

Корреляционный анализ

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»



Проведено 4 эксперимента. Для каждого построить график и посчитать коэффициенты корреляции Пирсона.

```
x = np.array([10,8, 13, 9,11,14, 6,4,12, 7,5])  
y = np.array([8.04, 6.95, 7.58, 8.81, 8.33, 9.96, 7.24, 4.26, 10.84, 4.82, 5.68 ])
```

```
x= np.array([ 10,8, 13, 9,11,14, 6,4,12, 7,5 ])  
y2 = np.array([ 9.14, 8.14, 8.74,8.77, 9.26, 8.10, 6.13, 3.10, 9.13, 7.26, 4.74])
```

```
x= np.array([ 10,8, 13, 9,11,14, 6,4,12, 7,5 ])  
y3 = np.array([7.46,6.77, 12.74, 7.11, 7.81, 8.84, 6.08, 5.39, 8.15, 6.42, 5.73])
```

```
x4 = np.array([8, 8, 8, 8, 8, 8, 8, 19, 8, 8, 8])  
y4 = np.array([6.58, 5.76, 7.71, 8.84, 8.47, 7.04, 5.25,12.5, 5.56, 7.91, 6.89])
```

```
x0= np.array([ 10, 8, 13, 9, 11, 14, 6, 4, 12, 7,5, 15, 16, 18 ])  
y0 = np.array([ 9.14, 8.14, 8.74,8.77, 9.26, 8.10, 6.13, 3.10, 9.13, 7.26, 4.74, 6.5, 5, 2.9])
```

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»



```
import scipy.stats as stats  
import numpy as np  
import matplotlib.pyplot as plt
```

```
plt.scatter(x, y)  
plt.title('r= вписать значение коэффициента')  
plt.xlabel('x')  
plt.ylabel('y')  
plt.show()
```



Задача 2

Провести двусторонний тест и ответить на вопрос, есть ли статистически значимые различия между средними 2х нормально распределенных генеральных совокупностей, представленных следующими независимыми выборками:

```
a = np.array([12, 10, 11, 19, 13, 11, 17, 15, 19, 14, 21, 18, 21, 11, 17, 14, 15, 17, 20, 19])
```

```
b = np.array([11, 13, 18, 15, 17, 18, 10, 21, 26, 15, 11, 12, 15, 17, 10, 18, 18, 12, 21, 20])
```

Уровень статистической значимости принять за 5%

1 . Используйте функцию в Python:

2. Имея p-value из функции рассчитать наблюдаемое значение критерия.

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»



```
stats.ttest_ind(a, b, equal_var=True)
```



```
[41] a = np.array([12,10, 11, 19, 13, 11, 17, 15, 19, 14, 21, 18, 21, 11, 17, 14, 15, 17, 20, 19])  
b = np.array([11, 13, 18, 15, 17, 18, 10, 21, 26, 15, 11, 12, 15, 17, 10, 18, 18, 12,21, 20])
```

```
[44] s1=np.std(a, ddof=1)  
s1  
  
3.5850567051229985
```

```
[45] s2 = np.std(b, ddof=1)  
s2  
  
4.290748922483052
```

```
[46] len(a)  
  
20
```

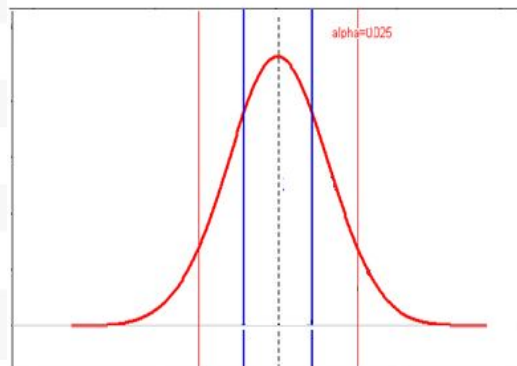
```
[47] len(b)  
  
20
```

```
[48] np.sqrt(s1**2/20+s2**2/20)  
  
1.250263130199736
```

```
(np.mean(a)-np.mean(b))/1.250263130199736  
-0.1599663264228627
```

```
[53] stats.ttest_ind(a,b, equal_var = True)  
  
Ttest_indResult(statistic=-0.1599663264228627, pvalue=0.8737549039369696)
```

```
stats.t.ppf(1-0.8737549039369696/2, 38)  
0.15996632632476554
```





Используя функцию `stats.ttest_ind`, проведите односторонний тест. Проверить, что $\mu_1 > \mu_2$

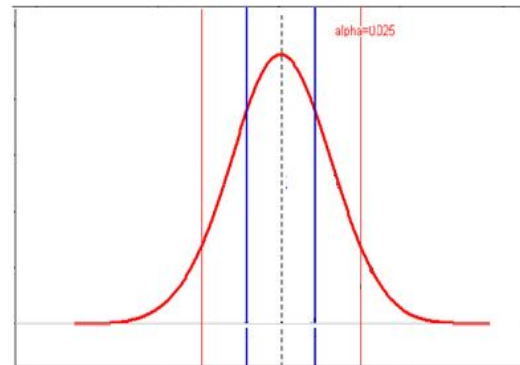
```
stats.ttest_ind(a, b, alternative='greater', equal_var = True)
```

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»

Используя функцию `stats.t.ppf` найдите, расчетное значение Стьюдента
Определите табличное значение критерия Стьюдента





```
stats.t.ppf(1-0.936/2,38)  
0.08083173800421259
```



```
stats.ttest_ind(a, b, alternative='greater', equal_var=True)
```

```
Ttest_indResult(statistic=0.08024979095341923, pvalue=0.4682355453874213)
```



Используя функцию `stats.ttest_ind`, проведите односторонний тест.
Проверить, что $\mu_2 < \mu_1$

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»



```
stats.ttest_ind(a, b, alternative='less', equal_var=True)  
Ttest_indResult(statistic=0.08024979095341923, pvalue=0.5317644546125787)
```



Спортсмен стреляет по мишени, разделенной на три сектора. Вероятность попадания в 1й сектор = 0.4, во второй = 0.3. Какова вероятность попадания в 1 или 2й сектор?

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика



В группе 21 студент: 5 отличников, 10 хорошистов, 6 слабых. На предстоящем экзамене отлично могут получить только отличники, хорошисты с равной вероятностью могут получить хорошо и отлично, слабые –с одинаковой вероятностью могут получить хорошо, удовлетворительно и неуд. На экзамен вызывается наугад 1 студент. Какова вероятность, что он получит хорошо или отлично

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика



Некоторое изделие выпускается 2мя заводами. При этом объем продукции 2го завода в 3 раза превосходит объем продукции 1го. Доля брака у 1го -2%, у 2го -1%. Изделие выпущено заводами за одинаковый промежуток времени. Изделия перемешали и отправили в продажу. Какова вероятность, что приобретенное изделие со 2го завода, если оно бракованное

Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика