Sec “9”

Confidence Interval (CI)

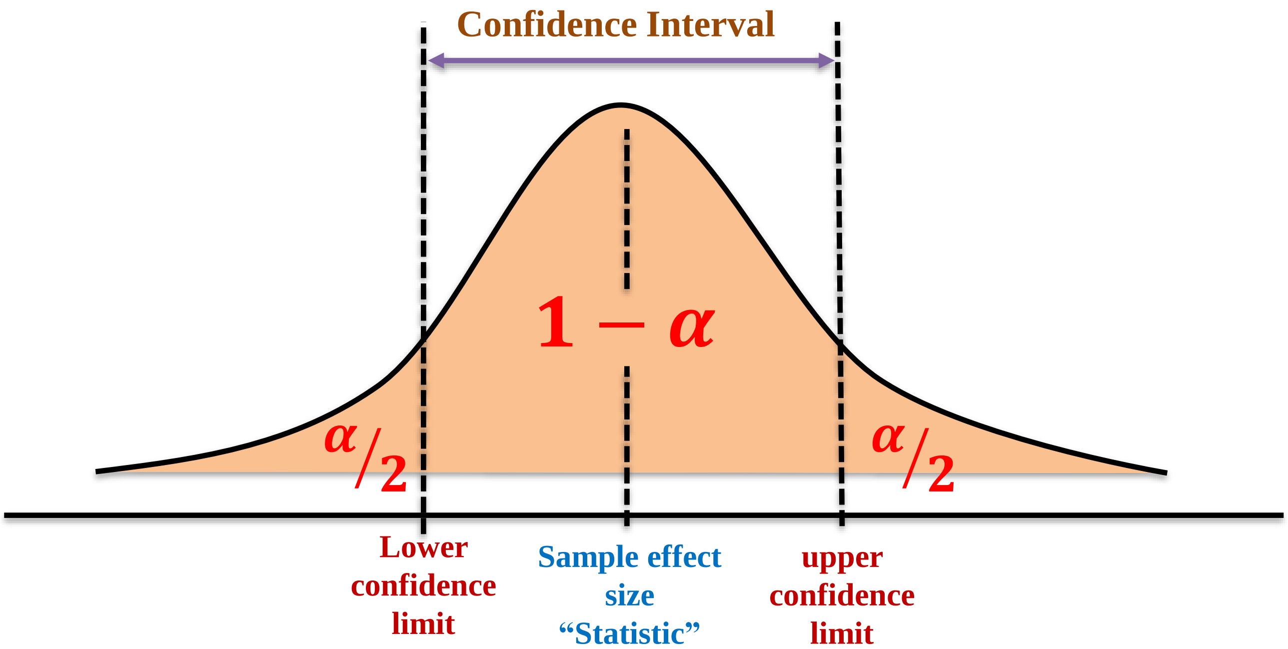
(Single mean, Single proportion)

Confidence Intervals (CI)

- A confidence interval, in statistics, refers to the probability that a population parameter will fall between two set values for a certain proportion of times.

