Chapter 7: Statistical Intervals for a Single Sample

Course Name: PROBABILITY & STATISTICS

Lecturer: Duong Thi Hong

Hanoi, 2022

Content

Introduction

- In the previous chapter we illustrated how a parameter can be estimated from sample data. However, it is important to understand how good is the estimate obtained.
- Bounds that represent an interval of plausible values for a parameter are an example of an interval estimate.

Content

1.1 Development of the Confidence Interval and Its Basic Properties

Suppose that $X_1, X_2, ..., X_n$ is a random sample from a normal distribution with unknown mean μ and known variance σ^2. Then, the sample mean $\bar{X}$ is normally distributed with mean μ and variance σ^2/n. Therefore,

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

is a standard normal distribution.

Confidence Interval

A **confidence interval** (CI) estimate for μ is an interval of the form $l \leq \mu \leq u$, where the endpoints l and u are computed from the sample data. Because different samples will produce different values of l and u, these end-points are values of random variables L and U, respectively. Suppose that we can determine values of L and U such that the following probability statement is true:

$$P(L \leq \mu \leq U) = 1 - \alpha, \quad \text{for } 0 \leq \alpha \leq 1$$

1.1 Development of the Confidence Interval and Its Basic Properties

- The endpoints or bounds l and u are called lower- and upper-confidence limits, respectively.
- $1 - \alpha$ is called the **confidence coefficient**.
- Since $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ has a standard normal distribution, we may write

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

This results in

$$P(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

1.1 Development of the Confidence Interval and Its Basic Properties

Confidence Interval on the Mean, Variance Known

If $\bar{x}$ is the sample mean of a random sample of size n from a normal population with known variance σ^2, a $100(1-\alpha)\%$ CI on μ is given by

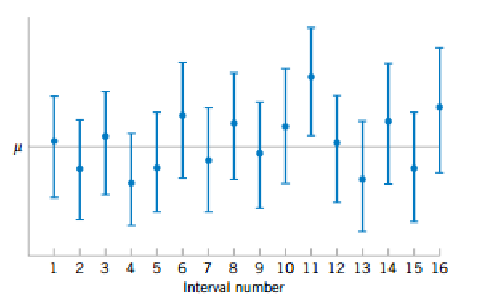
$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

where $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution.

Note: $\Phi(z_\alpha) = P(Z \leq z_\alpha) = 1 - \alpha$

Confidence Interval on the Mean, Variance Known

The situation is illustrated in following figure, which shows several $100(1-\alpha)\%$ confidence intervals for the mean μ of a normal distribution. The dots at the center of the intervals indicate the point estimate of μ (that is, $\bar{x}$). Notice that one of the intervals fails to contain the true value of μ. If this were a 95% confidence interval, in the long run only 5% of the intervals would fail to contain μ.



Confidence Interval on the Mean, Variance Known

Example 1

In a sample of 36 randomly selected women, it was found that their mean height was 65.3 inches. From previous studies, it is assumed that the standard deviation $\sigma = 2.5$. Construct the 90% confidence interval for the population mean. Let $z_{0.05} = 1.64$.

Confidence Interval on the Mean, Variance Known

Example 1

In a sample of 36 randomly selected women, it was found that their mean height was 65.3 inches. From previous studies, it is assumed that the standard deviation $\sigma = 2.5$. Construct the 90% confidence interval for the population mean. Let $z_{0.05} = 1.64$.

Answer: We known $\sigma = 2.5$, $\bar{x} = 65.3$, $n = 36$, $\alpha = 0.1$. Therefore,

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = z_{0.05}\frac{2.5}{\sqrt{36}} = 0.683$$

Hence, the 90% confidence interval for the population mean is

$$(65.3 - 0.683, 65.3 + 0.683) = (64.617, 65.983)$$

Confidence Interval on the Mean, Variance Known

Example 2

An economist is interested in studying the incomes of consumers in a particular region. The population standard deviation is known to be \$1000. A random sample of 64 individuals resulted in an average income of \$21000. What is the width of the 95% confidence interval?. Let $z_{0.025} = 1.96$.

Confidence Interval on the Mean, Variance Known

Example 2

An economist is interested in studying the incomes of consumers in a particular region. The population standard deviation is known to be \$1000. A random sample of 64 individuals resulted in an average income of \$21000. What is the width of the 95% confidence interval?. Let $z_{0.025} = 1.96$.

Answer: We known $\sigma = 1000$, $\bar{x} = 21000$, $n = 64$, $\alpha = 0.05$. The with of the 95% confidence interval on mean is

$$2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2 * z_{0.025}\frac{1000}{\sqrt{64}} = 490$$

Confidence Interval on the Mean, Variance Known

Note

The length of a confidence interval is a measure of the precision of estimation. From the preceeding discussion, we see that precision is inversely related to the confidence level. It is desirable to obtain a confidence interval that is short enough for decision-making purposes and that also has adequate confidence. One way to achieve this is by choosing the sample size n to be large enough to give a CI of specified length or precision with prescribed confidence.

