SOC 3811/5811: BASIC SOCIAL STATISTICS

Hypothesis Testing

Hypothesis Tests

We will consider four different sorts of hypothesis tests, all of which follow the same logic

1. **Hypothesis Tests for <u>Proportions</u>**

   Use $\hat{p}$ to infer p, the population proportion

2. **Hypothesis Tests for <u>Means</u>**

   Use $\bar{Y}$ to infer μ_Y, the population mean of Y

3. **Hypothesis Tests for <u>Differences in Proportions</u>**

   Use $\hat{p}_1 - \hat{p}_2$ to infer the difference between two population proportions, p_1 and p_2

4. **Hypothesis Tests for <u>Differences in Means</u>**

   Use $\bar{Y}_1 - \bar{Y}_2$ to infer the difference between two population means, μ_{Y1} and μ_{Y2}

How to Test a Hypothesis

1. State the null (H_0) and alternative (H_1) hypotheses
2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further
3. Choose an α probability level … that is, a probability associated with incorrectly rejecting the null hypothesis
4. Determine the "critical value" … that is, how large the test statistic must be in order to reject the null hypothesis at the given α level
5. Calculate the test statistic
6. Compare the test statistic to the critical value
   - If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
   - If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

Hypotheses about a Mean: Example #1

A 2001 census of penguins found that there were 12.1 penguins per square mile in Antarctica

Now researchers are interested in seeing how the size of the population of penguins in Antarctica has changed

> They randomly sampled 200 square mile sections in Antarctica, and observed the number of penguins that lived on each section
>
> In their sample they observed a mean of 11.7 penguins per square mile, with $s=2.3$ penguins

Is this evidence sufficient to confidently conclude that the size of the penguin population has changed?

Hypotheses about a Mean: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether the mean is *different* from a particular value (as opposed to being either *higher* or *lower* than a particular value), this is a two-sided hypothesis test

H_0: $\mu = 12.1$

H_1: $\mu \neq 12.1$

Hypotheses about a Mean: Example #1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The random sample is small (<30) and symmetrically distributed with no outliers, and the population of measurements is bell-shaped (Not True)

--- or ---

The size of the random sample is large ($\geq$30), regardless of the shape of the distribution of the measurements in the population (True)

Hypotheses about a Mean: Example #1

3. Choose an α probability level … that is, a probability associated with incorrectly rejecting the null hypothesis

Again, the precise α level we select is arbitrary, and varies by discipline (and sometimes by journal or author)

Let's go with α=0.01 in this example

Hypotheses about a Mean: Example #1

4. Determine the "critical value" … that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

Because we are dealing with sample means, our test statistic will be a t-score (not a Z-score)

What critical value t* corresponds to having 1%/2=0.5% of the area under the normal curve above +t* and 0.5% of the area under the normal curve below –t*?

