9. Large-Sample Tests of Hypotheses

Adams Wai Kin Kong
School of Computer Science and Engineering
Nanyang Technological University, Singapore
adamskong@ntu.edu.sg

9.3 A Large-Sample Test of Hypothesis for the Difference Between Two Population Means

A Large-Sample Test of Hypothesis for the Difference Between Two Population Means (1 of 5)

If n is large, $(\bar{x}_1 - \bar{x}_2)$ follows an approximate normal distribution (CLT) with mean $(\mu_1 - \mu_2)$ and standard error

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{estimated by} \quad SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Using this information, we can test e.g., the hypothesis.

$$H_o:\ (\mu_1 - \mu_2) = 0$$
$$H_A:\ (\mu_1 - \mu_2) \neq 0$$

The formal testing procedure is described below.

Large-Sample Statistical Test for $(\mu_1 - \mu_2)$

1. Null hypothesis: $H_0 : (\mu_1 - \mu_2) = D_0$ where D_0 is some specified difference that you wish to test. For many tests, you will hypothesize that there is no difference between μ_1 and μ_2; that is, $D_0 = 0$.

A Large-Sample Test of Hypothesis for the Difference Between Two Population Means (3 of 5)

2. Alternative hypothesis:

One-Tailed Test	Two-Tailed **Test**
$H_a: (\mu_1 - \mu_2) > D_0$	$H_a: (\mu_1 - \mu_2) \neq D_0$
[or $H_a: (\mu_1 - \mu_2) < D_0$]	

3. Test statistic: $z \approx \dfrac{(\bar{x}_1 - \bar{x}_2) - D_0}{\text{SE}} = \dfrac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

A Large-Sample Test of Hypothesis for the Difference Between Two Population Means (4 of 5)

4. Rejection region: Reject H_0 when

One-Tailed Test

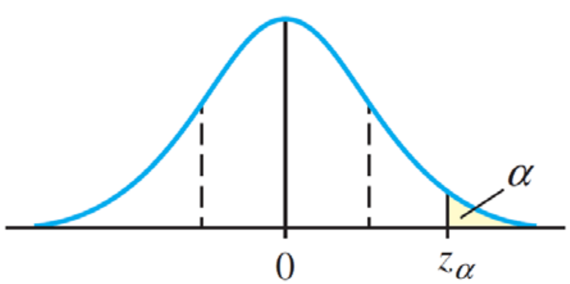
$$z > z_{\alpha}$$

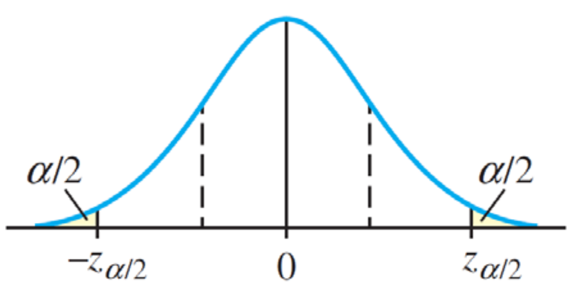
[or $z < -z_{\alpha}$ when the alternative hypothesis is $H_a : (\mu_1 - \mu_2) < D_0$]

or when p-value < α

Two-Tailed Test

$$z > z_{\alpha/2} \quad \text{or} \quad z < -z_{\alpha/2}$$





A Large-Sample Test of Hypothesis for the Difference Between Two Population Means (5 of 5)

Assumptions: The samples are randomly and independently selected from the two populations with $n_1 \geq 30$ and $n_2 \geq 30$.

Example 9.9

To determine whether car ownership affects a student's academic achievement, random samples of 100 car owners and 100 nonowners were drawn from the student body.

The grade point average for the $n_1 = 100$ nonowners had an average and variance equal to $\bar{x}_1 = 2.70$ and $s_1^2 = 0.36$ while $\bar{x}_2 = 2.54$ and $s_2^2 = 0.40$ for the $n_2 = 100$ car owners. Do the data present sufficient evidence to indicate a difference in the mean achievements between car owners and nonowners? Using $\alpha = 0.05$ for the test.