1.2 Choice of Sample Size

Sample Size for Specified Error on the Mean, Variance Known

If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error $|\bar{x}-\mu|$ will not exceed a specified amount E when the sample size is

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

$2E = 2z_{\alpha/2}\sigma/\sqrt{n}$ is the length of the resulting confidence interval.

Sample Size for Specified Error on the Mean, Variance Known

Example 3

A nurse at a local hospital is interested in estimating the birth weight of infants. How large a sample must she select if she desires to be 98% confident that the true mean is within 4 ounces of the sample mean? The standard deviation of the birth weights is known to be 6 ounces. Let $z_{0.01} = 2.33$.

Sample Size for Specified Error on the Mean, Variance Known

Example 3

A nurse at a local hospital is interested in estimating the birth weight of infants. How large a sample must she select if she desires to be 98% confident that the true mean is within 4 ounces of the sample mean? The standard deviation of the birth weights is known to be 6 ounces. Let $z_{0.01} = 2.33$.

Answer: We known $\sigma = 6, E = 4, \alpha = 0.02$. Thus, the required sample size:

$$n = \left(\frac{z_{0.01}\sigma}{E}\right)^2 = \left(\frac{2.33 \times 6}{4}\right)^2 = 12.215 \approx [12.215] = 13$$

One-Sided Confidence Bounds on the Mean, Variance Known

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

$$\mu \leq u = \bar{x} + z_{\alpha}\sigma/\sqrt{n}$$

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - z_{\alpha}\sigma/\sqrt{n} = l \leq \mu$$

1.3 One-Sided Confidence Bounds

One-Sided Confidence Bounds on the Mean, Variance Known

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

$$\mu \leq u = \bar{x} + z_{\alpha}\sigma/\sqrt{n}$$

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - z_{\alpha}\sigma/\sqrt{n} = l \leq \mu$$

Example 4

The diameter of holes for a cable harness is known to have a normal distribution with $\sigma = 0.01$ inch. A random sample of size 15 yields an average diameter of 1.5 inch. Find a 98% upper-confidence bound for the population mean. Let $z_{0.02} = 2.05$.

1.3 One-Sided Confidence Bounds

One-Sided Confidence Bounds on the Mean, Variance Known

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

$$\mu \leq u = \bar{x} + z_\alpha \sigma/\sqrt{n}$$

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - z_\alpha \sigma/\sqrt{n} = l \leq \mu$$

Example 4

The diameter of holes for a cable harness is known to have a normal distribution with $\sigma = 0.01$ inch. A random sample of size 15 yields an average diameter of 1.5 inch. Find a 98% upper-confidence bound for the population mean. Let $z_{0.02} = 2.05$.

Answer: We known $\sigma = 0.01, \bar{x} = 1.5, \alpha = 0.02, n = 15$. Then a 98% upper-confidence bound for μ is: $\mu \leq \bar{x} + z_\alpha \sigma/\sqrt{n} = 1.5 + (2.05 \times 0.01)/\sqrt{15} = 1.495$

1.4 Large-Sample Confidence Interval for μ

Large-Sample Confidence Interval on the Mean

When n is large, the quantity

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has an approximate standard normal distribution. Consequently,

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

is a **large sample confidence interval** for μ, with confidence level of approximately $100(1-\alpha)\%$.

Large-Sample Confidence Interval on the Mean

Example 5

A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed. A random sample of 49 specimens has a mean compressive strength of 3250(psi) and a variance $25(psi)^2$. Construct approximately a 95% two-sided confidence interval on mean compressive strength? Let $z_{0.025} = 1.96$

Large-Sample Confidence Interval on the Mean

Example 5

A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed. A random sample of 49 specimens has a mean compressive strength of 3250(psi) and a variance $25(psi)^2$. Construct approximately a 95% two-sided confidence interval on mean compressive strength? Let $z_{0.025} = 1.96$

Answer: We known $n = 49, \bar{x} = 3250, s^2 = 25, \alpha = 0.05$. Hence, a 95% two-sided confidence interval on mean is

$$3250 - 1.96\frac{5}{\sqrt{49}} \leq \mu \leq 3250 + 1.96\frac{5}{\sqrt{49}}$$

$$3248.6 \leq \mu \leq 3251.4$$

Question 1

The standard IQ test has a mean of 106 and a standard deviation of 12. We want to be 90% certain that we are within 4 IQ points of the true mean. Determine the required sample size.

Select one:

- a. 25
- b. 6
- c. 34
- d. 130

Question 2

A nurse at a local hospital is interested in estimating the birth weight of infants. How large a sample must she select if she desires to be 95% confident that the true mean is within 4 ounces of the sample mean? The standard deviation of the birth weights is known to be 7 ounces.

Select one:

- a. 7
- b. 3
- c. 6
- d. 12

Question 3

The Hilbert Drug Store owner plans to survey a random sample of his customers with the objective of estimating the mean dollars spent on pharmaceutical products during the past three months. He has assumed that the population standard deviation is known to be $14.50. Given this information, what would be the required sample size if we want the total width of the two-side confidence interval on mean to be $4 at 95 percent confidence?