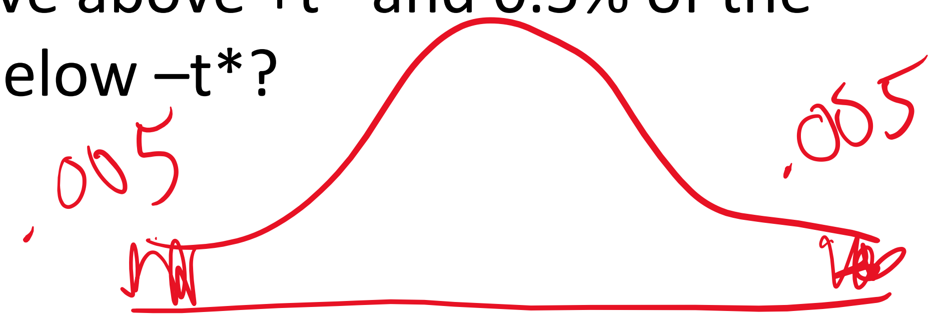


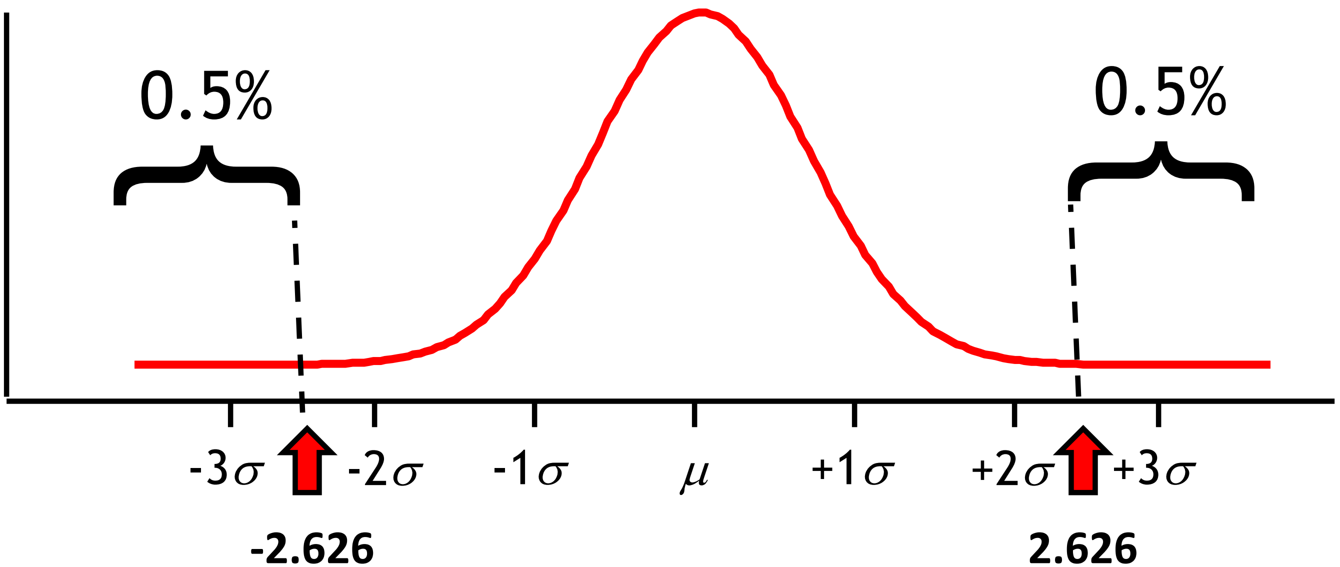
TABLE D

t distribution critical values

	Upper-tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300

Hypotheses about a Mean: Example #1

With α=0.01, N=200, and a two-sided test ...



... so our critical value t* is 2.626

H_0: $\mu = 12.1$ ... fail to reject H_0 if $|t| \leq 2.626$

H_1: $\mu \neq 12.1$ ... reject H_0 if $|t| > 2.626$

Hypotheses about a Mean: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing…

…subtract the population value that is assumed to be true under the null from the observed (sample) value

…divide that figure by the standard deviation of the sampling distribution of the statistic in question

For single means:

$$t = \frac{\overline{Y} - \mu_Y}{s_Y/\sqrt{n}}$$

Hypotheses about a Mean: Example #1

In our example:

H_0: $\mu = 12.1$

H_1: $\mu \neq 12.1$

$\overline{Y} = 11.7$

$s_Y = 2.3$

n=200

$$t = \frac{\overline{Y} - \mu_Y}{s_Y/\sqrt{N}} = \frac{11.7 - 12.1}{2.3/\sqrt{200}} = -2.459$$

Hypotheses about a Mean: Example #1

In our example:

H_0: $\mu = 12.1$

H_1: $\mu \neq 12.1$

$\overline{Y} = 11.7$

s_Y=2.3

n=200

$$t = \frac{\overline{Y} - \mu_Y}{s_Y/\sqrt{N}} = \frac{11.7 - 12.1}{2.3/\sqrt{200}} = -2.459$$

Hypotheses about a Mean: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value t* is 2.626

