# Solutions to Assignment of Chapter 4

D. Ding

March 2022

**4.1** If $x_1$ and $x_2$ are the values of a random sample of size $n = 2$ from a population having a uniform distribution with parameters 0 and $\theta > 0$, find the number $k$ so that the interval $\big[0,\ k(x_1+x_2)\big]$ is a $(1-\alpha)100\%$ confidence interval for the parameter $\theta$ when (1) $0 < \alpha \leq \frac{1}{2}$; and (2) $\frac{1}{2} < \alpha < 1$.

**Solution**. Since $\big[0,\ k(x_1+x_2)\big]$ is a confidence interval for $\theta$, we have

$$\mathbb{P}\big(0 < \theta < k(X_1 + X_2)\big) = 1 - \alpha,$$

or,

$$\mathbb{P}\Big((X_1 + X_2) > \frac{\theta}{k}\Big) = 1 - \alpha.$$

Since the joint density of $X_1$ and $X_2$ is given by

$$f(x_1, x_2; \theta) = \begin{cases} 1/\theta^2, & 0 \leq x_1 \leq \theta \text{ and } 0 \leq x_2 \leq \theta, \\ 0, & \text{elsewhere,} \end{cases}$$

we see that, if $0 < \alpha \leq 0.5$, then $k \geq 1$; and if $0.5 < \alpha < 1$, then $0.5 < k < 1$.

(1) If $0 < \alpha \leq 0.5$, then we have

$$\alpha = \mathbb{P}\Big((X_1 + X_2) \leq \frac{\theta}{k}\Big) = \int_0^\theta \int_0^{(\theta/k)-x_2} \frac{1}{\theta^2} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{1}{2k^2}.$$

Thus, we get $k = 1/\sqrt{2\alpha}$;

(2) If $0.5 < \alpha < 1$, then we have

$$1 - \alpha = \mathbb{P}\Big((X_1 + X_2) > \frac{\theta}{k}\Big) = \int_{\theta(1-k)/k}^\theta \int_{(\theta/k)-x_2}^\theta \frac{1}{\theta^2} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{(1-2k)^2}{2k^2}.$$

Thus, we get

$$2k^2(1 - \alpha) = (1 - 2k)^2.$$

Solving this equation of $k$ we obtain

$$k = \frac{1}{2 - \sqrt{2(1-\alpha)}}$$

since $0.5 < k < 1$. $\square$.

**4.2** A district official intends to use the mean of a random sample of 150 sixth graders to estimate the mean score of all the sixth graders in the district if they took a certain arithmetic test. Suppose that the official gets the sample mean 61.8 with a standard deviation of 9.4.

(1) What can the official assert with probability 0.95 about the maximum error?

(2) Use all the given information to construct a 99% confidence interval for the mean score of all the sixth graders in the district.

**Solution**. (1) By the statistical table of normal distribution, $z_{\alpha/2} = z_{0.025} = 1.96$. Since $n = 150 > 30$, according to the Central limit theorem we have that, approximately,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Thus, as Theorem 11.1, we have

$$\mathbb{P}\Big(\big|\bar{X} - \mu\big| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\Big) = 1 - \alpha.$$

Substituting $s = 9.6$ for $\sigma$, $z_{\alpha/2} = z_{0.025} = 1.96$ and $n = 150$ into the above formula, we get that, with probability 0.95,

$$\text{error} = \big|\bar{X} - \mu\big| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1.50.$$

Thus, the official can assert the maximum error of their estimate will be less than 1.50 with probability 0.95.

(2) By the statistical table of normal distribution, $z_{\alpha/2} = z_{0.005} = 2.575$. Substituting the given data into the large-sample confidence interval formula (II):

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}},$$

we get

$$59.82 \leq \mu \leq 63.78.$$

Hence [59.82, 63.78] is a 99% confidence interval for the mean score of all the sixth graders in the district. □

**4.3** A researcher intends to use the mean of a random sample of size $n = 120$ to estimate the mean blood pressure of women in their fifties. Suppose that the researcher gets the sample mean of 141.8 mm of mercury with a standard deviation of 10.5 mm of mercury. Construct a 99% confidence interval for the population mean.