NANYANG TECHNOLOGICAL UNIVERSITY

Example 9.9 – Solution (1 of 5)

To detect a difference, if it exists, between the mean academic achievements for nonowners of cars μ_1 and car owners μ_2, you will test the null hypothesis that there is no difference between the means against the alternative hypothesis that $(\mu_1 - \mu_2) \neq 0$; that is,

$$H_0 : (\mu_1 - \mu_2) = D_0 = 0 \qquad \text{versus} \qquad H_A : (\mu_1 - \mu_2) \neq 0$$

NANYANG TECHNOLOGICAL UNIVERSITY

Example 9.9 – Solution (2 of 5)

Substituting into the formula for the test statistic, you get

$$z \approx \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.70 - 2.54}{\sqrt{\frac{.36}{100} + \frac{.40}{100}}} = 1.84$$

The critical value approach: Using a two-tailed test with significance level $\alpha = .05$, you place $\alpha/2 = 0.025$ in each tail of the z distribution and reject H_0 if $z > 1.96$ or $z < -1.96$.

NANYANG TECHNOLOGICAL UNIVERSITY

Example 9.9 – Solution (3 of 5)

Since $-1.96<1.84<1.96$, H_0 cannot be rejected.

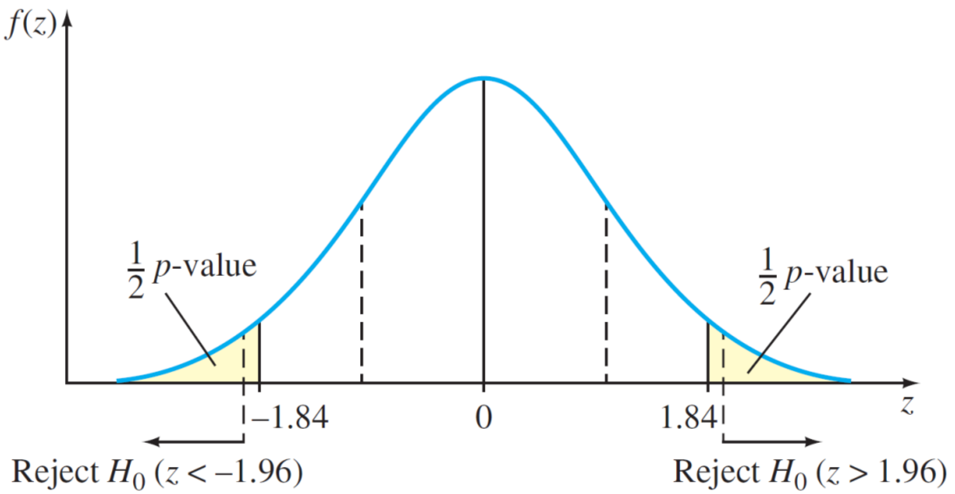


Figure 9.12

Example 9.9 – Solution (4 of 5)

That is, there is insufficient evidence to declare a difference in the average academic achievements for the two groups.

The *p*-value approach: Calculate the *p*-value, the probability that z is greater than $z = 1.84$ plus the probability that z is less than $z = -1.84$, as shown in Figure 9.12:

$$
\begin{aligned}
p\text{-value} &= P(z > 1.84) + P(z < -1.84) \\
&= (1 - .9671) + .0329 \\
&= .0658
\end{aligned}
$$

Example 9.9 – Solution (5 of 5)

The p-value lies between 0.10 and 0.05, so you can reject H_0 at the 0.10 level but not at the 0.05 level of significance.

Since the p-value of 0.0658 exceeds the specified significance level $\alpha = 0.05$, H_0 cannot be rejected.