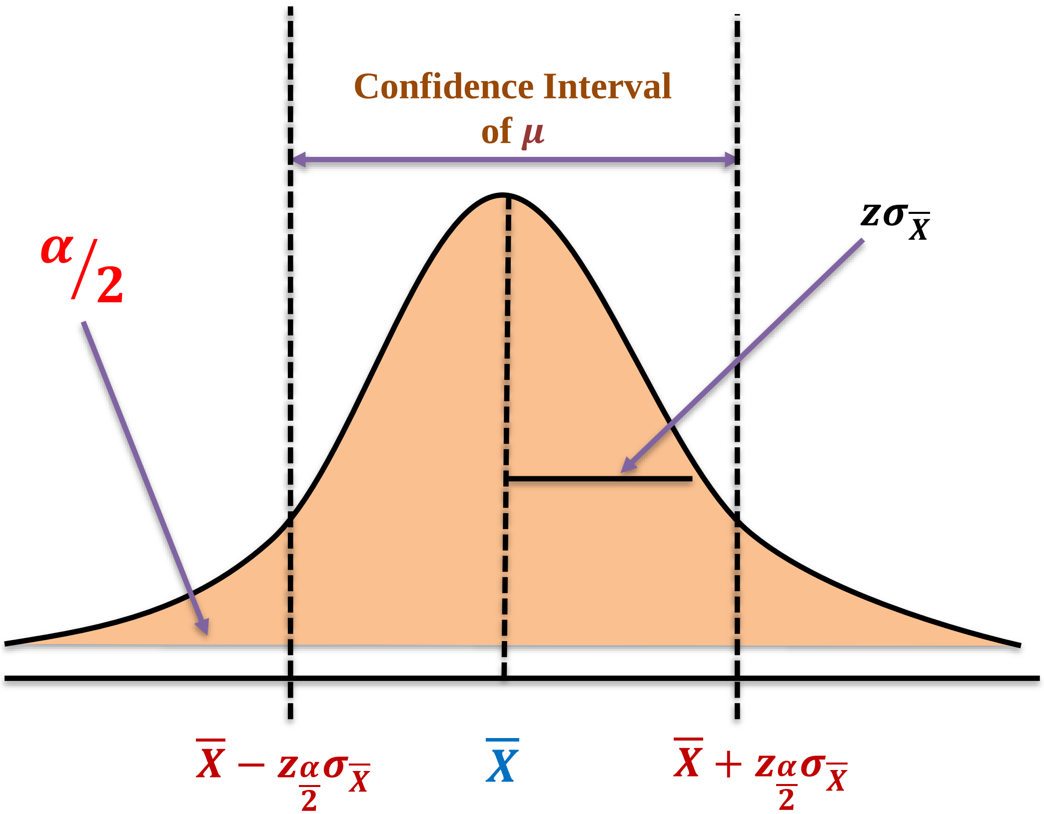
$$i.e., a < \mu < b$$

- The estimated range being calculated from a given set of sample data.
- It is often expressed a % whereby a population means lies between an upper and lower interval.
- A confidence interval can take any number of probabilities, with the most common being a 95% or 99% confidence level "α".



Confidence Interval of the Population Mean "μ"

- A confidence interval for a mean gives us a range of plausible values for the population mean.



- The Lower limit of the confidence interval:

$$\overline{X} - z_{\frac{\alpha}{2}}\sigma_{\overline{X}}$$

- The upper limit of the confidence interval:

$$\overline{X} + z_{\frac{\alpha}{2}}\sigma_{\overline{X}}$$

$\therefore$ The CI of μ is given by $\frac{\sigma}{\sqrt{n}}$

$$\overline{X} \pm z_{\frac{\alpha}{2}}\sigma_{\overline{X}}$$

Also can be written as follows

$$\overline{X} - z_{\frac{\alpha}{2}}\sigma_{\overline{X}} \le \mu \le \overline{X} + z_{\frac{\alpha}{2}}\sigma_{\overline{X}}$$

❑ How to Calculate a Confidence Interval for a Population Mean?

Step 1: Find the number of observations $\boldsymbol{n}$, calculate their sample mean $\overline{\boldsymbol{X}} = \frac{\sum_{i=1}^{n} x_i}{n}$, and sample standard deviation $\boldsymbol{s} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n-1}}$, if the population's standard deviation is unknown.

- **Step 2: D**ecide what Confidence Interval we want. i.e., $(1-\alpha)\%$ confidence interval for μ "The area in which the population mean μ is located". $[(1-\alpha)\%] \rightarrow$ given in examples$\rightarrow \alpha \rightarrow \frac{\alpha}{2}$

- **Step 3:**

Population Standard Deviation "σ"

If σ known

Use: Z- table.

CI: $\boldsymbol{\overline{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}$

If σ unknown$\rightarrow$ s known

If $n \geq 30$

Use: Z- table.

