



GeekBrains

Теория вероятностей и математическая статистика

Вебинары



GeekBrains

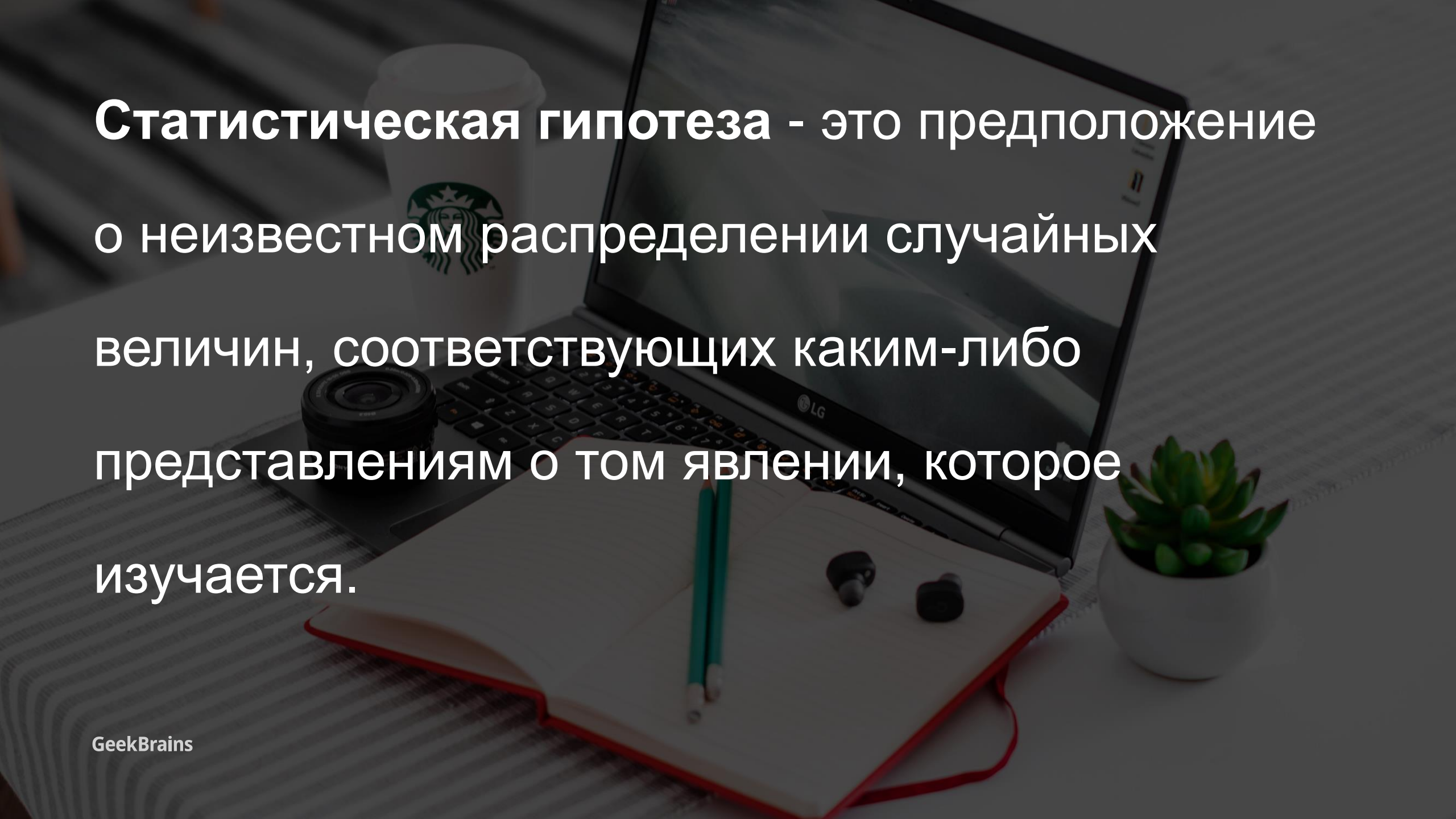
Урок 5

Теория вероятности и математическая статистика

Проверка статистических гипотез. Р-значения. Доверительные интервалы. A/B-тестирование

На этом уроке мы изучим:

1. Что такое статистическая гипотеза
2. Нулевые и альтернативные гипотезы
3. Статистические критерии для проверки гипотез
4. Доверительные интервалы
5. A / B тестирование

A grayscale background image of a desk setup. It includes a laptop with an LG logo, a Starbucks coffee cup, a camera lens, an open notebook with two pencils, and a small potted succulent.

Статистическая гипотеза - это предположение о неизвестном распределении случайных величин, соответствующих каким-либо представлениям о том явлении, которое изучается.

Нулевая гипотеза H_0 : $\mu = \mu_0$

Нулевая гипотеза - это утверждение о свойствах генеральной совокупности, которое кажется правдоподобным, но требует проверки.

