CIVL - 7904/8904

TRAFFIC FLOW THEORY

LECTURE -8

Chi-square Test

- How to determine the interval from a continuous distribution

$$I = \frac{Range}{1 + 3.322(logN)}$$

I-> Range of the class interval

N-> Number of observations

Range-> Total range

Find range = 32/(1+3.322log (200)) = 3.7

Consider range as 3.7 mph

Form Normal Distribution Table

Standard Error of the Mean

- As we take sample mean as the population mean and same for s.d, there is a dispersion called as "standard error of the mean"

$$s_{\overline{x}} = \frac{s}{\sqrt{N}}$$

- $s_{\overline{x}}$ -> Standard error of the mean (mph)
- s ->S. D. of the sample of individual speeds
- N->Number of individual speeds observed

Required Sample Size

$$n = \left(\frac{s}{s_{\bar{x}}}\right)^2$$

Or

$$n = \left(\frac{ts}{\varepsilon}\right)^2$$

- ε -> User specified allowable error
- n-> sample size
- t-> Coefficient of the standard error of the mean that represents user specified probability level

Ranges of the Population Mean

- $\mu - z\sigma < \overline{U} \leq \mu + z\sigma$
- Find out the range of speed for following confidence
  - 68.2%
  - 95%
  - 99%

Testing The Difference Between Means (Large Independent Samples)

Difference Between Means

- It is unlikely that any two samples of speed measurements will have exactly the same mean
  - Even when both are taken from the same population
- There may be some difference between means may be due to chance
  - While in other situations the differences are significant
- It is a critical task to explore is there are distinct differences between sample mean speeds

Two Sample Hypothesis Testing

In a two-sample hypothesis test, two parameters from two populations are compared.

- For a two-sample hypothesis test,

1. the **null hypothesis** H_o is a statistical hypothesis that usually states there is no difference between the speeds of two populations. The null hypothesis always contains the symbol $\leq$, $=$, or $\geq$.
2. the **alternative hypothesis** H_a is a statistical hypothesis that is true when H_o is false. The alternative hypothesis always contains the symbol $>$, $\neq$, or $<$.

Two Sample Hypothesis Testing

To write a null and alternative hypothesis for a two-sample hypothesis test, translate the claim made about the population parameters from a verbal statement to a mathematical statement.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 \leq \mu_2 \\ H_a: \mu_1 > \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 \geq \mu_2 \\ H_a: \mu_1 < \mu_2 \end{cases}$$

Regardless of which hypotheses used, $\mu_1 = \mu_2$ is always assumed to be true.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 > \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 < \mu_2 \end{cases}$$

Two Sample z-Test

Three conditions are necessary to perform a z-test for the difference between two population means μ_1 and μ_2.

1. The samples must be randomly selected.
2. The samples must be independent. Two samples are **independent** if the sample selected from one population is not related to the sample selected from the second population.
3. Each sample size must be at least 30, or, if not, each population must have a normal distribution with a known standard deviation.

Two Sample z-Test

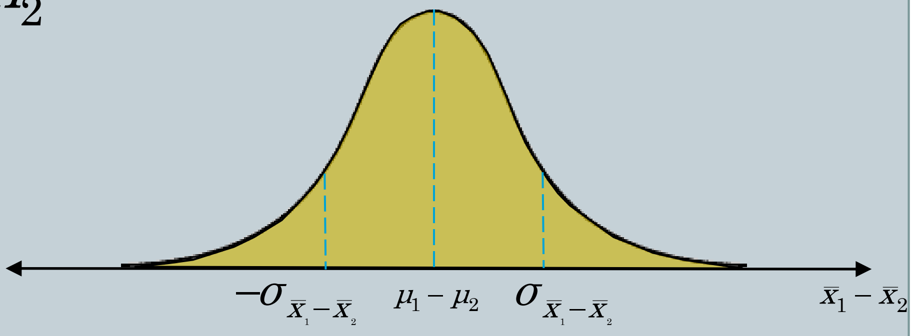
If these requirements are met, the sampling distribution for $\bar{x}_1 - \bar{x}_2$ (the difference of the sample means) is a normal distribution with mean and standard error of

$$\mu_{\bar{x}_1 - \bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$$

and

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Sampling distribution for $\bar{x}_1 - \bar{x}_2$



Two Sample z-Test

Two-Sample z-Test for the Difference Between Means

A two-sample z-test can be used to test the difference between two population means μ_1 and μ_2 when a large sample (at least 30) is randomly selected from each population and the samples are independent. The test statistic is and the standardized test statistic is

$$\bar{x}_1 - \bar{x}_2$$

The population standard deviation and z can be computed as follows.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} \quad \text{where} \quad \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Two Sample z-Test for the Means

Using a Two-Sample z-Test for the Difference Between Means (Large Independent Samples)

In Words	*In Symbols*
1. State the claim mathematically. Identify the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Sketch the sampling distribution.	
4. Determine the critical value(s).	
5. Determine the rejection regions(s).	