We obtained a test statistic of -2.459

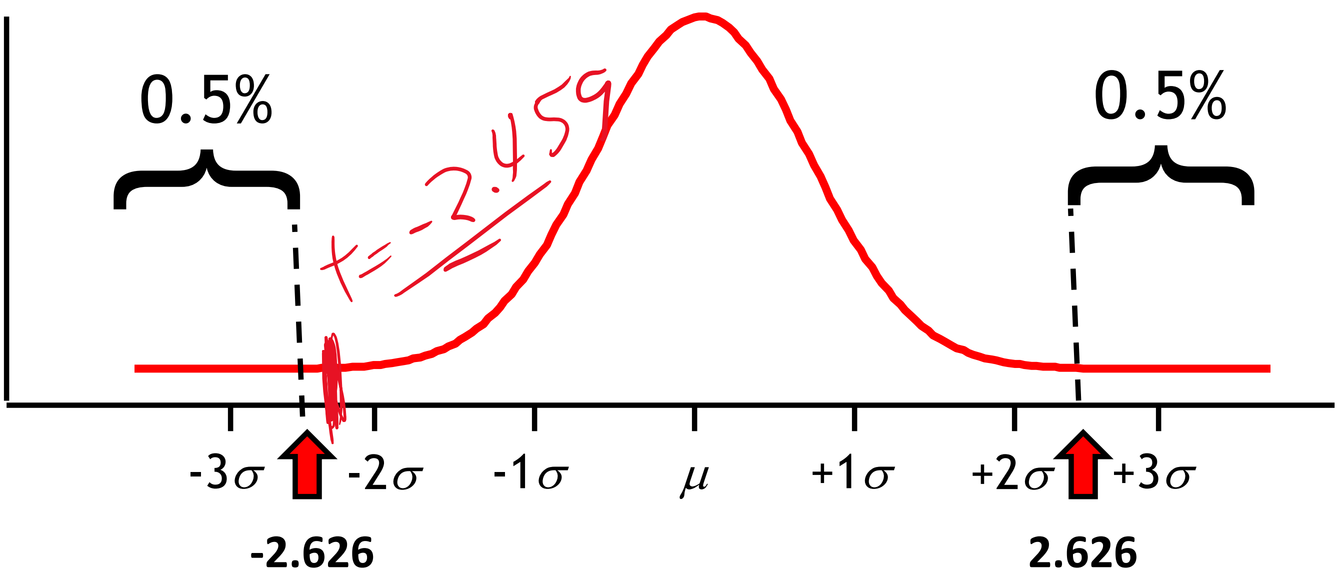
H_0: $\mu = 12.1$ … fail to reject H_0 if $|t| \leq 2.626$

H_1: $\mu \neq 12.1$ … reject H_0 if $|t| > 2.626$

Conclusion? Fail to Reject H_0

Hypotheses about a Mean: Example #1

With α=0.01, N=200, and a two-sided test …



… so our critical value t* is 2.626

H_0: μ = 12.1 … fail to reject H_0 if $|t| \leq 2.626$

H_1: $\mu \neq$ 12.1 … reject H_0 if $|t| > 2.626$

Hypotheses about a Mean: Example #1

A 2001 census of penguins found that there were 12.1 penguins per square mile in Antarctica

Now researchers are interested in seeing how the size of the population of penguins in Antarctica has changed

- They randomly sampled 200 square mile sections in Antarctica, and observed the number of penguins that lived on each section
- In their sample they observed a mean of 11.7 penguins per square mile, with s=2.3 penguins

Is this evidence sufficient to confidently conclude that the size of the penguin population has changed? NO

Worksheet

Does the average undergraduate student bring more than \$40 to class?

Using data from a survey conducted on the first day of my undergraduate research methods class, I found that the mean amount of money that the **75** students had on them was **\$43** with a standard deviation of **\$78** (Assume for now that students in my class represent a random sample of all undergraduate students)

Use α=0.05

Hypotheses about Differences in Proportions: Example #1

Are men and women equally likely to smoke?

To find out, I will use data from the General Social Surveys (GSS)

Among men: 2,839 out of 7,014 smoke (p=0.405, or 40.5%)

Among women: 2,886 out of 9,359 smoke (p=0.308, or 30.8%)

Is this sufficient evidence to confidently conclude that men and women are <u>not</u> equally likely to smoke?

Hypotheses about Differences in Proportions: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether or not the difference in proportions is different from a particular value (zero), this is a two-sided hypothesis test

H_0: $p_M = p_F$ (or, $p_M - p_F = 0$)

H_1: $p_M \neq p_F$ (or, $p_M - p_F \neq 0$)

Hypotheses about Differences in Proportions: Example #1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

The two samples must be independent and representative of the two populations from which they were drawn (True)

As with confidence intervals, $n_1 \hat{p}_1$, $n_1(1- \hat{p}_1)$,
$n_2 \hat{p}_2$, and $n_2(1- \hat{p}_2)$ must all exceed 5 (and preferably 10) (True)