NANYANG TECHNOLOGICAL UNIVERSITY

Hypothesis Testing and Confidence Intervals

NANYANG TECHNOLOGICAL UNIVERSITY

Hypothesis Testing and Confidence Intervals (1 of 1)

If the hypothesized value *lies outside* of the confidence limits (confidence Interval), the null hypothesis is rejected at the α level of significance.

Example 9.10

Construct a 95% confidence interval for the difference in average academic achievements between car owners and nonowners.

Using the confidence interval, can you conclude that there is a difference in the population means for the two groups of students?

Example 9.10 – Solution (1 of 3)

We know that for the difference in two population means, the confidence interval is approximated as

$$(\bar{x}_1 - \bar{x}_2) \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(2.70 - 2.54) \pm 1.96\sqrt{\frac{.36}{100} + \frac{.40}{100}}$$

$$0.16 \pm 0.17$$

Example 9.10 – Solution (2 of 3)

or $-0.01 < (\mu_1 - \mu_2) < 0.33$.

This interval gives you a range of possible values for the difference in the population means.

Since the hypothesized difference, $(\mu_1 - \mu_2) = 0$, is contained in the confidence interval, you should not reject H_0.

Example 9.10 – Solution (3 of 3)

You cannot tell from the interval whether the difference in the means is negative (−), positive (+), or zero (0)—the latter of the three would indicate that the two means are the same.

There is not enough evidence to indicate that there is a difference in the average achievements for car owners versus nonowners. The conclusion is the same one reached in Example 9.9.

9.4 A Large-Sample Test of Hypothesis for a Binomial Proportion

If n is large, $\hat{p}$ follows an approximate normal distribution (CLT) with a mean p and a standard error

$$SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{pq}{n}}$$

Large-Sample Statistical Test for *p*

1. Null hypothesis: $H_0 : p = p_0$
2. Alternative hypothesis:

One-Tailed Test

$H_a : p > p_0$

(or, $H_a : p < p_0$)

Two-Tailed Test

$H_a : p \neq p_0$

3. Test statistic: $z = \dfrac{\hat{p} - p_0}{\text{SE}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$ with $\hat{p} = \dfrac{x}{n}$

where x is the number of successes in n binomial trials.

4. Rejection region: Reject H_0 when

One-Tailed Test

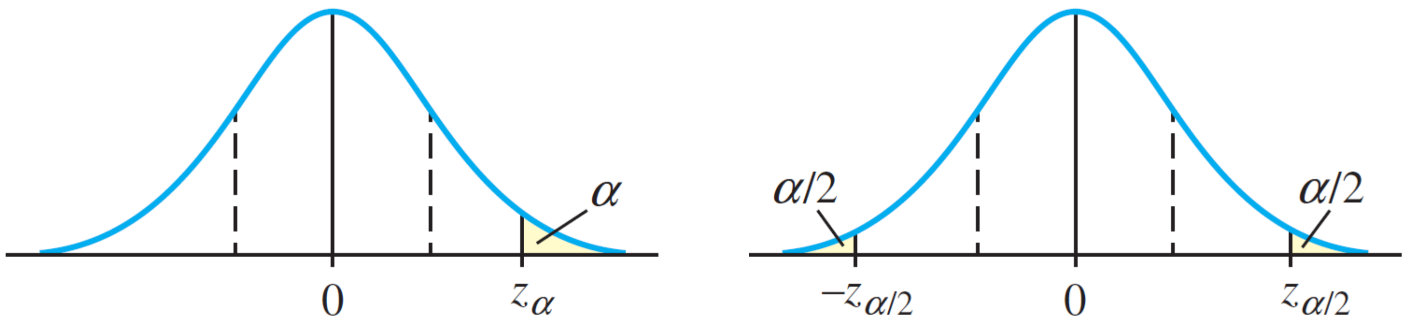
$z > z_\alpha$

(or, $z < -z_\alpha$ when the alternative hypothesis is $H_a : p < p_0$)

Two-Tailed Test

$z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$

or when p-value < α



Assumption: The sampling satisfies the assumptions of a binomial experiment (see Section 5.2), and n is large enough so that the sampling distribution of $\hat{p}$ can be approximated by a normal distribution ($np_0 > 5$ and $nq_0 > 5$).

