



# ECON1310

## Introductory Statistics for Social Sciences

### Tutorial 7: CONFIDENCE INTERVALS I

Tutor: Francisco Tavares Garcia

# LBRT 01 marks and solutions are out!

## ECON1310 - Week 6: CML 3 (1st Attempt) Closing Today

Posted on: Monday, 9 January 2023 08:55:00 o'clock AEST

Dear Students,

Welcome to Week 6!

1. Your **marks and answers** to both attempts of **LBRT #1** are now available under **My Grades** on Blackboard. Please feel free to email me at [cml.1310@uq.edu.au](mailto:cml.1310@uq.edu.au) if you have any questions.
2. **CML 3 (1st Attempt)** will close at **4pm today** (9 January). Please read the **CML Information Sheet** carefully, especially the **CML rules** (located under the CML Administrative Folder). Remember to **CHECK, SAVE and SUBMIT** your CML before the closing time, as the quiz does NOT auto-submit. You will be able to **view your answers** to CML 3 (1st Attempt) **after the closing time at 4pm today** through the My Grades tab. **Instructions** on how to access your answers are located on page 7 of the CML Information Sheet.
3. **CML 3 (2nd Attempt)** will be open at **9am this Wednesday** (11 January) and close at **4pm this Friday** (13 January).
4. **CML 4 (1st Attempt)** will also be open at **9am this Wednesday** (11 January) and close at **4pm next Monday** (16 January).
5. **LBRT #2 (First Attempt)** will open at 9am on Tuesday, 17 January, 2023 and close at 4pm on Wednesday, 18 January, 2023. I will post some further details about LBRT#2 in an announcement later this week.

Feel free to email me for clarification on any of the above.

Best of luck!

Dominic

## ECON1310

### Tutorial 7 – Week 8

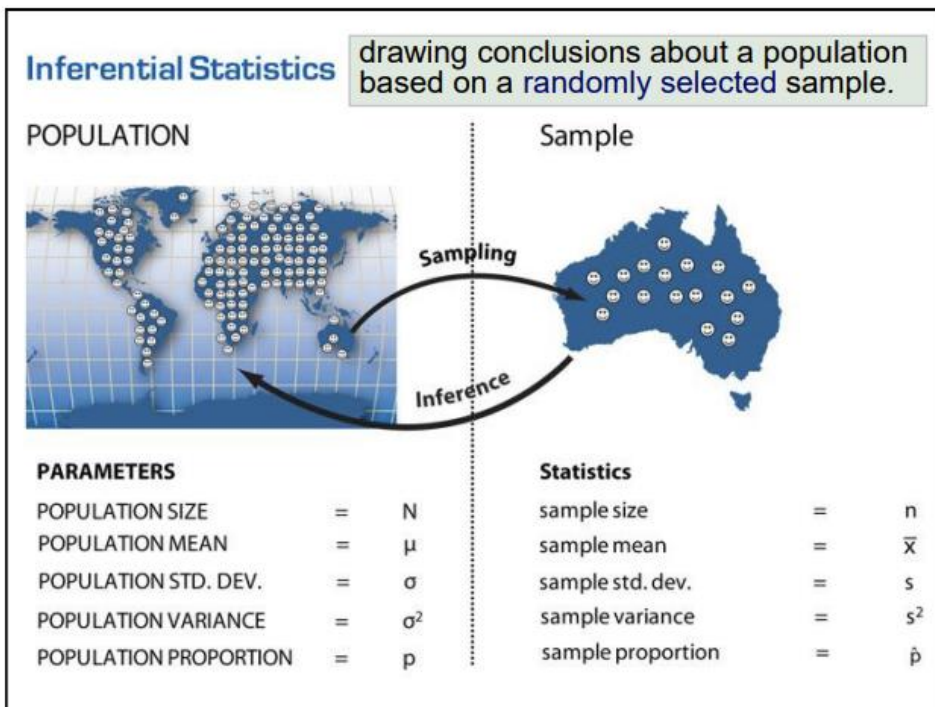
#### CONFIDENCE INTERVALS I

At the end of this tutorial you should be able to

- Describe the difference between a point estimate and interval estimate,
- Determine when it is appropriate to use the  $Z$  statistic for interval estimation and when it is appropriate to use the  $t$  statistic,
- Use the  $t$  distribution tables,
- Calculate confidence intervals for population means using  $Z$  statistics and  $t$  statistics,
- Calculate confidence intervals for population proportions.

- Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

(Poll)



1. What symbol would you give to the value 114 engineers? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐ n

2. What symbol would you give to the value 11.78 years? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐ n

3. What symbol would you give to the value 99% confidence interval? (Single Choice) \*

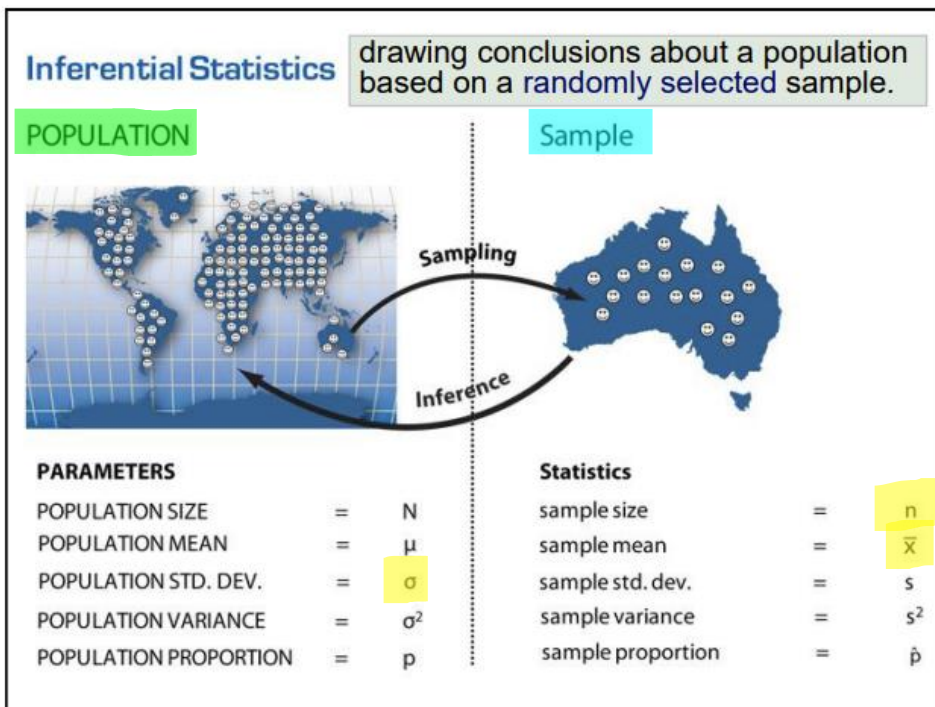
- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐ n

4. What symbol would you give to the value 3.2 years? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐ n

- Q1. A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

(Poll)



1. What symbol would you give to the value 114 engineers? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐  $s$   
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☒  $n$

2. What symbol would you give to the value 11.78 years? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐  $s$   
☐  $\mu$  (mu)  
☒  $\bar{x}$  (x bar)  
☐  $n$

3. What symbol would you give to the value 99% confidence interval? (Single Choice) \*

- ☒ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐  $s$   
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐  $n$

4. What symbol would you give to the value 3.2 years? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☒  $\sigma$  (sigma)  
☐  $s$   
☐  $\mu$  (mu)  
☐  $\bar{x}$  (x bar)  
☐  $n$

**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☐ Population Mean (Pelican) ( $\sigma$  is known)
- ☐ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☐ Population Proportion (Freaky fish) (proportion)



(Poll)

2. What table will we use? (Single Choice) \*

- ☐ Z table
- ☐ t table

3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

- ☐ 1%
- ☐ 5%
- ☐ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%

**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☒ Population Mean (Pelican) ( $\sigma$  is known)
- ☐ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☐ Population Proportion (Freaky fish) (proportion)

(Poll)

2. What table will we use? (Single Choice) \*

- ☒ Z table
- ☐ t table



3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

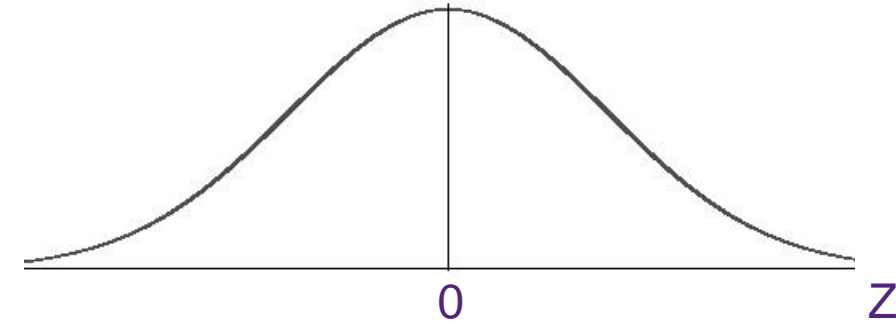
- ☒ 1%
- ☐ 5%
- ☐ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$



## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

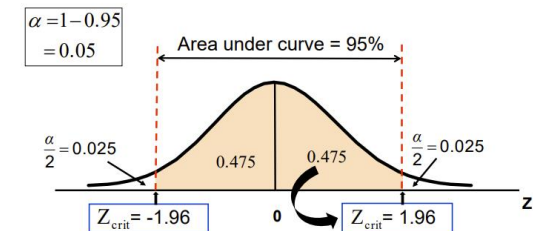
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

## Finding the Critical Value, $Z_{\text{crit}}$

Consider a 95% confidence interval

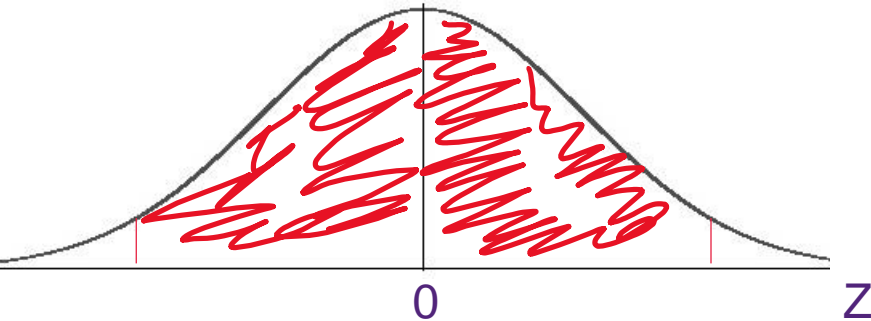







**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

0.99



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$

1. Confidence Interval for  $\mu$  ( $\sigma$  known)  
 (refer to the  transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

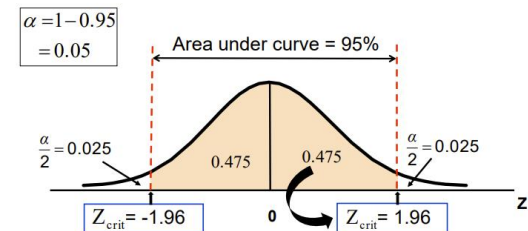
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

Finding the Critical Value,  $Z_{\text{crit}}$

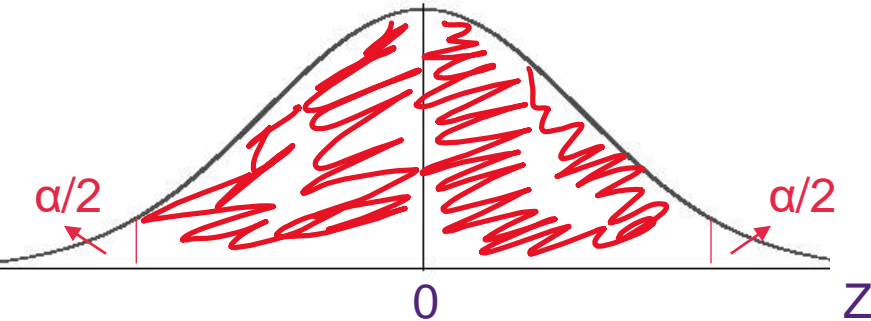
Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?

0.99



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$

1. Confidence Interval for  $\mu$  ( $\sigma$  known)  
 (refer to the  transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

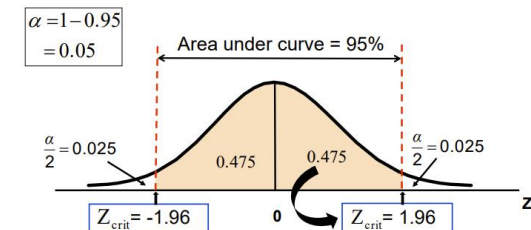
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

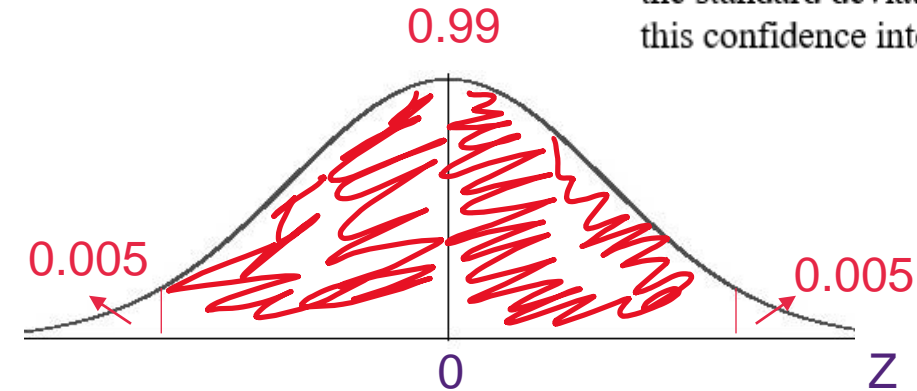
Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

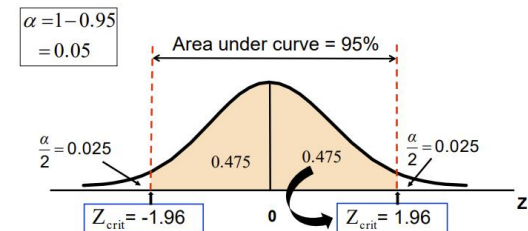
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

