

Univariate Statistical Analysis

Lecture 6

Confidence Interval for Population Mean (Chapter 9, Section 9.3)

Today

- Define Confidence Interval
- Confidence Interval for a Population Mean

Revisit Lecture 5 Example 1 - Finding Area

The life of AA brand batteries is normally distributed with an average of 25 hours and a standard deviation of 4 hours. What is the probability that a randomly selected battery can last for:

- between 23 and 28 hours
- between 13 and 37 hours

Example 1 - Finding Area

Cont'

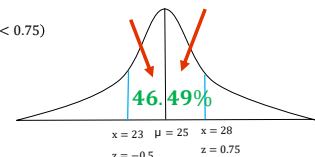
Given $\mu = 25$, $\sigma = 4$ and $z = \frac{x - \mu}{\sigma}$

a. $P(23 \leq x < 28)$

$$z = \frac{x - \mu}{\sigma} = \frac{23 - 25}{4} = \frac{-2}{4} = -0.5000$$

$$z = \frac{x - \mu}{\sigma} = \frac{28 - 25}{4} = \frac{3}{4} = 0.7500$$

So, $P(23 \leq x < 28) = P(-0.5 \leq z < 0.75)$
 $= P(z < 0.75) - P(z \leq -0.5)$
 $= 0.7734 - 0.3085$
 $= 0.4649$



Example 1 - Finding Area

Cont'

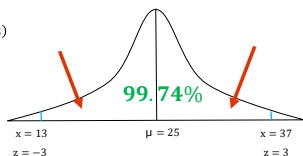
Given $\mu = 25$, $\sigma = 4$ and $z = \frac{x - \mu}{\sigma}$

b. $P(13 \leq x < 37)$

$$z = \frac{x - \mu}{\sigma} = \frac{13 - 25}{4} = \frac{-12}{4} = -3$$

$$z = \frac{x - \mu}{\sigma} = \frac{37 - 25}{4} = \frac{12}{4} = 3$$

So, $P(13 \leq x < 37) = P(-3 \leq z < 3)$
 $= P(z < 3) - P(z \leq -3)$
 $= 0.9987 - 0.0013$
 $= 0.9974$



Example 1 - Finding Area

Cont'

a. between 23 and 28 hours

$P(23 \leq x < 28) = 0.4649$

There is a 46.49% chance of getting a battery, which can last for between 23 and 28 hours.

b. between 13 and 37 hours

$P(13 \leq x < 37) = 0.9974$

There is a 99.74% chance of getting a battery, which can last for between 13 and 37 hours.

That is an example of showing how we use the sample data to describe the population scenario.

Given the **x-values** → z-scores → **area (probability)**

The reverse way is sorted of the process of constructing the Confidence Interval. Given the **area (probability)** → z-scores → **the interval (lower and upper bound values)**

Confidence Interval Population Mean μ z score

Sample data likely include the population parameter. Hence, based on some confidence levels, we can predict or construct the (mean) interval for the population.

The One-Sample z Confidence Interval for μ

The general formula for a confidence interval for a population mean μ when

1. $\bar{x}$ is the sample mean from a **simple random sample**,
2. the **sample size n is large** (generally $n \geq 30$), and
3. σ , the **population standard deviation, is known**

is

$$\bar{x} \pm (z \text{ critical value}) \left(\frac{\sigma}{\sqrt{n}} \right)$$

Confidence Interval, μ z score

$$\mu = \bar{x} \pm z_{\alpha} \left(\frac{\sigma}{\sqrt{n}} \right)$$

μ = Population Mean

$\bar{x}$ = Sample Mean

σ = Population Standard Deviation

n = Number of Samples

α = Significant Level = 1 – Confidence Level = 1 - CL

z_{α} = Critical Value

Example 2 - Scores

Scores on a test are normally distributed with a population standard deviation of 15. A random sample of 30 scores on the test has a mean of 50. Estimate the population mean with

- a. 90% confidence interval
- b. 95% confidence interval
- c. 99% confidence interval

Example 2 - Scores

Estimate the population mean with

a. 90% confidence interval
Given $\sigma = 15$, $n = 30$, $\bar{x} = 50$,
CI = 90% = 0.9 $\rightarrow \alpha = 1 - 0.9 = 0.1$

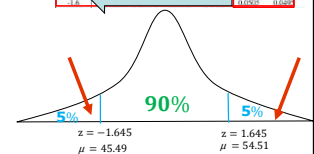
Critical value = $z_{\frac{\alpha}{2}} = z_{0.05} = -1.645$

$$\mu = \bar{x} \pm z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$\begin{aligned} &= 50 \pm (1.645) \left(\frac{15}{\sqrt{30}} \right) \\ &= 50 \pm (1.645)(2.74) \\ &= 50 \pm 4.51 \end{aligned}$$