Example 9.11

Regardless of age, about 20% of American adults participate in fitness activities at least twice a week. Does this percentage decrease as people get older? In a local survey of n = 100 adults over 40 years old, a total of 15 people indicated that they participated in a fitness activity at least twice a week. Do these data indicate that the participation rate for adults over 40 years of age is significantly less than the 20% figure? Calculate the p-value and use it to draw the appropriate conclusions. Using $\alpha = 0.1$

NANYANG TECHNOLOGICAL UNIVERSITY

Example 9.11 – Solution (1 of 3)

Assuming that

- the sampling procedure satisfies the requirements of a binomial experiment;
- the true value *of* p *is* $p_0 = 0.2$

$H_0 : p = 0.2$ versus $H_a : p < 0.2$

The observed value of

$$\hat{p} = \frac{x}{n} = \frac{15}{100} = 0.15$$

Example 9.11 – Solution (2 of 3)

The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{.15 - .20}{\sqrt{\frac{(.20)(.80)}{100}}} = -1.25$$

The p-value associated with this test is found as the area under the standard normal curve to the left of $z = -1.25$ as shown in Figure.

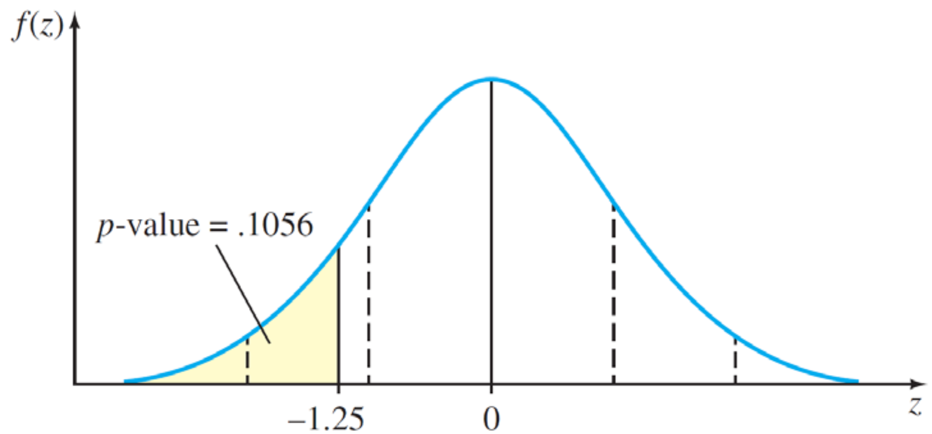


Figure 9.13

Example 9.11 – Solution (3 of 3)

Therefore,

$$p\text{-value} = P(z < -1.25) = 0.1056$$

If 0.1056 is greater than 0.10, you would not reject H_0.

There is insufficient evidence to conclude that the percentage of adults over age 40 who participate in fitness activities twice a week is less than 20%.

Statistical Significance and Practical Importance

Statistical Significance and Practical Importance (1 of 4)

In statistical language, the word *significant* does not necessarily mean practically "important," but only that the results could not have occurred by chance.

For example, suppose that in Example 9.11, the researcher had used n = 400 instead of n = 100 adults in her experiment and had observed the same sample proportion.

Statistical Significance and Practical Importance (2 of 4)

The test statistic is now

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{.15 - .20}{\sqrt{\frac{(.20)(.80)}{400}}} = -2.50$$

with

$$p\text{-value} = P(z < -2.50) = .0062$$

Now H_0 is rejected and the results are *highly significant.*

There is sufficient evidence to indicate that the percentage of adults over age 40 who participate in physical fitness activities is less than 20%.