Select one:

- a. 202
- b. 16
- c. 231
- d. 163

Question 4

Past experience indicates that the standard deviation in the time it takes for a "fast lube" operation to actually complete the lube and oil change for customers is 3.00 minutes. The manager wishes to estimate the mean time with 99% confidence and a total width of the two-side confidence interval on mean to be 1 minute. Given this, what must the sample size be?

Select one:

- a. Nearly 500
- b. None of the above.
- c. Approximately 366
- d. About 239

Question 5

If a manager believes that the required sample size is too large for a situation in which she desires to estimate the mean income of blue collar workers in a state, which of the following would lead to a reduction in sample size?

Select one:

- a. Reduce the level of confidence
- b. Somehow reduce the variation in the population (reduce variance)
- c. All of the above.
- d. Allow a higher margin of error

Question 6

A university is interested in estimating the mean time that students spend at the student recreation center per week. A previous study indicated that the standard deviation in time is about 30 minutes per week. If the officials wish to estimate the mean time within 8 minutes with a 90 percent confidence, what should the sample size be?

Select one:

- a. Can't be determined without knowing how many students there are at the university.
- b. 39
- c. 62
- d. 44
- e. 302

Question 7

In order to set rates, an insurance company is trying to estimate the number of sick days that full time workers at an auto repair shop take per year. A previous study indicated that the standard deviation was 3.2 days. How large a sample must be selected if the company wants to be 95% confident that the true mean differs from the sample mean by no more than 2 day? Let $z_{0.05} = 1.96$.

Select one:

- a. 141
- b. 31
- c. 9
- d. 10

Question 8

In order to efficiently bid on a contract, a contractor wants to be 99% confident that his error is less than two hours in estimating the average time it takes to install tile flooring. Previous contracts indicate that the standard deviation is 5 hours. How large a sample must be selected? Let $z_{0.005} = 2.58$.

Select one:

- a. 42
- b. 43
- c. 41
- d. 40

Question 9

In order to fairly set flat rates for auto mechanics, a shop foreman needs to estimate the average time it takes to replace a fuel pump in a car. How large a sample must he select if he wants to be 99% confident that the true average time is within 8 minutes of the sample average? Assume the standard deviation of all times is 21 minutes. Let $z_{0.005} = 2.58$.

Select one:

- a. 46
- b. 45
- c. 47
- d. 48

Question 10

A Professor at Hanoi Medical University is interested in estimating the birth weight of infants. How large a sample must he select if he desires to be 99% confident that the true mean is within 0.1 kilograms of the sample mean? A past experience indicates that the standard deviation of the birth weights is known to be 0.7 kilograms. Let $z_{0.005} = 2.58$.

Select one:

- a. 300
- b. 327
- c. 319
- d. 301

Question 11

Suppose a 95% confidence interval for μ turns out to be (1000, 1900). Give a definition of what it means to be "95% confident" in an inference.

Select one:

- a. 95% of the observations in the sample fall in the given interval.
- b. In repeated sampling, 95% of the intervals constructed would contain the population mean.
- c. 95% of the observations in the entire population fall in the given interval.
- d. In repeated sampling, the population parameter would fall in the given interval 95% of the time.

Question 12

Suppose a 99% confidence interval for population mean turns out to be (1500, 2200). To make more useful inferences from the data, it is desired to reduce the width of the confidence interval. Which of the following will result in a reduced interval width?

Select one:

- a. Both increase the confidence level and decrease the sample size.
- b. Decrease the confidence level.
- c. Increase the sample size.
- d. Both increase the sample size and decrease the confidence level.

Question 13

An economist is interested in studying the incomes of consumers in a particular region. The population standard deviation is known to be \$1000. A random sample of 59 individuals resulted in an average income of \$21000. What is the width of the 90% confidence interval?

Select one:

- a. \$364.30
- b. \$232.60
- c. \$465.23
- d. \$428.32

Question 14

A 99% confidence interval estimate can be interpreted to mean that

(i) if all possible samples are taken and confidence interval estimates are developed, 99% of them would include the true population mean somewhere within their interval.

(ii) we have 99% confidence that we have selected a sample whose interval does include the population mean.

Select one:

- a. (i)
- b. Both of (i) and (ii)
- c. (ii)
- d. Neither (i) nor (ii)

Question 15

If you were constructing a 99% confidence interval of the population mean based on a sample of $n = 12$ where the standard deviation of the sample $s = 3.25$, the critical value of t will be

Select one:

- a. 3.1058
- b. 2.4922.
- c. 2.7874.
- d. 2.7969.

Question 16

It is desired to estimate the average total compensation of CEOs. Data were randomly collected from 32 CEOs and the 95% confidence interval was calculated to be ($3 212 540, $6 020 240). Which of the following interpretations is correct?

Select one:

- a. 95% of the sampled total compensation values fell between $3 212 540 to $6 020 240.
- b. We are 95% confident that the mean of the sampled CEOs falls in the interval $3 212 540 to $6 020 240.
- c. In the population of CEOs, 95% of them will have total compensations that fall in the interval $3 212 540 to $6 020 240.
- d. We are 95% confident that the average total compensation of all CEOs falls in the interval $3 212 540 to $6 020 240.