Hypotheses about Differences in Proportions: Example #1

3. Choose an α probability level … that is, a probability associated with incorrectly rejecting the null hypothesis

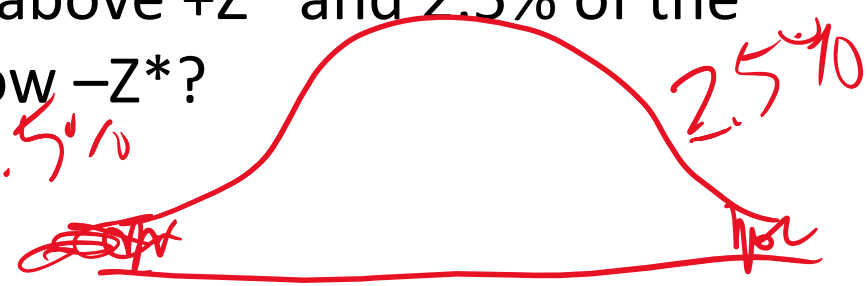
Let's go with α=0.05 in this example

Hypotheses about Differences in Proportions: Example #1

4. Determine the "critical value" … that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

Because we are dealing with sample proportions, our test statistic will be a Z-score

What critical value Z* corresponds to having 5%/2=2.5% of the area under the normal curve above +Z* and 2.5% of the area under the normal curve below –Z*?



Standard Normal Probabilities

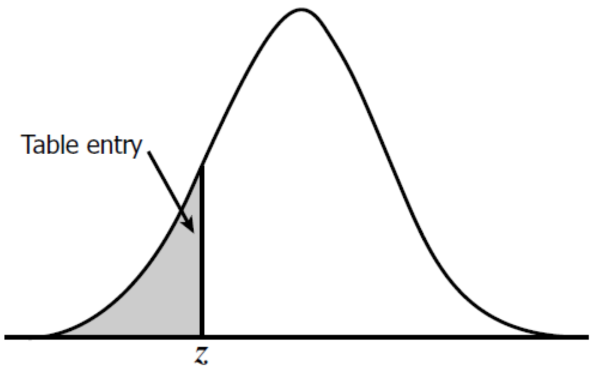
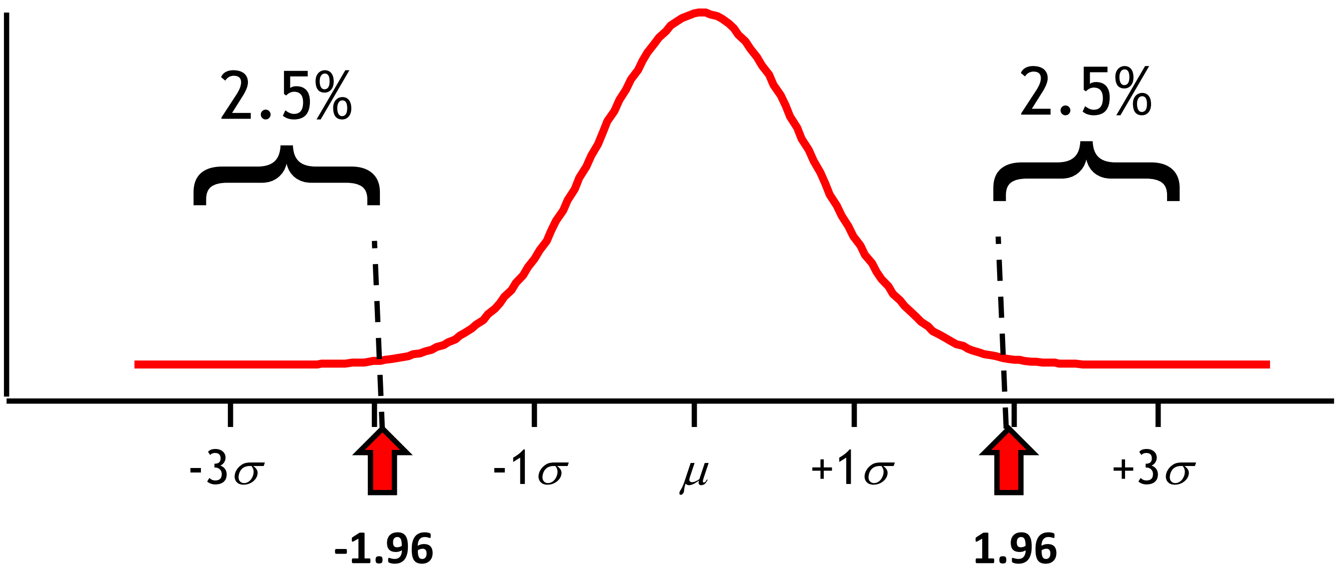