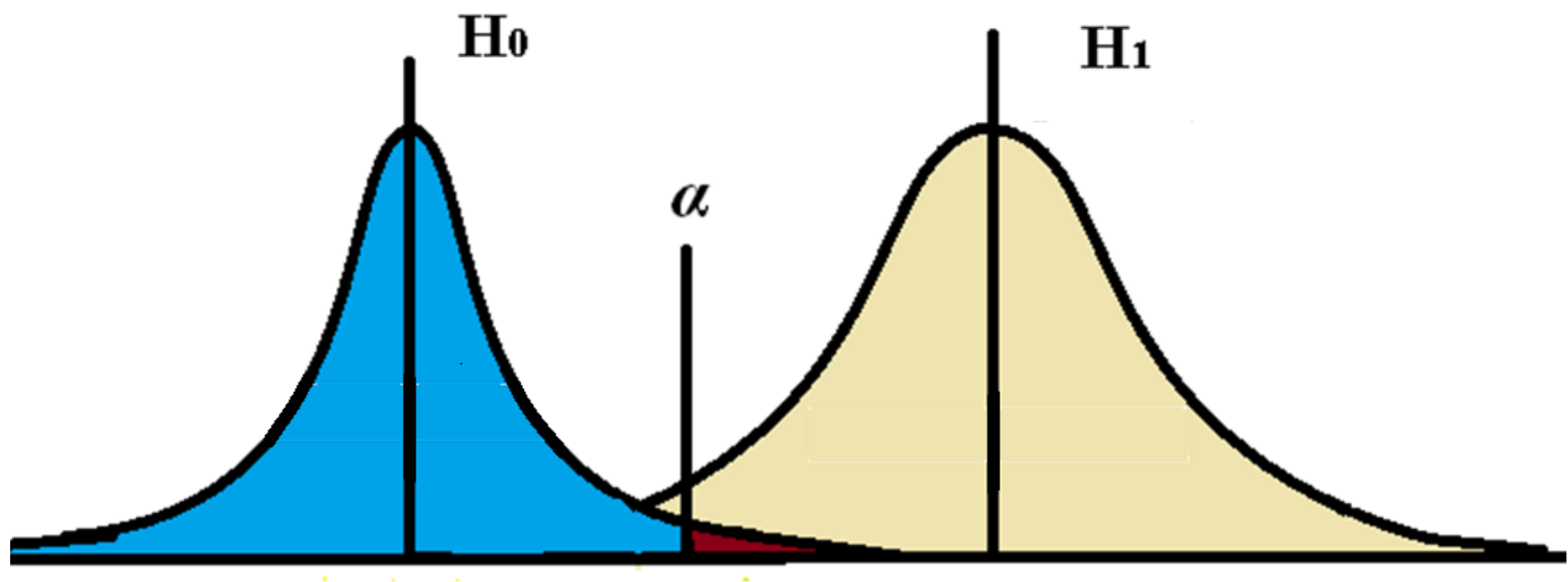
Альтернативная гипотеза H_1

Новое утверждение всегда вкладывается в альтернативную гипотезу H_1

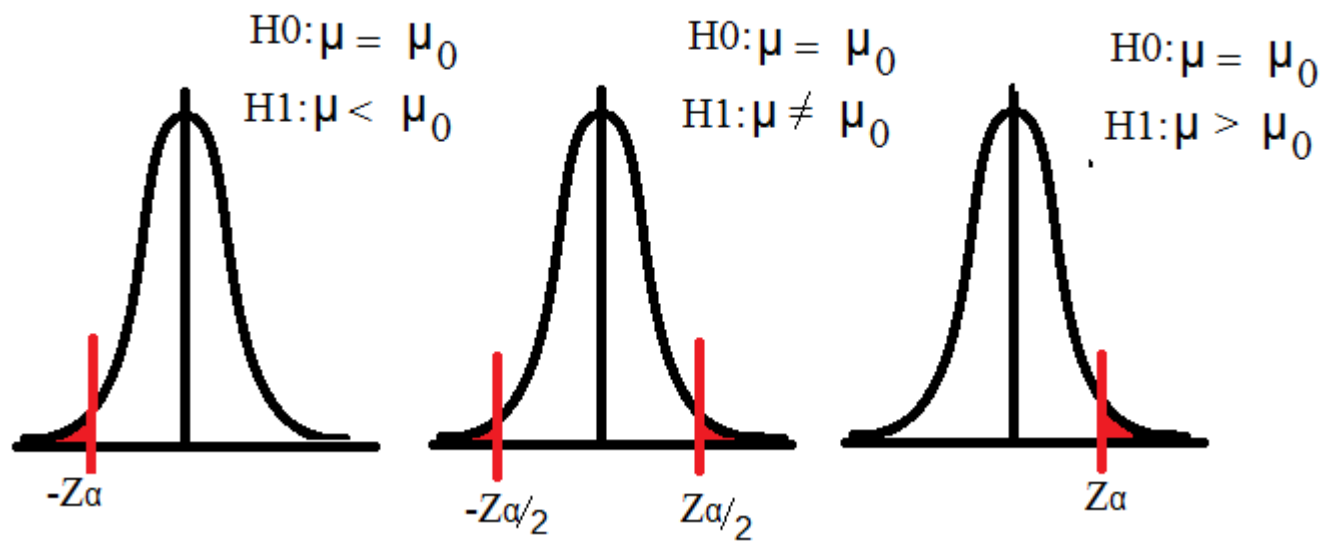
Альтернативной гипотезой является любая действительная гипотеза, отличная от нулевой.

