STATISTICAL INFERENCES (2cr)
Chapter 9 One- and Two-Sample Estimation Problems

Zhong Guan

Math, IUSB

Outline

Outline

Outline

1. 9.10 Estimating a Proportion
   - Estimating Single Proportion
   - Sample Size for Estimating p
2. 9.11 Difference of Two Proportions

Do you know how to read the polls' results?

- If the the opinion poll says candidate A leads B 3%: 45% vs 42%, what does it mean?
- The poll gives sample proportions not the population proportions.
- You need to find the margin of error.

Do you know how to read the polls' results?

- If the the opinion poll says candidate A leads B 3%: 45% vs 42%, what does it mean?
- The poll gives sample proportions not the population proportions.
- You need to find the margin of error.

Do you know how to read the polls' results?

- If the the opinion poll says candidate A leads B 3%: 45% vs 42%, what does it mean?
- The poll gives sample proportions not the population proportions.
- You need to find the margin of error.

Do you know how to read the polls' results?

- If the the opinion poll says candidate A leads B 3%: 45% vs 42%, what does it mean?
- The poll gives sample proportions not the population proportions.
- You need to find the margin of error.

How does the poll work?

- The poll collects the opinions of n voters.
- Y of the n voters will vote for A.
- The sample proportion is $\hat{p} = Y/n$ which is a point estimate of the population proportion p.
- Y has a binomial distribution $b(n, p)$.
- How to construct a confidence interval for p?

How does the poll work?

- The poll collects the opinions of n voters.
- Y of the n voters will vote for A.
- The sample proportion is $\hat{p} = Y/n$ which is a point estimate of the population proportion p.
- Y has a binomial distribution $b(n, p)$.
- How to construct a confidence interval for p?

How does the poll work?

- The poll collects the opinions of n voters.
- Y of the n voters will vote for A.
- The sample proportion is $\hat{p} = Y/n$ which is a point estimate of the population proportion p.
- Y has a binomial distribution $b(n, p)$.
- How to construct a confidence interval for p?

How does the poll work?

- The poll collects the opinions of n voters.
- Y of the n voters will vote for A.
- The sample proportion is $\hat{p} = Y/n$ which is a point estimate of the population proportion p.
- Y has a binomial distribution $b(n, p)$.
- How to construct a confidence interval for p?

How does the poll work?

- The poll collects the opinions of n voters.
- Y of the n voters will vote for A.
- The sample proportion is $\hat{p} = Y/n$ which is a point estimate of the population proportion p.
- Y has a binomial distribution $b(n, p)$.
- How to construct a confidence interval for p?

Normal Approximation of Sampling Distribution of $\hat{p}$:

- By the C.L.T., if n is large, then

$$P\left[-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}\right] \approx 1-\alpha.$$

- Since $p \approx \hat{p}$,

$$P\left[-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \leqslant z_{\alpha/2}\right] \approx 1-\alpha.$$

Normal Approximation of Sampling Distribution of $\hat{p}$:

- By the C.L.T., if n is large, then

$$P\left[-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}\right] \approx 1-\alpha.$$

- Since $p \approx \hat{p}$,

$$P\left[-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \leqslant z_{\alpha/2}\right] \approx 1-\alpha.$$

Confidence Interval for Population Proportion p:

- If sample size n is large, i.e. $n\hat{p} > 5$ and $n\hat{q} = n(1 - \hat{p}) > 5$,
- the $100(1 - \alpha)\%$ confidence interval for the population proportion p is

$$\overline{X} \pm E, \quad E = z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}.$$

- E is called the margin of error, and $z_{\alpha/2}$ satisfies

Confidence Interval for Population Proportion p:

- If sample size n is large, i.e. $n\hat{p} > 5$ and $n\hat{q} = n(1-\hat{p}) > 5$,
- the $100(1-\alpha)\%$ confidence interval for the population proportion p is

$$\overline{X} \pm E, \quad E = z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}.$$

- E is called the margin of error, and $z_{\alpha/2}$ satisfies

Confidence Interval for Population Proportion p:

- If sample size n is large, i.e. $n\hat{p} > 5$ and $n\hat{q} = n(1 - \hat{p}) > 5$,
- the $100(1 - \alpha)\%$ confidence interval for the population proportion p is

$$\overline{X} \pm E, \quad E = z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}.$$

- E is called the margin of error, and $z_{\alpha/2}$ satisfies

Confidence Interval for Population Proportion p:

- If sample size n is large, i.e. $n\hat{p} > 5$ and $n\hat{q} = n(1 - \hat{p}) > 5$,
- the $100(1 - \alpha)\%$ confidence interval for the population proportion p is

$$\overline{X} \pm E, \quad E = z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}.$$