Statistical Significance and Practical Importance (3 of 4)

However, is this drop in activity really *important*?

Suppose that physicians would be concerned only about a drop in physical activity of more than 10%.

If there had been a drop of more than 10% in physical activity, this would imply that the true value of p was less than 0.10. What is the largest possible value of p?

Statistical Significance and Practical Importance (4 of 4)

Using a 90% confidence interval,

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.15 + 1.645\sqrt{\frac{0.15 \times 0.85}{400}} = 0.15 \pm 0.029$$

The confidence is [0.121 0.179].

The physical activity for adults aged 40 and older has dropped from 20%, but you cannot say that it has dropped below 10%. So, the results, although *statistically significant*, are not *practically important*.

A Large-Sample Test of Hypothesis for the Difference Between Two Binomial Proportions

A Large-Sample Test of Hypothesis for the Difference Between Two Binomial Proportions (1 of 5)

When random and independent samples are selected from two *binomial* populations, the focus of the experiment may be the difference $(p_1 - p_2)$ in the proportions of individuals or items possessing a specified characteristic in the two populations. In this situation, you can use the difference in the sample proportions $(\hat{p}_1 - \hat{p}_2)$ along with its standard error,

$$\mathrm{SE} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

in the form of a *z* statistic to test for a significant difference in the two population proportions.

Large-Sample Statistical Test for $(p_1 - p_2)$

1. Null hypothesis: $H_0 : (p_1 - p_2) = 0$ or equivalently $H_0 : p_1 = p_2$

2. Alternative hypothesis:

One-Tailed Test	**Two-Tailed Test**
$H_a : (p_1 - p_2) > 0$ [or $H_a : (p_1 - p_2) < 0$]	$H_a : (p_1 - p_2) \neq 0$

NANYANG TECHNOLOGICAL UNIVERSITY

3. Test statistic: $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - 0}{\text{SE}} = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}} = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\dfrac{pq}{n_1} + \dfrac{pq}{n_2}}}$

where $\hat{p}_1 = x_1 / n_1$ and $\hat{p}_2 = x_2 / n_2$. Since the common value of $p_1 = p_2 = p$ (used in the standard error) is unknown, it is estimated by

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

and the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\dfrac{\hat{p}\hat{q}}{n_1} + \dfrac{\hat{p}\hat{q}}{n_2}}} \quad \text{or} \quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

4. Rejection region: Reject H_0 when

One-Tailed Test

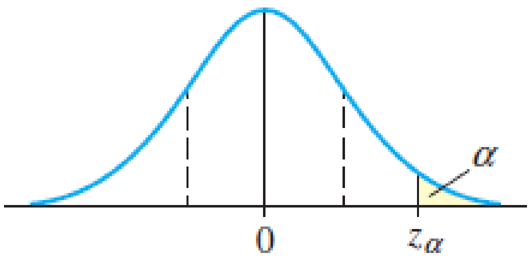
$z > z_\alpha$

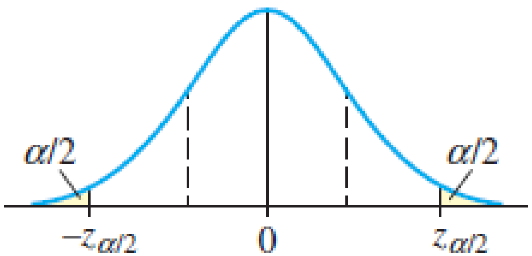
[or, $z < -z_\alpha$ when the alternative hypothesis is $H_a : (p_1 - p_2) < p_0$]

Two-Tailed Test

$z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$

or when p-value < α





A Large-Sample Test of Hypothesis for the Difference Between Two Binomial Proportions (5 of 5)

Assumptions: Samples are selected in a random and independent manner from two binomial populations, and n_1 and n_2 are large enough so that the sampling distribution of $(\hat{p}_1 - \hat{p}_2)$ can be approximated by a normal distribution.