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681

Hypotheses about Differences in Proportions: Example #1

With α=0.05 and a two-sided test …



… so our critical value Z* is 1.96

H_0: $p_M = p_F$ … fail to reject H_0 if $|Z| \leq 1.96$

H_1: $p_M \neq p_F$ … reject H_0 if $|Z| > 1.96$

Hypotheses about Differences in Proportions: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing…

…subtract the population value that is assumed to be true under the null from the observed (sample) value

…divide that figure by the standard deviation of the sampling distribution of the statistic in question

For differences in proportions:

2

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$$

where

1

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

Hypotheses about Differences in Proportions: Example #1

In our example:

$\hat{p}_M = 0.405$ $\quad n_M = 7{,}014$

$\hat{p}_F = 0.308$ $\quad n_F = 9{,}359$

$$\hat{p} = \frac{n_M \hat{p}_M + n_F \hat{p}_F}{n_M + n_F} = \frac{7014 \cdot .405 + 9359 \cdot .308}{7014 + 9359} = .35$$

$$Z = \frac{\hat{p}_M - \hat{p}_F - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_M} + \frac{\hat{p}(1-\hat{p})}{n_F}}} = \frac{.405 - .308 - 0}{\sqrt{\frac{.35(1-.35)}{7014} + \frac{(.35)(1-.35)}{9359}}} = 12.9$$

Hypotheses about Differences in Proportions: Example #1

In our example: $\hat{p}_M = 0.405$ $n_M = 7{,}014$

$\hat{p}_F = 0.308$ $n_F = 9{,}359$

$$\hat{p} = \frac{n_M\hat{p}_M + n_F\hat{p}_F}{n_M + n_F} = \frac{7{,}014(0.405) + 9{,}359(0.308)}{7{,}014 + 9{,}359} = 0.350$$

$$Z = \frac{\hat{p}_M - \hat{p}_F - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_M} + \frac{\hat{p}(1-\hat{p})}{n_F}}} = \frac{0.405 - 0.308 - 0}{\sqrt{\frac{0.350(1-0.350)}{7{,}014} + \frac{0.350(1-0.350)}{9{,}359}}} = 12.9$$

Hypotheses about Differences in Proportions: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value Z* is 1.96

We obtained a test statistic of 12.9

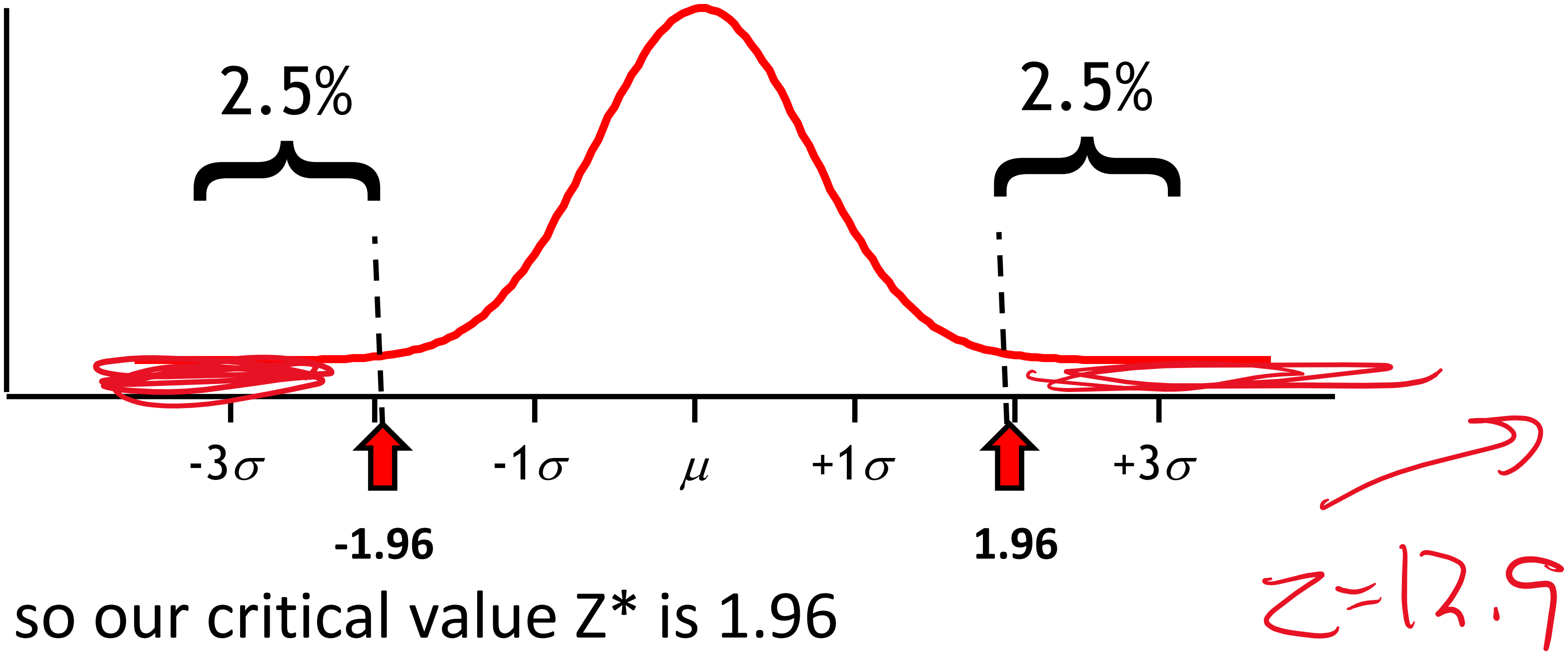
H_0: $p_M = p_F$ … fail to reject H_0 if $|Z| \leq 1.96$

