

# Тестирование гипотез

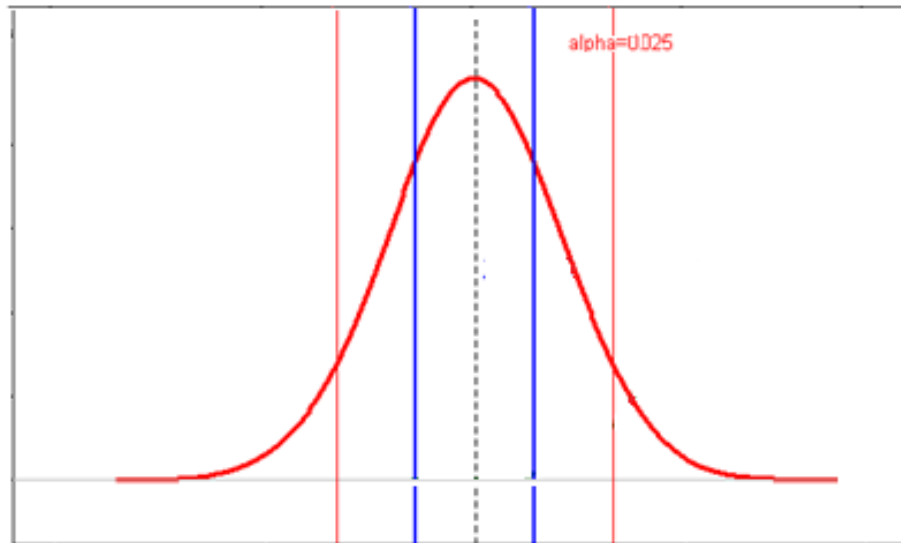
Используемая литература для подготовки задач к семинару:

Гусак А.А. «Теория вероятностей»

Гланц С. «Медико-биологическая статистика»



Показать на графике p value





Показать область ошибки 1го рода, показать область ошибки 2го рода



Ниже приведены диаметры коронарных артерий после приема нифедипина и плацебо. Позволяют ли приводимые ниже данные утверждать, что нифедипин влияет на диаметр коронарных артерий?

1. Выполнить расчеты в Python
2. Сделайте расчет в ручную
3. Сравните критерий Стьюдента и p-value со значениями, полученными в Python

```
x=np.array([2.5, 2.2, 2.6, 2, 2.1, 1.8,2.4, 2.3, 2.7, 2.7, 1.9])  
y= np.array([2.5, 1.7, 1.5, 2.5, 1.4, 1.9, 2.3, 2.0, 2.6, 2.3, 2.2])  
.
```

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Используемая литература для подготовки задач к семинару:  
Гланц С. «Медико-биологическая статистика»

✓ [2] import scipy.stats as stats  
import numpy as np

✓ [7] stats.t.ppf(0.975, 20)  
2.0859634472658364


✓ [8] stats.t.cdf(2.086, 20)  
0.9750018227712799

✓ [14] stats.t.cdf(1.32834847578, 20)  
0.9004886710051916

✓ [17] 1-0.9004886710051916  
0.09951132899480841

✓ [18] 1-2\*0.09951132899480841  
0.8009773420103832

✓ [19] pv=1-0.8009773420103832  
pv  
0.19902265798961682



```
x=np.array([2.5, 2.2, 2.6, 2, 2.1, 1.8, 2.4, 2.3, 2.7, 2.7, 1.9])
x
array([2.5, 2.2, 2.6, 2. , 2.1, 1.8, 2.4, 2.3, 2.7, 2.7, 1.9])

y= np.array([2.5, 1.7, 1.5, 2.5, 1.4, 1.9, 2.3, 2.0, 2.6, 2.3, 2.2])
y
array([2.5, 1.7, 1.5, 2.5, 1.4, 1.9, 2.3, 2. , 2.6, 2.3, 2.2])

import scipy.stats as stats
stats.ttest_ind(x,y)
Ttest_indResult(statistic=1.3283484757831465, pvalue=0.19902265798859647)

d1=np.var(x, ddof=1)
d1
0.10090909090909095

d2 =np.var(y, ddof=1)
d2
0.17163636363636364

m1= np.mean(x)
m1
2.2909090909090906

m2= np.mean(y)
m2
2.0818181818181816

D=0.5*(d1+d2)
D
0.1362727272727273

(m1-m2)/np.sqrt(2*D/len(x))
1.3283484757831465
```



Ниже приведены данные из исследования Фреба и Уайта, посвященному исследованию состоянию легких. Мы возьмем данные для группы людей, которые работали в накурленном помещении и для людей, выкуривающих небольшое число сигарет в день.