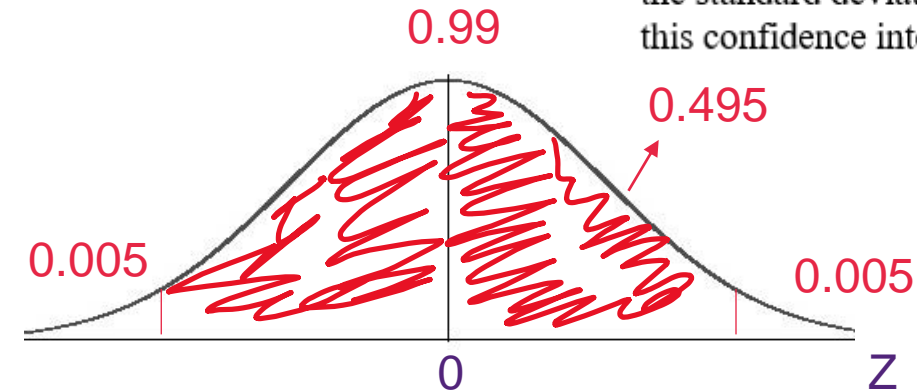
## Finding the Critical Value, $Z_{\text{crit}}$

Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

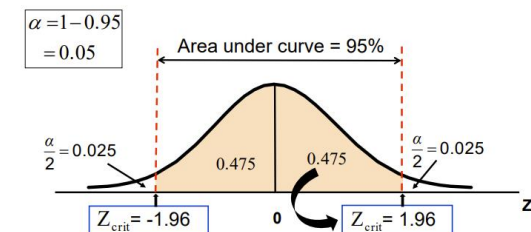
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

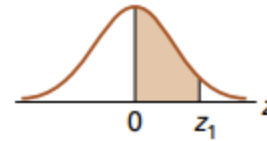
## Finding the Critical Value, $Z_{\text{crit}}$

Consider a 95% confidence interval



**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).



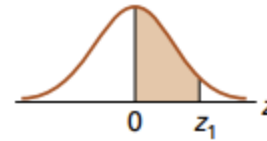
0.495

$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998									
4.0	.49997									
4.5	.499997									
5.0	.4999997									
6.0	.499999999									



**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).



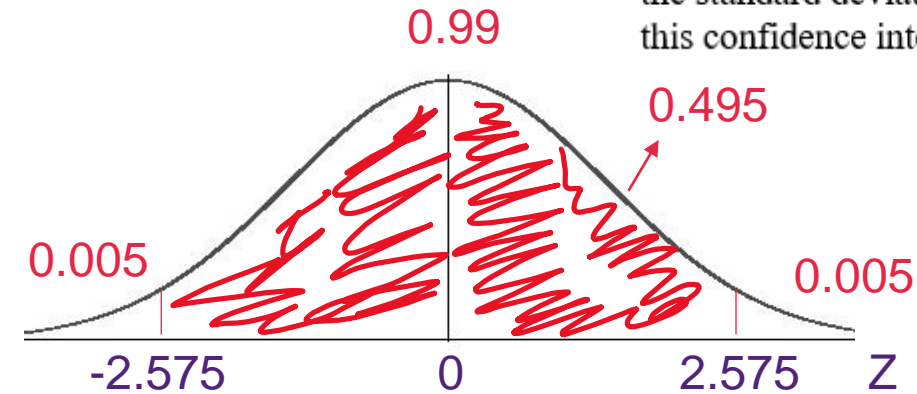
2.575

$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998									
4.0	.49997									
4.5	.499997									
5.0	.4999997									
6.0	.499999999									

0.495



**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{\text{crit}} = 2.575$

### 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

#### Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

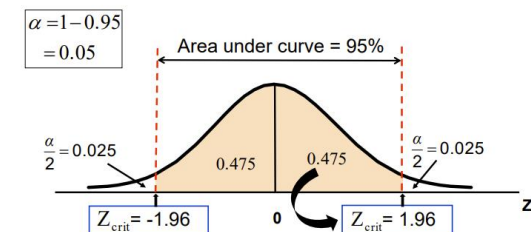
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{\text{crit}} * \frac{\sigma}{\sqrt{n}}$$

13

### Finding the Critical Value, $Z_{\text{crit}}$

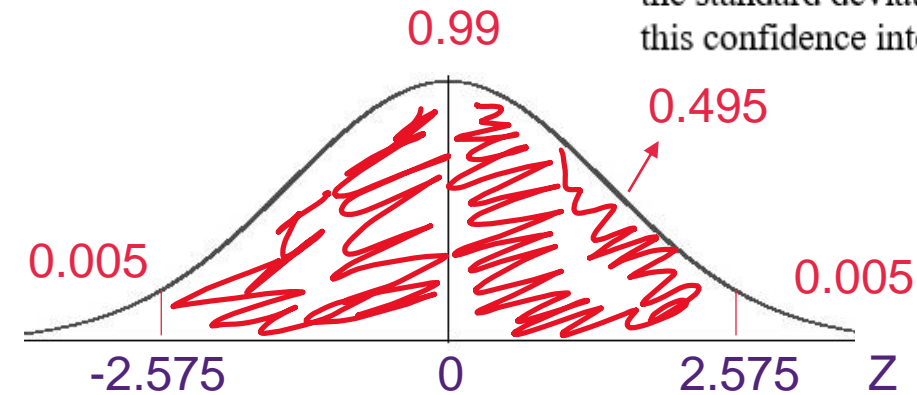
Consider a 95% confidence interval







**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

## Summary: Rearranged useful formulae



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

$$\mu = \bar{X} - Z \frac{\sigma}{\sqrt{n}} \quad (\text{use correct sign of } Z)$$

$$\bar{X} = \mu + Z \frac{\sigma}{\sqrt{n}} \quad (\text{use correct sign of } Z)$$

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

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

### Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging  $Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$

The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

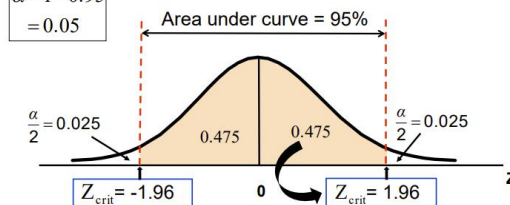
$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

## Finding the Critical Value, $Z_{crit}$

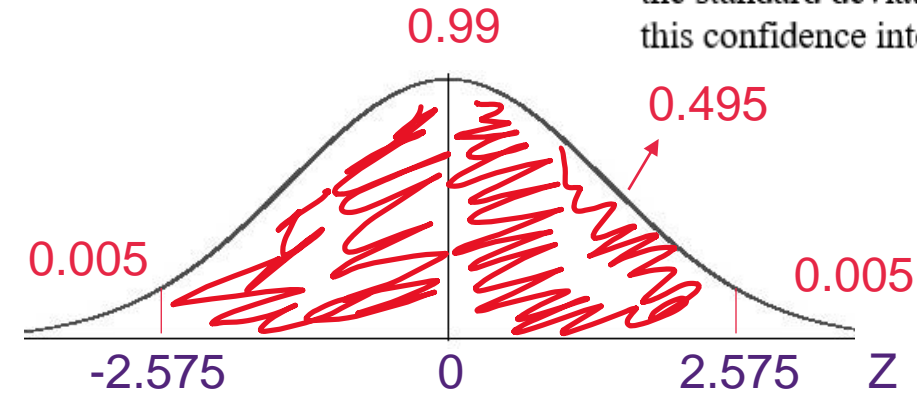
Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$






**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

$$\bar{X} \pm Z_{crit} * \sigma_{\bar{X}} = \bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}} = ?$$

1. Confidence Interval for  $\mu$  ( $\sigma$  known)  
 (refer to the  transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging  $Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$

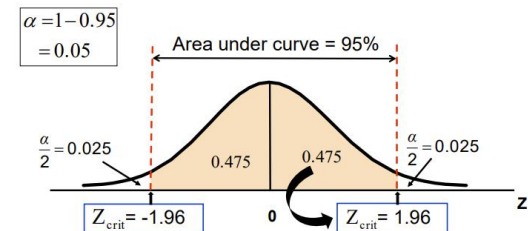
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

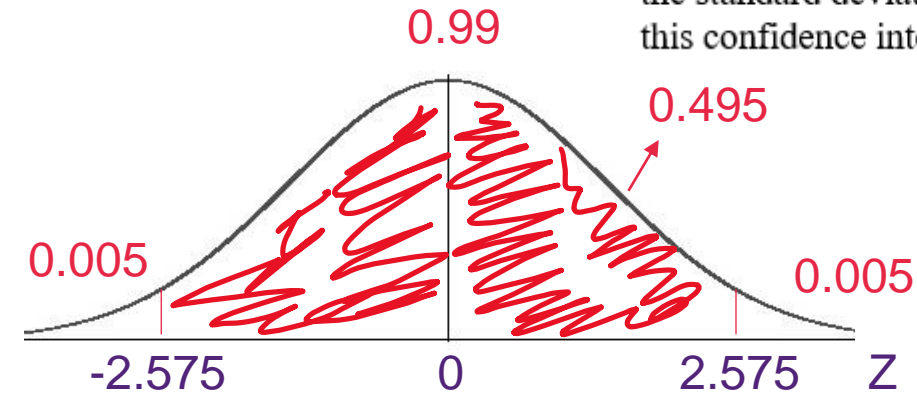
Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

$$\bar{X} \pm Z_{crit} * \sigma_{\bar{X}} = \bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}} = 11.78 \pm 2.575 * \frac{3.2}{\sqrt{114}}$$

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging  $Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$

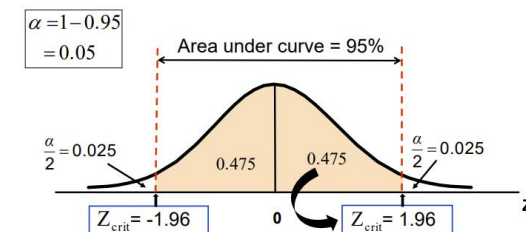
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

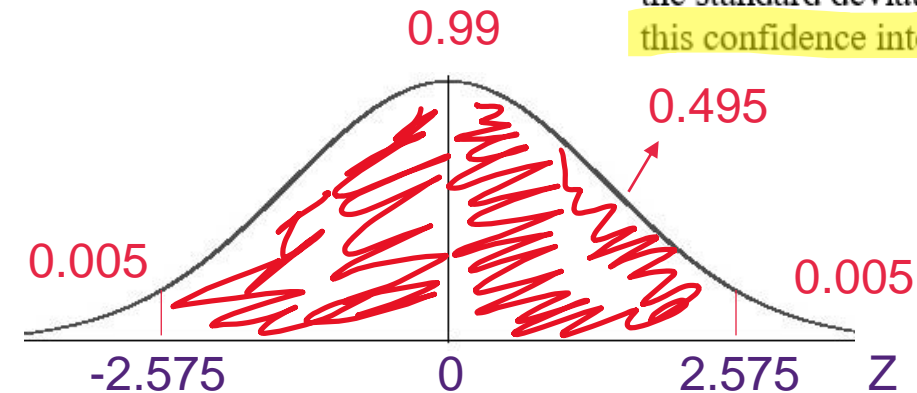
## Finding the Critical Value, $Z_{crit}$

Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. **How would you interpret this confidence interval?**



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

$$\bar{X} \pm Z_{crit} * \sigma_{\bar{X}} = \bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}} = 11.78 \pm 2.575 * \frac{3.2}{\sqrt{114}}$$

$$11.008 < \mu < 12.552$$

### 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

#### Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging  $Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$

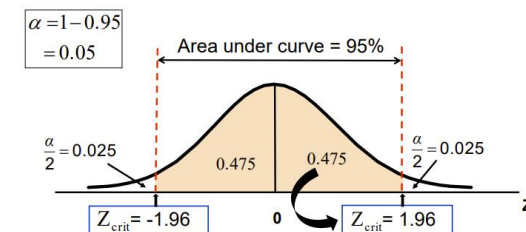
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

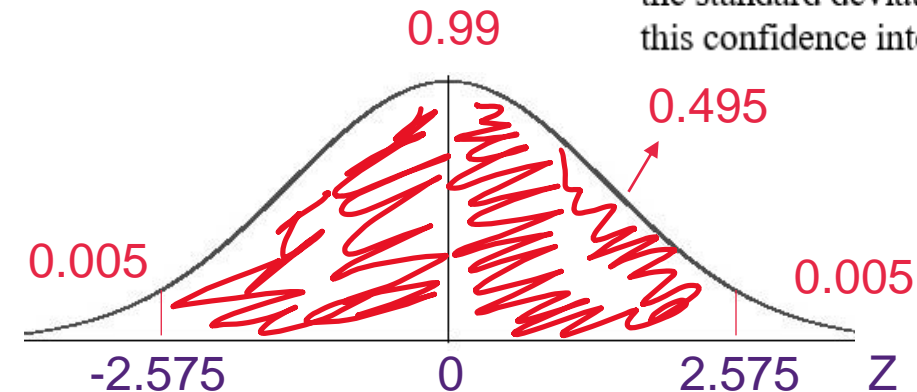
### Finding the Critical Value, $Z_{crit}$

Consider a 95% confidence interval





**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $\text{LOC} = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

$$\bar{X} \pm Z_{crit} * \sigma_{\bar{X}} = \bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}} = 11.78 \pm 2.575 * \frac{3.2}{\sqrt{114}}$$

$$11.008 < \mu < 12.552$$

Based on the sample mean, the mean number of years all engineers have spent with their company is estimated, with 99% confidence, to be between 11.008 and 12.552 years.