Question 17

Compute the critical value $z_{\alpha/2}$ that corresponds to a 94% level of confidence.

Select one:

- a. 2.33
- b. 1.88
- c. 1.645
- d. 1.96

Question 18

In a recent study of 49 eighth graders, the mean number of hours per week that they watched television was 18.6 with a population standard deviation of 6.8 hours. Find the 95% confidence interval for the population mean.

Select one:

- a. (16.7, 20.5)
- b. (18.3, 20.9)
- c. (19.1, 20.4)
- d. (17.5, 21.7)

Question 19

In a sample of 25 randomly selected women, it was found that their mean height was 65.2 inches. From previous studies, it is assumed that the standard deviation, σ, is 2.4. Construct the 95% confidence interval for the population mean.

Select one:

- a. (59.7, 66.5)
- b. (58.1, 67.3)
- c. (64.3, 66.1)
- d. (61.9, 64.9)

Question 20

A random sample of 169 students has a grade point average with a mean of 6.6 and with a population standard deviation of 0.8. Construct a 98% confidence interval for the population mean, μ.

Select one:

- a. (6.71, 8.01)
- b. (5.43, 7.79)
- c. (2.31, 3.88)
- d. (6.46, 6.74)

Question 21

A random sample of 42 students has a mean annual earnings of $1200 and a population standard deviation of $230. Construct a 95% confidence interval for the population mean, μ.

Select one:

- ○ a. ($1110, $2330)
- ○ b. ($210, $110)
- ○ c. ($1087, $2346)
- ○ d. ($1130, $1270)

Question 22

A random sample of 68 fluorescent light bulbs has a mean life of 600 hours with a population standard deviation of 25 hours. Construct a 95% confidence interval for the population mean.

Select one:

- a. (536.9, 653.1)
- b. (512.0, 768.0)
- c. (594.1, 605.9)
- d. (539.6, 551.2)

Question 23

A group of 65 randomly selected students has a mean age of 20.5 years with a population standard deviation of 2.7. Construct a 98% confidence interval for the population mean.

Select one:

- a. (21.1, 23.7)
- b. (18.8, 26.3)
- c. (19.7, 21.3)
- d. (19.8, 25.1)

Question 24

A group of 55 bowlers showed that their average score was 190 with a population standard deviation of 8. Find the 99% confidence interval of the mean score of all bowlers.

Select one:

- a. (186.5, 197.5)
- b. (189.5, 194.5)
- c. (187.2, 192.8)
- d. (188.5, 195.6)

Question 25

In a random sample of 120 computers, the mean repair cost was \$55 with a population standard deviation of \$12. Construct a 99% confidence interval for the population mean.

Select one:

- a. (\$18, \$62)
- b. (\$18, \$54)
- c. (\$53, \$65)
- d. (\$52, \$58)

Question 26

The grade point averages for 11 randomly selected students in a statistics class are listed below.

2.4 3.2 1.8 1.9 2.9 4.0 3.3 0.9 3.6 0.8 2.2

What is the effect on the width of the confidence interval if the sample size is increased to 15?

Select one:

- a. It is impossible to tell without more information.
- b. The width decreases.
- c. The width remains the same.
- d. The width increases.

Content

2. Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

When the population standard deviation σ is not know:
- n is reasonably large ($n \geq 30$), we can even handle the case of unknown variance for the large-sample-size situation in Section 1.4.

2. Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

When the population standard deviation σ is not know:
- n is reasonably large ($n \geq 30$), we can even handle the case of unknown variance for the large-sample-size situation in Section 1.4.
- However, when the sample is small:

- We need to replace the population standard deviation σ by a sample standard deviation s.

2. Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

When the population standard deviation σ is not know:
- n is reasonably large ($n \geq 30$), we can even handle the case of unknown variance for the large-sample-size situation in Section 1.4.
- However, when the sample is small:

- We need to replace the population standard deviation σ by a sample standard deviation s.
- Use t-distribution instead of normal distribution.

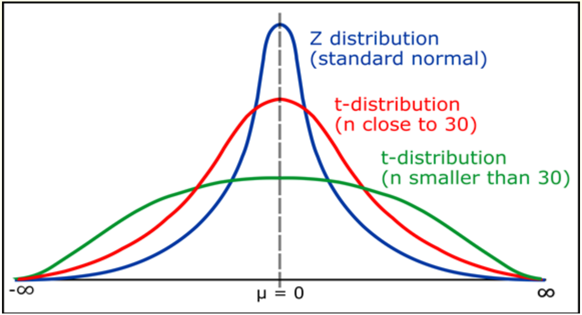
What is t-distribution?

The t probability density function is

$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\Gamma(k/2)} \cdot \frac{1}{[(x^2/k)+1]^{(k+1)/2}}, \quad -\infty < x < \infty$$

where, $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt, s > 0$, k is the number of degrees of freedom.

The mean and variance of the t distribution are zero and $k/(k-2)$ (for $k > 2$), respectively.



t Distribution

Let $X_1, X_2, ..., X_n$ is a random sample from a normal distribution with unknown mean μ and unknown variance σ^2. The random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a t distribution with $n - 1$ degrees of freedom.