Two Sample z-Test for the Means

Using a Two-Sample z-Test for the Difference Between Means (Large Independent Samples)

In Words	*In Symbols*
6. Find the standardized test statistic.	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$
7. Make a decision to reject or fail to reject the null hypothesis.	If z is in the rejection region, reject H_o. Otherwise, fail to reject H_o.
8. Interpret the decision in the context of the original claim.	

One tailed and Two tailed Tests

- H1: $\mu_1 \neq \mu_2$-> Two tailed test
- H1: $\mu_1 > \mu_2$-> Right tailed test
- H1: $\mu_1 < \mu_2$-> Left tailed test

- At significance level of 0.05
  - For a two tailed test the z value is 1.96
  - For one tailed test the z value is 1.64

- In other words
  - For a two-tailed test accept H0 if z is within $\mp$ 1.96
  - For a one-tailed test accept H0 if z is within $\mp$1.64

Example –Two Tailed Test (1)

- Examine if there are significant differences between average speeds obtained from two samples at 0.05 significance level

Parameter	Study-1	Study-2
Mean	30.8	32
Standard Deviation	6.2	5.4
Sample Size	100	200

Example –Two Tailed Test (2)

- Hypothesis

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$$

- Compute standard deviation of the difference between means

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \text{sqrt}[(6.2^2/100)+(5.4^2/200)] = 0.76$$

Example –Two Tailed Test (3)

- Compute $z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$
- =32-30.8/0.76=1.57
- The z value is within ∓ 1.96, so accept Ho
- Conclusion: There appears no statistical difference between means drawn from two samples
- If the question was for different significance level then (two-tailed)
  - Use z =1 for alpha=0.32
  - Use z = 1.65 for alpha = 0.10
  - Use z = 1.96 for alpha = 0.05
  - Use z = 2.58 for alpha = 0.01

Z-scores for one-tailed and two-tailed

Cumulative Relative Frequencies and P Values Associated with Z Values in a Standard Gaussian Distribution

	One-Tailed	Two-Tailed	
Z	External Probability*	External Probability	Internal Probability
0	.5000	1	0
±0.500	.3090	.6170	.3830
±0.674	.2500	.5000	.5000
±1.000	.1585	.3170	.6830
±1.282	.1000	.2000	.8000
±1.500	.0670	.1340	.8660
±1.645	.0500	.1000	.9000
±1.960	.0250	.0500	.9500
±2.000	.0230	.0460	.9540
±2.240	.0125	.0250	.9750
±2.500	.0060	.0120	.9880
±2.576	.0050	.0100	.9900
±3.000	.0073	.0027	.9973
±3.290	.0005	.0010	.9990

* For positive Z values only. For negative Z values, the external probability (for values higher than −Z) is 1 minus the cited result. For example, at Z = −1.282, the one-tailed external probability is 1 − .1 = .9.

Example –One Tailed Test (1)

- Examine if there are significant differences between average speeds obtained from two samples at 0.01 significance level. Perform a one-tailed test.

Parameter	Study-1	Study-2
Mean	28	26.5
Standard Deviation	5.1	4.8
Sample Size	50	75

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 > \mu_2 \end{cases}$$

Example –One Tailed Test (2)

- Compute standard deviation of the difference between means
  = sqrt[(5.1²/50)+(4.8²/75)] =1.02

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

- Compute z = 28-26.5/1.02 = 1.47

- Z value is within ∓ 2.3
- Accept Ho

Testing The Difference Between Means (Small Independent Samples)

Two Sample *t*-Test

If samples of size less than 30 are taken from normally-distributed populations, a *t*-test may be used to test the difference between the population means μ_1 and μ_2.

Three conditions are necessary to use a *t*-test for small independent samples.

1. The samples must be randomly selected.
2. The samples must be independent. Two samples are **independent** if the sample selected from one population is not related to the sample selected from the second population.
3. Each population must have a normal distribution.