CI: $\boldsymbol{\overline{X} \pm z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}}$

If $n < 30$

Use: **t-** table.

CI: $\boldsymbol{\overline{X} \pm t_{\frac{\alpha}{2},v} \frac{s}{\sqrt{n}}}$,

$\boldsymbol{v = n - 1}$

Exercise

(1) An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with a standard deviation of 40 hours. If a sample of 30 bulbs has an average life of 780 hours, find a 96% confidence interval for the population mean of all bulbs produced by this firm.

Solution:

- **Step 1:** $n = 30, \quad \overline{X} = 780, \quad \sigma = 40.$
- **Step 2:** $(1-\alpha)\% = 96\% \rightarrow 1-\alpha = 0.96 \rightarrow \alpha = 1 - 0.96$

$$\alpha = 0.04$$

- **Step 3:** $\sigma \rightarrow$ known, so we use z-table

$z_{\frac{\alpha}{2}} = z_{\frac{0.04}{2}} = z_{0.02} = 2.05$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244

$0.0202 - 0.02 = 0.0002$ ✓✓

$0.02 - 0.0197 = 0.0003$

CI of μ, with known σ is given by

$$\overline{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 780 \pm (2.05)\left(\frac{40}{\sqrt{30}}\right)$$

$= 780 \pm 14.9711 \rightarrow$

$780 - 14.9711 \leq \mu \leq 780 + 14.9711 \rightarrow 765.0289 \leq \mu \leq 794.971$

$$\therefore 765 \leq \mu \leq 795$$

(2) Many cardiac patients wear an implanted pacemaker to control their heartbeat. A plastic connector module mounts on the top of the pacemaker. Assuming a standard deviation of 0.0015 inch and an approximately normal distribution, find a 95% confidence interval for the mean of the depths of all connector modules made by a certain manufacturing company. A random sample of 75 modules has an average depth of 0.310 inch.

Solution:

- **Step 1:** $n = 75, \quad \overline{X} = 0.310, \quad \sigma = 0.0015.$
- **Step 2:** $(1-\alpha)\% = 95\% \rightarrow 1-\alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\boxed{\alpha = 0.05}$$

- **Step 3:** $\sigma \rightarrow$ known, so we use z-table

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244

$$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$$

CI of μ, with known σ is given by

$$\overline{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 0.310 \pm (1.96)\left(\frac{0.0015}{\sqrt{75}}\right)$$

$$= 0.310 \pm 0.000173 \rightarrow$$

$$0.310 - 0.000173 \leq \mu \leq 0.310 + 0.000173$$

$$\therefore\ 0.3098 \leq \mu \leq 0.3102$$

(3) A machine produces metal pieces that are cylindrical in shape. A sample of pieces is taken, and the diameters are found to be 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01, and 1.03 centimeters. Find a 99% confidence interval for the mean diameter of pieces from this machine, assuming an approximately normal distribution.

Solution:

- **Step 1:** $n = 9, \quad \overline{X} = \frac{1.01+\cdots+1.03}{9} = 1.0056,$

$$s = \sqrt{\frac{(1.01-1.0056)^2+\cdots+(1.03-1.0056)^2}{9-1}} = 0.02455.$$

- **Step 2:** $(1-\alpha)\% = 99\% \rightarrow 1-\alpha = 0.99 \rightarrow \alpha = 1-0.99$

$$\boxed{\alpha = 0.01}$$

Step 3: $\sigma \rightarrow$ unknown, $n < 30$, so we use t-table

$t_{\frac{\alpha}{2},v} = t_{\frac{\alpha}{2},n-1} = t_{\frac{0.01}{2},8} = t_{0.005,8} = 3.355$