Объемы выборок одинаковые – по 200 человек. Для людей, работающих в накурленном помещении средняя скорость середины выдоха составляет 2,72,  $\text{std} = 0.71$ , а выкуривающих небольшое число сигарет 2,63,  $\text{std} = 0.73$ . Можно ли считать одинаковой максимальную объемную скорость середины выдоха одинаковой в обеих группах?

v	Уровень значимости $\alpha$									
	0,5	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001	
1	1,000	3,078	6,314	12,706	31,821	63,656	127,321	318,289	636,578	
2	0,816	1,886	2,920	4,303	6,965	9,925	14,089	22,328	31,600	
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,214	12,924	
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610	
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869	
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959	
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408	
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041	
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781	
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587	
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437	
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318	
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221	
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140	
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073	
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015	
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965	
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922	
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883	
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850	
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819	
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792	
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768	
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745	
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725	
26	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707	
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,689	
28	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674	
29	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,660	
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646	
31	0,682	1,309	1,696	2,040	2,453	2,744	3,022	3,375	3,633	
32	0,682	1,309	1,694	2,037	2,449	2,738	3,015	3,365	3,622	
33	0,682	1,308	1,692	2,035	2,445	2,733	3,008	3,356	3,611	
34	0,682	1,307	1,691	2,032	2,441	2,728	3,002	3,348	3,601	
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591	
36	0,681	1,306	1,688	2,028	2,434	2,719	2,990	3,333	3,582	
37	0,681	1,305	1,687	2,026	2,431	2,715	2,985	3,326	3,574	
38	0,681	1,304	1,686	2,024	2,429	2,712	2,980	3,319	3,566	
39	0,681	1,304	1,685	2,023	2,426	2,708	2,976	3,313	3,558	
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551	

v	Уровень значимости $\alpha$									
	0,5	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001	
42	0,680	1,302	1,682	2,018	2,418	2,698	2,963	3,296	3,538	
44	0,680	1,301	1,680	2,015	2,414	2,692	2,956	3,286	3,526	
46	0,680	1,300	1,679	2,013	2,410	2,687	2,949	3,277	3,515	
48	0,680	1,299	1,677	2,011	2,407	2,682	2,943	3,269	3,505	
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496	
52	0,679	1,298	1,675	2,007	2,400	2,674	2,932	3,255	3,488	
54	0,679	1,297	1,674	2,005	2,397	2,670	2,927	3,248	3,480	
56	0,679	1,297	1,673	2,003	2,395	2,667	2,923	3,242	3,473	
58	0,679	1,296	1,672	2,002	2,392	2,663	2,918	3,237	3,466	
60	0,679	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460	
62	0,678	1,295	1,670	1,999	2,388	2,657	2,911	3,227	3,454	
64	0,678	1,295	1,669	1,998	2,386	2,655	2,908	3,223	3,449	
66	0,678	1,295	1,668	1,997	2,384	2,652	2,904	3,218	3,444	
68	0,678	1,294	1,668	1,995	2,382	2,650	2,902	3,214	3,439	
70	0,678	1,294	1,667	1,994	2,381	2,648	2,899	3,211	3,435	
72	0,678	1,293	1,666	1,993	2,379	2,646	2,896	3,207	3,431	
74	0,678	1,293	1,666	1,993	2,378	2,644	2,894	3,204	3,427	
76	0,678	1,293	1,665	1,992	2,376	2,642	2,891	3,201	3,423	
78	0,678	1,292	1,665	1,991	2,375	2,640	2,889	3,198	3,420	
80	0,678	1,292	1,664	1,990	2,374	2,639	2,887	3,195	3,416	
90	0,677	1,291	1,662	1,987	2,368	2,632	2,878	3,183	3,402	
100	0,677	1,290	1,660	1,984	2,364	2,626	2,871	3,174	3,390	
120	0,677	1,289	1,658	1,980	2,358	2,617	2,860	3,160	3,373	
140	0,676	1,288	1,656	1,977	2,353	2,611	2,852	3,149	3,361	
160	0,676	1,287	1,654	1,975	2,350	2,607	2,847	3,142	3,352	
180	0,676	1,286	1,653	1,973	2,347	2,603	2,842	3,136	3,345	
200	0,676	1,286	1,653	1,972	2,345	2,601	2,838	3,131	3,340	
$\infty$	0,675	1,282	1,645	1,960	2,327	2,576	2,808	3,091	3,291	

J. H. Zar. Biostatistical analysis (2 ed.). Prentice-Hall, Englewood Cliffs, N. J., 1984.



Утверждается, что средний рост мужчин национальности X 179,5.

Была взята выборка из 100 человек, по которой получилось среднее арифметическое 178,5. Проверить это утверждение с помощью одностороннего теста, если известно, что стандартное отклонение генеральной совокупности 3 см.

А уровень статистической значимости принять за 1%

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036





# Задачи на повторение



Вероятность рождения мальчиков 0.515. Найти наивероятнейшее число девочек из 600 новорожденных.

Если число  $np + p$  не является целым числом, то  $k_0$  равно целой части этого значения

Если  $np + p$  – целое число, то наивероятнейшее число исходов  $k_0$  принимает два значения  $np - q$  и  $np + p$