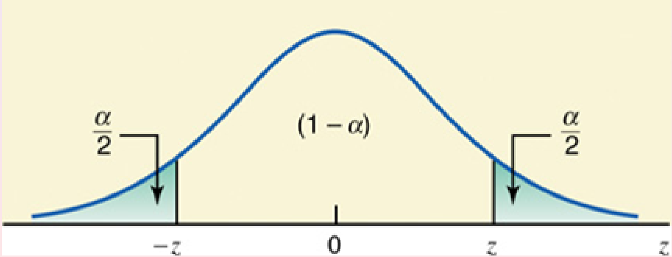
- E is called the margin of error, and $z_{\alpha/2}$ satisfies

Confidence Interval for Population Proportion p:

- If sample size n is large, i.e. $n\hat{p} > 5$ and $n\hat{q} = n(1 - \hat{p}) > 5$,
- the $100(1 - \alpha)\%$ confidence interval for the population proportion p is

$$\overline{X} \pm E, \quad E = z_{\alpha/2}\sqrt{\hat{p}(1 - \hat{p})/n}.$$

- E is called the margin of error, and $z_{\alpha/2}$ satisfies



A Little Better Confidence Interval for p:

- Since $-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}$ is equivalent to

$$\frac{|\overline{X}-p|}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2} \quad \text{or} \quad \frac{(\overline{X}-p)^2}{p(1-p)/n} \leqslant z^2_{\alpha/2}$$

- That is $H(p) = (\overline{X} - p)^2 - z^2_{\alpha/2} p(1-p)/n \leqslant 0$
- Let $a = 1 + \frac{z^2_{\alpha/2}}{n}$, $b = -2\overline{X} - \frac{z^2_{\alpha/2}}{n}$, $c = \overline{X}^2$. $H(p) = ap^2 + bp + c$.
- The solutions of $H(p) = 0$ are

$$\hat{p}_L = \frac{-b-\sqrt{b^2-4ac}}{2a}, \quad \hat{p}_U = \frac{-b+\sqrt{b^2-4ac}}{2a}.$$

So an approximate $100(1-\alpha)\%$ CI for p is $[\hat{p}_L, \quad \hat{p}_U]$.

A Little Better Confidence Interval for p:

- Since $-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}$ is equivalent to

$$\frac{|\overline{X}-p|}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2} \quad \text{or} \quad \frac{(\overline{X}-p)^2}{p(1-p)/n} \leqslant z^2_{\alpha/2}$$

- That is $H(p) = (\overline{X} - p)^2 - z^2_{\alpha/2} p(1-p)/n \leqslant 0$

- Let $a = 1 + \frac{z^2_{\alpha/2}}{n}$, $b = -2\overline{X} - \frac{z^2_{\alpha/2}}{n}$, $c = \overline{X}^2$. $H(p) = ap^2 + bp + c$.
- The solutions of $H(p) = 0$ are

$$\hat{p}_L = \frac{-b-\sqrt{b^2-4ac}}{2a}, \quad \hat{p}_U = \frac{-b+\sqrt{b^2-4ac}}{2a}.$$

So an approximate $100(1-\alpha)\%$ CI for p is $[\hat{p}_L, \quad \hat{p}_U]$.

A Little Better Confidence Interval for p:

- Since $-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}$ is equivalent to

$$\frac{|\overline{X}-p|}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2} \quad \text{or} \quad \frac{(\overline{X}-p)^2}{p(1-p)/n} \leqslant z^2_{\alpha/2}$$

- That is $H(p) = (\overline{X} - p)^2 - z^2_{\alpha/2} p(1-p)/n \leqslant 0$
- Let $a = 1 + \frac{z^2_{\alpha/2}}{n}$, $b = -2\overline{X} - \frac{z^2_{\alpha/2}}{n}$, $c = \overline{X}^2$.
$H(p) = ap^2 + bp + c$.
- The solutions of $H(p) = 0$ are

$$\hat{p}_L = \frac{-b-\sqrt{b^2-4ac}}{2a}, \quad \hat{p}_U = \frac{-b+\sqrt{b^2-4ac}}{2a}.$$

So an approximate $100(1-\alpha)\%$ CI for p is $[\hat{p}_L, \quad \hat{p}_U]$.

A Little Better Confidence Interval for p:

- Since $-z_{\alpha/2} \leqslant \frac{\overline{X}-p}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2}$ is equivalent to

$$\frac{|\overline{X}-p|}{\sqrt{p(1-p)/n}} \leqslant z_{\alpha/2} \quad \text{or} \quad \frac{(\overline{X}-p)^2}{p(1-p)/n} \leqslant z^2_{\alpha/2}$$

- That is $H(p) = (\overline{X} - p)^2 - z^2_{\alpha/2}p(1-p)/n \leqslant 0$
- Let $a = 1 + \frac{z^2_{\alpha/2}}{n}$, $b = -2\overline{X} - \frac{z^2_{\alpha/2}}{n}$, $c = \overline{X}^2$.
  $H(p) = ap^2 + bp + c$.
- The solutions of $H(p) = 0$ are

$$\hat{p}_L = \frac{-b-\sqrt{b^2-4ac}}{2a}, \quad \hat{p}_U = \frac{-b+\sqrt{b^2-4ac}}{2a}.$$

So an approximate $100(1-\alpha)\%$ CI for p is $[\hat{p}_L, \quad \hat{p}_U]$.