Чаще всего для α выбирают значения 0.01 (1%), 0.05 (5%), 0.1 (10%).

- Выбираем статистический критерий
- Определяем критический регион
- Рассчитываем статистический критерий на основе данных выборки
- Принимаем решение о том, примем мы нулевую гипотезу или отвергнем



Варианты альтернативной гипотезы:



При тестировании гипотезы возможны ошибки:

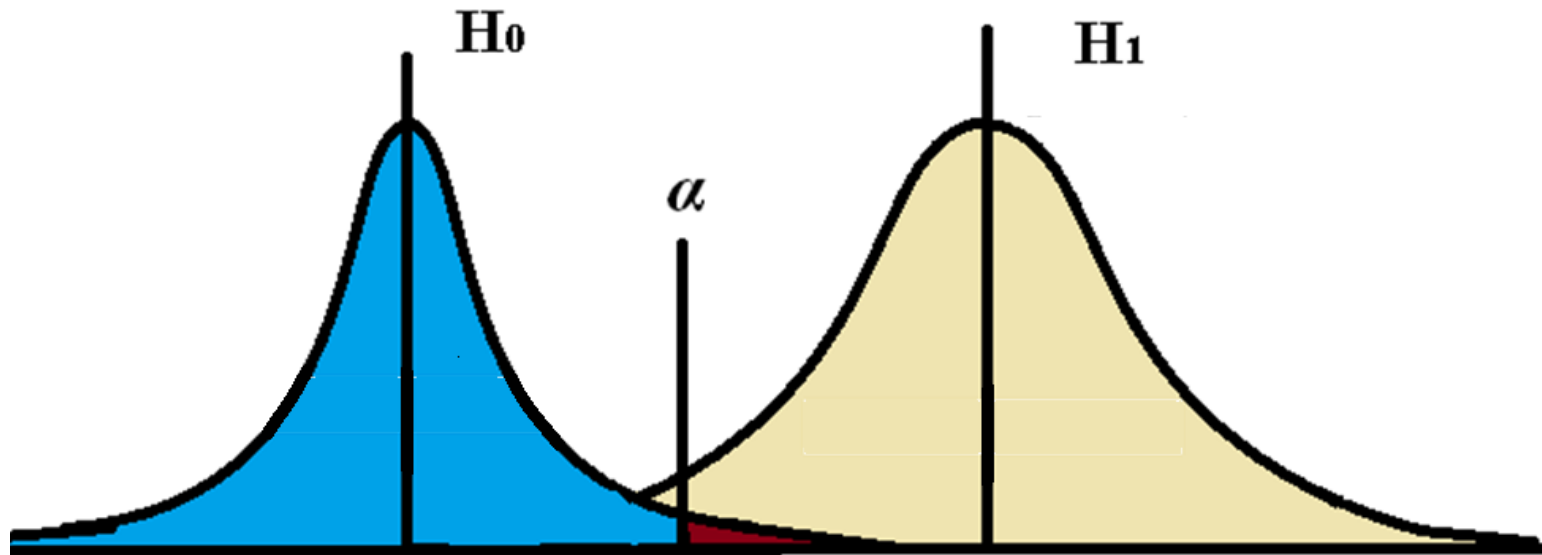
-**Ошибка I рода** : мы отвергаем H_0 , когда она верна