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging 
$$Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$$

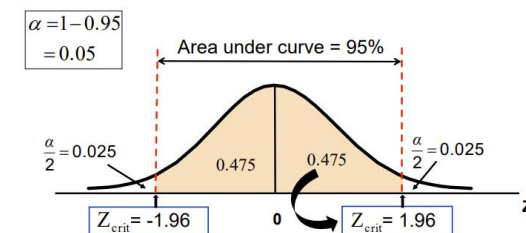
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

## Finding the Critical Value, $Z_{crit}$

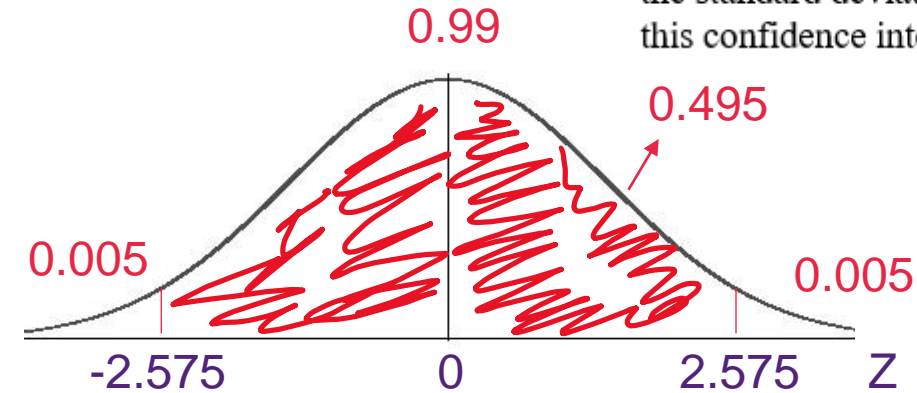
Consider a 95% confidence interval







**Q1.** A sample of 114 engineers was surveyed and it was found that the average time they had spent with their current company was 11.78 years. Determine the 99% confidence interval for the mean number of years engineers have spent with their current company. Historically the standard deviation of time spent with the company is 3.2 years. How would you interpret this confidence interval?



$n = 114$  engineers  
 $\bar{X} = 11.78$  years  
 $\sigma = 3.2$  years  
 $LOC = 99\% = 0.99$   
 $\alpha = 1 - 0.99 = 0.01$   
 $Z_{crit} = 2.575$

$$\bar{X} \pm Z_{crit} * \sigma_{\bar{X}} = \bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}} = 11.78 \pm 2.575 * \frac{3.2}{\sqrt{114}}$$

$$11.008 < \mu < 12.552$$

Based on the sample mean, the mean number of years all engineers have spent with their company is estimated, with 99% confidence, to be between 11.008 and 12.552 years.

## 1. Confidence Interval for $\mu$ ( $\sigma$ known) (refer to the transformation in Lecture 6)

Assumptions

- Population standard deviation  $\sigma$  is known
- Population variable is normally distributed
- If population variable is not normally distributed, use a large sample with  $n \geq 30$ .

by rearranging  $Z = \frac{(\bar{X} - \mu)}{(\frac{\sigma}{\sqrt{n}})}$

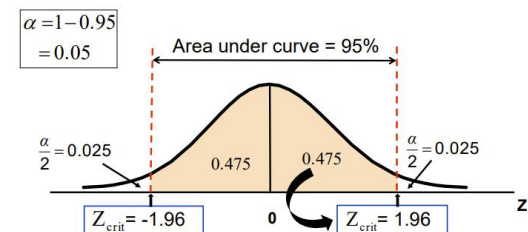
The confidence interval estimate for the population mean  $\mu$ , built around a random sample mean  $\bar{X}$  is:

$$\bar{X} \pm Z_{crit} * \frac{\sigma}{\sqrt{n}}$$

13

## Finding the Critical Value, $Z_{crit}$

Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?

(Poll)

### Inferential Statistics

drawing conclusions about a population based on a randomly selected sample.

#### POPULATION



#### Sample



Sampling

Inference

#### PARAMETERS

POPULATION SIZE	=	N
POPULATION MEAN	=	$\mu$
POPULATION STD. DEV.	=	$\sigma$
POPULATION VARIANCE	=	$\sigma^2$
POPULATION PROPORTION	=	p

#### Statistics

sample size	=	n
sample mean	=	$\bar{x}$
sample std. dev.	=	s
sample variance	=	$s^2$
sample proportion	=	$\hat{p}$

1. What symbol would you give to the value 80 randomly selected people? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐ p  
☐  $\hat{p}$  (p hat)  
☐ n

2. What symbol would you give to the value 90%? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐ s  
☐ p  
☐  $\hat{p}$  (p hat)  
☐ n

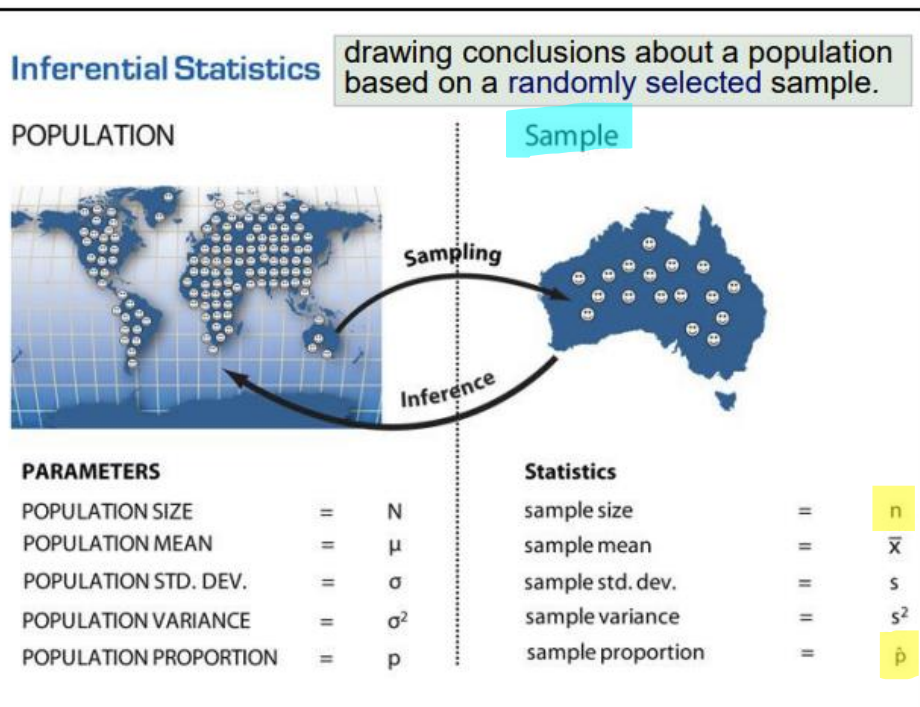
3. What is the value of  $\hat{p}$  (p hat)? (Single Choice)

- ☐ 0.48  
☐ 0.6  
☐ 0.8  
☐ 48  
☐ 80



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?

## (Poll)



1. What symbol would you give to the value 80 randomly selected people? (Single Choice) \*

- ☐ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐  $s$   
☐  $p$   
☐  $\hat{p}$  (p hat)  
☒  $n$

2. What symbol would you give to the value 90%? (Single Choice) \*

- ☒ Level of Confidence (LOC)  
☐  $\sigma$  (sigma)  
☐  $s$   
☐  $p$   
☐  $\hat{p}$  (p hat)  
☐  $n$

3. What is the value of  $\hat{p}$  (p hat)? (Single Choice)

- ☐ 0.48  
☒ 0.6  
☐ 0.8  
☐ 48  
☐ 80

**Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?

1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☐ Population Mean (Pelican) ( $\sigma$  is known)
- ☐ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☐ Population Proportion (Freaky fish) (proportion)



(Poll)

2. What table will we use? (Single Choice) \*

- ☐ Z table
- ☐ t table

3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

- ☐ 1%
- ☐ 5%
- ☐ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%

**Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?

1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☐ Population Mean (Pelican) ( $\sigma$  is known)
- ☐ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☒ Population Proportion (Freaky fish) (proportion)

(Poll)

2. What table will we use? (Single Choice) \*

- ☒ Z table
- ☐ t table

3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

- ☐ 1%
- ☐ 5%
- ☒ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

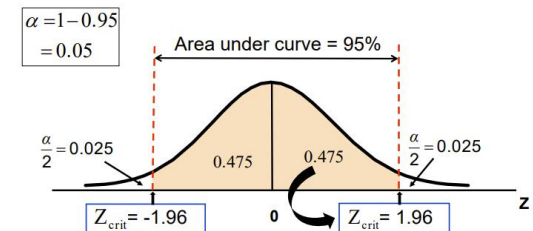
$\hat{p}$  = is the sample proportion

$n$  = sample size



Finding the Critical Value,  $Z_{\text{crit}}$

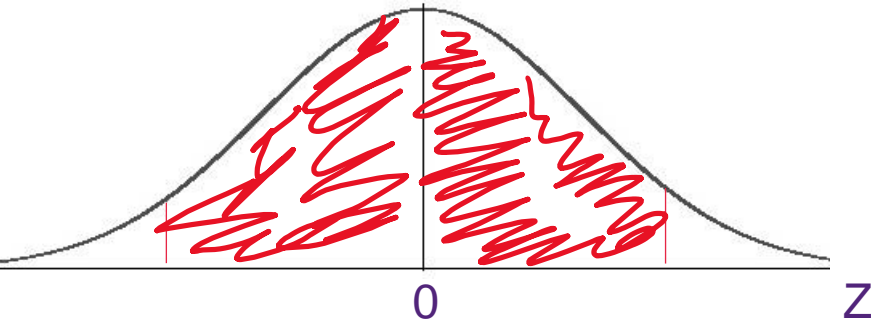
Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



0.90



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

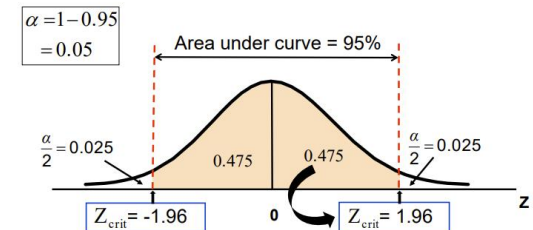
$\hat{p}$  = is the sample proportion

$n$  = sample size



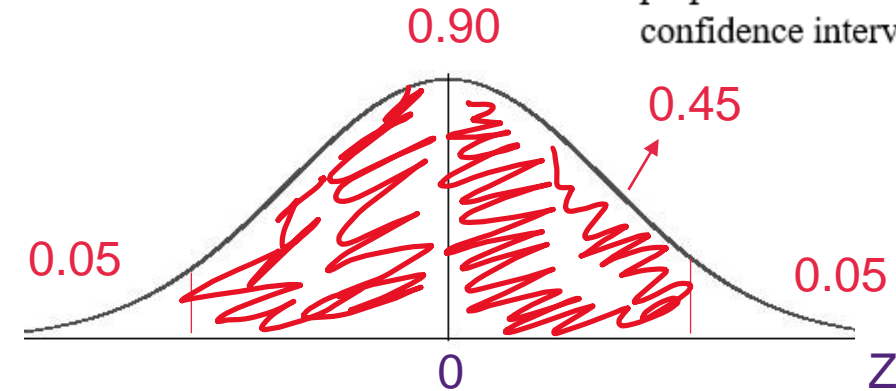
Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval





- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

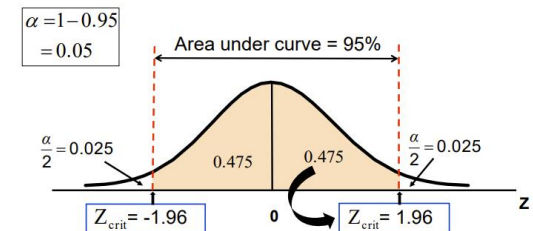
$\hat{p}$  = is the sample proportion

$n$  = sample size



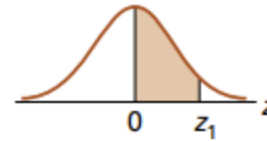
Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval



**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).



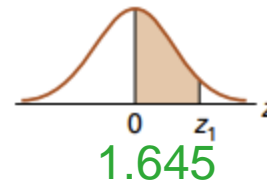
0.45

$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767



**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).

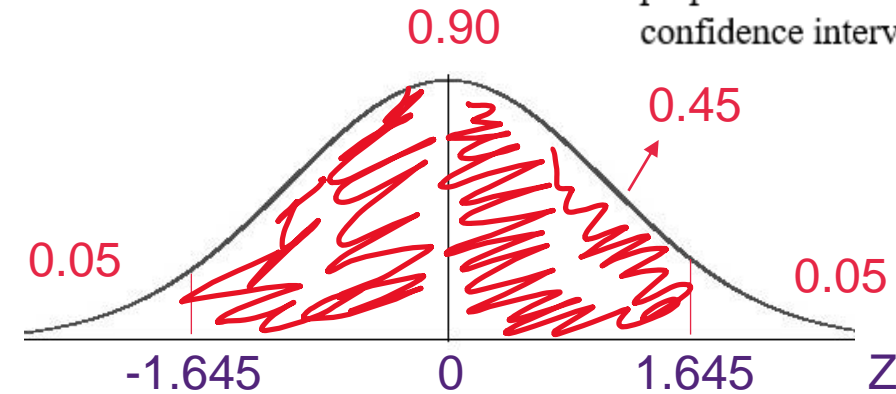


0.45

$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

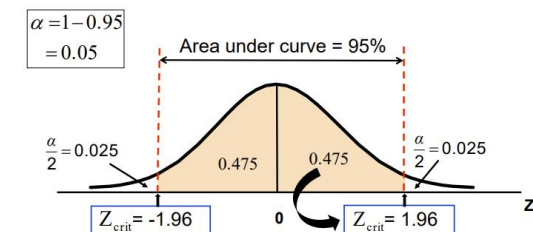
$\hat{p}$  = is the sample proportion

$n$  = sample size

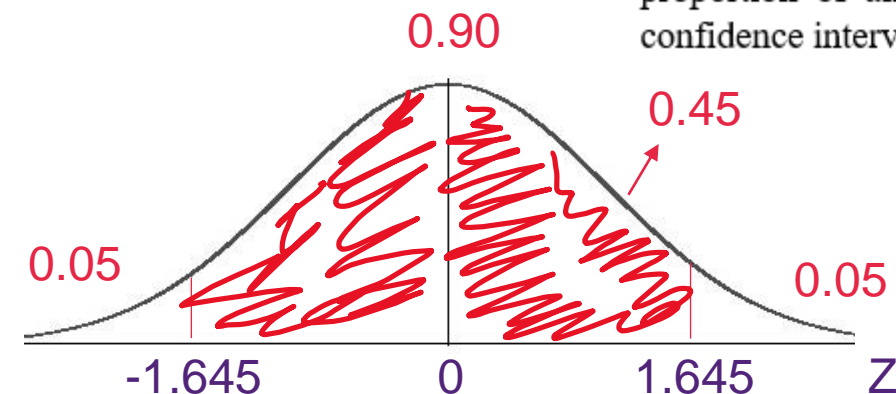


Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = ?$$

The confidence interval limits for a population proportion are:

Lower limit:  $\hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Upper limit:  $\hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{crit}$  = critical value of Z for the level of confidence

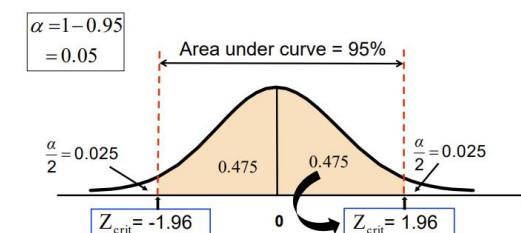
$\hat{p}$  = is the sample proportion

$n$  = sample size

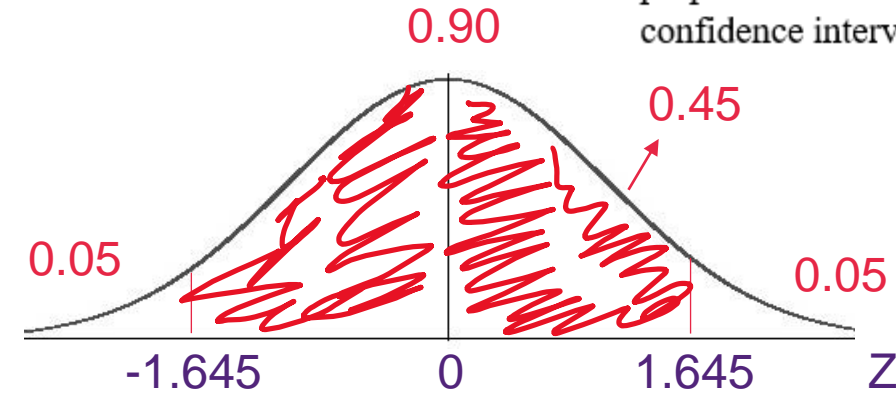


Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.645 * \sqrt{\frac{0.6(1-0.6)}{80}} =$$

The confidence interval limits for a population proportion are:

Lower limit:  $\hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Upper limit:  $\hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{crit}$  = critical value of Z for the level of confidence

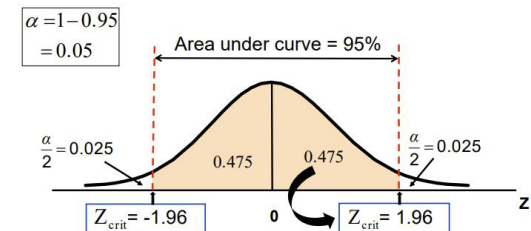
$\hat{p}$  = is the sample proportion

$n$  = sample size

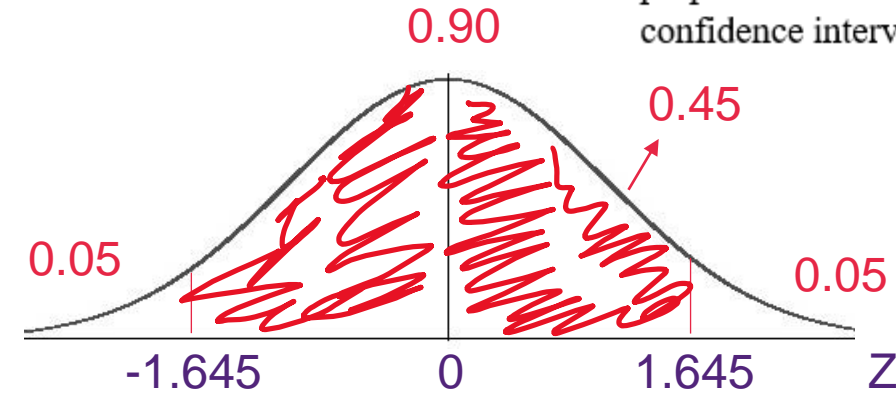


Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.645 * \sqrt{\frac{0.6(1-0.6)}{80}} =$$

$$0.5099 < p < 0.6901$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

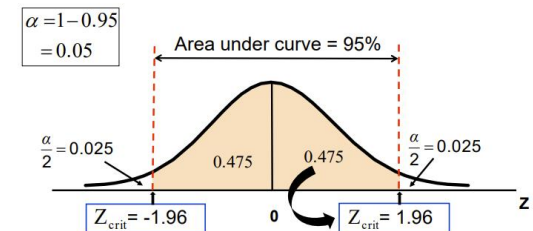
$\hat{p}$  = is the sample proportion

$n$  = sample size



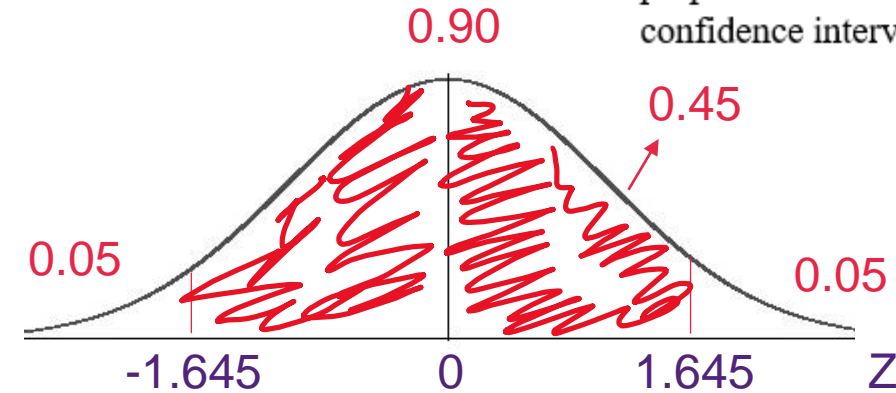
Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval





- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.645 * \sqrt{\frac{0.6(1-0.6)}{80}} =$$

$$0.5099 < p < 0.6901$$

Based on the sample proportion, we have 90% of confidence that the proportion of voters who would vote for Mrs Wilson is estimated to be between 51% and 69%.

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

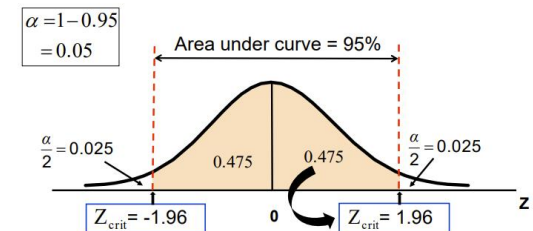
$\hat{p}$  = is the sample proportion

$n$  = sample size

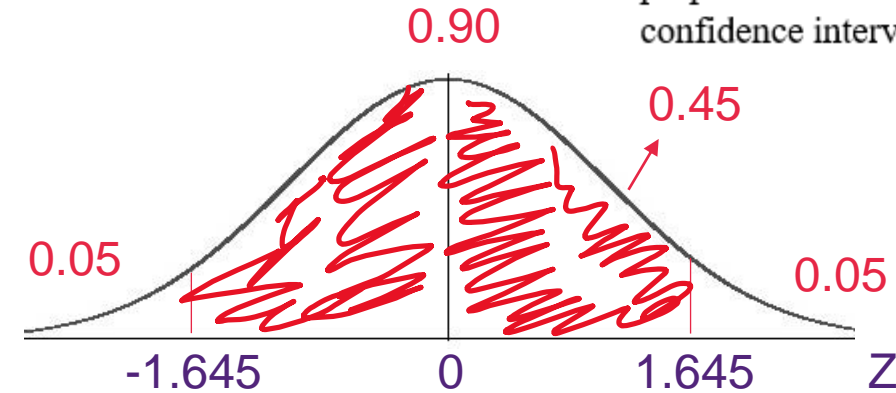


Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval



- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.645 * \sqrt{\frac{0.6(1-0.6)}{80}} =$$

$$0.5099 < p < 0.6901$$

Based on the sample proportion, we have 90% of confidence that the proportion of voters who would vote for Mrs Wilson is estimated to be between 51% and 69%.

The confidence interval limits for a population proportion are:

**Lower limit:**  $\hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**Upper limit:**  $\hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{crit}$  = critical value of Z for the level of confidence

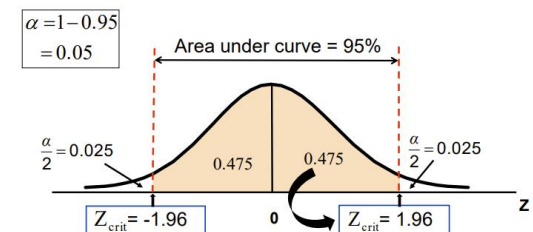
$\hat{p}$  = is the sample proportion

$n$  = sample size



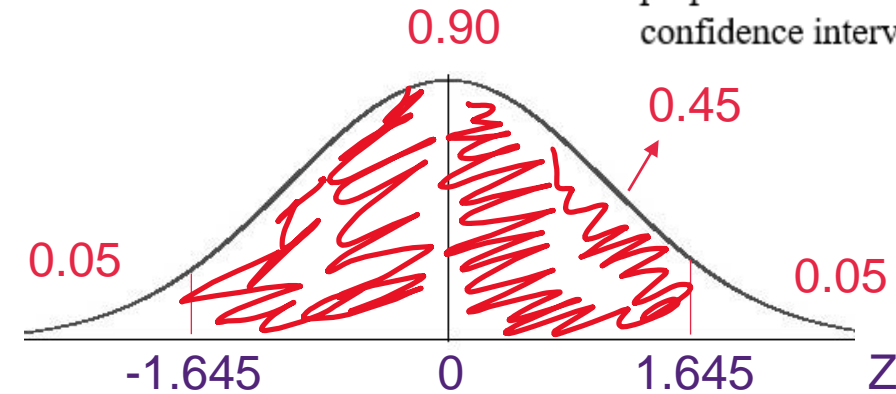
Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval





- Q2.** Mrs Wilson is considering running for mayor in her town. She decides to conduct a survey before completing the formalities of putting her name forward. Of the 80 randomly selected people surveyed, 48 say they will vote for her. Find a 90% confidence interval for the proportion of all voters in this community who will vote for Mrs Wilson. From your confidence interval can you determine whether or not she should run for mayor?



$$\begin{aligned} n &= 80 \text{ people} \\ \hat{p} &= 48/80 = 0.6 \\ \text{LOC} &= 90\% = 0.9 \\ \alpha &= 1 - 0.9 = 0.1 \end{aligned}$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.645 * \sqrt{\frac{0.6(1-0.6)}{80}} =$$

$$0.5099 < p < 0.6901$$

Based on the sample proportion, we have 90% of confidence that the proportion of voters who would vote for Mrs Wilson is estimated to be between 51% and 69%.

Since she should be confident of winning over 50% of the votes, she should run for mayor.

The confidence interval limits for a population proportion are:

**Lower limit:**  $\hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**Upper limit:**  $\hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{crit}$  = critical value of Z for the level of confidence

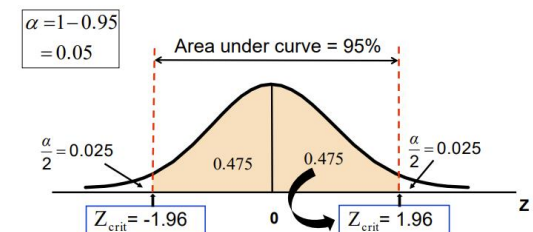
$\hat{p}$  = is the sample proportion

$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval



- Q3.** i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?

## (Poll)

1. What is the mean of a t distribution? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

2. Are values in the t distribution positive or negative? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

3. What is the parameter of the t distribution? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

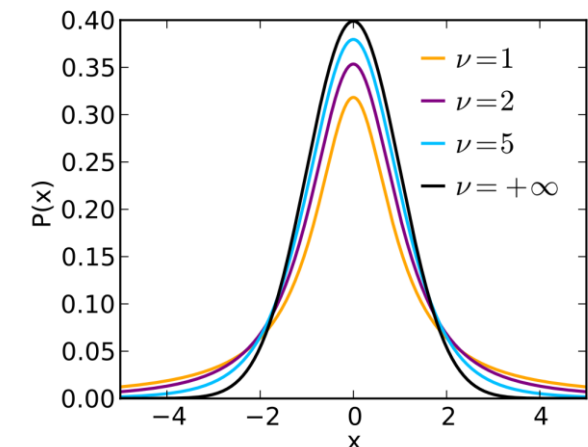
4. How do we calculate the degrees of freedom? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

- Q3.** i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?

## (Poll)

- Family of symmetrical bell-shaped curves
- Mean is 0
- Always above the horizontal axis
- Infinite in both directions
- Parameter: Degrees of freedom



Source: Wikipedia, Skbkakas

1. What is the mean of a t distribution? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☒ 0
- ☐ (n - 1)

2. Are values in the t distribution positive or negative? (Single Choice) \*

- ☒ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

3. What is the parameter of the t distribution? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☒ Degrees of freedom
- ☐ 0
- ☐ (n - 1)

4. How do we calculate the degrees of freedom? (Single Choice) \*

- ☐ Always positive (above the horizontal axis)
- ☐ Negative and positive (both sides of the horizontal axis)
- ☐ Degrees of freedom
- ☐ 0
- ☒ (n - 1)

- Q3.** i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?

## (Poll)

1. Which distribution is taller at the mean? (Single Choice) \*

- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal

3. Which distribution has more variability ( $\sigma > 1$ )? (Single Choice) \*

- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal

2. Which distribution is fatter tails? (Single Choice) \*

- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal

4. When does the t distribution approaches a Z distribution? (Single Choice) \*

- ☐ When the degrees of freedom decreases. (smaller sample)
- ☐ When the degrees of freedom increases. (bigger sample)

- Q3. i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?

(Poll)

1. Which distribution is taller at the mean? (Single Choice) \*

- ☒ Z distribution
- ☐ t distribution
- ☐ It's the equal

3. Which distribution has more variability ( $\sigma > 1$ )? (Single Choice) \*

- ☐ Z distribution
- ☒ t distribution
- ☐ It's the equal

2. Which distribution is fatter tails? (Single Choice) \*

- ☐ Z distribution
- ☒ t distribution
- ☐ It's the equal

4. When does the t distribution approaches a Z distribution? (Single Choice) \*

- ☐ When the degrees of freedom decreases. (smaller sample)
- ☒ When the degrees of freedom increases. (bigger sample)

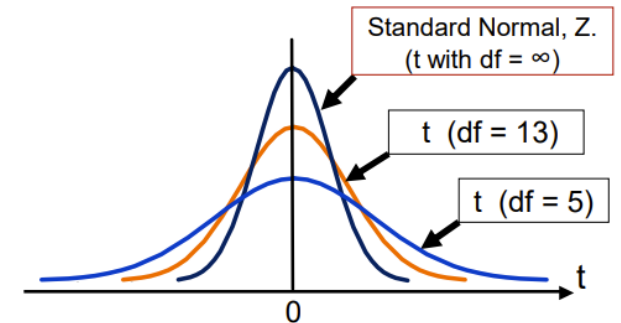
Student t distribution

- Not as high in the middle.
- Has fatter tails.
- Has more variability ( $\sigma > 1$ ).
- As  $df \uparrow$ , it approaches Z-dist.

