Hypothesis Testing

aka, The Scientific Method

What is a Hypothesis?

An educated guess about a population parameter μ, p, σ^2, σ.

Population Mean μ
Population Proportion p
Population Variance σ^2
Population Standard Deviation σ

Example The mean lifespan of California resident's is 77.8 years.

Example The proportion of college students who take a full load is not 40%.

Example The mean pulse rate of adult males is at least 72 bpm (beats per minute).

Example The standard deviation of IQ scores for a population is less than 15.

Example The variance of IQ Scores for a population is no more than 125.

Example The probability that the Dodgers will win the World Series in 2021 is more than one-half.

There are two types of Hypothesis

Def Null Hypothesis A hypothesis that involves a condition of equality.

$$=, \geq, \leq$$

$$H_0$$

Def Alternate Hypothesis A hypothesis that does not involve a condition of equality.

$$\neq, >, <$$

$$H_1$$

Example 52% of newborn babies are boys.

Example 64% of college students believe Big Foot is real.

Example Earthquake occur at a mean depth that is not the equal to 5.0 km.

Example Insomnia treatment with Zopiclone induces a mean wake up time less than 103.2 minutes.

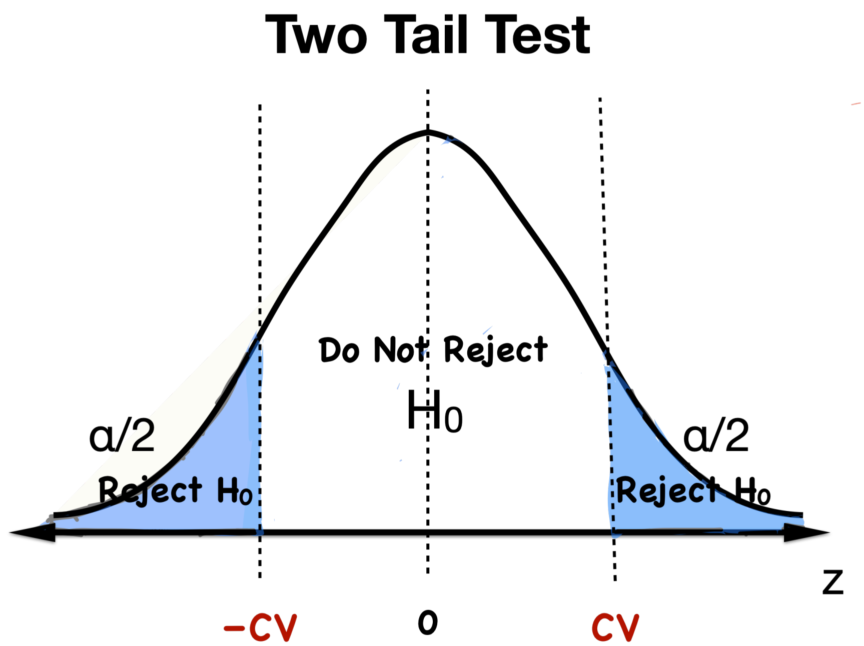
Example The mean IQ scores for College Professors is no more than 125.

Hypothesis testing is based on taking a sample to **compute a test statistic** which is used to form a conclusion on a **decision rule**. There are three types of Hypothesis Testing formats that are based on your original hypothesis.

Decision Rule for a **Two Tail Test**

$$H_0: \text{ parameter} = \text{value}$$

$$H_1: \text{ parameter} \neq \text{value}$$



α represents the **level of significance** and is the sum of the percent of the bell that is shaded in the tails. Critical Values are determined by our z-table.

Significance level α

A probability value used as a cutoff for determining when the sample evidence constitutes significant evidence against the Null Hypothesis. The level of significance α is the probability of mistakenly rejecting the null hypothesis when it is actually true.

$$\alpha = p(\text{rejecting } H_0 \text{ when it is actually true})$$

Type I and Type II Errors

When testing a hypothesis (claim) we arrive at a conclusion of either rejecting it or failing to reject it (accept it). Our conclusions are sometimes correct or sometimes wrong, even when we follow our procedure correctly. In essence, our hypothesis testing is based on likelihood and is not certain.