Уровень значимости α –вероятность ошибки I рода

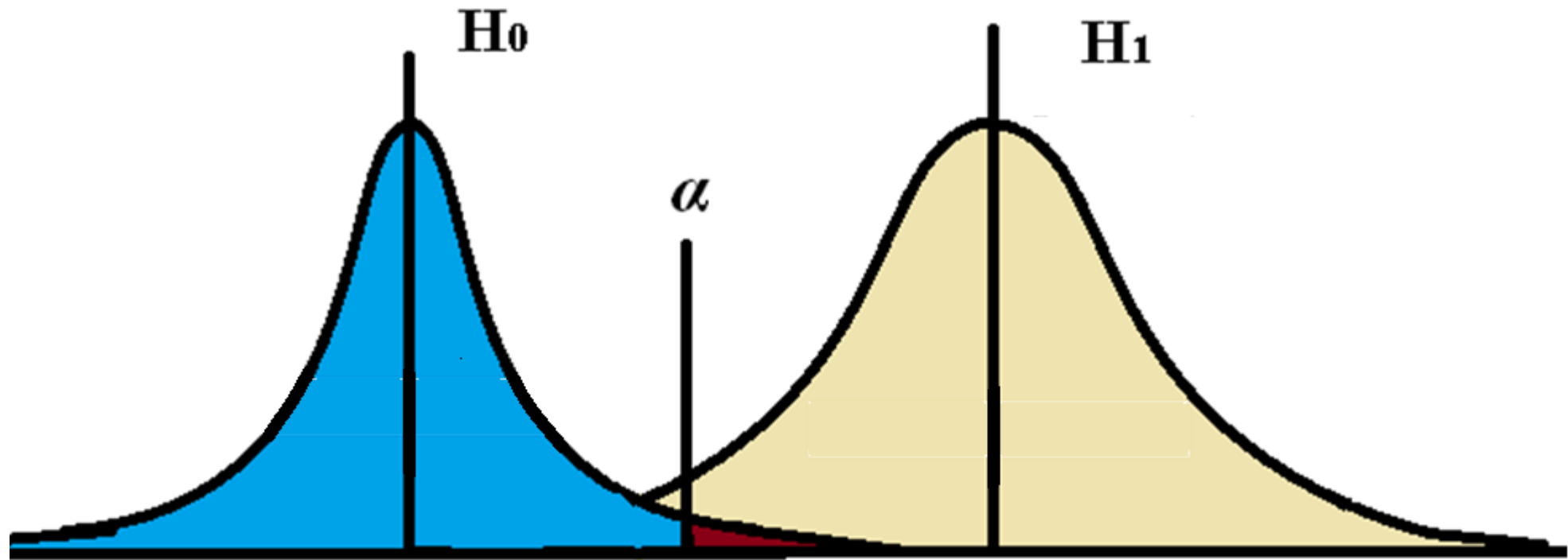
-**ошибка II рода** : мы принимаем H_0 , когда она неверна

- β - вероятность ошибки II рода

-Величина $(1 - \beta)$ –мощность теста –вероятность отклонить H_0 , когда верна H_1



Пока не будет доказано, что **нулевая гипотеза** - ложная, она считается истинной.



Этапы проверки гипотез:

- 1) формулировка основной и альтернативной гипотез
- 2) выбор статистического критерия, на основе которого будет проводиться проверка
- 3) выбор уровня значимости α
- 4) определение границ области данной гипотезы
- 5) подведение итогов и формулировка вывода

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	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936

Тестирование гипотезы о средней арифметической нормально распределенной популяции. Среднее квадратичное отклонение известно.

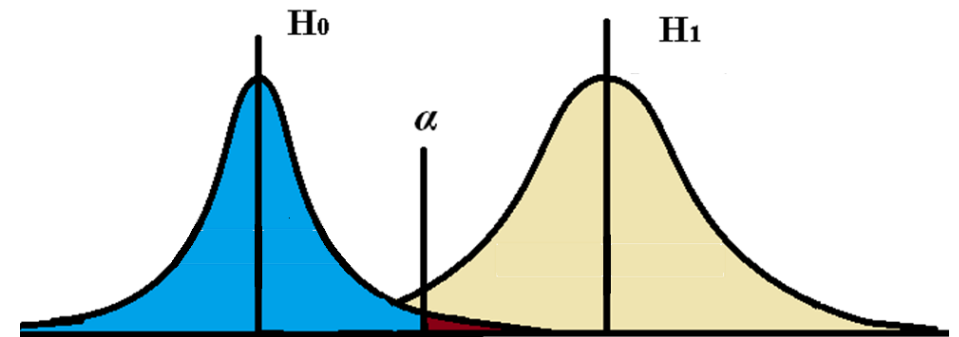
Пример:

Утверждается, что шарики для подшипников имеют диаметр 10мм. Используя односторонний критерий $\alpha=0,05$, проверить эту гипотезу, если в выборке из $n=16$ шариков, среднее оказалось равным 10,3 мм, а дисперсия нам известна и равна 1

- 1) H_0 : диаметр = 10 мм, $\mu = \mu_0$
 H_1 : диаметр = 10,3 мм, $\mu > \mu_0$