#### t-distribution

is **NOT** a normal distribution, even though it looks similar. It is **another** mathematical function.

it does approach Z distribution as n increases.



- Q3.** i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?

(Poll)

1. Which distribution we use when  $\sigma$  is known? (Single Choice) \*

- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal

2. Which distribution we use when  $\sigma$  is unknown? (Single Choice) \*

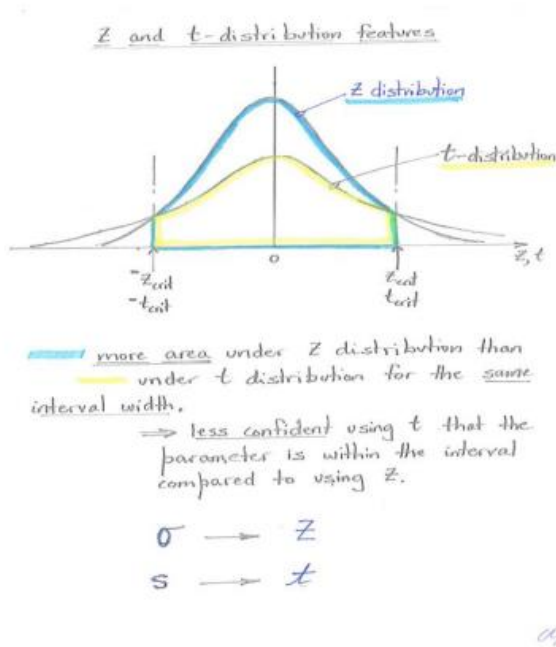
- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal

3. Which distribution is more flexible in regard to the population being normally distributed? (Single Choice) \*

- ☐ Z distribution
- ☐ t distribution
- ☐ It's the equal



- Q3.** i) What are the important characteristics of the Student t distribution?
- ii) How does the Student t distribution compare with the standard normal distribution?
- iii) When is the Student t distribution used in estimation?



1. Which distribution we use when  $\sigma$  is known? (Single Choice) \*

- ☒ Z distribution
- ☐ t distribution
- ☐ It's the equal

2. Which distribution we use when  $\sigma$  is unknown? (Single Choice) \*

- ☐ Z distribution
- ☒ t distribution
- ☐ It's the equal

3. Which distribution is more flexible in regard to the population being normally distributed? (Single Choice) \*

- ☐ Z distribution
- ☒ t distribution
- ☐ It's the equal

## Fun facts

In the English-language literature, the distribution takes its name from William Sealy Gosset's 1908 paper in *Biometrika* under the pseudonym "Student".<sup>[9]</sup> One version of the origin of the pseudonym is that Gosset's employer preferred staff to use pen names when publishing scientific papers instead of their real name, so he used the name "Student" to hide his identity. Another version is that Guinness did not want their competitors to know that they were using the t-test to determine the quality of raw material.<sup>[10][11]</sup>



The Brewer Who Secretly Revolutionized Statistics | Great Minds: William Gosset  
183K views • 1 year ago

SciShow

When you have a study with a small sample size, how do you know that the results...

Subtitles

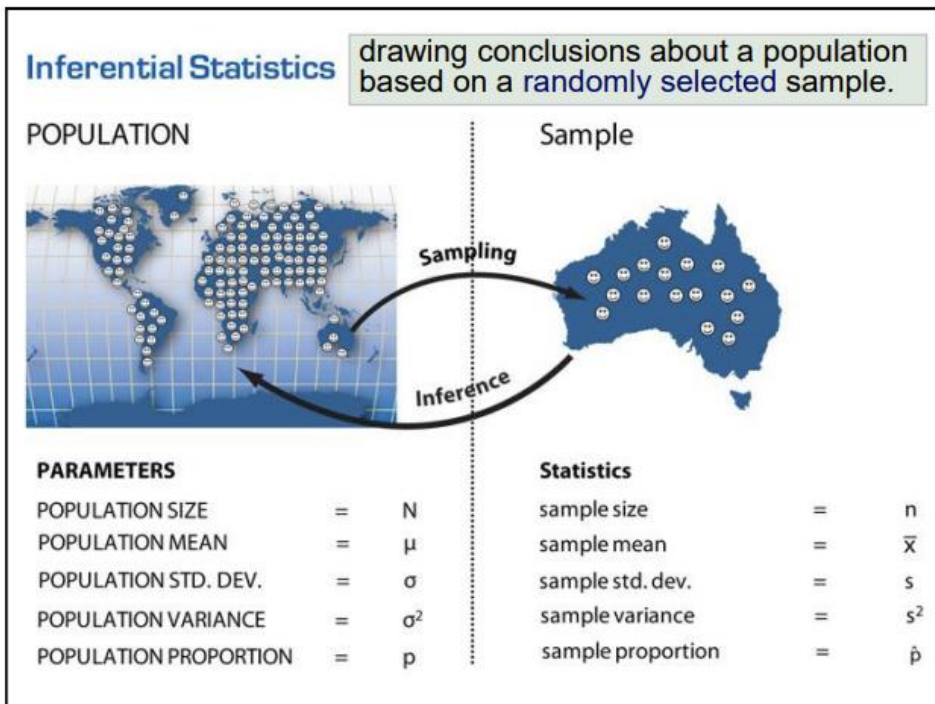
t-distribution

3 moments

**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?

(Poll)



1. What symbol would you give to the random sample of 70? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

2. What symbol would you give to the value 175.9 square metres? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

3. What symbol would you give to the value 38 square metres? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

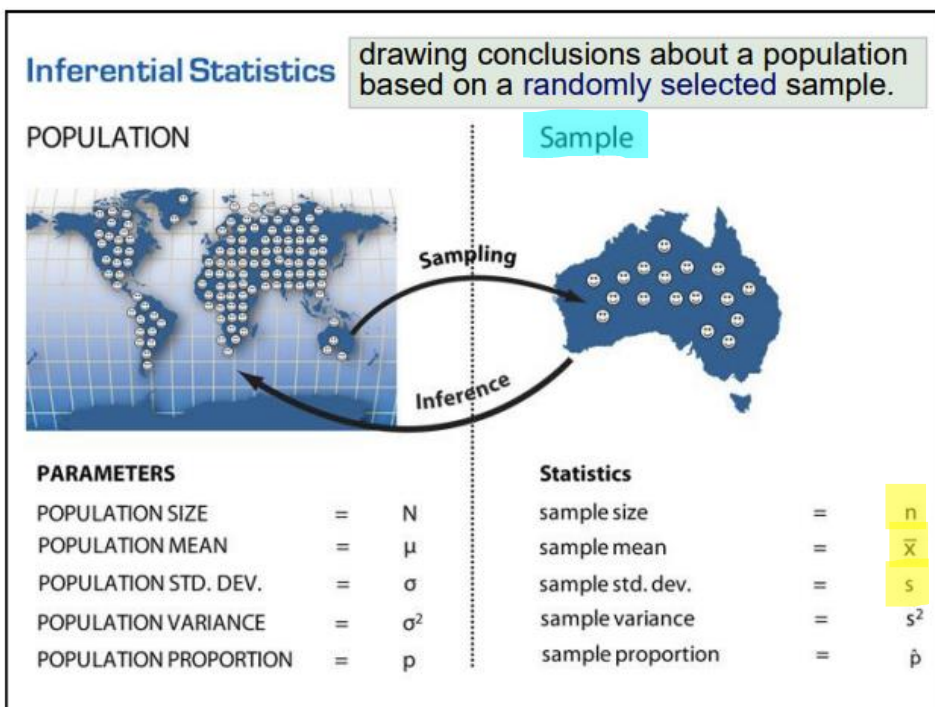
4. What symbol would you give to the value 99% confidence interval? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?

(Poll)



1. What symbol would you give to the random sample of 70? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☒ n

2. What symbol would you give to the value 175.9 square metres? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☒  $\bar{x}$  (x bar)
- ☐ n

3. What symbol would you give to the value 38 square metres? (Single Choice) \*

- ☐ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☒ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

4. What symbol would you give to the value 99% confidence interval? (Single Choice) \*

- ☒ Level of Confidence (LOC)
- ☐  $\sigma$  (sigma)
- ☐ s
- ☐  $\mu$  (mu)
- ☐  $\bar{x}$  (x bar)
- ☐ n

- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.
- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
  - Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☐ Population Mean (Pelican) ( $\sigma$  is known)
- ☐ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☐ Population Proportion (Freaky fish) (proportion)

2. What table will we use? (Single Choice) \*

- ☐ Z table
- ☐ t table

3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

- ☐ 1%
- ☐ 5%
- ☐ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%

(Poll)



**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



1. What type of problem is it? (Single Choice) \*



- ☐ Population Mean (Seagull ) (no sample)
- ☐ Population Mean (Pelican) ( $\sigma$  is known)
- ☒ Population Mean (Shag) ( $\sigma$  is unknown but  $s$  is known)
- ☐ Population Proportion (Freaky fish) (proportion)

(Poll)

2. What table will we use? (Single Choice) \*

- ☐ Z table
- ☒ t table

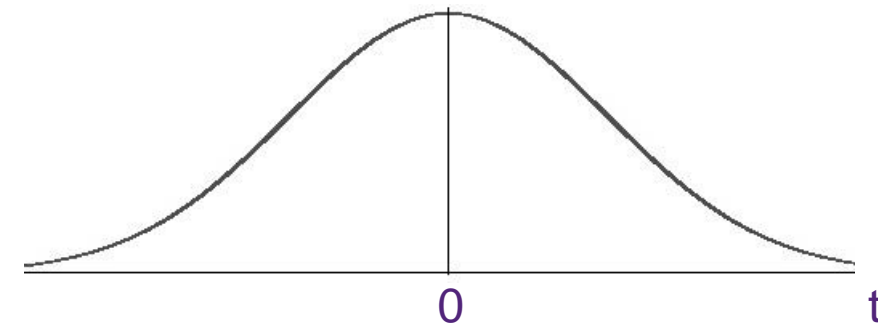
3. What is the value of  $\alpha$  (alpha)? (Single Choice) \*

- ☒ 1%
- ☐ 5%
- ☐ 10%
- ☐ 90%
- ☐ 95%
- ☐ 99%

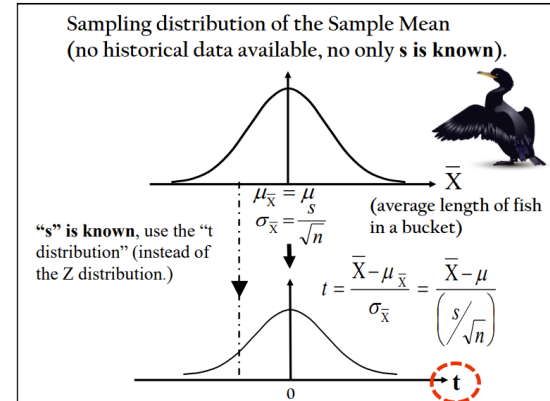


**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$



**Confidence Interval Estimate for  $\mu$ , ( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

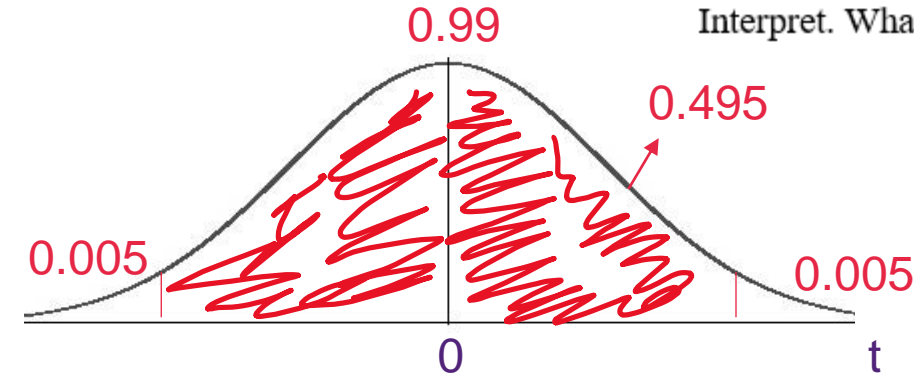
**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

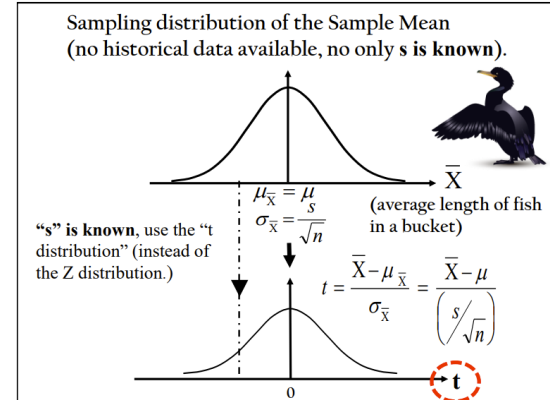
- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



n = 70 family houses  
 $\bar{X}$  = 175.9 square metres  
s = 38 square metres  
LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{crit} = t_{\alpha/2, df} = ?$



**Confidence Interval Estimate for  $\mu$ ,  
( $\sigma$  unknown, and only have s).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

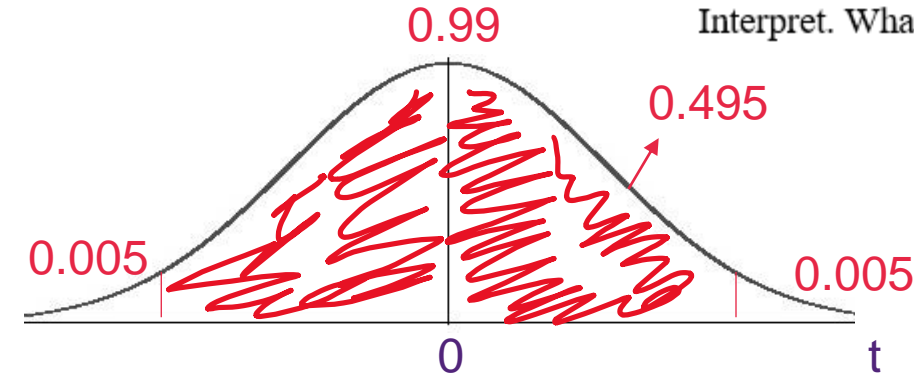
**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- n - 1 degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

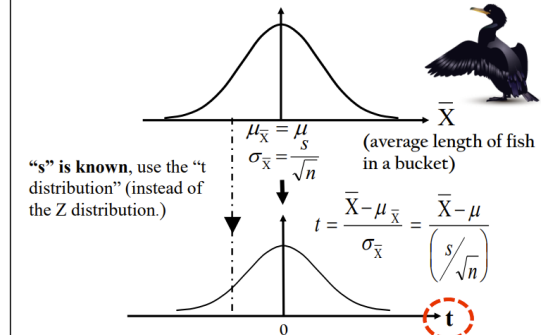
**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{\alpha/2, df} = t_{0.01/2, 70-1} = t_{0.005, 69}$

Sampling distribution of the Sample Mean  
(no historical data available, no only  $s$  is known).