$$45.49 \leq \mu \leq 54.51$$

z	0.00	0.01	0.02	0.03	0.04	0.05
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
-3.2	0.0007	0.0007	0.0007	0.0006	0.0006	0.0006
-3.1	0.0010	0.0010	0.0009	0.0009	0.0009	0.0009
-3.0	0.0013	0.0013	0.0013	0.0013	0.0012	0.0012
-2.9	0.0019	0.0018	0.0018	0.0018	0.0017	0.0017
-2.8	0.0026	0.0025	0.0025	0.0025	0.0024	0.0024
-2.7	0.0035	0.0034	0.0034	0.0033	0.0033	0.0032
-2.6	0.0047	0.0046	0.0045	0.0044	0.0044	0.0043
-2.5	0.0062	0.0060	0.0059	0.0058	0.0057	0.0056
-2.4	0.0082	0.0080	0.0079	0.0077	0.0076	0.0075
-2.3	0.0107	0.0104	0.0103	0.0102	0.0101	0.0100
-2.2	0.0139	0.0136	0.0135	0.0134	0.0133	0.0132
-2.1	0.0179	0.0174	0.0173	0.0171	0.0170	0.0169
-2.0	0.0228	0.0222	0.0221	0.0219	0.0218	0.0217
-1.9	0.0287	0.0281	0.0279	0.0277	0.0276	0.0275
-1.8	0.0359	0.0351	0.0349	0.0347	0.0346	0.0345
-1.7	0.0448	0.0439	0.0437	0.0435	0.0434	0.0433



We are 90% confident that the **population mean score** is between **45.49** and **54.51**.

Example 2 - Scores

Estimate the population mean with

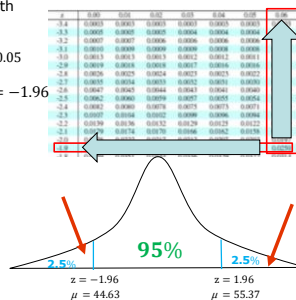
b. 95% confidence interval
Given $\sigma = 15$, $n = 30$, $\bar{x} = 50$,
CI = 95% = 0.95 $\rightarrow \alpha = 1 - 0.95 = 0.05$

Critical value = $z_{\frac{\alpha}{2}} = z_{0.025} = -1.96$

$$\mu = \bar{x} \pm z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$\begin{aligned} &= 50 \pm (1.96) \left(\frac{15}{\sqrt{30}} \right) \\ &= 50 \pm (1.96)(2.74) \\ &= 50 \pm 5.37 \end{aligned}$$

$$44.63 \leq \mu \leq 55.37$$



We are 95% confident that the **population mean score** is between **44.63** and **55.37**.

Example 2 - Scores

Estimate the population mean with

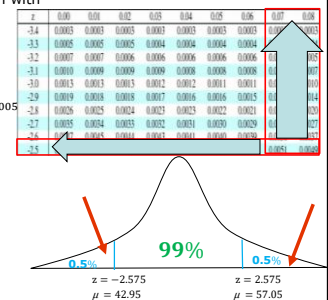
c. 99% confidence interval
Given $\sigma = 15$, $n = 30$, $\bar{x} = 50$,
CI = 99% = 0.99 $\rightarrow \alpha = 1 - 0.99 = 0.01$

Critical value = $z_{\frac{\alpha}{2}} = z_{0.005} = -2.575$

$$\mu = \bar{x} \pm z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$\begin{aligned} &= 50 \pm (2.575) \left(\frac{15}{\sqrt{30}} \right) \\ &= 50 \pm (2.575)(2.74) \\ &= 50 \pm 7.05 \end{aligned}$$

$$42.95 \leq \mu \leq 57.05$$



We are 99% confident that the **population mean score** is between **42.95** and **57.05**.

Example 2 – Scores**Cont'**

	90% C.I.	95% C.I.	99% C.I.
Z-value	1.645	1.96	2.575
Population Mean	$45.49 \leq \mu \leq 54.51$	$44.63 \leq \mu \leq 55.37$	$42.95 \leq \mu \leq 57.05$

Stop and Think

Assume other factors are remained the same, what's happen when C.I. increases?

Can we have 100% C.I. possible in real life?

Practical Question 1 – z score

Suppose that a random sample of 30 observations from a population whose standard deviation is 50 with a sample mean of 1,000. Estimate the population mean with

- 90% confidence interval
- 95% confidence interval
- 99% confidence interval

Answer:

- $984.98 \leq \mu \leq 1015.02$
- $982.11 \leq \mu \leq 1017.89$
- $976.49 \leq \mu \leq 1023.51$

Confidence Interval Population Mean μ *t* Distribution

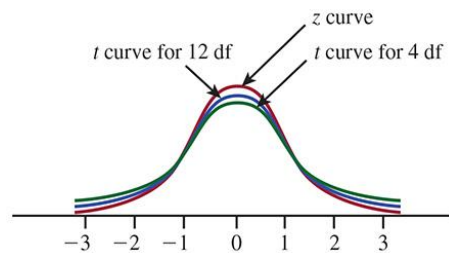
To be realistic, it is hard to know the population standard deviation, σ . Therefore, we use sample standard deviation, s , to estimate the population standard deviation, σ .

Hence, we need another distribution when the σ is UNKNOWN
→ *t* distributions.