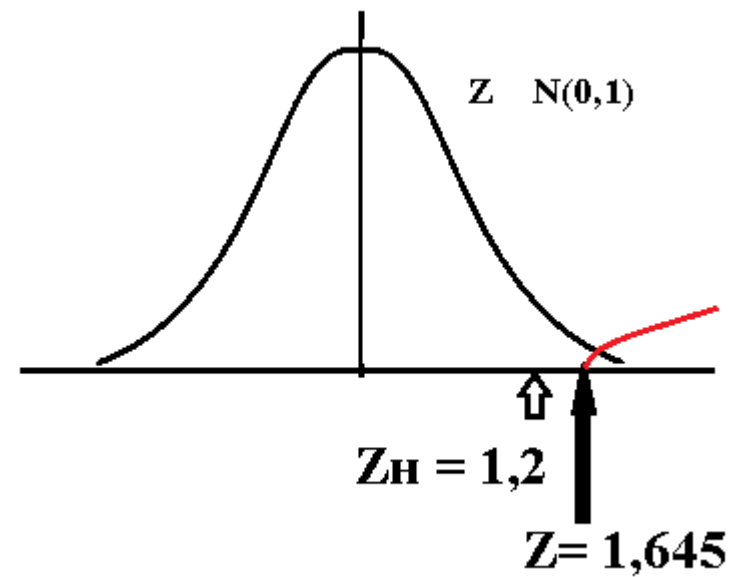
2) Критерий Z –тест

Применяется для статистической проверки гипотез, основанных на нормальном распределении.

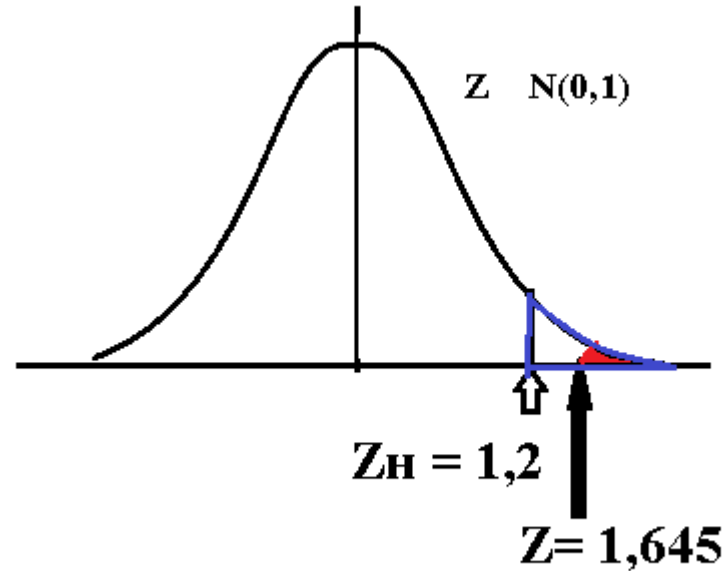


$$Z_{\text{н}} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$\alpha = 0,05 \Rightarrow P(Z_H > Z)$
 $Z_H = (10,3 - 10) / 1/4 = 1,2$
5) $Z = 1,654$
 H_0 верна



P-значение – это вероятность в распределении тест-статистика, которая лежит за пределами наблюдаемого значения



P-значение $>$ альфа, мы принимаем нулевую гипотезу

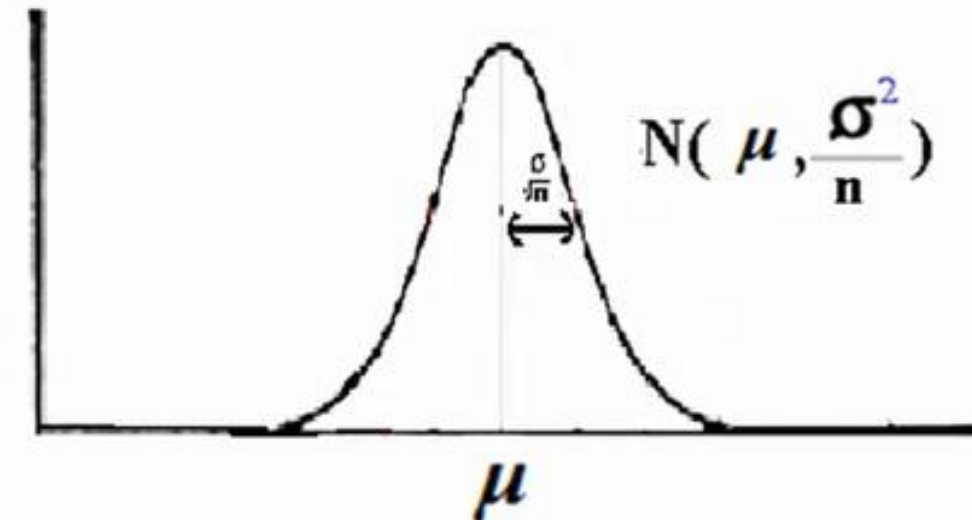
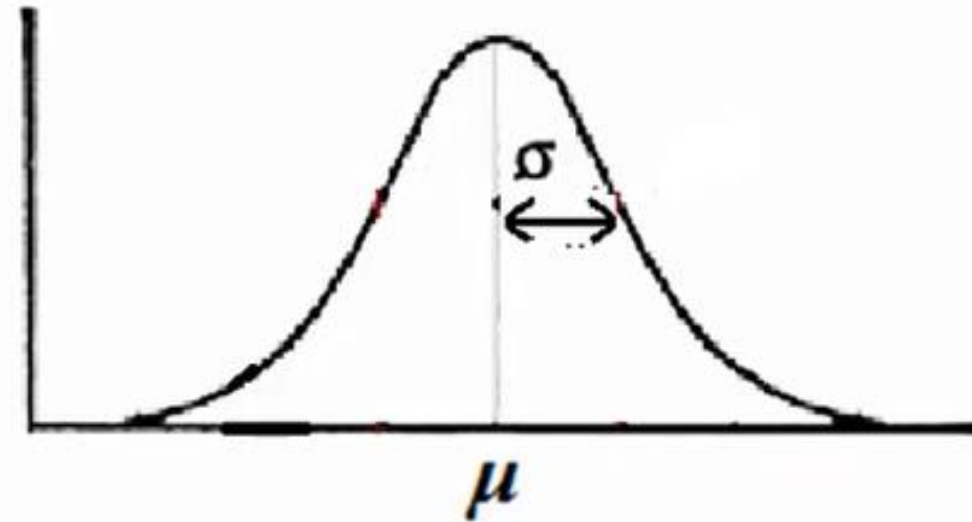
P-значение $<$ альфа, наблюдаемое значение лежит в пределах критического региона мы отвергаем нулевую гипотезу