Type I Error- The mistake of rejecting the null hypothesis H_0 when it is actually true.

$$\alpha = p(\text{rejecting } H_0 \text{ when it is actually true})$$

Type II Error - The mistake of failing to reject (accepting) the null hypothesis H_0 when it's actually false.

$$B = p(\text{failing to reject } H_0 \text{ when it is actually false})$$

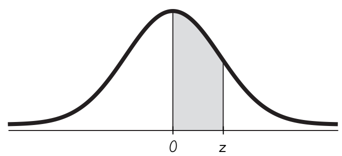


TABLE A-2 Standard Normal (z) Distribution

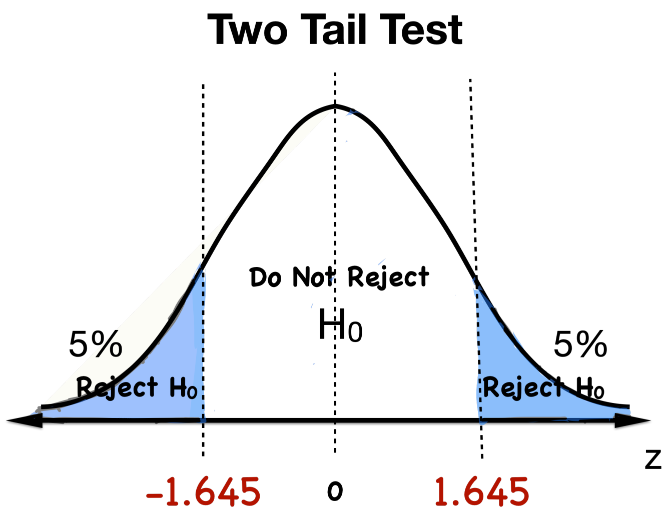
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495 *	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949 *	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.10 and higher	.4999									

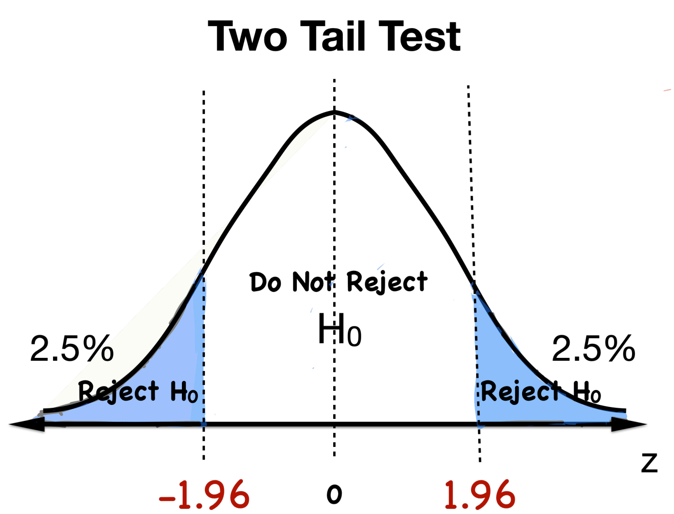
NOTE: For values of z above 3.09, use 0.4999 for the area.
*Use these common values that result from interpolation:

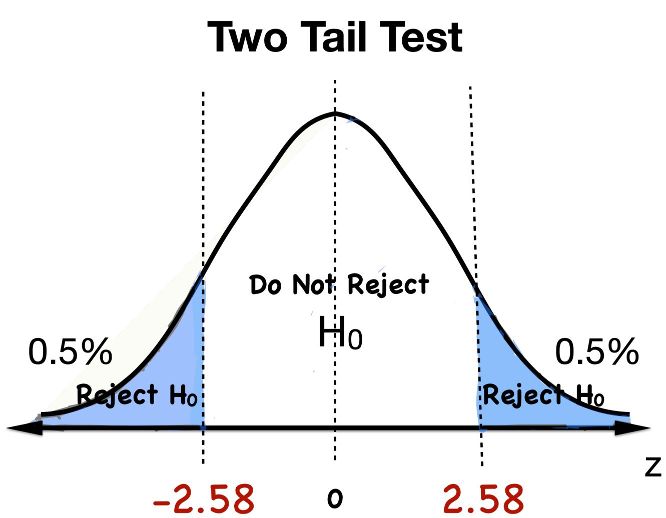
z score	Area
1.645	0.4500
2.575	0.4950

α

10% level of significance, 5% level of significance, and the 1% level of significance