That is $n_1\hat{p}_1$, $n_1\hat{q}_1$, $n_2\hat{p}_2$, and $n_2\hat{q}_2$ should all be greater than 5.

Example 9.12

The records of a hospital show that 52 men in a sample of 1000 men versus 23 women in a sample of 1000 women were admitted because of heart disease.

Do these data present sufficient evidence to indicate a higher rate of heart disease among men admitted to the hospital? Use $\alpha = 0.05$.

Example 9.12 – Solution (1 of 5)

Assume that the number of patients admitted for heart disease has an approximate binomial probability distribution for both men and women with parameters p_1 and p_2, respectively.

Then, because you wish to determine whether $p_1 > p_2$, you will test the null hypothesis $p_1 = p_2$ —that is, $H_0 : (p_1 - p_2) = 0$ —against the alternative hypothesis : $H_a : p_1 > p_2$ or, equivalently, $H_a : (p_1 - p_2) > 0$.

Example 9.12 – Solution (2 of 5)

To conduct this test, use the z statistic and approximate the standard error using the pooled estimate of p. Since H_a implies a one-tailed test, you can reject H_0 only for large values of z. Thus, for $\alpha = .05$, you can reject H_0 if $z > 1.645$.

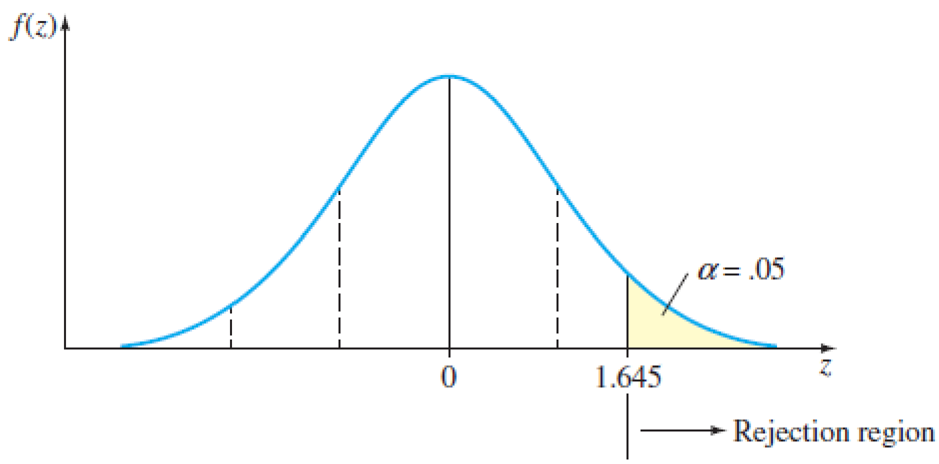


Figure 9.14

Example 9.12 – Solution (3 of 5)

The pooled estimate of *p* required for the standard error is

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{52 + 23}{1000 + 1000} = .0375$$

and the test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.052 - .023}{\sqrt{(.0375)(.9625)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 3.41$$

Example 9.12 – Solution (4 of 5)

Since the calculated value of *z* falls in the rejection region, you can reject the hypothesis that $p_1 = p_2$. The data present sufficient evidence to indicate that the percentage of men entering the hospital because of heart disease is higher than that of women.

NANYANG TECHNOLOGICAL UNIVERSITY

Example 9.12 – Solution (5 of 5)

How *much higher* is the proportion of men than women entering the hospital with heart disease? A 95% lower one-sided confidence bound will help you find the lowest likely value for the difference.

$$(\hat{p}_1 - \hat{p}_2) - 1.645\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(.052 - .023) - 1.645\sqrt{\frac{.052(.948)}{1000} + \frac{.023(.977)}{1000}}$$

$$.029 - .014$$

or $(p_1 - p_2) > .015$. The proportion of men is roughly 1.5% higher than women.

NANYANG TECHNOLOGICAL UNIVERSITY