В статистике нам часто приходится делать выводы. На ряду с тестированием гипотезы ,так же есть метод построения доверительных интервалов

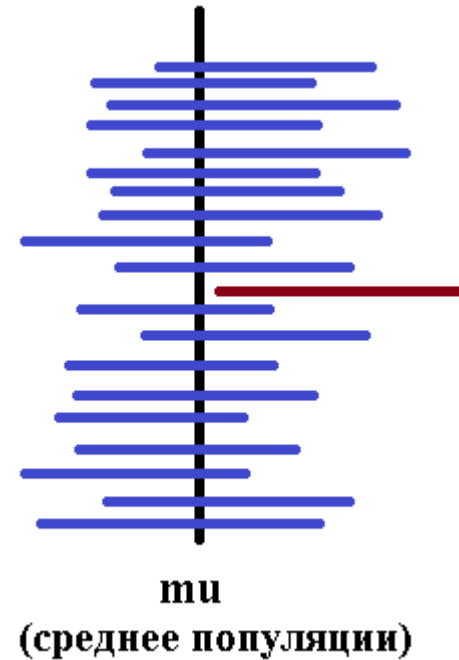
μ – среднее
арифметическое
популяции (нам не
известное)

Для каждой выборки мы можем
построить доверительный
интервал ,например ,
95% (это $1 - \alpha$)

$$\bar{X} \pm Z_{\alpha/2} \cdot \sigma / \sqrt{n}$$



95% доверительный интервал означает, что только 95 % построенных интервалов будут захватывать среднее арифметическое популяции (например, 20 выборок и для каждой построен конфиденциальный интервал, из 20 только 19 попадать на среднее арифметическое популяции)



95 доверительный интервал значит, что с 95 % вероятностью интервал захватит наше оцениваемое значение

(здесь среднее арифметическое популяции)

И это не то же самое, что среднее арифметическое попадет в наш интервал с 95 % вероятностью

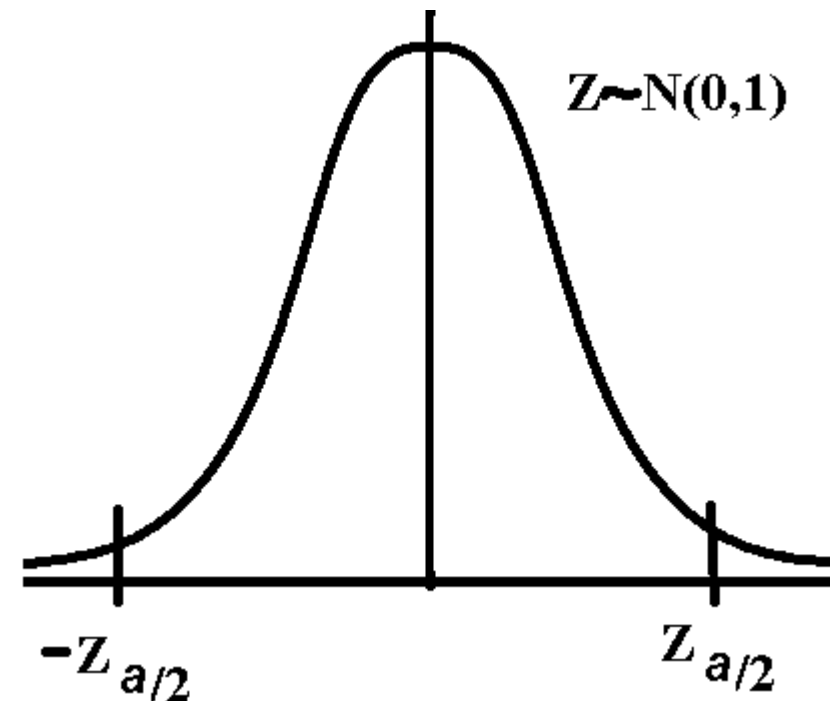
ДОВЕРИТЕЛЬНЫЙ ИНТЕРВАЛ ДЛЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ С ИЗВЕСТНЫМ СРЕДНИМ КВАДРАТИЧНЫМ ОТКЛОНЕНИЕМ ПОПУЛЯЦИИ

