_0$   $7$   $\frac{1}{90}$  (пен  $H_1$ :  $X_0$   $2$   $\frac{1}{90}$ ) выбирают односторон-
нюю критическую область и критическое 
значение статистики находет из усмовия

 $P(t_{kp}) = P(t_{1-2k}) = 1-2k$  (\*)

Три конкурирующей шпотеде  $H_2$ :  $X_0 \neq Y_0$ 
выбирают двустронного критическую област и критическое значение статистики находет

 $P(t_{kp}) = P(t_{1-k}) = 1-k$  (\*\*)

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

```
import math
x = 85
y = 78
o1 = 100
02 = 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
⊱ []
```

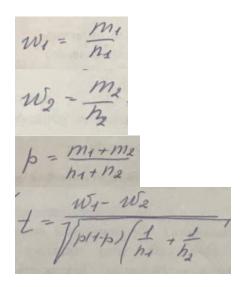
Вывод: так как фактически наблюдаемое значение  $t=4.0>t_{\kappa p}$  (при любой из взятых конкурирующих гипотез, то гипотеза  $H_0$  отвергается, то есть на 5%ном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

## Задание 2

```
Задача 2

E/p по воений мажилите вогносинем студений 2-х групп В 1-ой группе общо приденопиено 100 задах, гу которых правиньно решени 58. Во 2-ой группе ну 120 задах верню решено 65. На уровне значиност \mathcal{L}=0,02 проверию инпотеду о том, что мажрись одинаново усвоен студентамии общем групп.
```

# Математическая модель:



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 ')

# Результат программы:

Вывод: t = 0.565. Для уровня значимости  $\alpha = 0.02$  ( $\gamma = 0.98$ ) по таблицам находим  $t_{\kappa p} = 2,33$ . Так как  $|t| < t_{\kappa p}$ , то гипотезу  $H_0$  об одинаковой усваиваемости материала принимаем.

### Задание 3

```
Задага 3.

Теранулировка

Контрономую работу по волоний мачематики по индивидуамыми вариай ам выпамени ошудент тервого куреа. В 3-ой группи были тредеотеено 105 задаг, пу котрых верно решено 60 задаг. Во 2-ой группи вено задаг 140 и верно решено 69. В Трийей группи было 125 задаг, верно решено 63. В 4-ой группи всего задаг 160 и верно решено 63. В 4-ой группи всего задаг 160 верно решено 63. В 4-ой группи всего задаг 160 верно решено 63. В 4-ой группи всего задаг 160 верно решено 63. В 4-ой группи всего задаг 160 верно решено 105.

То уровне значимости в 20,05 выпесных, агожено им считах, что размичих в усвоении учетного материала от удентыми 4-х группи первого куреа существенных.
```

### Математическая модель:

$$W_1 = \frac{m_1}{n_2}$$
,  $W_2 = \frac{m_2}{n_2}$ ,  $W_\ell = \frac{m_\ell}{n_\ell}$ ,

 $v_{ij} = m_i - v_{ij} = v_{ij}$ ,  $v_{ij} = v_{ij}$ ,

 $v_{ij} = v_{ij}$ ,  $v_{ij}$ 

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

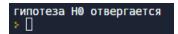
# import math

m1 = 60m2 = 69m3 = 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 H0 отвергается')
else:
 print('гипотеза H0 принимается')
```

Результат программы:



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  $\alpha=0.05$ 

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