Гаязов Ильназ Дамирович, группа 09-012

Задание 2, раздел 1

*Проверка статистических гипотез*

Имеется две выборки из нормального распределения (i-ые значения которых считаются связанными) и дополнительные параметры: требуемый уровень значимости α = 0,025 и вид нулевой гипотезы ­*К: увеличится*.

1. Предположение о выборках, при которых возможно применение критерия:

Рассматриваем в качестве исследуемой выборки Ui = X­i – Yi являющееся реализацией случайной величины U = X-Y, где Xi и Yi соответствуют связанным значениям из первой и второй выборки. Предполагается, что i-ые наблюдения Xi, Yi в выборках являются наблюдениями характеристик одного и того же i-го объекта. X, Y имеют нормальное распределение с математическими ожиданиями µ1, µ2. Тогда, распределение U нормальное с математическим ожиданием µ­1 - µ2.

Таким образом, можем применить одновыборочный критерий Стьюдента.

Проверим гипотезу, что матожидание случайной величины U –

К0: µ1 - µ­­2 = 0

Возьмём в качестве альтернативы К1 = К:

К1: µ1 - µ­­2 <0

2. Вид критической области, значение критической константы.

Значение критической константы равно Cкрит = -1.996.

Критическая область: {T: T <-1.996}.

3. Вычисление статистики и вывод о принятии гипотезы:

Статистика Стьюдента: T = -29.09

**Вывод**: так как T <Скрит, следовательно, на уровне значимости α гипотеза К0 отвергается.

Таким образом, математическое ожидание распределения второй выборки µ­2 увеличится по сравнению с µ1.

4. p – значение.

p = 4.932837558499242e-40

5. Код на Python.

*# импорт библиотеки для работы с Excel***import** openpyxl  
**import** math  
**from** scipy.stats **import** t  
*# открыть файл для чтения*book = openpyxl.open(**'r2z1.xlsx'**, read\_only=**True**)  
sheet = book.active  
*# коприуем данные из листа в список*data = []  
**for** row **in** range(2, sheet.max\_row + 1):  
 data.append(round(sheet[row][0].value - sheet[row][1].value, 2))  
  
*# Student's statistic*T, m0 = 0, 0  
sum, sum2 = 0, 0  
n = len(data)  
**for** x **in** data:  
 sum += x  
 sum2 += x\*\*2  
Sx = math.sqrt((sum2/n - (sum/n) \*\* 2))

T = (sum/n - m0)\*math.sqrt(n - 1) / Sx  
print(**"Статистика Стьюдента: "** + str(round(T, 2)))  
  
alpha = 0.025  
*# calculate the critical value*cv = t.ppf(alpha, n - 1)  
*# calculate the p-value*p = t.cdf(T, n - 1)  
print(**"The critical value "**, round(cv, 3))  
print(**"The p-value "**, p)