Известно, что генеральная совокупность распределена нормально со средним квадратичным 5. Найти доверительный интервал для оценки среднего арифметического с надежностью 0,95, если выборочная средняя $M = 24,15$, а объем выборки 100

95% интервал это интервал $(1-\alpha)$

Найдем Z по таблице

$$\bar{X} \pm Z_{\alpha/2} * \sigma / \sqrt{n}$$

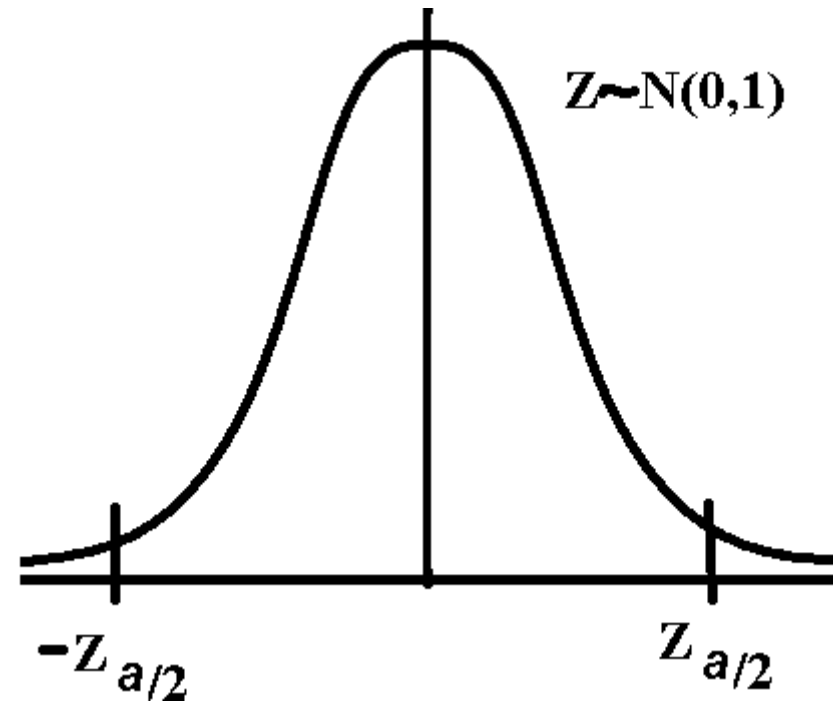


x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936

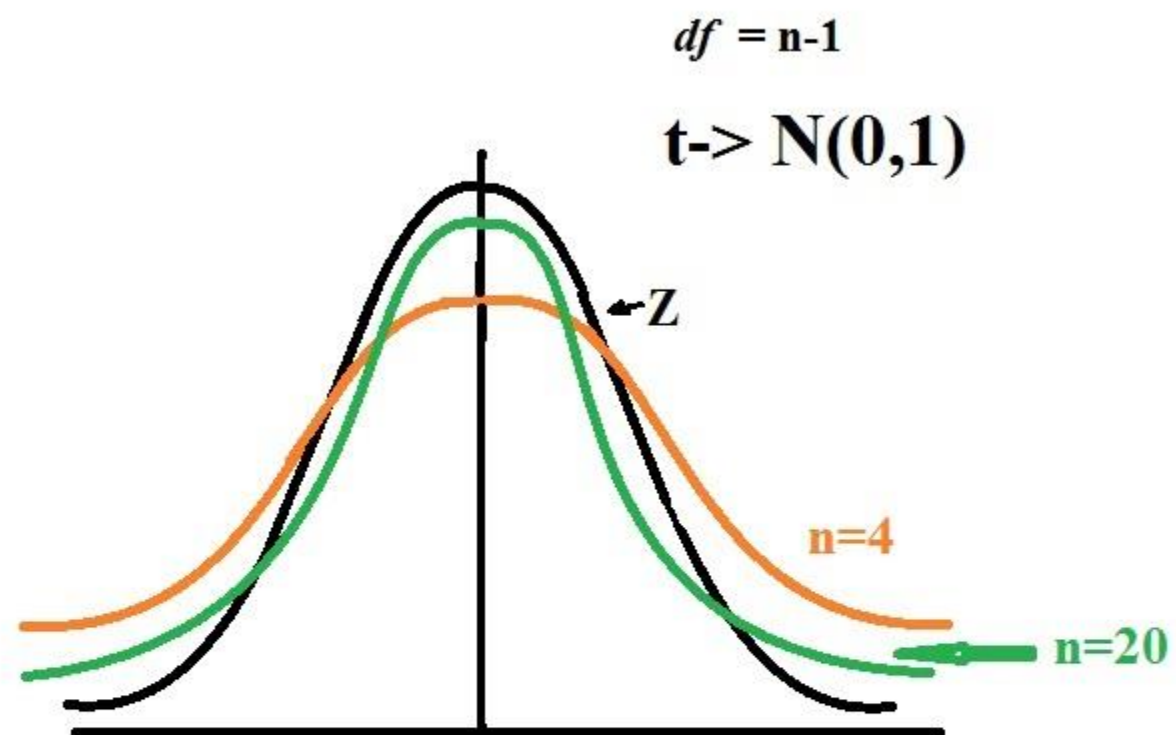
$$\bar{X} \pm Z_{\alpha/2}^* \sigma / \sqrt{n}$$