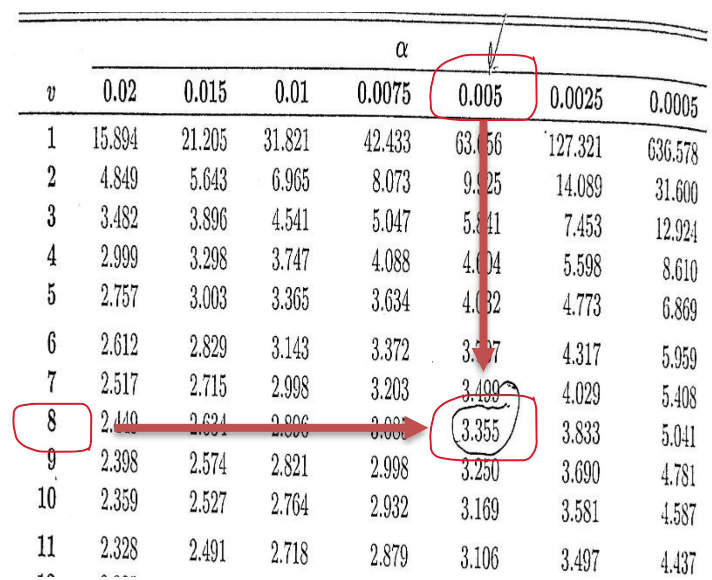
CI of μ, with unknown σ, $n < 30$ is given by

$\overline{X} \pm t_{\frac{\alpha}{2},v}\frac{s}{\sqrt{n}} = 1.0056 \pm (3.355)\left(\frac{0.02455}{\sqrt{9}}\right)$

$= 1.0056 \pm 0.0275 \rightarrow$

$1.0056 - 0.0275 \le \mu \le 1.0056 + 0.0275$

$\therefore 0.9781 \le \mu \le 1.0331$



v	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
1	15.894	21.205	31.821	42.433	[illegible]	127.321	636.578
2	4.849	5.643	6.965	8.073	[illegible]	14.089	31.600
3	3.482	3.896	4.541	5.047	[illegible]	7.453	12.924
4	2.999	3.298	3.747	4.088	[illegible]	5.598	8.610
5	2.757	3.003	3.365	3.634	[illegible]	4.773	6.869
6	2.612	2.829	3.143	3.372	[illegible]	4.317	5.959
7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	[illegible]	[illegible]	[illegible]	[illegible]	3.355	3.833	5.041
9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	2.328	2.491	2.718	2.879	3.106	3.497	4.437

(4) A random sample of 12 graduates of a certain secretarial school typed an average of 79.3 words per minute with a standard deviation of 7.8 words per minute. Assuming a normal distribution for the number of words typed per minute, find a 95% confidence interval for the average number of words typed by all graduates of this school.

Solution:

- **Step 1:** $n = 12, \quad \overline{X} = 79.3, \; s = 7.8.$
- **Step 2:** $(1 - \alpha)\% = 95\% \rightarrow 1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\alpha = 0.05$$

- **Step 3:** $\sigma \rightarrow$ unknown, $n < 30$, so we use t-table

$$t_{\frac{\alpha}{2},v} = t_{\frac{\alpha}{2},n-1} = t_{\frac{0.05}{2},11} = t_{0.025,11} = 2.201$$

v	0.40	0.30	0.20	0.15	0.10	0.05	0.025
				α			
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.[illegible]03
3	0.277	0.584	0.978	1.250	1.638	2.353	3.[illegible]82
4	0.271	0.569	0.941	1.190	1.533	2.132	2.[illegible]76
5	0.267	0.559	0.920	1.156	1.476	2.015	2.[illegible]71
6	0.265	0.553	0.906	1.134	1.440	1.943	2.[illegible]47
7	0.263	0.549	0.896	1.119	1.415	1.895	2.[illegible]65
8	0.262	0.546	0.889	1.108	1.397	1.860	2.[illegible]06
9	0.261	0.543	0.883	1.100	1.383	1.833	2.[illegible]62
10	0.260	0.542	0.879	1.093	1.372	1.812	2.[illegible]28
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
		[illegible]	0.870	1.079	1.350	1.771	2.160