**Solution**. By the Statistical table III, we get $z_{\alpha/2} = z_{0.005} = 2.575$. Substituting the data: $n = 120$, $\bar{x} = 141.8$ and $s = 10.5$ into the large-sample confidence interval formula (II):

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}},$$

we get

$$139.3318 \leq \mu \leq 144.2682.$$

Hence, [139.33, 144.27] is a 99% confidence interval for the mean blood pressure $\mu$. □

**4.4** A study of the annual growth of certain cactuses showed that 64 of them, selected at random in a desert region, grew on the average 52.80 mm with a standard deviation of 4.5 mm. Construct a 99% confidence interval for the true average annual growth of the given kind of cactus.

**Solution**. Since $n = 64 > 30$, we can use the large-sample confidence interval formula (II). From the statistical table of the normal distribution, we get $z_{\alpha/2} = z_{0.005} = 2.575$. Substituting the data: $n = 64$, $\bar{x} = 52.80$, and $s = 4.5$ into the large-sample confidence interval formula (II):

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}},$$

we get

$$51.35 \le \mu \le 54.25.$$

So, the 99% confidence interval for the true average annual growth of the given kind of cactus is $[51.35,\ 54.25]$. □

**4.5** The length of the skulls of 10 fossil skeletons of an extinct species of birds has a mean of 5.68 cm and a standard deviation of 0.29 cm. Assume that such measurements are normally distributed.

(1) Find a 95% confidence interval for the mean length of skulls of this species of birds.

(2) Construct a 95% confidence interval for the true variance of skulls of this species of birds.

**Solution**. (1) From the statistical table of t-distribution (Table IV), we have $t_{\alpha/2,n-1} = t_{0.025,9} = 2.262$. Substituting the data: $n = 10$, $\bar{x} = 5.65$, and $s = 0.29$ into the small-sample confidence interval formula (III):

$$\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}},$$

we obtain

$$5.47 \le \mu \le 5.89.$$

So we can assert with 95% confidence interval for the mean length is $[5.47,\ 5.89]$.

(2) From the statistical table of chi-square distribution, we have

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,9} = 19.023 \quad \text{and} \quad \chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,9} = 2.700.$$

Substituting the date: $n = 10$ and $s = 0.29$ into the formula (VIII):

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}},$$

we obtain

$$0.04 \le \sigma^2 \le 0.28.$$

Thus, the 95% confidence interval for the true variance is $[0.04,\ 0.28]$. □

**4.6** Assume that $\bar{X}$ and $S$ are the mean and the standard deviation of a random sample of size $n$ from a normal population. Show that, if $\bar{X}$ is to be used an estimator of the mean of the population, then the probability is $1-\alpha$ that the error will be less than $t_{\alpha/2,n-1}S/\sqrt{n}$.

**Proof**. Since the population has a normal distribution $N(\mu, \sigma^2)$, by Proposition 2.4.3 (Theorem 8.13), we have

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$

Thus, we get

$$\mathbb{P}\Big(\big|\bar{X} - \mu\big| \le t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\Big) = \mathbb{P}\big(\big|T\big| \le t_{\alpha/2,n-1}\big) = 1 - \alpha.$$

Hence, if $\bar{X}$ is to be used to estimate $\mu$, then the probability is $1 - \alpha$ that the error will be less than $t_{\alpha/2,n-1}S/\sqrt{n}$. □

**4.7** Prove Proposition 4.1.5 (Theorem 11.3).

**Proof**. According to Problem 4.6, we have

$$\mathbb{P}\Big(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \le \mu \le \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\Big) = \mathbb{P}\big(|T| \le t_{\alpha/2,n-1}\big) = 1-\alpha.$$

Thus, we have that

$$\Big[\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\Big]$$

is a $(1-\alpha)100\%$ confidence interval of the mean $\mu$. The theorem is proved. □

**4.8** A food inspector, examining 12 jars of a certain brand of peanut butter, obtained the following percentages of impurities:

2.3, 1.9, 2.1, 2.8, 2.3, 3.6, 1.4, 1.8, 2.1, 3.2, 2.0, 1.9.

Assume that such measurements are normally distributed.