$$24,15 \pm (1,96 * 5 / 10)$$

$$[23,17; 25,13]$$



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



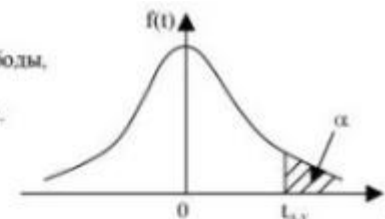
$$\bar{X} \pm t_{\alpha/2}^* \sigma / \sqrt{n}$$

Найдем t для построения 95% -го доверительного интервала, n=10

Пример: $t_{0.05,20} = 1.725$; v – число степеней свободы,

$P(T > 1.725) = 0.05$; α – уровень значимости.

$P(|T| > 1.725) = 0.10$.



$v = n - 1$, где n – объем выборки

α v	0.4	0.25	0.10	0.05	0.025	0.01	0.005	0.001	.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.31	636.6
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.6
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.214	12.94
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.859
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.405
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.050	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.040	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.040	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	2.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.255	0.680	1.296	1.676	2.009	2.403	2.678	3.262	3.495
60	0.255	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.254	0.679	1.292	1.664	1.990	2.374	2.639	3.195	3.415
100	0.254	0.678	1.290	1.660	1.984	2.365	2.626	3.174	3.389
120	0.254	0.677	1.289	1.658	1.980	2.358	2.467	3.160	3.366
200	0.254	0.676	1.286	1.653	1.972	2.345	2.601	3.131	3.339
500	0.253	0.675	1.283	1.648	1.965	2.334	2.586	3.106	3.310
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

A / B тестирование - маркетинговый метод, используемый для оценки эффективности веб-страниц и управления ими.

Этот метод также называется разделенным тестированием (с английского. сплит-тестирование – “разделенное тестирование”).

При А / В тестировании сравнивают страницы А и В, имеющие разные элементы дизайна (например, разный цвет кнопки заказа товара).
На каждую страницу случайным образом запускают 50% аудитории сайта и затем сравнивают, какая страница показывает наибольший процент конверсии.

- 1) Для чего нужны такие методы, как тестирование гипотезы и построение доверительного интервала
- 2) Когда пользуются таблицей Стьюдента?
- 3) В чем заключается ошибка первого и второго рода?
- 4) Если увеличивать значение альфа, вероятность какой ошибки тогда растет?
- 5) Если P - значение меньше альфа, то какая гипотеза верна?

- 1 Известно, что генеральная совокупность распределена нормально со средним квадратическим отклонением, равным 16. Найти доверительный интервал для оценки математического ожидания с надежностью 0.95, если выборочная средняя $M = 80$, а объем выборки $n = 256$.

2 В результате 10 независимых измерений некоторой величины X , выполненных с одинаковой точностью, получены опытные данные:

6.9, 6.1, 6.2, 6.8, 7.5, 6.3, 6.4, 6.9, 6.7, 6.1

Предполагая, что результаты измерений подчинены нормальному закону распределения вероятностей, оценить истинное значение величины X при помощи доверительного интервала, покрывающего это значение с доверительной вероятностью 0,95.

3,4 задачи решать через тестирование гипотезы

3. Утверждается, что шарики для подшипников, изготовленные автоматическим станком, имеют средний диаметр 17 мм. Используя односторонний критерий с $\alpha=0,05$, 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.

4. 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%?

ИТОГИ

1. Что такое статистическая гипотеза
2. Нулевые и альтернативные гипотезы
3. Статистические критерии для проверки гипотез
4. Доверительные интервалы
5. А / В тестирование