Two Sample *t*-Test

Two-Sample *t*-Test for the Difference Between Means

A **two-sample *t*-test** is used to test the difference between two population means μ_1 and μ_2 when a sample is randomly selected from each population. Performing this test requires each population to be normally distributed, and the samples should be independent. The standardized test statistic is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}.$$

If the population variances are equal, then information from the two samples is combined to calculate a **pooled estimate of the standard deviation** $\hat{\sigma}$.

$$\hat{\sigma} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Continued.

Two Sample t-Test

Two-Sample t-Test (Continued)

The standard error for the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$ Variances equal

and d.f.= $n_1 + n_2 - 2$.

If the population variances are not equal, then the standard error is

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$ Variances not equal

and d.f = smaller of $n_1 - 1$ or $n_2 - 1$.

Normal or *t*-Distribution?

Are both sample sizes at least 30?
- **Yes** → Use the *z*-test.
- **No** → Are both populations normally distributed?
  - **No** → You cannot use the *t*-test.
  - **Yes** → Are both population standard deviations known?
    - **Yes** → Use the *t*-test.
    - **No** → Are the population variances equal?
      - **Yes** → Use the *t*-test with $\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ and d.f = $n_1 + n_2 - 2$.
      - **No** → Use the *t*-test with $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ and d.f = smaller of $n_1 - 1$ or $n_2 - 1$.

Two Sample t-Test for the Means

Using a Two-Sample t-Test for the Difference Between Means (Small Independent Samples)

In Words	*In Symbols*
1. State the claim mathematically. Identify the null and alternative hypotheses.	State H_o and H_a.
2. Specify the level of significance.	Identify α.
3. Identify the degrees of freedom and sketch the sampling distribution.	d.f. = $n_1 + n_2 - 2$ or d.f. = smaller of $n_1 - 1$ or $n_2 - 1$.
4. Determine the critical value(s).	

Continued.

Two Sample t-Test for the Means

Using a Two-Sample t-Test for the Difference Between Means (Small Independent Samples)

In Words	*In Symbols*
5. Determine the rejection regions(s).	
6. Find the standardized test statistic.	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$
7. Make a decision to reject or fail to reject the null hypothesis.	If t is in the rejection region, reject H_o. Otherwise, fail to reject H_o.
8. Interpret the decision in the context of the original claim.	

Example- Two tailed test

- Examine if there are significant differences between average speeds obtained from two samples at 0.01 significance level

Parameter	Study-1	Study-2
Mean	48	44.5
Standard Deviation	4.3	4.1
Sample Size	15	18

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$$

Example- Two tailed test

- Estimate $\hat{\sigma} = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ =4.19
- SD $\sigma_{\bar{x}_1 - \bar{x}_2} = \hat{\sigma} \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$ = 1.46
- t-value $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}.$ = -2.38
- Reject hypothesis as the t-value is not in the range of $\mp$ 2.131

Testing The Difference Between Means (Dependent Samples)

Independent and Dependent Samples

Two samples are **independent** if the sample selected from one population is not related to the sample selected from the second population. Two samples are **dependent** if each member of one sample corresponds to a member of the other sample. Dependent samples are also called **paired samples** or **matched samples**.

Independent Samples

Dependent Samples

Independent and Dependent Samples

Example:
Classify each pair of samples as independent or dependent.

Sample 1: The weight of 24 students in a first-grade class
Sample 2: The height of the same 24 students

These samples are dependent because the weight and height can be paired with respect to each student.

Sample 1: The average price of 15 new trucks
Sample 2: The average price of 20 used sedans

These samples are independent because it is not possible to pair the new trucks with the used sedans. The data represents prices for different vehicles.

t-Test for the Difference Between Means

To perform a two-sample hypothesis test with dependent samples, the difference between each data pair is first found:

$d = x_1 - x_2$ Difference between entries for a data pair.

The test statistic is the mean of these differences. $\bar{d}$

$$\bar{d} = \frac{\sum d}{n}.$$

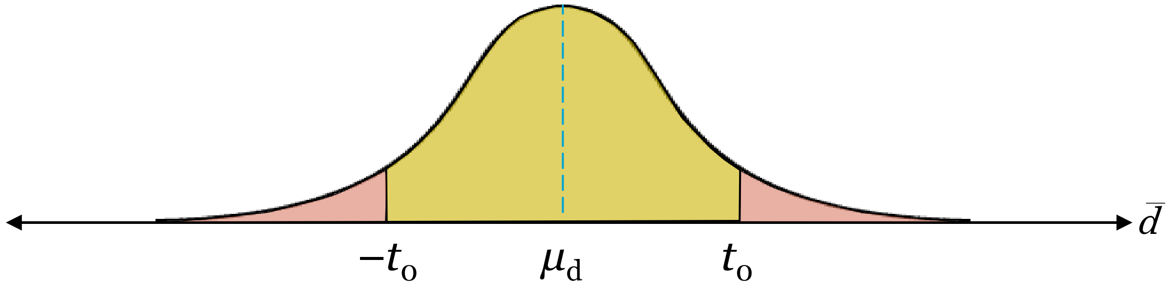
Mean of the differences between paired data entries in the dependent samples.