2.1 t Distribution

Let $t_{\alpha,k}$ be the value of the random variable T with k degrees of freedom above which we find an area (or probability) α. Then, $t_{1-\alpha,k} = -t_{\alpha,k}$. Appendix Table V provides percentage points of the t distribution.
To illustrate the use of the table, note that the t-value with 10 degrees of freedom having an area of 0.05 to the right is $t_{0.05,10} = 1.812$. That is,

$$P(T > t_{0.05,10}) = P(T > 1.812) = 0.05$$

Therefore, $t_{0.95,10} = -t_{0.05,10} = -1.812$

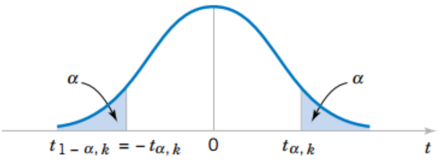


Figure 8-5 Percentage points of the t distribution.

2.2 t Confidence Interval on μ

We know that the distribution of $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ is t with $n-1$ degrees of freedom. Letting $t_{\alpha/2,n-1}$ be the upper $100\alpha/2$ percentage point of the t distribution with $n-1$ degrees of freedom, we may write

$$P(-t_{\alpha/2,n-1} < T < t_{\alpha/2,n-1}) = 1-\alpha$$

Hence,

$$P(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \le \mu \le \bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}) = 1-\alpha$$

2.2 t Confidence Interval on μ

Confidence Interval on the Mean, Variance Unknown

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2, a $100(1-\alpha)\%$ **confidence interval on** μ is given by

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

where $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the t distribution with $n-1$ degrees of freedom.

2.2 t Confidence Interval on μ

One-Sided Confidence Bounds on the Mean, Variance Known

A $100(1-\alpha)\%$ **upper-confidence bound** for μ is

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

and a $100(1-\alpha)\%$ **lower-confidence bound** for μ is

$$\bar{x} - t_{\alpha,n-1}\frac{s}{\sqrt{n}} \leq \mu$$

Example 6

The brightness of a television picture tube can be evaluated by measuring the amount of current required to achieve a particular brightness level. A sample of 12 tubes results in $\bar{x} = 316.8$ and $s = 15.2$. Find (in microamps) a 99% confidence interval on mean current required. Let $t_{0.005,11} = 3.106$

2.2 t Confidence Interval on μ

Example 6

The brightness of a television picture tube can be evaluated by measuring the amount of current required to achieve a particular brightness level. A sample of 12 tubes results in $\bar{x} = 316.8$ and $s = 15.2$. Find (in microamps) a 99% confidence interval on mean current required. Let $t_{0.005,11} = 3.106$

Answer: We known $n = 12, \bar{x} = 316.8, s = 15.2, \alpha = 0.01$ (variance unknown model). A 99% confidence interval on mean is

$$316.8 - 3.106\frac{15.2}{\sqrt{12}} \leq \mu \leq 316.8 + 3.106\frac{15.2}{\sqrt{12}}$$

$$303.171 \leq \mu \leq 330.429$$

Question 1

The grade point averages for 10 randomly selected high school students are listed below and has mean of 2.54 and standard deviation of 1.11.

2.9 0.9 4.0 3.6 0.8 2.0 3.2 1.8 3.3 2.9

Assume the grade point averages are normally distributed. Find a 98% confidence interval for the true mean.

Select one:

- a. (3.11, 4.35)
- b. (2.12, 3.14)
- c. (1.55, 3.53)
- d. (0.67, 1.81)

Question 2

A local bank needs information concerning the checking account balances of its customers. A random sample of 18 accounts was checked. The mean balance was \$600.70 with a standard deviation of \$196.20. Find a 98% confidence interval for the true mean. Assume that the account balances are normally distributed.

Select one:

- a. (\$481.85, \$719.55)
- b. (\$438.23, \$726.41)
- c. (\$513.17, \$860.33)
- d. (\$487.31, \$563.80)

Question 3

A sample of 28 teachers had mean annual earnings of \$3450 with a standard deviation of \$600. Construct a 95% confidence interval for the population mean, μ. Assume the population has a normal distribution.

Select one:

- a. (\$1324, \$3567)
- b. (\$2803, \$3437)
- c. (\$2135, \$4567)
- d. (\$3218, \$3682)

Question 4

A random sample of 15 students has a grade point average of 2.86 with a standard deviation of 0.78. Construct the confidence interval for the population mean at a significant level of 10% . Assume the population has a normal distribution.

Select one:

- a. (2.28, 3.66)
- b. (2.41, 3.42)
- c. (2.37, 3.56)
- d. (2.51, 3.21)

Question 5

Construct a 95% confidence interval for the population mean, μ. Assume the population has a normal distribution. A sample of 28 randomly selected students has a mean test score of 82.5 with a standard deviation of 9.2.

Select one:

- a. (56.12, 78.34)
- b. (66.35, 69.89)
- c. (78.93, 86.07)
- d. (77.29, 85.71)

Question 6

Construct a 95% confidence interval for the population mean, μ. Assume the population has a normal distribution. A random sample of 24 fluorescent light bulbs has a mean life of 665 hours with a standard deviation of 24 hours.