(1) What can the inspector assert with 95% confidence about the maximum error, if she uses the sample mean $\bar{x}$ of the above sample as an estimate of the mean of population sampled, i.e. the true average percentage of impurities in this brand of peanut butter?

(2) Construct a 90% confidence interval for the variance of the population sampled, i.e. the percentage of impurities in this brand of peanut butter.

**Solution**. From By using Excel

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Average | Variance |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.3 | 1.9 | 2.1 | 2.8 | 2.3 | 3.6 | 1.4 | 1.8 | 2.1 | 3.2 | 2.0 | 1.9 | 2.2833 | 0.3906 |

we get that $\bar{x} = 2.833$ and $s = \sqrt{0.3906} = 0.6250$.

(1) By the statistical table of $t$-distribution, we have that $t_{\alpha/2,n-1} = t_{0.025,11} = 2.201$. By Problem 4.6, with probability 0.95, we have

$$\text{error} = \big|\bar{x} - \mu\big| \le t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = 0.397.$$

Thus, the inspector can assert the maximum error of her estimate will be less than 0.397 with probability 0.95.

(2) From the statistical table of chi-square distribution,

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.05,11} = 19.675 \quad \text{and} \quad \chi^2_{1-\alpha/2,n-1} = \chi^2_{0.95,11} = 4.575.$$

Substituting the date: $n = 12$ and $s = 0.625$ into the formula (VIII),

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}},$$

we obtain the 90% confidence interval for $\sigma^2$ by

$$0.2184 \le \sigma^2 \le 0.939.$$

Thus, the 90% confidence interval for the true variance the variance of the population sampled is $[0.2184, 0.939]$. □

**4.9** Two independent random samples of size $n_1 = 16$ and $n_2 = 25$ from normal populations with $\sigma_1 = 4.84$ and $\sigma_2 = 3.5$ have the means $\bar{X}_1 = 18.2$ and $\bar{x}_2 = 23.4$. Find a 90% confidence interval for $\mu_1 - \mu_2$, the difference between two means of two populations.

**Solution**. Substituting the data: $n_1 = 16$ and $n_2 = 25$, $\bar{X}_1 = 18.2$ and $\bar{x}_2 = 23.4$, and $\sigma_1 = 4.84$ and $\sigma_2 = 3.5$ into the formula (IV):

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},$$

we get that a 90% confidence interval for the difference between the means is given by

$$-7.485 \le \mu_1 - \mu_2 \le -2.915.$$

□

**4.10** Show Proposition 4.2.1 (Theorem 11.4).

**Proof**. According to Problem 2.3, we have that $\bar{X}_1 - \bar{X}_2$ is a random variable having a normal distribution with the mean $\mu_1 - \mu_2$ and the variance $(\sigma_1^2/n_1) + (\sigma_2^2/n_2)$. Thus, we get

$$\begin{aligned} &\mathbb{P}\Big((\bar{X}_1 - \bar{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\Big) \\ = \;& \mathbb{P}\big(|Z| \le z_{\alpha/2}\big) \;=\; 1 - \alpha, \end{aligned}$$

where

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}.$$

Thus, we have that

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

is a $(1 - \alpha)100\%$ confidence interval of $\mu_1 - \mu_2$. The theorem is proved. □

**4.11** (Exercise 11.9) Show that the pooled estimator $S_p^2$ is an unbiased estimator of $\sigma^2$, and find its variance under the conditions of Proposition 4.2.2 (Theorem 5).

**Proof**. Note that the pooled estimator $S_p^2$ is given

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2},$$

where $S_1^2$ and $S_2^2$ be the sample variances of two independent random samples of size $n_1$ and $n_2$ from two normal populations with same variance $\sigma^2$. According to Proposition 2.3.7 (Theorem 8.11) we have

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1) \quad \text{and} \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1),$$

and since two random samples are independent, $S_1^2$ and $S_2^2$ are independent. Now, according to Proposition 2.3.1, we get that

$$\mathbb{E}\Big[\frac{(n_k - 1)S_k^2}{\sigma^2}\Big] = n_k - 1 \quad \text{and} \quad \mathrm{var}\Big(\frac{(n_k - 1)S_k^2}{\sigma^2}\Big) = 2(n_k - 1).$$