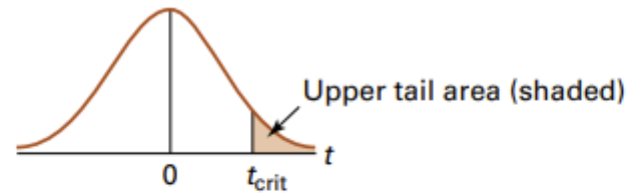
**Confidence Interval Estimate for  $\mu$ ,  
( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

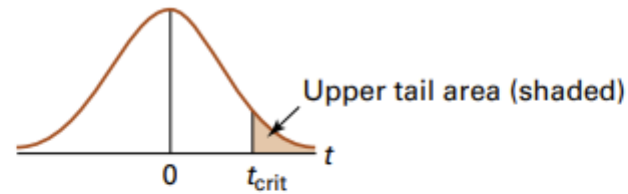
where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the  $t$  distribution with:

- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- $t$  distribution assumptions must be satisfied



df	Upper tail areas					
	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$
51	1.298	1.675	2.008	2.402	2.676	3.258
52	1.298	1.675	2.007	2.400	2.674	3.255
53	1.298	1.674	2.006	2.399	2.672	3.251
54	1.297	1.674	2.005	2.397	2.670	3.248
55	1.297	1.673	2.004	2.396	2.668	3.245
56	1.297	1.673	2.003	2.395	2.667	3.242
57	1.297	1.672	2.002	2.394	2.665	3.239
58	1.296	1.672	2.002	2.392	2.663	3.237
59	1.296	1.671	2.001	2.391	2.662	3.234
60	1.296	1.671	2.000	2.390	2.660	3.232
61	1.296	1.670	2.000	2.389	2.659	3.229
62	1.295	1.670	1.999	2.388	2.657	3.227
63	1.295	1.669	1.998	2.387	2.656	3.225
64	1.295	1.669	1.998	2.386	2.655	3.223
65	1.295	1.669	1.997	2.385	2.654	3.220
66	1.295	1.668	1.997	2.384	2.652	3.218
67	1.294	1.668	1.996	2.383	2.651	3.216
68	1.294	1.668	1.995	2.382	2.650	3.214
69	1.294	1.667	1.995	2.382	2.649	3.213
70	1.294	1.667	1.994	2.381	2.648	3.211
71	1.294	1.667	1.994	2.380	2.647	3.209
72	1.293	1.666	1.993	2.379	2.646	3.207
73	1.293	1.666	1.993	2.379	2.645	3.206
74	1.293	1.666	1.993	2.378	2.644	3.204
75	1.293	1.665	1.992	2.377	2.643	3.202

 $t_{0.005, 69}$

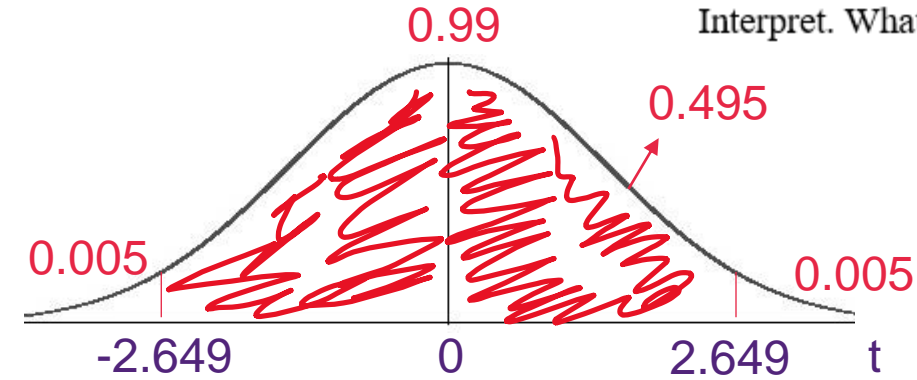

 $t_{0.005, 69}$ 

df	Upper tail areas					
	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$
51	1.298	1.675	2.008	2.402	2.676	3.258
52	1.298	1.675	2.007	2.400	2.674	3.255
53	1.298	1.674	2.006	2.399	2.672	3.251
54	1.297	1.674	2.005	2.397	2.670	3.248
55	1.297	1.673	2.004	2.396	2.668	3.245
56	1.297	1.673	2.003	2.395	2.667	3.242
57	1.297	1.672	2.002	2.394	2.665	3.239
58	1.296	1.672	2.002	2.392	2.663	3.237
59	1.296	1.671	2.001	2.391	2.662	3.234
60	1.296	1.671	2.000	2.390	2.660	3.232
61	1.296	1.670	2.000	2.389	2.659	3.229
62	1.295	1.670	1.999	2.388	2.657	3.227
63	1.295	1.669	1.998	2.387	2.656	3.225
64	1.295	1.669	1.998	2.386	2.655	3.223
65	1.295	1.669	1.997	2.385	2.654	3.220
66	1.295	1.668	1.997	2.384	2.652	3.218
67	1.294	1.668	1.996	2.383	2.651	3.216
68	1.294	1.668	1.995	2.382	2.650	3.214
69	1.294	1.667	1.995	2.382	2.649	3.213
70	1.294	1.667	1.994	2.381	2.648	3.211
71	1.294	1.667	1.994	2.380	2.647	3.209
72	1.293	1.666	1.993	2.379	2.646	3.207
73	1.293	1.666	1.993	2.379	2.645	3.206
74	1.293	1.666	1.993	2.378	2.644	3.204
75	1.293	1.665	1.992	2.377	2.643	3.202



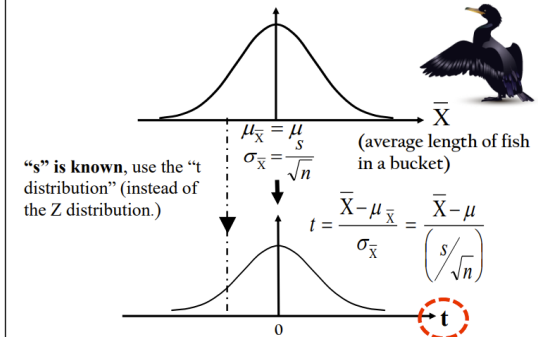
**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{\alpha/2, df} = t_{0.01/2, 70-1} = t_{0.005, 69} = 2.649$

Sampling distribution of the Sample Mean  
(no historical data available, no only  $s$  is known).



**Confidence Interval Estimate for  $\mu$ ,  
( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

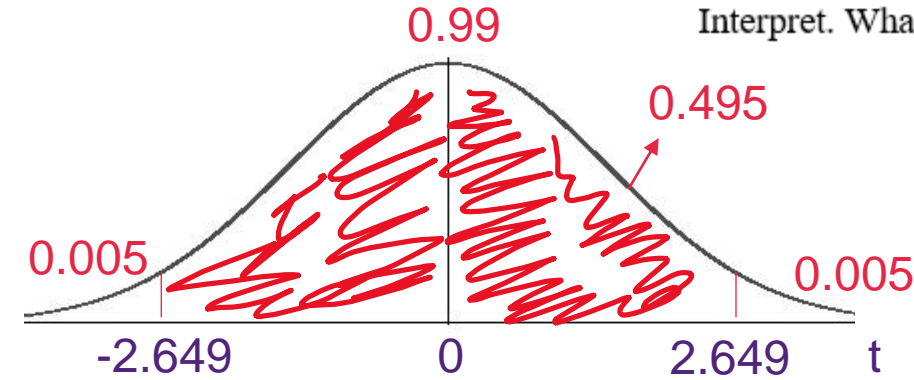
**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the  $t$  distribution with:

- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- $t$  distribution assumptions must be satisfied

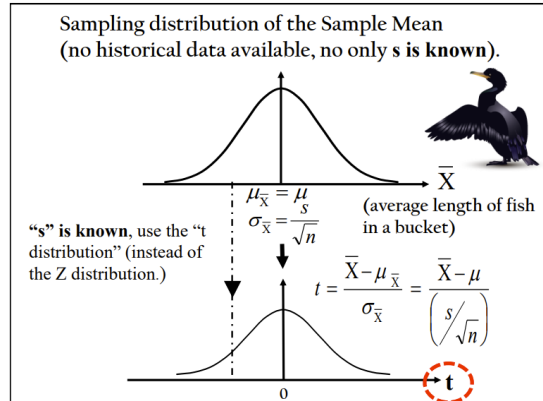
**Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{\alpha/2, df} = t_{0.01/2, 70-1} = t_{0.005, 69} = 2.649$

$$\bar{X} \pm t_{\alpha/2, df} * s_{\bar{X}} = \bar{X} \pm t_{\alpha/2, df} * \frac{s}{\sqrt{n}} = ?$$



**Confidence Interval Estimate for  $\mu$ , ( $\sigma$  unknown, and only have  $s$ ).**

Lower limit:  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

Upper limit:  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

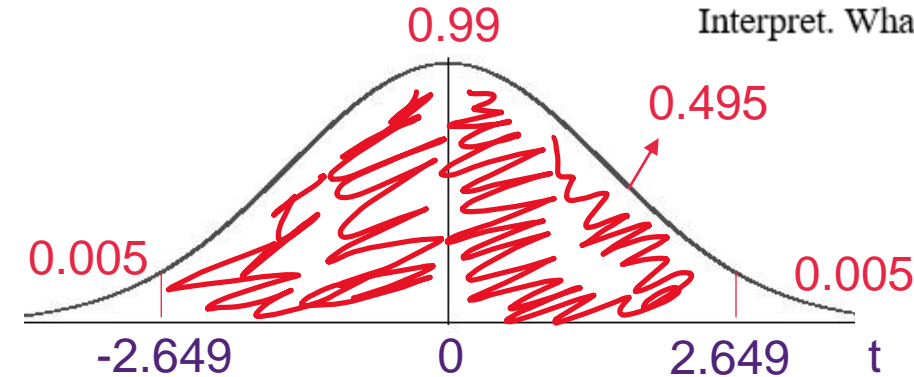
where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

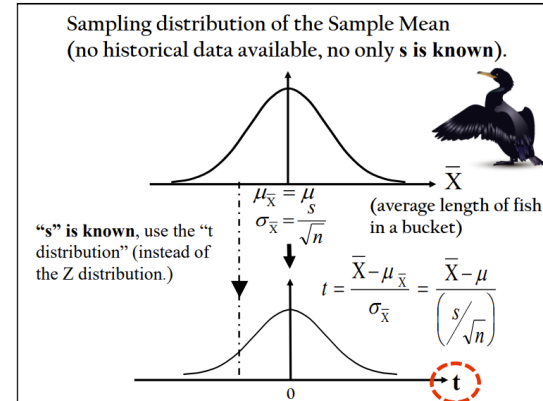
- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{\alpha/2, df} = t_{0.01/2, 70-1} = t_{0.005, 69} = 2.649$

$$\bar{X} \pm t_{\alpha/2, df} * s_{\bar{X}} = \bar{X} \pm t_{\alpha/2, df} * \frac{s}{\sqrt{n}} = 175.9 \pm 2.649 * \frac{38}{\sqrt{70}} =$$

$$163.8686 < \mu < 187.9315$$



**Confidence Interval Estimate for  $\mu$ , ( $\sigma$  unknown, and only have  $s$ ).**

Lower limit:  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

Upper limit:  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

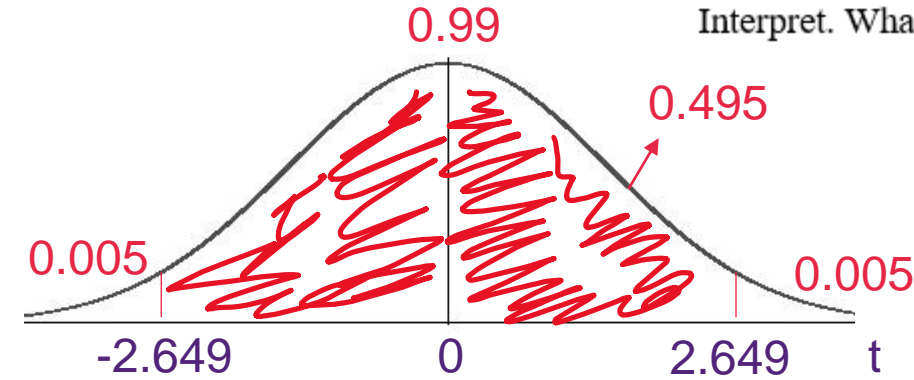
where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?