Select one:

- a. (531.2, 612.9)
- b. (628.5, 661.5)
- c. (654.9, 675.1)
- d. (876.2, 981.5)

Question 7

Construct a 99% confidence interval for the population mean, μ. Assume the population has a normal distribution. A group of 29 randomly selected students has a mean age of 20.4 years with a standard deviation of 3.5 years.

Select one:

- a. (18.6, 22.2)
- b. (16.3, 26.9)
- c. (17.2, 23.6)
- d. (19.9, 24.9)

Question 8

Construct a 96% confidence interval for the population mean, μ. Assume the population has a normal distribution. A study of 31 bowlers showed that their average score was 187 with a standard deviation of 8.

Select one:

- a. (186.3, 197.7)
- b. (115.4, 158.8)
- c. (183.9, 190.1)
- d. (222.3, 256.1)

Question 9

Construct a 90% confidence interval for the population mean, μ. Assume the population has a normal distribution. In a recent study of 22 eighth graders, the mean number of hours per week that they watched television was 20.5 with a standard deviation of 4.6 hours.

Select one:

- a. (18.81, 22.19)
- b. (19.62, 23.12)
- c. (5.87, 7.98)
- d. (17.47, 21.73)

Question 10

A random sample of 10 parking meters in a beach community showed the following incomes for a day. Assume the incomes are normally distributed.

\$6.30 \$6.75 \$4.25 \$3.60 \$4.50 \$2.80 \$8.00 \$3.00 \$2.60 \$5.20

Find the 95% confidence interval for the true mean.

Select one:

- a. (\$4.81, \$6.31)
- b. (\$2.11, \$5.34)
- c. (\$3.39, \$6.01)
- d. (\$1.35, \$2.85)

Content

1. Confidence Interval on the Mean of a Normal Distribution, Variance Known
2. Confidence Interval on the Mean of a Normal Distribution, Variance Unknown
3. **Large-Sample Confidence Interval for a Population Proportion**

3. Large-Sample Confidence Interval for a Population Proportion

A random sample of size n has been taken from a large (possibly infinite) population and that $X(\leq n)$ observations in this sample belong to a class of interest. Then $\hat{P} = \frac{X}{n}$ is a point estimator of the proportion of the population p that belongs to this class. Note that n and p are the parameters of a binomial distribution. When n large, we known that

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{P} - p}{\sqrt{p(1-p)/n}}$$

is approximately standard normal.
To construct the confidence interval on p, note that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) \cong 1 - \alpha$$

3. Large-Sample Confidence Interval for a Population Proportion

so

$$P(\hat{P} - z_{\alpha/2}\sqrt{p(1-p)/n} \leq p \leq \hat{P} + z_{\alpha/2}\sqrt{p(1-p)/n}) \cong 1 - \alpha$$

The quantity $\sqrt{p(1-p)/n}$ is called the **standard error of the point estimator** $\hat{P}$.

Approximate Confidence Interval on a Binomial Proportion

If $\hat{p}$ is the proportion of observations in a random sample of size n that belongs to a class of interest, an approximate $100(1-\alpha)\%$ confidence interval on the proportion p of the population that belongs to this class is

$$\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

3. Large-Sample Confidence Interval for a Population Proportion

Since $\hat{P}$ is the point estimator of p, we can define the error in estimating p by $\hat{P}$ as $E = \mid p - \hat{P} \mid$. Note that we are approximately $100(1-\alpha)\%$ confident that this error is less than $z_{\alpha/2}\sqrt{p(1-p)/n}$. In situations where the sample size can be selected, we may choose n to be $100(1-\alpha)\%$ confident that the error is less than some specified value E.

Sample Size for a Specified Error on a Binomial Proportion

If we set $E = z_{\alpha/2}\sqrt{p(1-p)/n}$ and solve for n, the appropriate sample size is:

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

Else, since $\max p(1-p) = 0.25$, we are at least $100(1-\alpha)\%$ confident:

$$n = (\frac{z_{\alpha/2}}{E})^2 \times 0.25$$

3. Large-Sample Confidence Interval for a Population Proportion

Example 7

Of 1000 randomly selected cases of lung cancer, 750 resulted in death within 10 years.

a) Calculate a 95% two-sided confidence interval on the death rate from lung cancer.

b) What sample size is needed to be 95% confident that the error in estimating the true value of p is less than 0.04? Let $z_{0.025} = 1.96$.

3. Large-Sample Confidence Interval for a Population Proportion

Example 7

Of 1000 randomly selected cases of lung cancer, 750 resulted in death within 10 years.
a) Calculate a 95% two-sided confidence interval on the death rate from lung cancer.
b) What sample size is needed to be 95% confident that the error in estimating the true value of p is less than 0.04? Let $z_{0.025} = 1.96$.