H_1: $p_M \neq p_F$ … reject H_0 if $|Z| > 1.96$

Conclusion? Reject H_0

Hypotheses about Differences in Proportions: Example #1

With α=0.05 and a two-sided test …



… so our critical value Z* is 1.96

H_0: $p_M = p_F$ … fail to reject H_0 if $|Z| \leq 1.96$

H_1: $p_M \neq p_F$ … reject H_0 if $|Z| > 1.96$

Hypotheses about Differences in Proportions: Example #1

Are men and women equally likely to smoke?

To find out, I will use data from the General Social Surveys (GSS)

Among men: 2,839 out of 7,014 smoke (p=0.405, or 40.5%)

Among women: 2,886 out of 9,359 smoke (p=0.308, or 30.8%)

Is this sufficient evidence to confidently conclude that men and women are <u>not</u> equally likely to smoke?

YES

Worksheet

Did high school age boys' condom use during last intercourse increase between 2009 and 2014?

2009 Youth Risk Behavior Survey

69% of high school age boys used a condom

n = 2,640

2014 National Survey of Sexual Health and Behavior

79% of high school age boys used a condom

n = 57

Use α=0.05

Hypotheses about Differences in Means: Example #1

Someone claims that women have better vocabularies than men

To test this claim we use data from the General Social Survey, which contains a short vocabulary test

In these data:

$\bar{Y}_M = 5.93$ $s_M = 2.200$ $n_M = 8{,}709$

$\bar{Y}_F = 6.01$ $s_F = 2.153$ $n_F = 11{,}530$

Is this evidence sufficient to confidently conclude that (in the population) women have better vocabularies than men?

Hypotheses about Differences in Means: Example #1

1. State the null (H_0) and alternative (H_1) hypotheses

Because we are asking whether the difference in means is less than a particular value (zero), this is a one-sided hypothesis test

H_0: $\mu_W \leq \mu_M$ (or $\mu_M — \mu_F \geq 0$)

H_1: $\mu_W > \mu_M$ (or $\mu_M — \mu_F < 0$)

Hypotheses about Differences in Means: Example #1

2. Check that the sample data conform to basic assumptions; if they do not, then do not go any further

Each of the two random samples are small (<30) and symmetrically distributed with no outliers, and the populations of measurements is bell-shaped (Not True)

--- or ---

The size of the random samples are each large (≥30), regardless of the shape of the distributions of the measurements in the populations (True)

Hypotheses about Differences in Means: Example #1

3. Choose an α probability level … that is, a probability associated with incorrectly rejecting the null hypothesis

Let’s go with α=0.05 in this example

Hypotheses about Differences in Means: Example #1

4. Determine the "critical value" … that is, how large the test statistic must be in order to reject the null hypothesis at the given α level

Because we are dealing with sample means, our test statistic will be a t-score (not a Z-score)

What critical value t* corresponds to having 5% of the area under the normal curve below –t*?