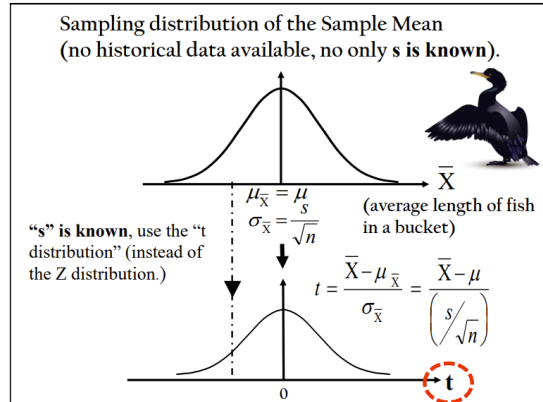


$n = 70$  family houses  
 $\bar{X} = 175.9$  square metres  
 $s = 38$  square metres  
 LOC = 99%  
 $\alpha = 1 - 0.99 = 0.01$   
 $t_{\alpha/2, df} = t_{0.01/2, 70-1} = t_{0.005, 69} = 2.649$

$$\bar{X} \pm t_{\alpha/2, df} * s_{\bar{X}} = \bar{X} \pm t_{\alpha/2, df} * \frac{s}{\sqrt{n}} = 175.9 \pm 2.649 * \frac{38}{\sqrt{70}} =$$

$$163.8686 < \mu < 187.9315$$

Based on the sample mean, the average size of single family houses is estimated with 99% confidence to be between 163.9 and 197.9 square metres.



**Confidence Interval Estimate for  $\mu$ ,  
 ( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- $n-1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

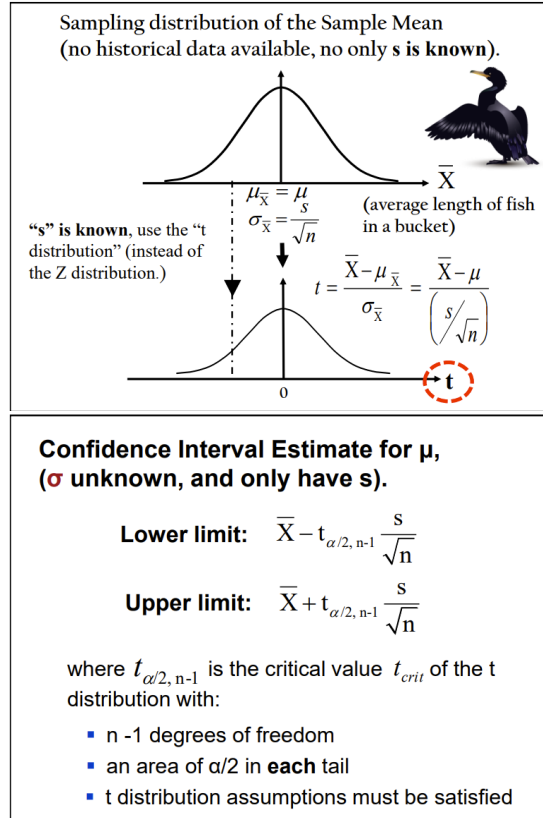
- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?

### Assumptions:

- House size of single family houses (variable) is normally distributed or at least not highly skewed.
- This is necessary since we're using t distribution (only s is known). Which requires sample to come from a normal distribution if sample size is small.
- More robust when sample size is large provided the population is not highly skewed.





- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

- Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?

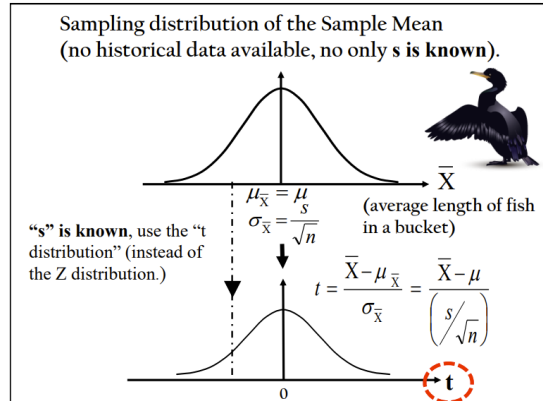


### Assumptions:

- House size of single family houses (variable) is normally distributed or at least not highly skewed.
- This is necessary since we're using t distribution (only s is known). Which requires sample to come from a normal distribution if sample size is small.
- More robust when sample size is large provided the population is not highly skewed.

### Increasing sample size ( $n \uparrow$ )

- Sampling error decreases ( $s_{\bar{X}} \downarrow$ ).
- Interval estimate more precise.
- CI width would be narrower.



### Confidence Interval Estimate for $\mu$ , ( $\sigma$ unknown, and only have s).

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- n - 1 degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

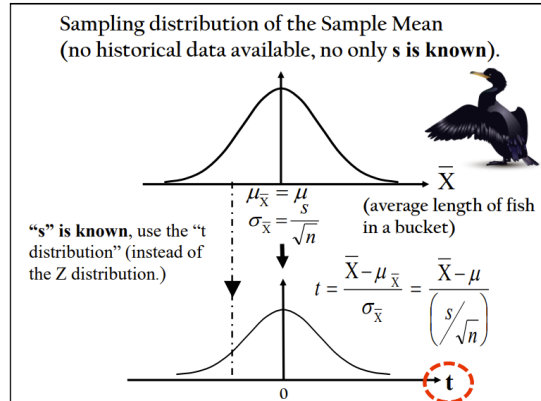
- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
LOC = 95%  
 $\alpha = 1 - 0.95 = 0.05$



**Confidence Interval Estimate for  $\mu$ ,  
( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

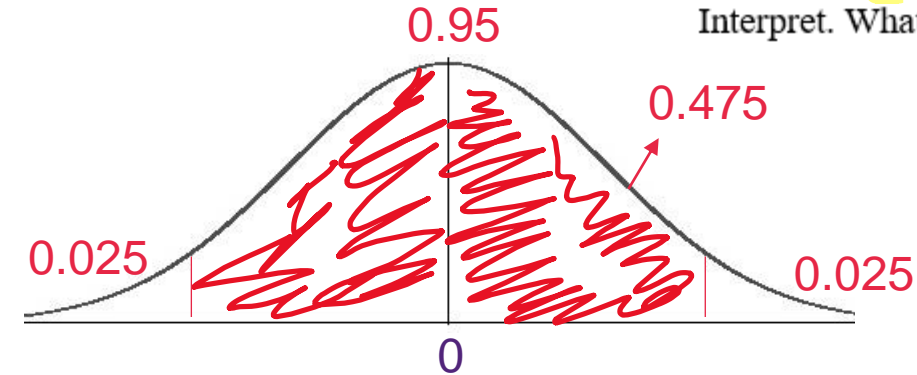
where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in **each** tail
- t distribution assumptions must be satisfied

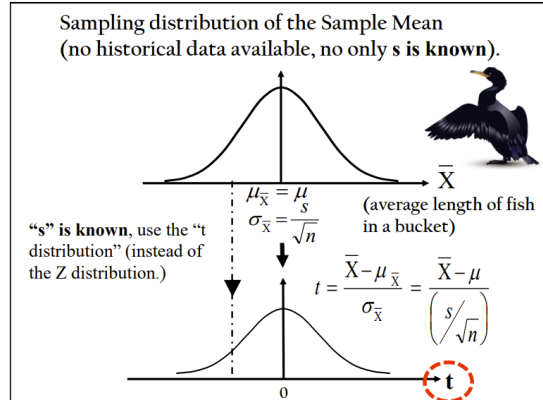
- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 $LOC = 95\%$   
 $\alpha = 1 - 0.95 = 0.05$   
 Air conditioning?



**Confidence Interval Estimate for  $\mu$ , ( $\sigma$  unknown, and only have  $s$ ).**

**Lower limit:**  $\bar{X} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

**Upper limit:**  $\bar{X} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

where  $t_{\alpha/2, n-1}$  is the critical value  $t_{crit}$  of the t distribution with:

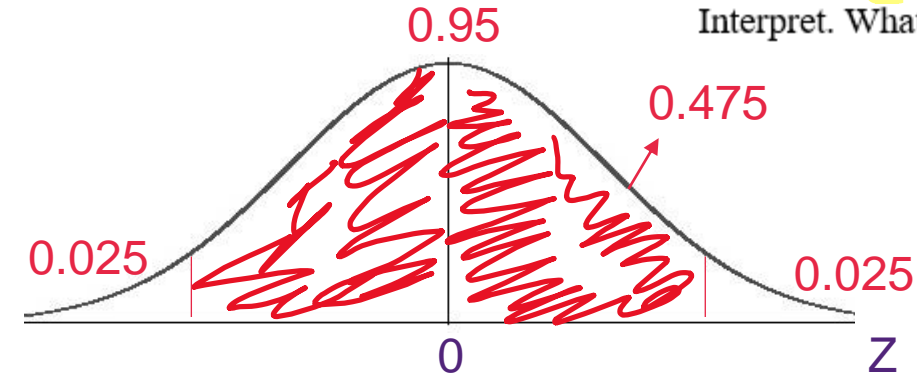
- $n - 1$  degrees of freedom
- an area of  $\alpha/2$  in each tail
- t distribution assumptions must be satisfied

- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.



$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses  
 LOC = 95%  
 $\alpha = 1 - 0.95 = 0.05$   
 $\hat{p} = ?$

The confidence interval limits for a population proportion are:

**Lower limit:**  $\hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**Upper limit:**  $\hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

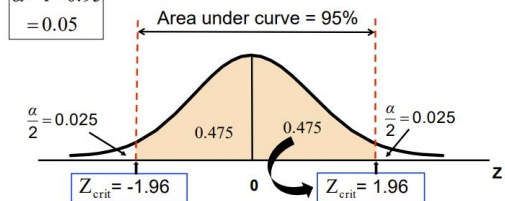
$n$  = sample size



Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$

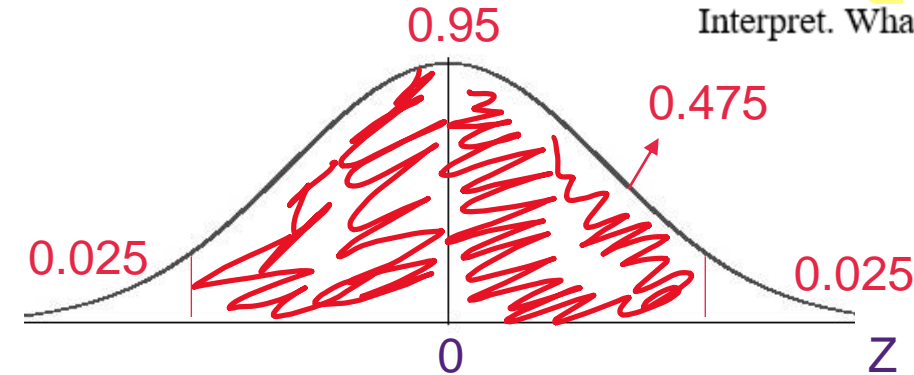


- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.



$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses

LOC = 95%

$$\alpha = 1 - 0.95 = 0.05$$

$$\hat{p} = 42/70 = 0.6$$

$$Z_{crit} = ?$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

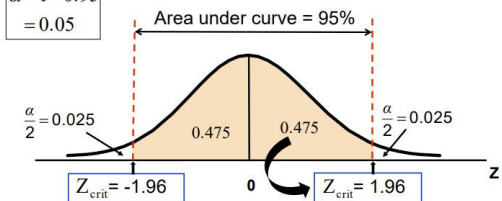
$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval

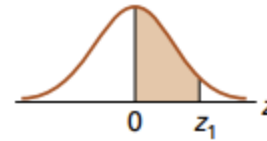
$$\alpha = 1 - 0.95 = 0.05$$





**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).

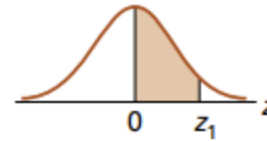


0.475

$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767

**TABLE A.5 Areas of the standard normal distribution  $\mu = 0, \sigma = 1$**

The entries in this table are the probabilities that a standard normal random variable is between 0 and  $z_1$  (the shaded area).



$z_1$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767

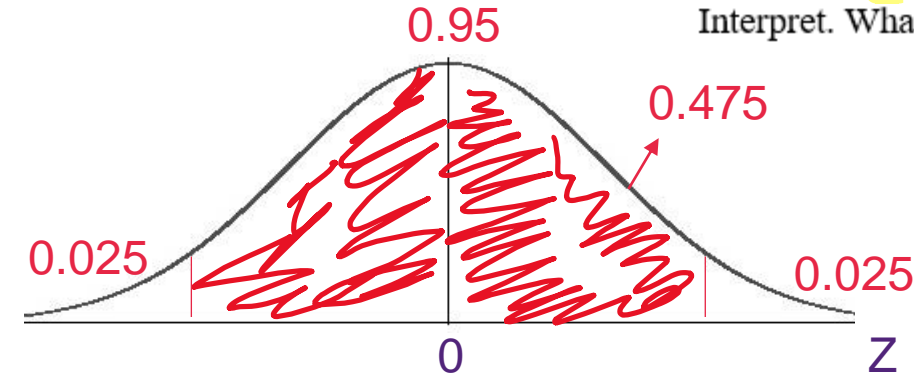
0.475

- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.



$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses

LOC = 95%

$$\alpha = 1 - 0.95 = 0.05$$

$$\hat{p} = 42/70 = 0.6$$

$$Z_{crit} = 1.96$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

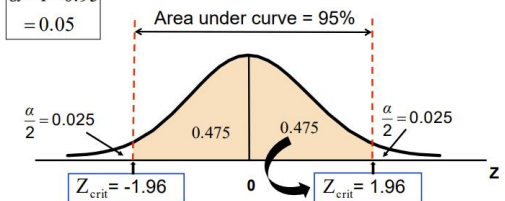
$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$

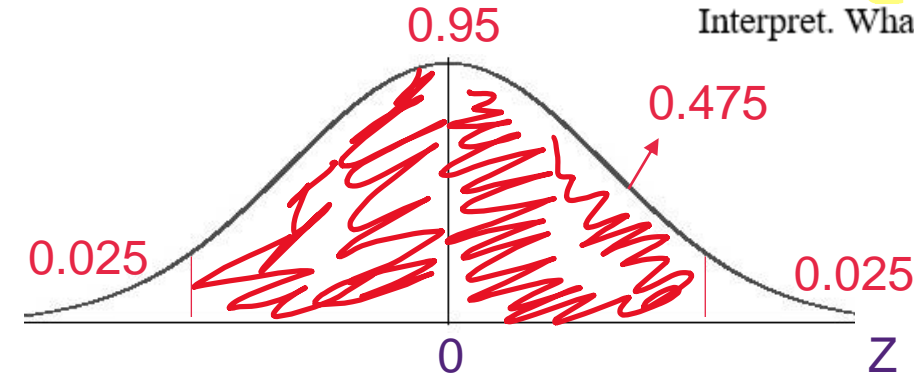


- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.