Answer. We known $n = 1000, \hat{p} = 750/1000 = 0.75$.
a) $\alpha = 0.05$. A 95% two-sided confidence interval on the death rate from lung cancer is

$$(0.75 - 1.96\sqrt{0.75 * 0.25/1000}; 0.75 + 1.96\sqrt{0.75 * 0.25/1000}) = (0.723, 0.777)$$

3. Large-Sample Confidence Interval for a Population Proportion

Example 7

Answer. We known $n = 1000, \hat{p} = 750/1000 = 0.75$.
b) $E = 0.04, \alpha = 0.05$. The required sample size

$$n = (\frac{1.96}{0.04})^2 \times 0.75 \times 0.25 \approx [450.19] = 451$$

3. Large-Sample Confidence Interval for a Population Proportion

Example 8

A local men's clothing store is being sold. The buyers are trying to estimate the percentage of items that are outdated. They will randomly sample among its 9800 items in order to determine the proportion of merchandise that is outdated. The current owners have never determined their outdated percentage and can not help the buyers. Approximately how large a sample do the buyers need in order to insure that they are 96% confident that the error is within 2%? Let $z_{0.02} = 2.05$.

3. Large-Sample Confidence Interval for a Population Proportion

Example 8

A local men's clothing store is being sold. The buyers are trying to estimate the percentage of items that are outdated. They will randomly sample among its 9800 items in order to determine the proportion of merchandise that is outdated. The current owners have never determined their outdated percentage and can not help the buyers. Approximately how large a sample do the buyers need in order to insure that they are 96% confident that the error is within 2%? Let $z_{0.02} = 2.05$.

Answer: $E = 0.02, \alpha = 0.04$. The required sample size is

$$n = (\frac{2.05}{0.02})^2 \times 0.25 \approx [2626.56] = 2627$$

3. Large-Sample Confidence Interval for a Population Proportion

Approximate OneSided Confidence Bounds on a Binomial Proportion

The approximate $100(1-\alpha)\%$ lower and upper confidence bounds are

$$\hat{p} - z_\alpha\sqrt{\hat{p}(1-\hat{p})/n} \leq p \text{ and } p \leq \hat{p} + z_\alpha\sqrt{\hat{p}(1-\hat{p})/n}$$

Question 1

The width of a confidence interval estimate for a proportion will be

Select one:

- a. narrower when the sample proportion is 0.70 than when the sample proportion is 0.20.
- b. narrower for 90% confidence than for 99% confidence.
- c. wider for a sample size of 150 than for a sample size of 120.
- d. narrower for 99% confidence than for 98% confidence.

Question 2

A confidence interval was used to estimate the proportion of statistics students that are females. A random sample of 200 statistics students generated the following 90% confidence interval: (0.48, 0.64). Based on the interval above, is the population proportion of females equal to 0.60?

Select one:

- a. Yes, and we are 90% sure of it.
- b. No. The proportion is 54.17%.
- c. Maybe. 0.60 is a believable value of the population proportion based on the information above.
- d. No, and we are 90% sure of it.

Question 3

An article a Florida newspaper reported on the topics that teenagers most want to discuss with their parents. The findings, the results of a poll, showed that 46% would like more discussion about the family's financial situation, 37% would like to talk about school, and 30% would like to talk about religion. These and other percentages were based on a national sampling of 549 teenagers. Estimate the proportion of all teenagers who want more family discussions about school. Use a 99% confidence level.

Select one:

- a. (0.628, 0.632)
- b. (0.577, 0.683)
- c. (0.318, 0.422)
- d. (0.368, 0.372)

Question 4

Many people think that a national lobby's successful fight against gun control legislation is reflecting the will of a minority of Americans. A random sample of 4000 citizens yielded 2250 who are in favor of gun control legislation. Estimate the true proportion of all Americans who are in favor of gun control legislation using a 90% confidence interval.

Select one:

- a. (0.4246, 0.4504)
- b. (0.1577, 0.9673)
- c. (0.5496, 0.5754)
- d. (0.4375, 0.8423)

Question 5

A university dean is interested in determining the proportion of students who receive some sort of financial aid. Rather than examine the records for all students, the dean randomly selects 200 students and finds that 118 of them are receiving financial aid. Use a 95% confidence interval to estimate the true proportion of students on financial aid.

Select one:

- a. (0.588, 0.592)
- b. (0.522, 0.658)
- c. (0.585, 0.595)
- d. (0.116, 1.064)

Question 6

Of 900 randomly selected cases of lung cancer, 360 resulted in death within five years. Construct a 95% two-sided confidence interval on the death rate from lung cancer.

Select one:

- a. (0.12, 0.95)
- b. (0.37, 0.43)
- c. (0.37, 0.95)
- d. (0.12, 0.43)

Question 7

A random sample of 60 suspension helmets used by motorcycle riders and automobile race-car drivers was subjected to an impact test, and on 15 of these helmets some damage was observed. Find a 95% two-sided confidence interval on the true proportion of helmets of this type that would show damage from this test.

Select one:

- a. (0.14, 0.36)
- b. (0.01, 0.68)
- c. (0.01, 0.36)
- d. (0.14, 0.68)

Question 8

A survey of 200 homeless persons showed that 35 were veterans. Construct a 90% confidence interval for the proportion of homeless persons who are veterans. Let $z_{0.05} = 1.65$.