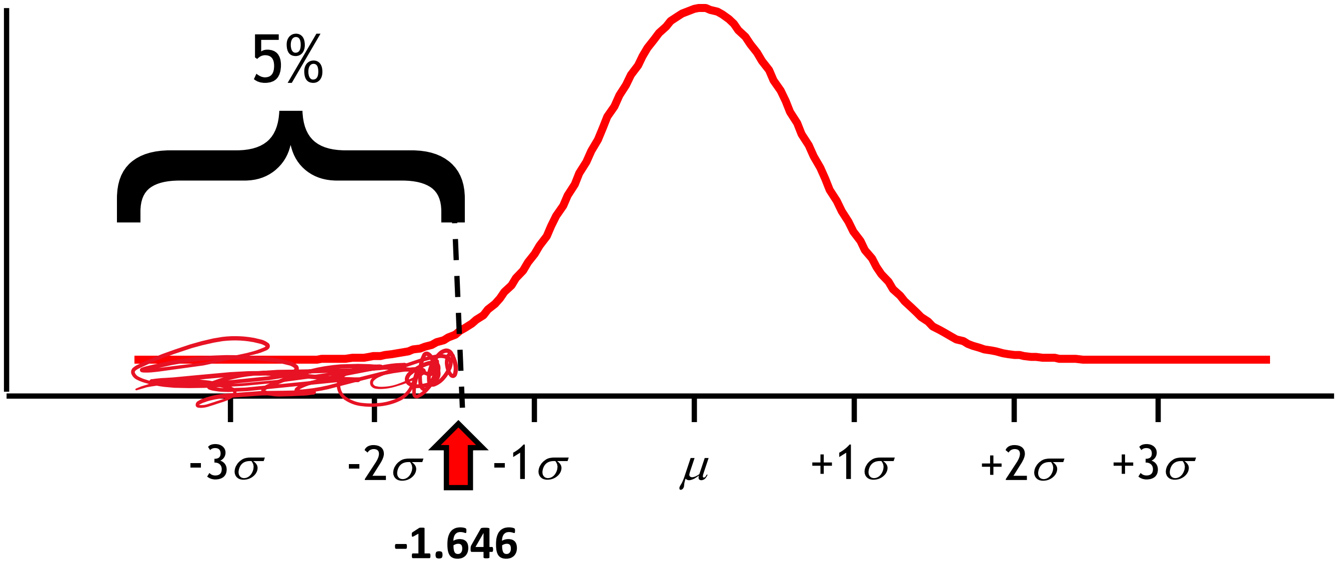
TABLE D

t distribution critical values

df	Upper-tail probability *p*											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300

Hypotheses about Differences in Means: Example #1

With α=0.05, n_M=8,709, n_F=11,530, & a one-sided test:



… so our critical value t* is -1.646

H_0: $\mu_W \leq \mu_M$ … fail to reject H_0 if $t \geq -1.646$

H_1: $\mu_W > \mu_M$ … reject H_0 if $t < -1.646$

Hypotheses about Differences in Means: Example #1

5. Calculate the test statistic

All test statistics for hypothesis testing…

…subtract the population value that is assumed to be true under the null from the observed (sample) value

…divide that figure by the standard deviation of the sampling distribution of the statistic in question

For differences in means:

$$t = \frac{\left(\bar{Y}_M - \bar{Y}_F\right) - \left(0\right)}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}}$$

Hypotheses about Differences in Means: Example #1

In our example:

$\overline{Y}_M = 5.93$ $s_M = 2.200$ $n_M = 8{,}709$

$\overline{Y}_F = 6.01$ $s_F = 2.153$ $n_F = 11{,}530$

$$t = \frac{\overline{Y}_M - \overline{Y}_F - 0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}} = \frac{5.93 - 6.01 - 0}{\sqrt{\frac{2.2^2}{8709} + \frac{2.153^2}{11530}}} = -2.58$$

Hypotheses about Differences in Means: Example #1

In our example:

$\overline{Y}_M = 5.93$ $s_M = 2.200$ $n_M = 8,709$

$\overline{Y}_F = 6.01$ $s_F = 2.153$ $n_F = 11,530$

$$t = \frac{\overline{Y}_M - \overline{Y}_F - 0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}} = \frac{5.93 - 6.01 - 0}{\sqrt{\frac{2.200^2}{8,709} + \frac{2.153^2}{11,530}}} = -2.58$$

Hypotheses about Differences in Means: Example #1

6. Compare the test statistic to the critical value

- If the test statistic is larger than the critical value, then reject H_0 (with probability of α of doing so even though H_0 should not actually be rejected)
- If the test statistic is less than or equal to the critical value, then do not reject H_0 (with probability of β of doing so even though H_0 should be rejected)