CI of μ, with unknown σ, $n < 30$ is given by

$$\overline{X} \pm t_{\frac{\alpha}{2},v} \frac{s}{\sqrt{n}} = 79.3 \pm (2.201)\left(\frac{7.8}{\sqrt{12}}\right)$$

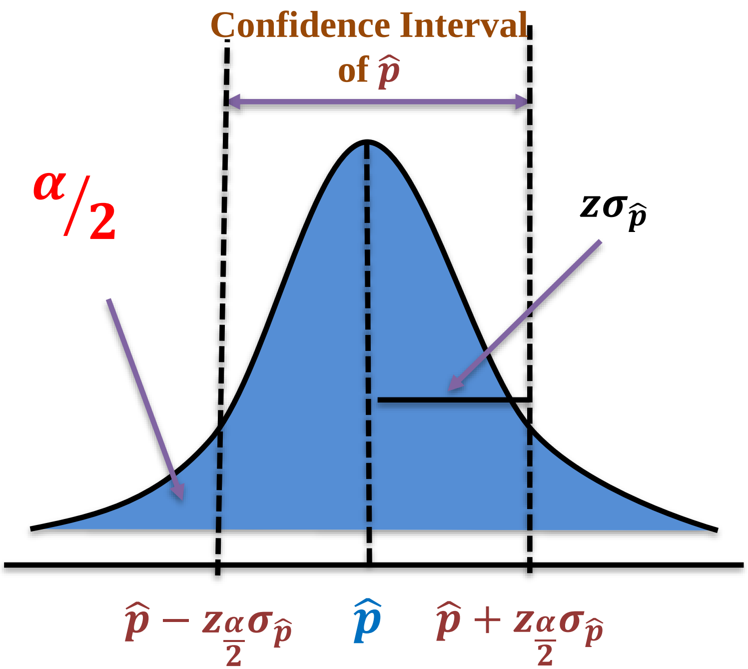
$$= 79.3 \pm 4.95592 \rightarrow$$

$$79.3 - 4.95592 \leq \mu \leq 79.3 + 4.95592$$

$$\therefore 74.34408 \leq \mu \leq 84.25592$$

Confidence Interval of the Population Proportion "p"

- A confidence interval for a proportion gives us a range of plausible values for the population proportion.



- The Lower limit of the confidence interval:

$$\hat{p} - z_{\frac{\alpha}{2}}\sigma_{\hat{p}} = \hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

- The upper limit of the confidence interval:

$$\hat{p} + z_{\frac{\alpha}{2}}\sigma_{\hat{p}} = \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

∴ The CI of p is given by

$$\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Also can be written as follows

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \le p \le \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

❑ How to Calculate a Confidence Interval for a Population Proportion?

Step 1: **F**ind the number of observations $\boldsymbol{n}$, calculate their <u>sample proportion</u> $\widehat{\boldsymbol{p}} = \frac{\boldsymbol{X}}{\boldsymbol{n}}$, where X : denote the number of success, and calculate $\widehat{\boldsymbol{q}} = \mathbf{1} - \widehat{\boldsymbol{p}}$.

Step 2: **D**ecide what Confidence Interval we want. i.e., $(1-\alpha)\%$ confidence interval for p "The area in which the population proportion p is located". $[(1-\alpha)\%] \rightarrow$ given in examples $\alpha \rightarrow \frac{\alpha}{2}$. Compute $z_{\frac{\alpha}{2}}$

Step 3: **T**he confidence interval for a population proportion p is given by:

$$\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Exercise

(1) Compute 95% confidence intervals for the proportion of defective items in a process when it is found that a sample of size 100 yields 8 defectives.