Select one:

- a. (0.03, 0.35)
- b. (0.13, 0.22)
- c. (0.13, 0.35)
- d. (0.03, 0.22)

Question 9

A manufacturer of electronic calculators is interested in estimating the fraction of defective units produced. A random sample of 1500 calculators contains 15 defectives. Compute a 99% upper-confidence bound on the fraction defective. Let $\mathbf{z}_{0.005} = 2.58$ and $\mathbf{z}_{0.01} = 2.33$.

Select one:

- a. $0.003 \leq p$
- b. $p \leq 0.016$
- c. $0.003 \leq p \leq 0.017$
- d. $0.004 \leq p \leq 0.016$

Question 10

The fraction of defective integrated circuits produced in a photolithography process is being studied. A random sample of 200 circuits is tested, revealing 8 defectives. Find a 95% two-sided confidence interval on the fraction of defective circuits produced by this particular tool.

Select one:

- a. (0.013, 0.085)
- b. (0.003, 0.085)
- c. (0.013, 0.067)
- d. (0.003, 0.067)

Question 11

A confidence interval was used to estimate the proportion of statistics students that are female. A random sample of 100 statistics students generated the following 99% confidence interval: (0.438, 0.642). Using the information above, what total size sample would be necessary if we wanted to estimate the true proportion to within 0.04 using 95% confidence?

Select one:

- a. 420
- b. 150
- c. 597
- d. 105

Question 12

A researcher at a major hospital wishes to estimate the proportion of the adult population of the United States that has high blood pressure. How large a sample is needed in order to be 95% confident that the sample proportion will not differ from the true proportion by more than 5%?

Select one:

- a. 267
- b. 385
- c. 378
- d. 755

Question 13

A pollster wishes to estimate the proportion of United States voters who favor capital punishment. How large a sample is needed in order to be 98% confident that the sample proportion will not differ from the true proportion by more than 4%?

Select one:

- a. 752
- b. 1068
- c. 849
- d. 2135

Question 14

A private opinion poll is conducted for a politician to determine what proportion of the population favors decriminalizing marijuana possession. How large a sample is needed in order to be 97% confident that the sample proportion will not differ from the true proportion by more than 7%?

Select one:

- a. 921
- b. 378
- c. 241
- d. 461

Question 15

A manufacturer of golf equipment wishes to estimate the number of left-handed golfers. How large a sample is needed in order to be 95% confident that the sample proportion will not differ from the true proportion by more than 2%? A previous study indicates that the proportion of left-handed golfers is 15%.

Select one:

- a. 1225
- b. 241
- c. 217
- d. 153

Question 16

A researcher wishes to estimate the number of households with two cars. How large a sample is needed in order to be 98% confident that the sample proportion will not differ from the true proportion by more than 6%? A previous study indicates that the proportion of households with two cars is 25%.

Select one:

- a. 283
- b. 779
- c. 1101
- d. 1448

Question 17

A local men's clothing store is being sold. The buyers are trying to estimate the percentage of items that are outdated. They will randomly sample among its 100000 items in order to determine the proportion of merchandise that is outdated. The current owners have never determined their outdated percentage and can not help the buyers. Approximately how large a sample do the buyers need in order to insure that they are 94% confident that the error is within 1%?

Select one:

- a. 8836
- b. 3684
- c. 7368
- d. 1842

Question 18

Many people think that a national lobby's successful fight against gun control legislation is reflecting the will of a minority of Americans. A previous random sample of 4000 citizens yielded 2500 who are in favor of gun control legislation. How many citizens would need to be sampled if a 94% confidence interval was desired to estimate the true proportion to within 5%?

Select one:

- a. 385
- b. 379
- c. 354
- d. 332

Question 19

The State Transportation Department is interested in estimating the proportion of vehicle owners that are operating vehicles without the required liability insurance. If they wish to estimate the population proportion within ± 0.08 and use 96 percent confidence, what is the largest random sample that they will need?

Select one:

- a. About 600
- b. About 320
- c. About 165
- d. About 2,401

Question 20

A manager wishes to estimate the proportion of parts in his inventory that are in proper working order. However, the sample size that he has been informed he will need exceeds his budget. Which of the following steps might he take to reduce the required sample size?

Select one:

- a. None of the others.
- b. Decrease the desired error
- c. Increase the confidence level
- d. Use a smaller point estimate

Question 21

A regional hardware chain is interested in estimating the proportion of their customers who own their own homes. There is some evidence to suggest that the proportion might be around 0.825. Given this, what sample size is required if they wish a 94 percent confidence level with a error of ± 0.025?

Select one:

- a. About 1,300
- b. About 817
- c. About 910
- d. About 100

Question 22

Suppose that an internal report submitted to the managers at a bank in Boston showed that with 95% confidence, the proportion of the bank's customers who also have accounts at one or more other banks is between 0.40 and 0.46. Given this information, what sample size was used to arrive at this estimate?

Select one:

- a. Can't be determined without more information.
- b. Approximately 344
- c. Approximately 1,066
- d. Approximately 700