Example 1: Poll Results in The News

Example 1: USA Today/Gallup Poll:
http://www.usatoday.com/
(a) For a poll result $\hat{p}$ of interest, find the sample size n, the margin of error, and calculate the 95% confidence interval for p.
(b) Find the margin of error for 99% confidence level and confidence interval.

Example 1: Poll Results in The News

Example 1: USA Today/Gallup Poll:
http://www.usatoday.com/
(a) For a poll result $\hat{p}$ of interest, find the sample size n, the margin of error, and calculate the 95% confidence interval for p.
(b) Find the margin of error for 99% confidence level and confidence interval.

Example 2:

A machine shop manufactures toggle levers. A lever is flawed if a standard nut cannot be screwed onto the threads. Let p equal the proportion of flawed toggle levers that the shop manufactures. If there were 24 flawed levers out of a sample of 642 that were selected randomly from the production line.

(a) Give a point estimate of p.

(b) Find a 95% confidence interval for p.

Solution of Example 2

Solution

(a) The point estimate is $\hat{p} = 24/642 = 0.037$.

(b) $\alpha = 0.05$, $z = 1.96$, $s_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})/n} = 0.0075$.

$E = 0.0147$ and 95% CI for p is $[0.0223, 0.0517]$.

If the machine shop wants to be 99% confident with a margin of error as small as 0.001, how large a sample should be selected?

What is the most conservative sample size?

Solution of Example 2

Solution

(a) The point estimate is $\hat{p} = 24/642 = 0.037$.
(b) $\alpha = 0.05$, $z = 1.96$, $s_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})/n} = 0.0075$.
$E = 0.0147$ and 95% CI for p is $[0.0223, 0.0517]$.
If the machine shop wants to be 99% confident with a margin of error as small as 0.001, how large a sample should be selected?
What is the most conservative sample size?

Outline

Determine Sample Size by Margin of Error

- If $\hat{p}$ is available,

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$E^2 = z^2_{\frac{\alpha}{2}} \frac{\hat{p}(1-\hat{p})}{n}$$

$$n = z^2_{\frac{\alpha}{2}} \frac{\hat{p}(1-\hat{p})}{E^2} = \left(\frac{z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})}}{E} \right)^2$$

- If $\hat{p}$ is not available, use the *most conservative* estimate

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E} \right)^2$$

Determine Sample Size by Margin of Error

- If $\hat{p}$ is available,

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$E^2 = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{n}$$

$$n = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{E^2} = \left(\frac{z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})}}{E} \right)^2$$

- If $\hat{p}$ is not available, use the *most conservative* estimate

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E} \right)^2$$

Determine Sample Size by Margin of Error

- If $\hat{p}$ is available,

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$E^2 = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{n}$$

$$n = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{E^2} = \left(\frac{z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})}}{E} \right)^2$$

- If $\hat{p}$ is not available, use the *most conservative* estimate

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E} \right)^2$$

Determine Sample Size by Margin of Error

- If $\hat{p}$ is available,

$$E = z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$E^2 = z^2_{\frac{\alpha}{2}}\frac{\hat{p}(1-\hat{p})}{n}$$

$$n = z^2_{\frac{\alpha}{2}}\frac{\hat{p}(1-\hat{p})}{E^2} = \left(\frac{z_{\frac{\alpha}{2}}\sqrt{\hat{p}(1-\hat{p})}}{E}\right)^2$$

- If $\hat{p}$ is not available, use the *most conservative* estimate

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E}\right)^2$$

Determine Sample Size by Margin of Error

- If $\hat{p}$ is available,

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$E^2 = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{n}$$

$$n = z_{\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{E^2} = \left(\frac{z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})}}{E} \right)^2$$

- If $\hat{p}$ is not available, use the *most conservative* estimate

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E} \right)^2$$

Example 3:

A machine shop manufactures toggle levers. A lever is flawed if a standard nut cannot be screwed onto the threads. Let p equal the proportion of flawed toggle levers that the shop manufactures. If there were 24 flawed levers out of a sample of 642 that were selected randomly from the production line.

In §6.7 Example 2, have found a point estimate of p: $\hat{p} = 0.037$ and the margin of error $E = 0.0147$ with 95% confidence.