Three conditions are required to conduct the test.

t-Test for the Difference Between Means

1. The samples must be randomly selected.
2. The samples must be dependent (paired).
3. Both populations must be normally distributed.

If these conditions are met, then the sampling distribution for $\bar{d}$ is approximated by a *t*-distribution with $n - 1$ degrees of freedom, where n is the number of data pairs.



t-Test for the Difference Between Means

The following symbols are used for the t-test for μ_d.

Symbol	Description
n	The number of pairs of data
d	The difference between entries for a data pair, $d = x_1 - x_2$
μ_d	The hypothesized mean of the differences of paired data in the population
$\bar{d}$	The mean of the differences between the paired data entries in the dependent samples $\bar{d} = \frac{\sum d}{n}$
s_d	The standard deviation of the differences between the paired data entries in the dependent samples $s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}}$

t-Test for the Difference Between Means

_t_-Test for the Difference Between Means

A *t*-test can be used to test the difference of two population means when a sample is randomly selected from each population. The requirements for performing the test are that each population must be normal and each member of the first sample must be paired with a member of the second sample.

The **test statistic** is

$$\bar{d} = \frac{\sum d}{n}$$

and the **standardized test statistic** is

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}.$$

The degrees of freedom are

d.f. = $n - 1$.

t-Test for the Difference Between Means

Using the t-Test for the Difference Between Means (Dependent Samples)

In Words	*In Symbols*
1. State the claim mathematically. Identify the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Identify the degrees of freedom and sketch the sampling distribution.	d.f. = $n - 1$
4. Determine the critical value(s).	

Continued.

t-Test for the Difference Between Means

Using a Two-Sample *t*-Test for the Difference Between Means (Small Independent Samples)

In Words	*In Symbols*
5. Determine the rejection region(s).	
6. Calculate $\bar{d}$ and s_d. Use a table.	$\bar{d} = \frac{\sum d}{n}$ $s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}}$
7. Find the standardized test statistic.	$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$

t-Test for the Difference Between Means

Using a Two-Sample *t*-Test for the Difference Between Means (Small Independent Samples)

In Words

8. Make a decision to reject or fail to reject the null hypothesis.

9. Interpret the decision in the context of the original claim.

In Symbols

If t is in the rejection region, reject H_o. Otherwise, fail to reject H_o.

t-Test for the Difference Between Means

Example:
The table shows the speeds of 6 locations before and after a horizontal curve design. At $\alpha = 0.05$, is there enough evidence to conclude that the speeds after the redesign are better than the speeds before ?

Student	1	2	3	4	5	6
Speed (before kmph)	85	96	70	76	81	78
Speed (after kmph)	88	85	89	86	92	89

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$ (Claim)

Continued.

t-Test for the Difference Between Means

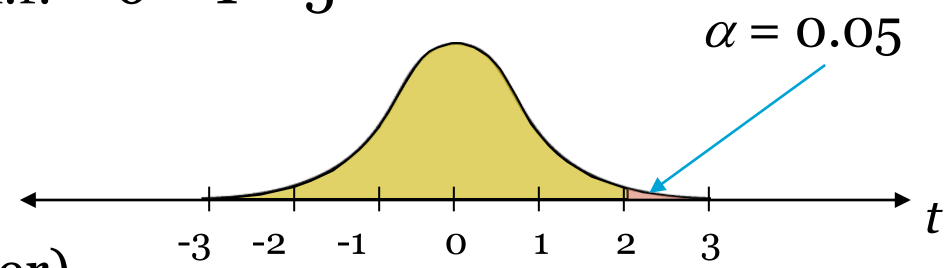
Example continued:

H_0: $\mu_d \le 0$

H_a: $\mu_d > 0$ (Claim)

d = (speed before) – (speed after)

d.f. = 6 – 1 = 5



$t_0 = 2.015$

Student	1	2	3	4	5	6
Speed (before)	85	96	70	76	81	78
Speed (after)	88	85	89	86	92	89
d	–3	11	–19	–10	–11	–11
d^2	9	121	361	100	121	121

$\sum d = -43$

$\sum d^2 = 833$

$$\bar{d} = \frac{\sum d}{n} = \frac{-43}{6} \approx -7.167$$

$$s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}} = \sqrt{\frac{6(833) - 1849}{6(5)}} \approx \sqrt{104.967} \approx 10.245$$

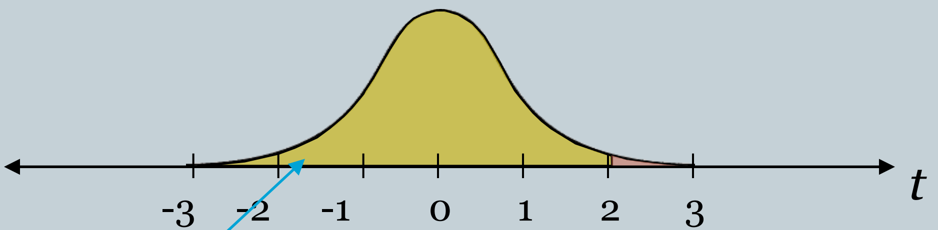
Continued.

t-Test for the Difference Between Means

Example continued:

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$ (Claim)



$t_0 = 2.015$

The standardized test statistic is

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-7.167 - 0}{10.245 / \sqrt{6}} \approx -1.714.$$

Accept H_0.

There is not enough evidence at the 5% level to support the claim that the speeds after the redesign of horizontal curve are better than the speeds before.

Testing The Difference Between Variance(Two Samples)

Variance Test

H_o: $\sigma^2_1 = \sigma^2_2$ or $\frac{\sigma^2_1}{\sigma^2_2} = 1$ vs. H_a: $\sigma^2_1 \neq \sigma^2_2$ or $\frac{\sigma^2_1}{\sigma^2_2} \neq 1$ (Two-Tailed)

vs. H_a: $\sigma^2_1 < \sigma^2_1$ or $\frac{\sigma^2_1}{\sigma^2_2} < 1$ (Left-Tailed)

vs. H_a: $\sigma^2_1 > \sigma^2_1$ or $\frac{\sigma^2_1}{\sigma^2_2} > 1$ (Right-Tailed)

Test Statistic

$$F = \frac{s^2_1}{s^2_2}$$

Where $s^2_1 > s^2_2$

Df, n1-1, and n2-1

Example

Parameter	Study-1	Study-2
Mean	30.8	32
Standard Deviation	6.2	5.4
Sample Size	100	200

F calculated = 1.31
F tabulated = 1.332

Null hypothesis is accepted

Testing The Difference Between Proportions

Two Sample *z*-Test for Proportions

A z-test is used to test the difference between two population proportions, p_1 and p_2.

Three conditions are required to conduct the test.

1. The samples must be randomly selected.
2. The samples must be independent.
3. The samples must be large enough to use a normal sampling distribution. That is,

$$n_1 p_1 \geq 5, \quad n_1 q_1 \geq 5,$$

$$n_2 p_2 \geq 5, \text{ and } n_2 q_2 \geq 5.$$

Two Sample z-Test for Proportions

If these conditions are met, then the sampling distribution for $\hat{p}_1 - \hat{p}_2$ is a normal distribution with mean

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$$

and standard error

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n2_1}\right)}, \text{ where } \bar{q} = 1 - \bar{p}.$$

A weighted estimate of p_1 and p_2 can be found by using

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}, \text{ where } x_1 = n_1\hat{p}_1 \text{ and } x_2 = n_2\hat{p}_2.$$

Two Sample z-Test for Proportions

Two Sample z-Test for the Difference Between Proportions

A two sample z-test is used to test the difference between two population proportions p_1 and p_2 when a sample is randomly selected from each population.

The **test statistic** is

$$\hat{p}_1 - \hat{p}_1$$

and the **standardized test statistic** is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \text{ and } \bar{q} = 1 - \bar{p}.$$

Note:

$n_1\bar{p}$, $n_1\bar{q}$, $n_2\bar{p}$, and $n_2\bar{q}$ must be at least 5.

Two Sample z-Test for Proportions

Using a Two-Sample z-Test for the Difference Between Proportions

In Words	*In Symbols*
1. State the claim. Identify the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the critical value(s).	Use Table 4 in Appendix B.
4. Determine the rejection region(s).	
5. Find the weighted estimate of p_1 and p_2.	$\bar{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

Continued.