Solution:

- **Step 1:** $n = 100$, X : number of defective items $= 8$,

$$\hat{p} = \frac{X}{n} = \frac{8}{100} = 0.08, \qquad \hat{q} = 1 - \hat{p} = 1 - 0.08 = 0.92$$

- **Step 2:** $(1-\alpha)\% = 95\% \rightarrow 1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95$

$$\alpha = 0.05$$

- **Step 3:**

$$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$$

CI of p, is given by

$$\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.08 \pm (1.96)\sqrt{\frac{(0.08)(0.92)}{100}} = 0.08 \pm 0.0532$$

$$\rightarrow 0.080 - 0.0532 \leq p \leq 0.080 + 0.0532$$

$$\therefore 0.0268 \leq p \leq 0.1332$$

(2) In a random sample of 1000 homes in a certain city, it is found that 228 are heated by oil. find 99% confidence intervals for the proportion of homes in this city that are heated by oil.

Solution:

- **Step 1:** $n = 1000$, X : number of homes heated by oil = 228,

$$\hat{p} = \frac{X}{n} = \frac{228}{1000} = 0.228, \qquad \hat{q} = 1 - \hat{p} = 1 - 0.228 = 0.772$$

- **Step 2:** $(1-\alpha)\% = 99\% \rightarrow 1 - \alpha = 0.99 \rightarrow \alpha = 1 - 0.99$ $\quad \boxed{\alpha = 0.01}$

- **Step 3:**

$$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005}$$

$$= \frac{2.57 + 2.58}{2} = 2.575$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084

$$0.0051 - 0.005 = 0.0001 \checkmark\checkmark$$

$$0.005 - 0.0049 = 0.0001 \checkmark\checkmark$$

CI of p, is given by

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= 0.228 \pm (2.575) \sqrt{\frac{(0.228)(0.772)}{1000}} = 0.228 \pm 0.0342 \rightarrow$$

$$0.228 - 0.0342 \leq p \leq 0.228 + 0.0342$$

$$\therefore 0.1938 \leq p \leq 0.2622$$

(3) A random sample of 200 voters in a town is selected, and 114 are found to support an annexation suit. Find the 96% confidence interval for the fraction of the voting population favoring the suit.

Solution:

- **Step 1:** $n = 200$, X: number of voters supported $= 114$, $\hat{p} = \frac{X}{n} = \frac{114}{200} = 0.57$, $\hat{q} = 1 - \hat{p} = 1 - 0.57 = 0.43$
- **Step 2:** $(1-\alpha)\% = 96\% \rightarrow 1 - \alpha = 0.96 \rightarrow \alpha = 1 - 0.96$

$\alpha = 0.04$

- **Step 3:** $z_{\frac{\alpha}{2}} = z_{\frac{0.04}{2}} = z_{0.02} = 2.05$

CI of p, is given by

$$\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.57 \pm (2.05\ )\sqrt{\frac{(0.57)(0.43)}{200}} = 0.57 \pm 0.0718 \rightarrow$$

$$0.57 - 0.0718 \leq p \leq 0.57 + 0.0718$$

$$\therefore 0.4982 \leq p \leq 0.6418$$

Problems:

1. The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint:

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

Assuming that the measurements represent a random sample from a normal population, find a 95% prediction interval for the drying time for the next trial of the paint.

2. A random sample of 10 chocolate energy bars of a certain brand has, on average, 230 calories per bar, with a standard deviation of 15 calories. Construct a 99% confidence interval for the true mean calorie content of this brand of energy bar. Assume that the distribution of the calorie content is approximately normal.

3. A random sample of 12 shearing pins is taken in a study of the Rockwell hardness of the pin head. Measurements on the Rockwell hardness are made for each of the 12, yielding an average value of 48.50 with a sample standard deviation of 1.5. Assuming the measurements to be normally distributed, construct a 90% confidence interval for the mean Rockwell hardness.

4. A manufacturer of MP3 players conducts a set of comprehensive tests on the electrical functions of its product. All MP3 players must pass all tests prior to being sold. Of a random sample of 500 MP3 players, 15 failed one or more tests. Find a 90% confidence interval for the proportion of MP3 players from the population that pass all tests.

5. A geneticist is interested in the proportion of African males who have a certain minor blood disorder. In a random sample of 100 African males, 24 are found to be afflicted. Compute a 99% confidence interval for the proportion of African males who have this blood disorder.