In our example, our critical value t* is -1.646

We obtained a test statistic of -2.58

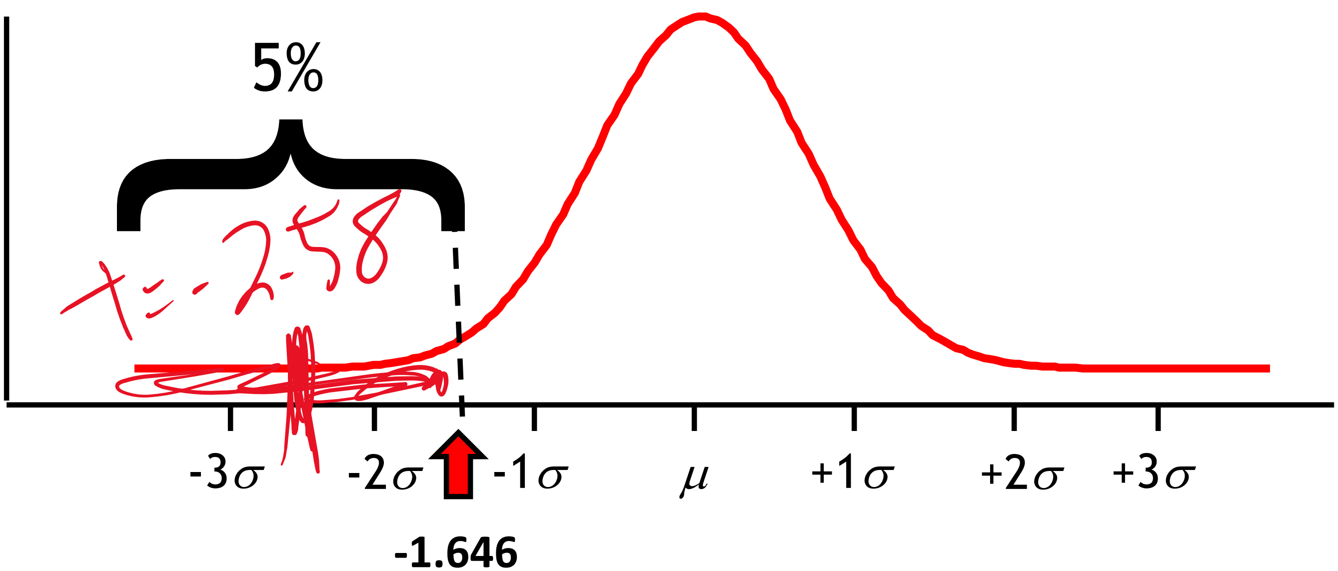
H_0: $\mu_W \leq \mu_M$ … fail to reject H_0 if $t \geq -1.646$

H_1: $\mu_W > \mu_M$ … reject H_0 if $t < -1.646$

Conclusion? Reject H_0

Hypotheses about Differences in Means: Example #1

With α=0.05, n_M=8,709, n_F=11,530, & a one-sided test:



… so our critical value t* is -1.646

H_0: $\mu_W \leq \mu_M$ … fail to reject H_0 if $t \geq -1.646$

H_1: $\mu_W > \mu_M$ … reject H_0 if $t < -1.646$

Hypotheses about Differences in Means: Example #1

Someone claims that women have better vocabularies than men

To test this claim we use data from the General Social Survey, which contains a short vocabulary test

In these data:

$\bar{Y}_M = 5.93$ $s_M = 2.200$ $n_M = 8{,}709$

$\bar{Y}_F = 6.01$ $s_F = 2.153$ $n_F = 11{,}530$

Is this evidence sufficient to confidently conclude that (in the population) women have better vocabularies than men?

YES

Worksheet

Do men and women tend to bring the same amount of money to class?

Using data from a survey conducted on the first day of my undergraduate research methods class, which included **21** men and **54** women, I found that men brought an average of **$38** with a standard deviation of **$59**, while women brought an average of **$45** with a standard deviation of **$85**

(Assume for now that students in my class represent a random sample of all undergraduate students)

Use α=0.05

Want More?

David Lane's Books

- http://onlinestatbook.com/2/logic_of_hypothesis_testing/logic_hypothesis.html
- http://davidmlane.com/hyperstat/logic_hypothesis.html

Lowry's Book (Chapter 7)

- http://vassarstats.net/textbook/

Dallal's Book

- http://www.jerrydallal.com/LHSP/sigtest.htm

Stat Trek's Discussion

- http://stattrek.com/hypothesis-test/hypothesis-testing.aspx