**Домашнее задание 5. Тестирование гипотез**

**Задача 1.**

*Когда используется критерий Стьюдента, а когда Z –критерий?*

**Критерий Стьюдента используется:**

1. Когда стандартное отклонение популяции неизвестно и должно быть оценено по образец данные.

2. Для сравнения средних значений двух групп.

3. Для сравнения средних значений двух выборок, если данные для двух групп распределены нормально.

**Закон Z-критерия используется:**

1. Когда известны среднее значение совокупности и стандартное отклонение, предполагается, что население распределено нормально

2. Для проверки гипотезы о равенстве среднего значения генеральной совокупности с известным значением.

3. Для проверки гипотезы о когда известны среднее значение совокупности и стандартное отклонение, что позволит оценить точность среднего значения измерения в генеральной совокупности, а также понять, насколько данные в выборке представляют собой нормальное распределение.

**Задачи 2,3 решать вручную**

**Задача 2.**

*Проведите тест гипотезы. Утверждается, что шарики для подшипников, изготовленные автоматическим станком, имеют средний диаметр 17 мм. Используя односторонний критерий с α=0,05, проверить эту гипотезу, если в выборке из n=100 шариков средний диаметр оказался равным 17.5 мм, а дисперсия известна и равна 4 кв. мм.*

*1. Определяем гипотезу:*

Основная гипотеза:

Альтернативная гипотеза:

ПКО:

*2. Определяем наблюдаемый критерий:*

Нам известны и среднее и дисперсия генеральной совокупности, соответственно для проверки гипотезы будем использовать z-критерий.

Дисперсия:

Среднее квадратическое отклонение генеральной совокупности:

Размер выборки: n=100

**import** numpy **as** np

**import** scipy.stats **as** stats

mu **=** 17.5

mu0 **=** 17

sigma **=** 2

n **=** 100

zn **=** (mu **-** mu0) **/** (sigma **/** np**.**sqrt(n))

print(f'Наблюдаемое значение z-критерия: {round(zn, 4)}')

Наблюдаемое значение z-критерия: 2.5

*3. Определяем уровень значимости:*

*4. Определяем табличный критерий:*

alpha = 0.05

zt = stats.norm.ppf(1 - alpha)

print(f'Табличное значение z-критерия: {round(zt, 4)}')

Табличное значение z-критерия: *1.6449*

zn < zt

False

*5. Так как , то нулевая гипотеза отвергается.*

**Задача 3.**

*Проведите тест гипотезы. Продавец утверждает, что средний вес пачки печенья составляет 200 г. Из партии извлечена выборка из 10 пачек. Вес каждой пачки составляет: 202, 203, 199, 197, 195, 201, 200, 204, 194, 190. Известно, что их веса распределены нормально. Верно ли утверждение продавца, если учитывать, что доверительная вероятность равна 99%? (Провести двусторонний тест.)*

*1. Определяем гипотезу:*

Основная гипотеза:

Альтернативная гипотеза:

ДКО:

*2. Определяем наблюдаемый критерий:*

Нам неизвестна дисперсия и мы проверяем гипотезу относительно математического ожидания нормально распределённой случайной величины будем использовать t-критерий.

Размер выборки: n=10

**import** numpy **as** np

**import** scipy.stats **as** stats

arr = np.array([202, 203, 199, 197, 195, 201, 200, 204, 194, 190])

Расчитаем наблюдаемое значение t-критерия :

n = 10

mu = 200

tn = (np.mean(arr) - mu) / (np.std(arr, ddof = 1) / n\*\*0.5)

print(f'Наблюдаемое значение t-критерия: {round(tn, 4)}')

Наблюдаемое значение t-критерия: -1.0651

*3. Определяем уровень значимости:*

*4. Определяем табличный критерий:*

alpha = 0.005

tt = stats.t.ppf(1 – alpha, n-1)

print(f'Табличное значение t-критерия: {round(tt, 4)}')

Табличное значение t-критерия: *3.2498*

np.abs(tn) < tt

True

*5. Так как , то нулевая гипотеза не отвергается.*

**Задачу 4 решать с помощью функции**

**Задача 4.**

*Есть ли статистически значимые различия в росте дочерей?*

*Рост матерей 172, 177, 158, 170, 178,175, 164, 160, 169, 165.*

*Рост взрослых дочерей: 173, 175, 162, 174, 175, 168, 155, 170, 160.*

**import** scipy.stats **as** stats

arr1 **=** [172, 177, 158, 170, 178, 175, 164, 160, 169, 165]

arr2 **=** [173, 175, 162, 174, 175, 168, 155, 170, 160, 163]

stats**.**ttest\_rel(arr1, arr2)

TtestResult(statistic=0.6648478531431979, pvalue=0.5228168632983574, df=9)

Так как , то нулевая гипотеза не отвергается на уровне значимости .

Статистически значимых различий в среднем росте матерей и дочерей нет.