Important Properties of *t* Distributions

- The *t* distribution corresponding to any particular number of degrees of freedom is bell shaped and centered at zero (just like the standard normal (*z*) distribution).
- Each *t* distribution is more spread out than the standard normal (*z*) distribution.
- As the number of degrees of freedom increases, the variability of the corresponding *t* distribution decreases.
- As the number of degrees of freedom increases, the corresponding sequence of *t* distributions approaches the standard normal (*z*) distribution.

Confidence Interval Population Mean μ *t* Distribution



Confidence Interval Population Mean μ *t* Distribution

The One-Sample *t* Confidence Interval for μ

The general formula for a confidence interval for a population mean μ based on a sample of size n when

- $\bar{x}$ is the sample mean from a **simple random sample**,
- the **population distribution is normal**, or the sample size n is large (generally $n \geq 30$), and
- σ , the **population standard deviation**, is **unknown**

is

$$\bar{x} \pm (t \text{ critical value}) \left(\frac{s}{\sqrt{n}} \right)$$

where the *t* critical value is based on $df = n - 1$. Appendix Table 3 gives critical values appropriate for each of the confidence levels 90%, 95%, and 99%, as well as several other less frequently used confidence levels.

Confidence Interval, μ *t* Distribution

$$\mu = \bar{x} \pm t_{df} \left(\frac{s}{\sqrt{n}} \right)$$

μ = Population Mean

$\bar{x}$ = Sample Mean

s = Sample Standard Deviation

n = Number of Samples

df = Degree of Freedom = $n - 1$ (The pieces of information required to work out n , e.g. traffic light)

t = Critical Value

Example 3 – Drive-Through Medicine (p.477)

Given $n = 38$, $\bar{x} = 26$ and $s = 1.57$, normally distributed. Estimate the population mean with a 95% confidence interval. $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$
Critical value = $t_{n-1} = t_{38-1} = t_{37}$
 $= 2.024$

$$\mu = \bar{x} \pm t_{df} \left(\frac{s}{\sqrt{n}} \right)$$

$$= 26 \pm (2.026) \left(\frac{1.57}{\sqrt{38}} \right)$$

$$= 26 \pm (2.026)(0.25)$$

$$= 26 \pm 0.51$$

$$25.49 \leq \mu \leq 26.51$$

$$t = -2.026$$

$$\mu = 25.49$$

$$t = 2.026$$

$$\mu = 26.51$$

We are 95% confident that the population mean is between 25.49 and 26.51.

Degrees of Freedom	.10	.05	.025	.01	.005
35	1.691	1.691	1.691	2.438	2.750
36	1.690	1.690	1.690	2.436	2.747
37	1.689	1.689	1.689	2.435	2.746
38	1.688	1.688	1.688	2.434	2.745
39	1.687	1.687	1.687	2.433	2.744
40	1.686	1.686	1.686	2.432	2.743

Example 4

Given $n = 5$, $\bar{x} = 30$ and $s = 1.45$, normally distributed. Estimate the population mean with a 90% confidence interval. $\frac{\alpha}{2} = \frac{0.1}{2} = 0.05$

Critical value = $t_{n-1} = t_{5-1} = t_4$
 $= 2.132$

$$\mu = \bar{x} \pm t_{df} \left(\frac{s}{\sqrt{n}} \right)$$

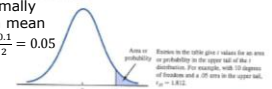
$$= 30 \pm (2.132) \left(\frac{1.45}{\sqrt{5}} \right)$$

$$= 30 \pm (2.132)(0.648)$$

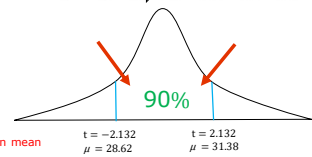
$$= 30 \pm 1.38$$

$$28.62 \leq \mu \leq 31.38$$

TABLE 2 t-Distribution



Degrees of Freedom	.10	.05	.025	.01	.005
1	1.676	1.676	1.676	3.078	3.078
2	1.385	1.385	1.385	2.353	2.353
3	1.250	1.250	1.250	2.353	2.353
4	1.190	1.190	1.190	2.353	2.353
5	1.156	1.156	1.156	2.353	2.353



We are 90% confident that the population mean is between 28.62 and 31.38.

$$t = -2.132$$

$$\mu = 28.62$$

$$t = 2.132$$

$$\mu = 31.38$$

Practical Question - t Distributions

Exercise 9.37 What is the appropriate t critical value for part a, c and e (page. 481)

a. $CI = 95\%$, $n = 17$
t critical value = 2.12

c. $CI = 99\%$, $n = 24$
t critical value = 2.807

e. $CI = 90\%$, $n = 13$
t critical value = 1.782

Conclusion

- Define Confidence Interval
- Confidence Interval for a Population Mean, μ
 - Population Standard deviation, σ is known
 - z score
 - $\mu = \bar{x} \pm z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}} \right)$
 - Population Standard deviation, σ is NOT known
 - t distribution
 - $\mu = \bar{x} \pm t_{df} \left(\frac{s}{\sqrt{n}} \right)$