$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses

LOC = 95%

$$\alpha = 1 - 0.95 = 0.05$$

$$\hat{p} = 42/70 = 0.6$$

$$Z_{crit} = 1.96$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = ?$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

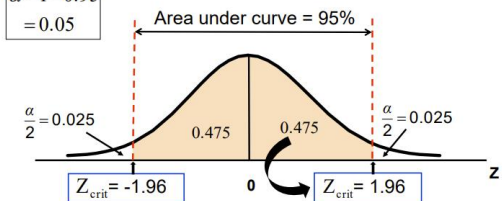
$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$





- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

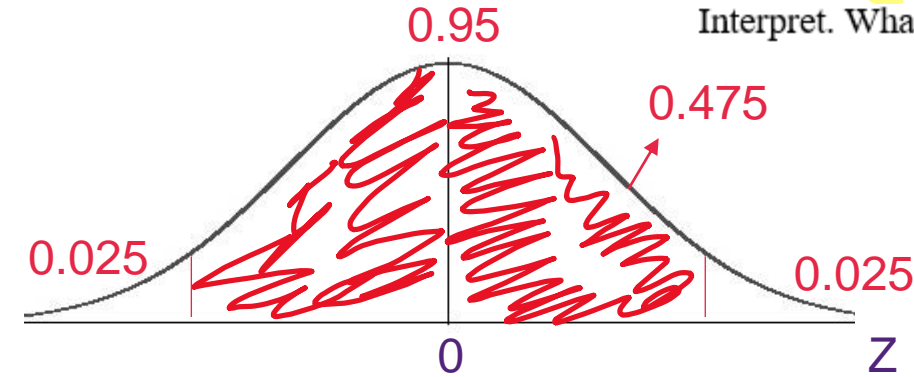


$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?

$$0.4852 < p < 0.7148$$

- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses

LOC = 95%

$$\alpha = 1 - 0.95 = 0.05$$

$$\hat{p} = 42/70 = 0.6$$

$$Z_{crit} = 1.96$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.96 * \sqrt{\frac{0.6(1-0.6)}{70}} =$$

$$0.4852 < p < 0.7148$$

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

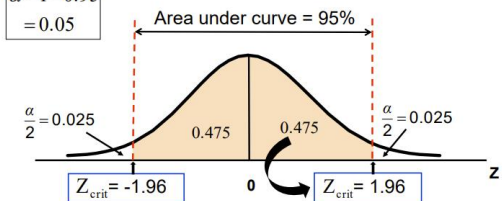
$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$





- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

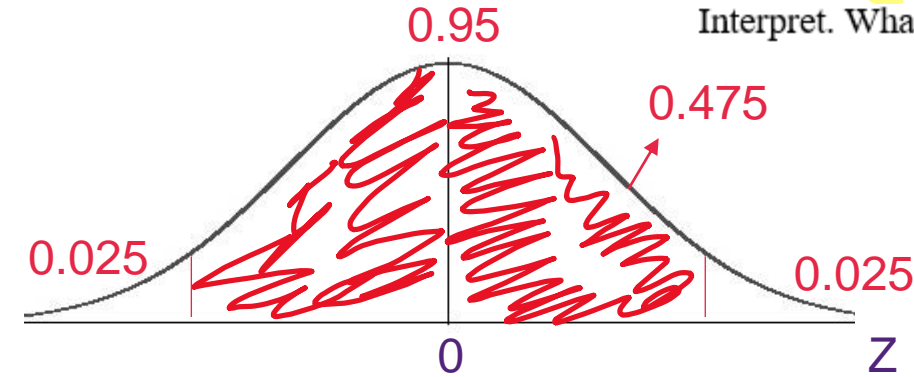


$$163.8686 < \mu < 187.9315$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?

$$0.4852 < p < 0.7148$$

- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an increase in the confidence level have on the estimate?



$n = 70$  family houses

LOC = 95%

$$\alpha = 1 - 0.95 = 0.05$$

$$\hat{p} = 42/70 = 0.6$$

$$Z_{crit} = 1.96$$

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}} = \hat{p} \pm Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.6 \pm 1.96 * \sqrt{\frac{0.6(1-0.6)}{70}} =$$

$$0.4852 < p < 0.7148$$

Based on the sample, with 95% confidence, the proportion of houses with air conditioning is estimated to be between 48.5% and 71.5%.

The confidence interval limits for a population proportion are:

$$\text{Lower limit: } \hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{Upper limit: } \hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$Z_{crit}$  = critical value of Z for the level of confidence

$\hat{p}$  = is the sample proportion

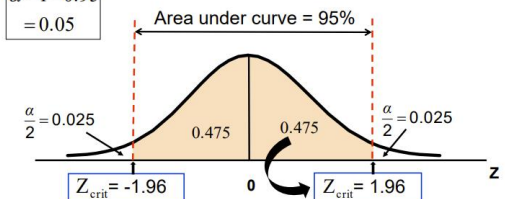
$n$  = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval

$$\alpha = 1 - 0.95 = 0.05$$



- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

$$0.4852 < p < 0.7148$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an **increase in the confidence level** have on the estimate?



The confidence interval limits for a population proportion are:

**Lower limit:**  $\hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**Upper limit:**  $\hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

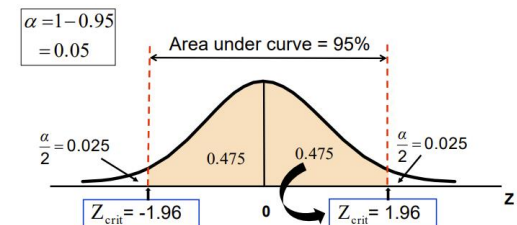
$\hat{p}$  = is the sample proportion

$n$  = sample size



Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval



- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

$$0.4852 < p < 0.7148$$

$$\text{LOC} \uparrow$$

$$\hat{p} \pm Z_{\text{crit}} * \sigma_{\hat{p}}$$

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an **increase in the confidence level** have on the estimate?



The confidence interval limits for a population proportion are:

Lower limit:  $\hat{p} - Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Upper limit:  $\hat{p} + Z_{\text{crit}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{\text{crit}}$  = critical value of Z for the level of confidence

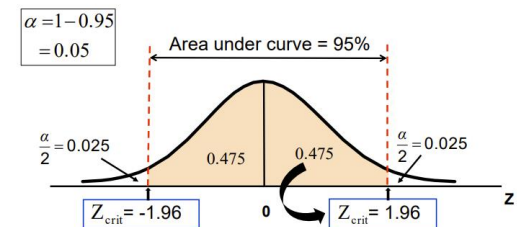
$\hat{p}$  = is the sample proportion

$n$  = sample size



Finding the Critical Value,  $Z_{\text{crit}}$

Consider a 95% confidence interval



- Q4.** The real estate assessor for a local government is studying various characteristics of single-family houses. A random sample of 70 reveals: average house area = 175.9 square metres with a standard deviation of 38.0 square metres, and 42 have air conditioning.

$$163.8686 < \mu < 187.9315$$

$$0.4852 < p < 0.7148$$

LOC↑

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}}$$

$Z_{crit}$  ↑

Width of CI ↑ → less precise

- a) Find the 99% confidence interval for the average house size. Interpret this confidence interval. What assumption has been made? What effect would increasing the sample size have on the precision of this interval?
- b) Find the 95% confidence interval for the proportion of houses that have air conditioning. Interpret. What effect would an **increase in the confidence level** have on the estimate?



The confidence interval limits for a population proportion are:

Lower limit:  $\hat{p} - Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Upper limit:  $\hat{p} + Z_{crit} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z_{crit}$  = critical value of Z for the level of confidence

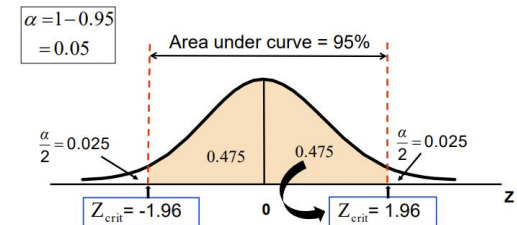
$\hat{p}$  = is the sample proportion

n = sample size



Finding the Critical Value,  $Z_{crit}$

Consider a 95% confidence interval



**Q5. True or False? Why?**

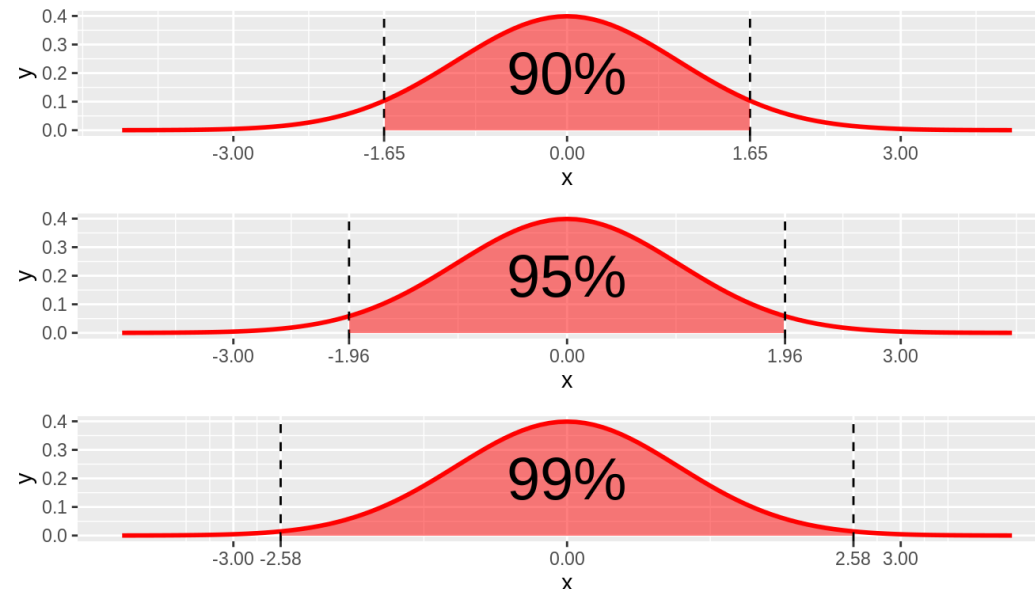
- (i) As the width of an estimating interval increases, the degree of confidence in it actually containing the population parameter being estimated also increases.
- (ii) Other things being equal, the confidence interval for the mean will be wider for 95% than for 90% confidence.
- (iii) If the 95% confidence interval for the mean was found to be from 3.7 to 9.2 then  $\underline{P}(3.7 < \bar{x} < 9.2) = 0.95$ .
- (iv) A 90% confidence interval built to estimate the population mean does not mean there is a 90% probability that the population mean lies within the interval.



## Q5. True or False? Why?

True

- (i) As the width of an estimating interval increases, the degree of confidence in it actually containing the population parameter being estimated also increases.
- (ii) Other things being equal, the confidence interval for the mean will be wider for 95% than for 90% confidence.
- (iii) If the 95% confidence interval for the mean was found to be from 3.7 to 9.2 then  $P(3.7 < \bar{x} < 9.2) = 0.95$ .
- (iv) A 90% confidence interval built to estimate the population mean does not mean there is a 90% probability that the population mean lies within the interval.



Source: **Introduction to Statistics and Data Science**  
*A modern dive into R and the tidyverse*  
Chester Ismay, Albert Y. Kim, Arend M. Kuyper, Elizabeth Tipton, and Kaitlyn G. Fitzgerald

[https://nulib.github.io/moderndive\\_book/](https://nulib.github.io/moderndive_book/)

## Q5. True or False? Why?

- True (i) As the width of an estimating interval increases, the degree of confidence in it actually containing the population parameter being estimated also increases.
- True (ii) Other things being equal, the confidence interval for the mean will be wider for 95% than for 90% confidence.
- (iii) If the 95% confidence interval for the mean was found to be from 3.7 to 9.2 then  $\underline{P}(3.7 < \bar{x} < 9.2) = 0.95$ .
- (iv) A 90% confidence interval built to estimate the population mean does not mean there is a 90% probability that the population mean lies within the interval.

LOC↑

$$\hat{p} \pm Z_{crit} * \sigma_{\hat{p}}$$

 $Z_{crit} \uparrow$ 

Width of CI ↑ → less precise

**Q5. True or False? Why?**

- True (i) As the width of an estimating interval increases, the degree of confidence in it actually containing the population parameter being estimated also increases.
- True (ii) Other things being equal, the confidence interval for the mean will be wider for 95% than for 90% confidence.
- False (iii) If the 95% confidence interval for the mean was found to be from 3.7 to 9.2 then  $P(3.7 < \bar{x} < 9.2) = 0.95$ .
- (iv) A 90% confidence interval built to estimate the population mean does not mean there is a 90% probability that the population mean lies within the interval.

Confidence Interval  $\neq$  Probability

**Q5. True or False? Why?**

- True (i) As the width of an estimating interval increases, the degree of confidence in it actually containing the population parameter being estimated also increases.
- True (ii) Other things being equal, the confidence interval for the mean will be wider for 95% than for 90% confidence.
- False (iii) If the 95% confidence interval for the mean was found to be from 3.7 to 9.2 then  $\underline{P}(3.7 < \bar{x} < 9.2) = 0.95$ .
- True (iv) A 90% confidence interval built to estimate the population mean does **not** mean there is a 90% probability that the population mean lies within the interval.

Confidence Interval  $\neq$  Probability

## ECON1310

### Tutorial 7 – Week 8

#### CONFIDENCE INTERVALS I

At the end of this tutorial you should be able to

- Describe the difference between a point estimate and interval estimate,
- Determine when it is appropriate to use the  $Z$  statistic for interval estimation and when it is appropriate to use the  $t$  statistic,
- Use the  $t$  distribution tables,
- Calculate confidence intervals for population means using  $Z$  statistics and  $t$  statistics,
- Calculate confidence intervals for population proportions.





THE UNIVERSITY  
OF QUEENSLAND  
AUSTRALIA

CREATE CHANGE

# Thank you

## Francisco Tavares Garcia

Academic Tutor | School of Economics

[tavaresgarcia.github.io](https://tavaresgarcia.github.io)

### Reference

Black et al. (2016), Australasian Business Statistics, 4th Edition, Wiley Australia.