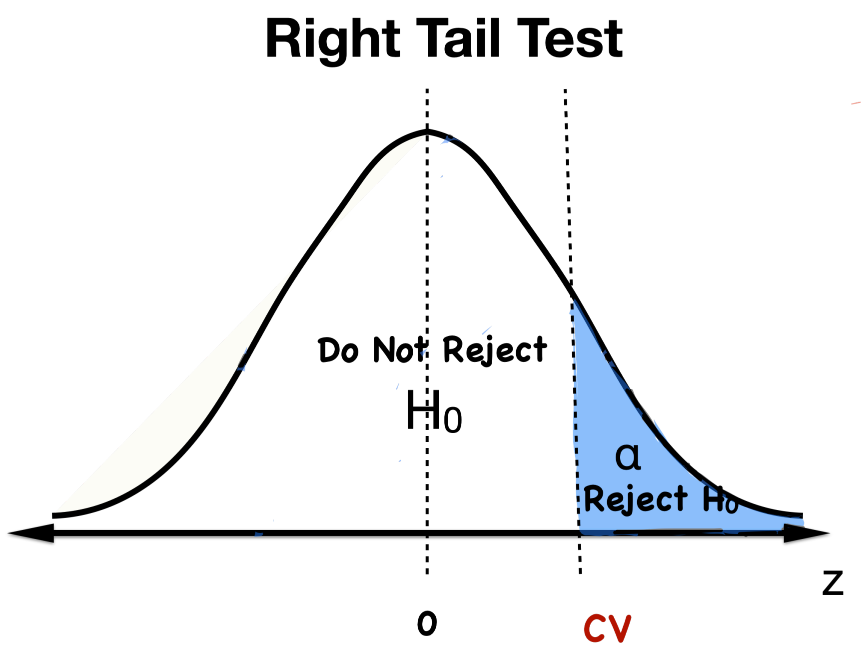




Decision Rule for a **Right Tail Test**

H_0: parameter$\leq$value

H_1: parameter $>$ value

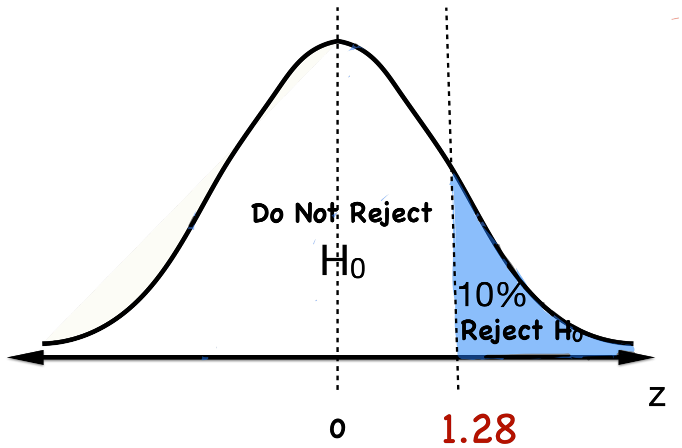


α represents the **level of significance** and is the percent of the bell that is shaded in the right tail. Critical Values are determined by our z-table.

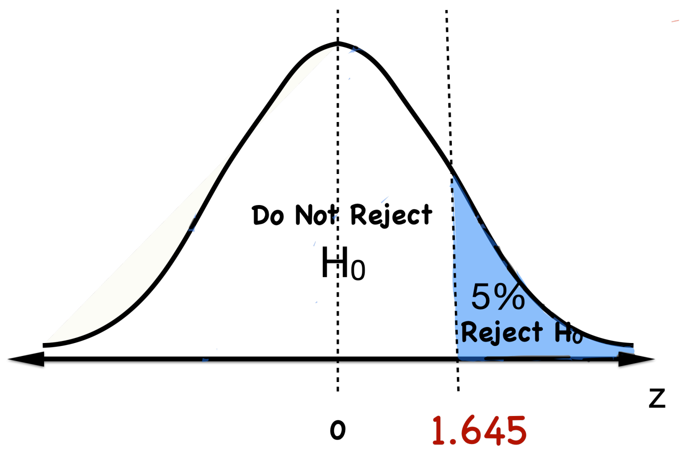
α

10% level of significance, 5% level of significance, and the 1% level of significance

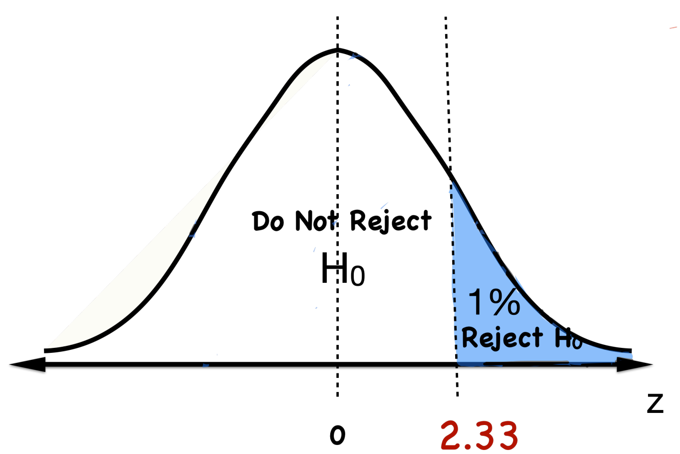
Right Tail Test



Right Tail Test



Right Tail Test



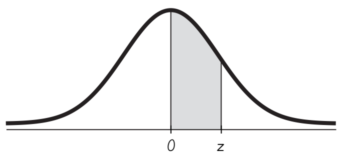


TABLE A-2 Standard Normal (z) Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495 *	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949 *	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.10 and higher	.4999									