Two Sample z-Test for Proportions

Using a Two-Sample z-Test for the Difference Between Proportions

In Words

6. Find the standardized test statistic.
7. Make a decision to reject or fail to reject the null hypothesis.
8. Interpret the decision in the context of the original claim.

In Symbols

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

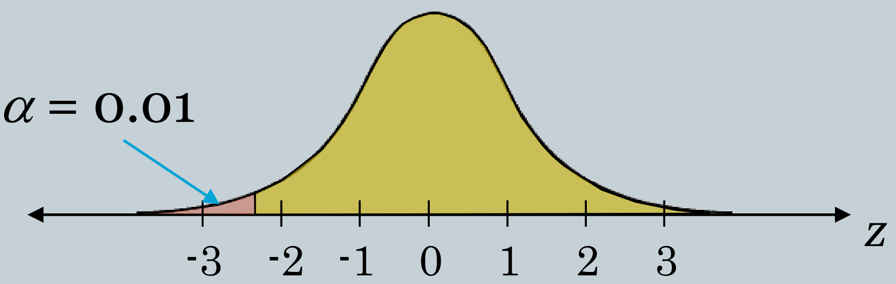
If z is in the rejection region, reject H_0. Otherwise, fail to reject H_0.

Two Sample z-Test for Proportions

Example:
A recent survey stated that male college students use freeways less than female college students. In a survey of 1245 male students, 361 said they use freeway. In a survey of 1065 female students, 341 said they use freeway. At $\alpha = 0.01$, can you support the claim that the proportion of female college students use freeway more than male students?

$H_0: p_1 \geq p_2$

$H_a: p_1 < p_2$ (Claim)



$-z_0 = -2.33$

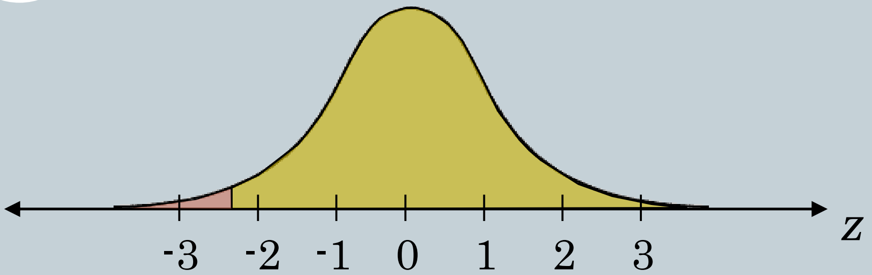
Continued.

Two Sample z-Test for Proportions

Example continued:

$H_0: p_1 \geq p_2$

$H_a: p_1 < p_2$ (Claim)



$-z_0 = -2.33$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{361 + 341}{1245 + 1065} = \frac{702}{2310} \approx 0.304$$

$$\bar{q} = 1 - \bar{p} = 1 - 0.304 = 0.696$$

Because 1245(0.304), 1245(0.696), 1065(0.304), and 1065(0.696) are all at least 5, we can use a two-sample z-test.

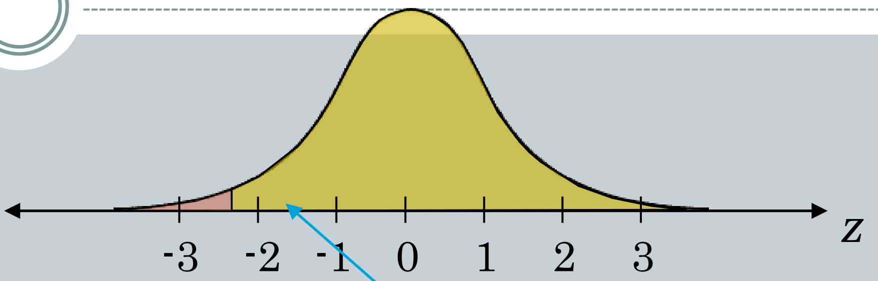
Continued.

Two Sample z-Test for Proportions

Example continued:

$H_0: p_1 \geq p_2$

$H_a: p_1 < p_2$ (Claim)



$-z_0 = -2.33$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.29 - 0.32) - 0}{\sqrt{(0.304)(0.696)\left(\frac{1}{1245} + \frac{1}{1065}\right)}} \approx -1.56$$

Fail to reject H_0.