(a) If the machine shop wants to be 99% confident with a margin of error as small as 0.001, how large a sample should be selected?

(b) What is the most conservative sample size?

Example 3:

A machine shop manufactures toggle levers. A lever is flawed if a standard nut cannot be screwed onto the threads. Let p equal the proportion of flawed toggle levers that the shop manufactures. If there were 24 flawed levers out of a sample of 642 that were selected randomly from the production line.
In §6.7 Example 2, have found a point estimate of p: $\hat{p} = 0.037$ and the margin of error $E = 0.0147$ with 95% confidence.
(a) If the machine shop wants to be 99% confident with a margin of error as small as 0.001, how large a sample should be selected?
(b) What is the most conservative sample size?

Example 3:

A machine shop manufactures toggle levers. A lever is flawed if a standard nut cannot be screwed onto the threads. Let p equal the proportion of flawed toggle levers that the shop manufactures. If there were 24 flawed levers out of a sample of 642 that were selected randomly from the production line.
In §6.7 Example 2, have found a point estimate of p: $\hat{p} = 0.037$ and the margin of error $E = 0.0147$ with 95% confidence.
(a) If the machine shop wants to be 99% confident with a margin of error as small as 0.001, how large a sample should be selected?
(b) What is the most conservative sample size?

Solution of Example 3

Solution

(a) $\alpha = 0.01$, $z_{\frac{\alpha}{2}} = 2.575$

$$n = \left(\frac{z}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{2.575}{0.001}\right)^2 (0.037)(1-0.037) = 236,256$$

(b) The most conservative sample size is

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{2E}\right)^2 = \left(\frac{2.575}{2(0.001)}\right)^2 = 1,657,657$$

Example 4:

- Consider two groups of women:

  Group 1: women who spend less than \$500 annually on clothes.

  Group 2: women who spend over \$1000 annually on clothes.

- Let p_1 and p_2 be the proportions of women in the two groups who believe that the clothes are too expensive, respectively.
- If $Y_1 = 1009$ out of $n_1 = 1230$ women from Group 1 and $Y_2 = 207$ out of $n_2 = 340$ from Group 2 believe that the clothes are too expensive.

(a) Find a point estimate of $p_1 - p_2$.

(b) Find an approximate 95% CI for $p_1 - p_2$.

Example 4:

- Consider two groups of women:

  Group 1: women who spend less than \$500 annually on clothes.

  Group 2: women who spend over \$1000 annually on clothes.

- Let p_1 and p_2 be the proportions of women in the two groups who believe that the clothes are too expensive, respectively.
- If $Y_1 = 1009$ out of $n_1 = 1230$ women from Group 1 and $Y_2 = 207$ out of $n_2 = 340$ from Group 2 believe that the clothes are too expensive.

(a) Find a point estimate of $p_1 - p_2$.

(b) Find an approximate 95% CI for $p_1 - p_2$.

Example 4:

- Consider two groups of women:
  - Group 1: women who spend less than $500 annually on clothes.
  - Group 2: women who spend over $1000 annually on clothes.
- Let p_1 and p_2 be the proportions of women in the two groups who believe that the clothes are too expensive, respectively.
- If $Y_1 = 1009$ out of $n_1 = 1230$ women from Group 1 and $Y_2 = 207$ out of $n_2 = 340$ from Group 2 believe that the clothes are too expensive.

(a) Find a point estimate of $p_1 - p_2$.

(b) Find an approximate 95% CI for $p_1 - p_2$.

Example 4:

- Consider two groups of women:
  - Group 1: women who spend less than $500 annually on clothes.
  - Group 2: women who spend over $1000 annually on clothes.
- Let p_1 and p_2 be the proportions of women in the two groups who believe that the clothes are too expensive, respectively.
- If $Y_1 = 1009$ out of $n_1 = 1230$ women from Group 1 and $Y_2 = 207$ out of $n_2 = 340$ from Group 2 believe that the clothes are too expensive.

(a) Find a point estimate of $p_1 - p_2$.

(b) Find an approximate 95% CI for $p_1 - p_2$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$,by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$,by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$, by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$,by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$,by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.

Solution of Example 4:

(a) Calculate the sample proportions

$$\hat{p}_1 = \frac{Y_1}{n_1} = 0.82, \quad \hat{p}_2 = \frac{Y_2}{n_2} = 0.61$$

The point estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = 0.21$.

(b) Since Y_1 and Y_2 are independent r.v.'s. $Y_1 \sim b(n_1, p_1)$ and $Y_2 \sim b(n_2, p_2)$,by the C.L.T.,

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \sim N(0,1)$$

From this we an approximate $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 \pm E$, where

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.056.$$

Thus $\hat{p}_1 - \hat{p}_2 \pm E = [0.154, 0.266]$.