NOTE: For values of z above 3.09, use 0.4999 for the area.
*Use these common values that result from interpolation:

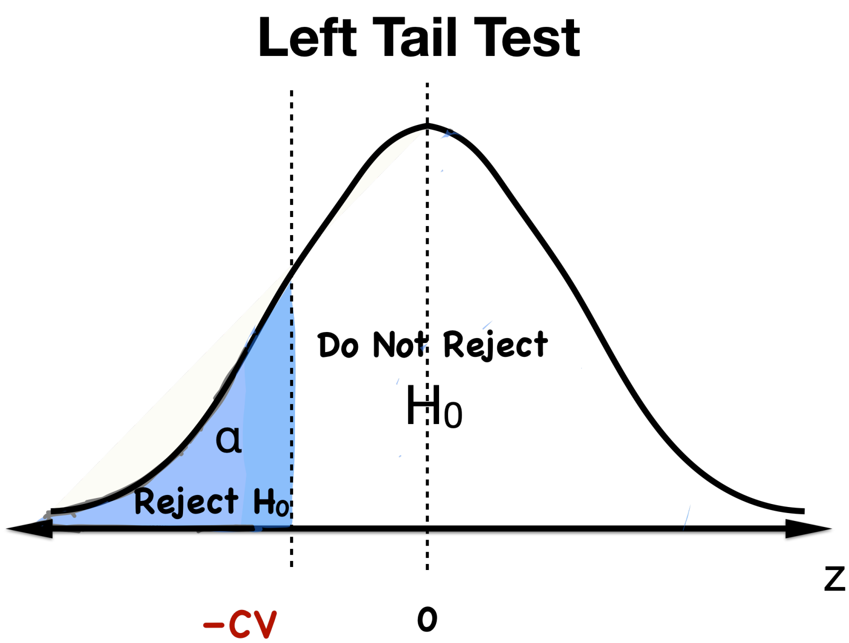
z score	Area
1.645	0.4500
2.575	0.4950

From Frederick C. Mosteller and Robert E. K. Rourke, *Sturdy Statistics*, 1973, Addison-Wesley Publishing Co., Reading, MA. Reprinted with permission of Frederick Mosteller.

Decision Rule for a **Left Tail Test**

H_0: parameter $\geq$value

H_1: parameter $<$ value

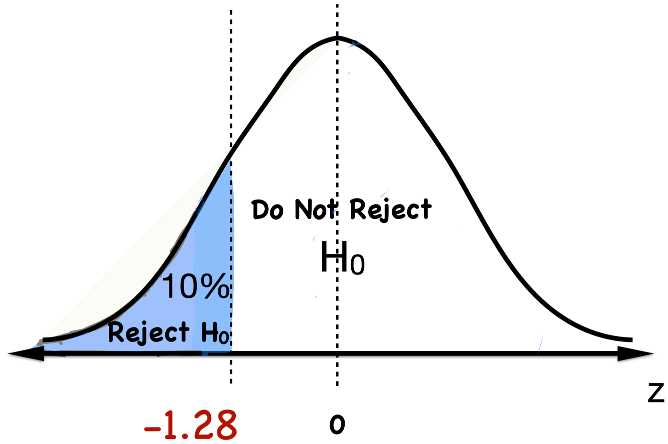


α represents the **level of significance** and is the percent of the bell that is shaded in the left tail. Critical Values are determined by our z-table.

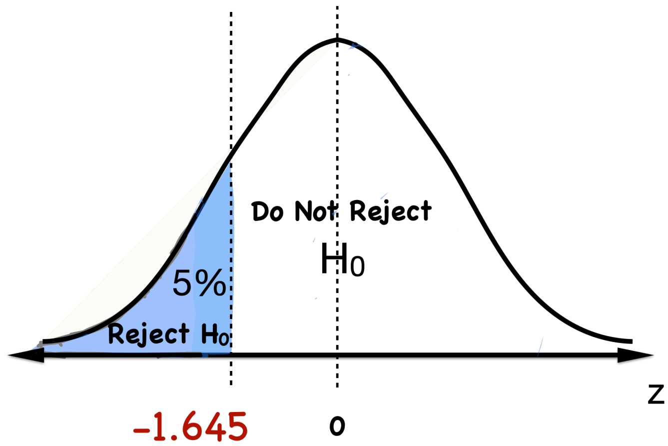
α

10% level of significance, 5% level of significance, and the 1% level of significance