Thus, we have

$$\mathbb{E}\big[S_p^2\big] = \mathbb{E}\Big[\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\Big] = \frac{(n_1-1)\sigma^2}{n_1+n_2-2} + \frac{(n_2-1)\sigma^2}{n_1+n_2-2} = \sigma^2.$$

This implies that the pooled estimator $S_p^2$ is an unbiased estimator of $\sigma^2$. Moreover, we have

$$\begin{aligned}
\operatorname{var}\big(S_p^2\big) &= \operatorname{var}\Big(\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\Big) = \frac{\sigma^4}{(n_1+n_2-2)^2}\operatorname{var}\Big(\frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2}\Big) \\
&= \frac{\sigma^4}{(n_1+n_2-2)^2}\Big[\operatorname{var}\Big(\frac{(n_1-1)S_1^2}{\sigma^2}\Big) + \operatorname{var}\Big(\frac{(n_2-1)S_2^2}{\sigma^2}\Big)\Big] \\
&= \frac{\sigma^4}{(n_1+n_2-2)^2}\Big[2(n_1-1) + 2(n_2-1)\Big] = \frac{2\sigma^4}{n_1+n_2-2}. \quad \square
\end{aligned}$$

**4.12** The following are the heat-producing capacities of coal from two mines (in millions of calories per ton):

Mine A : 8500, 8330, 8480, 7960, 8030.
Mine B : 7710, 7890, 7920, 8270, 7860.

Assume that the data constitute two independent random samples from normal populations with equal variances.

(1) Construct a 99% confidence interval for $\mu_1 - \mu_2$, the difference between the true average heat-producing capacities of coal from the mines.

(2) Construct a 90% confidence interval for $\sigma_1^2/\sigma_2^2$, the ratio of the variances of two populations sampled, i.e. the heat-producing capacities of coal from the mines.

**Solution**. By using Excel, we get

| | 1 | 2 | 3 | 4 | 5 | Average | Variance |
|---|---|---|---|---|---|---|---|
| Mine A | 8500 | 8330 | 8480 | 7960 | 8030 | 8260 | 63450 |
| Mine B | 7710 | 7890 | 7920 | 8270 | 7860 | 7930 | 42650 |

(1) Substituting the data: $n_1 = 5$ and $n_2 = 5$, $s_1 = \sqrt{63450} = 251.89$ and $s_2 = \sqrt{42650} = 206.52$ into the formula (VII):

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

we get that $s_p = 230.32$. From Table IV we have

$$t_{\alpha/2,n_1+n_2-2} = t_{0.005,8} = 3.355.$$

Substituting the data: $n_1 = 5$ and $n_2 = 5$, $\bar{x}_1 = 8260$ and $\bar{x}_2 = 7930$, and $s_p = 230.32$ into the formula (VI)

$$(\bar{x}_1 - \bar{x}_2) - t_0 s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_0 s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

we get the required 99% confidence interval by

$$-158.75 \le \mu_1 - \mu_2 \le 818.75.$$

(2) From the statistical table of $F$-distribution, we have

$$f_{\alpha/2,n_1-1,n_2-1} = f_{0.05,4,4} = 6.39.$$

Substituting the data: $n_1 = 5$ and $n_2 = 5$, $s_1^2 = 63450$ and $s_2^2 = 42650$ into the formula (IX),

$$\frac{s_1^2}{s_2^2} \cdot \frac{1}{f_{\alpha/2,n_1-1,n_2-1}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} \cdot f_{\alpha/2,n_2-1,n_1-1}$$

we get the required 90% confidence interval by

$$0.233 \le \frac{\sigma_1^2}{\sigma_2^2} \le 9.506. \qquad \square$$

**4.13** A study of two kinds of photocopying equipment shows that 61 failures of the first kind of equipment took on average 80.7 minutes to repair with a standard deviation of 19.4 minutes, whereas 61 failures of the second kind of equipment took on average 88.1 minutes to repair with a standard deviation of 18.8 minutes.

(1) Find a 99% confidence interval for $\mu_1 - \mu_2$, the difference between the true average amounts of time it takes to repair failures of the two kinds of equipment.

(2) Construct a 98% confidence interval for $\sigma_1^2/\sigma_2^2$, the ratio of the variances of two populations sampled.