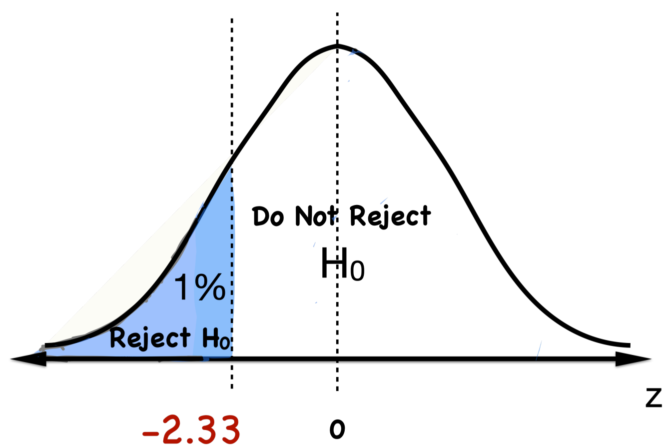
Left Tail Test



Left Tail Test



Left Tail Test



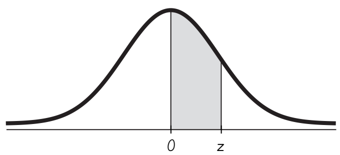


TABLE A-2 Standard Normal (z) Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495 *	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949 *	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.10 and higher	.4999									

NOTE: For values of z above 3.09, use 0.4999 for the area.
*Use these common values that result from interpolation:

z score	Area
1.645	0.4500
2.575	0.4950

From Frederick C. Mosteller and Robert E. K. Rourke, *Sturdy Statistics*, 1973, Addison-Wesley Publishing Co., Reading, MA. Reprinted with permission of Frederick Mosteller.

Hypothesis testing is based on taking a sample to **compute a test statistic** which is used to form a conclusion on a **decision rule**. There are three types of Hypothesis Testing formats that are based on the your original hypothesis.

Test Statistics

Hypothesis Test about a Proportion

$$ts = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Hypothesis Test about a Mean

$$ts = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Hypothesis Test about a Two Independent Proportions

$$ts = \frac{\overline{p_1} - \overline{p_2}}{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n_1} + \frac{\tilde{p}(1-\tilde{p})}{n_2}}}$$

Where $\tilde{p} = \frac{x_1+x_2}{n_1+n_2}$

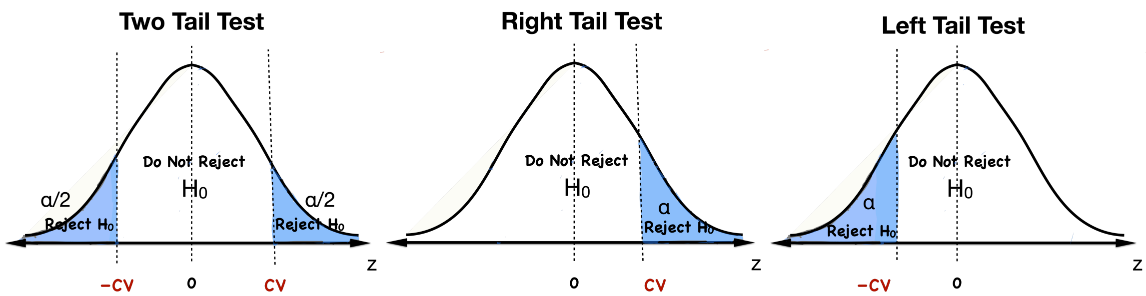
Hypothesis Test about a Two Independent Means

$$ts = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Traditional Hypothesis Testing Procedure

Step1- Set up your Hypothesis and **Step 2-** Create your **decision rule** based on the Hypothesis Testing set up (step 1).

$H_0: =$	$H_0: \leq$	$H_0: \geq$
$H_1: \neq$	$H_1: >$	$H_1: <$



Label which Hypothesis is the claim!

Step 3- Take your Sample.

Step 4- Compute your Test Statistic

$$ts = \frac{\bar{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \qquad ts = \frac{\bar{x}-\mu}{s/\sqrt{n}} \qquad ts = \frac{\overline{p_1}-\overline{p_2}}{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n_1}+\frac{\tilde{p}(1-\tilde{p})}{n_2}}} \qquad ts = \frac{\overline{x_1}-\overline{x_2}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

Step 5- State your **Conclusion.**
Either the **sample supports the claim** or the **sample does not support the claim**.