**Solution**. (1) Substituting the data: $n_1 = 61$ and $n_2 = 61$, $\bar{x}_1 = 80.7$ and $\bar{x}_2 = 88.14$, $s_1 = 19.4$ and $s_2 = 18.8$, and $z_{\alpha/2} = z_{0.005} = 2.575$ into the large-sample confidence interval formula (V):

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

we get that a 99% confidence interval for the difference between the means is given by

$$-16.31 \le \mu_1 - \mu_2 \le 1.51.$$

(2) From the statistical table of $F$-distribution we have that $f_{\alpha/2,n_1-1,n_2-1} = f_{0.01,60,60} = 1.84$. Substituting the data into the formula (IX),

$$\frac{s_1^2}{s_2^2} \cdot \frac{1}{f_{\alpha/2,n_1-1,n_2-1}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} \cdot f_{\alpha/2,n_2-1,n_1-1}$$

we get

$$0.58 \le \frac{\sigma_1^2}{\sigma_2^2} \le 1.96. \qquad \square$$

**4.14** A sample survey at a supermarket showed that 204 of 300 shoppers regularly use cents-off coupons. Construct a 95% confidence interval for the corresponding true proportion.

**Solution**. Substituting $z_{\alpha/2} = z_{0.025} = 1.96$, $n = 300$, and $\tilde{\theta} = 204/300 = 0.68$ into the confidence interval formula (X):

$$\tilde{\theta} - z_{\alpha/2}\sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n}} \le \theta \le \tilde{\theta} + z_{\alpha/2}\sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n}}$$

we get

$$0.627 \le \theta \le 0.733.$$

Thus, the 95% confidence interval for the true proportion is $[63\%, 73\%]$ if it is rounded to two decimals. $\square$

**4.15** Among 100 fish caught in a lake, 18 fish were inedible as a result of the chemical pollution of the environment. Construct a 99% confidence interval for the corresponding true proportion.

**Solution**. Substituting $z_{\alpha/2} = z_{0.005} = 2.575$, $n = 100$, and $\tilde{\theta} = 18/100 = 0.18$ into the confidence interval formula (X):

$$\tilde{\theta} - z_{\alpha/2}\sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n}} \leq \theta \leq \tilde{\theta} + z_{\alpha/2}\sqrt{\frac{\tilde{\theta}(1-\tilde{\theta})}{n}}$$

we get

$$0.081 \leq \theta \leq 0.279.$$

Thus, the 99% confidence interval for the true proportion is $[8\%, 28\%]$ if it is rounded to two decimals. □

**4.16** Among 500 marriage license applications chosen at random in a given year, there were 48 in which the woman was at least one year older than the man, and among 400 marriage license applications chosen at random six years later, there were 68 in which the woman was at least one year older than the man. Construct a 99% confidence interval for the difference between the corresponding true proportions of marriage license applications in which the woman was at least one year older than the man.

**Solution**. Substituting the data: $z_{\alpha/2} = z_{0.005} = 2.575$, $n_1 = 500$ and $n_2 = 400$, and $\tilde{\theta}_1 == 48/500 = 0.096$ and $\tilde{\theta}_2 = 68/400 = 0.170$ into the formula (XI)

$$(\tilde{\theta}_1 - \tilde{\theta}_2) - z_{\alpha/2}\Theta_{1,2} \leq \theta_1 - \theta_2 \leq (\tilde{\theta}_1 - \tilde{\theta}_2) + z_{\alpha/2}\Theta_{1,2},$$

where $\Theta_{1.2}$ is given by the formula (XII):

$$\Theta_{1,2} = \sqrt{\frac{\tilde{\theta}_1(1-\tilde{\theta}_1)}{n_1} + \frac{\tilde{\theta}_2(1-\tilde{\theta}_2)}{n_2}},$$

we get

$$\Theta_{1,2} = 0.0229,$$

and so

$$-0.133 \leq \theta_1 - \theta_2 \leq -0.015.$$

Thus, the 99% confidence interval for the difference between the corresponding true proportions is $[-13.3\%, \ -.5\%]$. Since $\theta_1 - \theta_2 \leq -0.015 < 0$, we conclude that there is a real difference between the actual proportions. □