AKTU

B. Tech II-Year

Gateway Classes

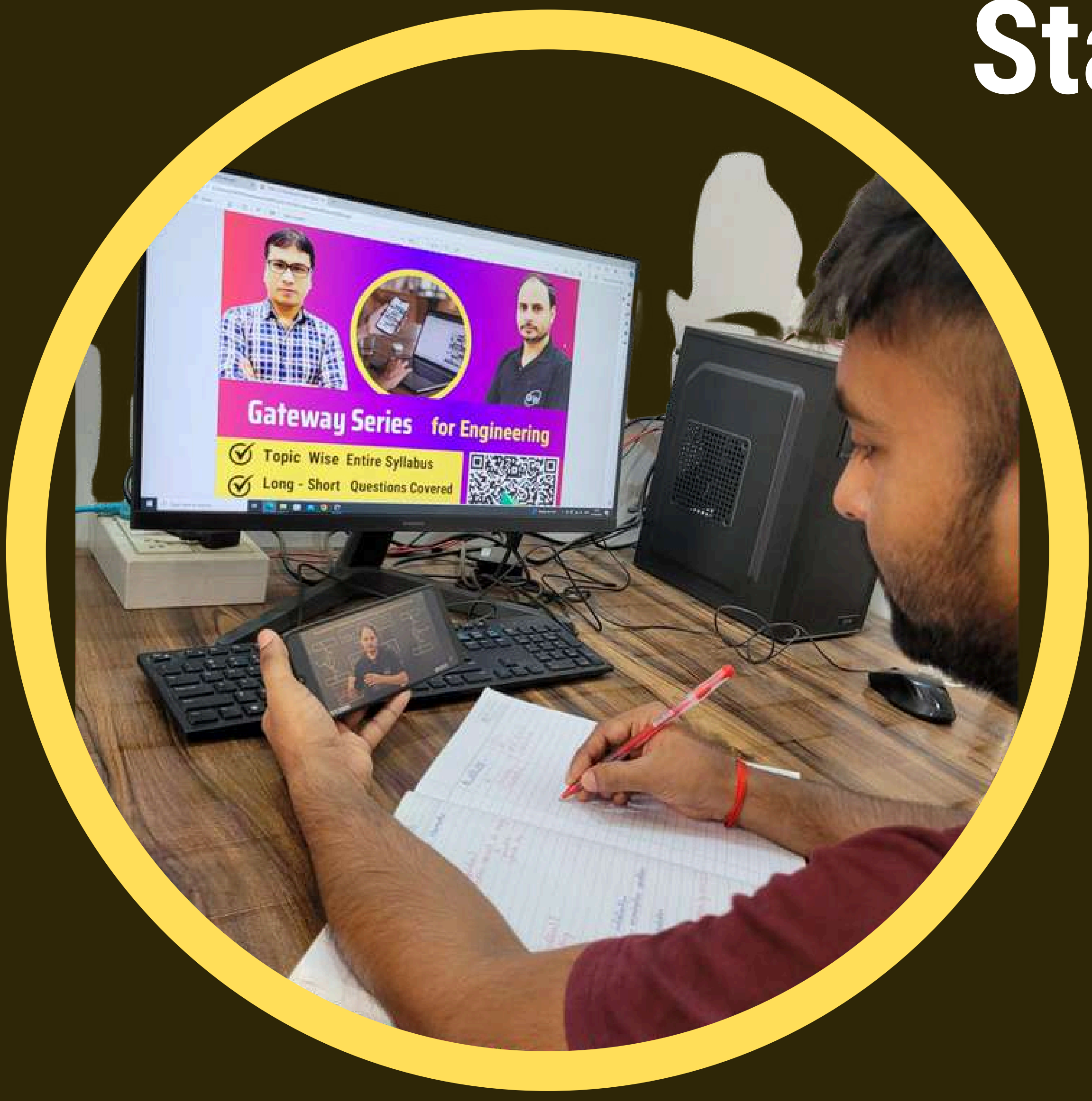
Semester -IV

ENGG.Mathematics-IV

BAS-403 ENGG- Mathematics-IV

UNIT-5 : ONE SHOT

Statistical Techniques III



Gateway Series for Engineering

- Topic Wise Entire Syllabus
- Long - Short Questions Covered
- AKTU PYQs Covered
- DPP
- Result Oriented Content

Download App

For Full Courses including Video Lectures

AKTU

B. Tech II-Year

Gateway Classes

GATEWAY CLASSES

BAS-403 ENGG. MATHEMATICS-IV

Unit-5-ONE SHOT

Introduction to Statistical Techniques III

Syallbus

: Statistical Techniques III Introduction of Sampling Theory, Hypothesis, Null hypothesis, Alternative hypothesis, Testing a Hypothesis, Level of significance, Confidence limits, Test of significance of difference of means, t-test, Z-test and Chi-square test, Statistical Quality Control (SQC) , Control Charts, Control Charts for variables (X and R Charts), Control Charts for Variables (p, np and C charts).


Engineering Mathematics IV
One Shot
STATISTICAL TECHNIQUES-II
UNIT-4
WATCH NOW

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-01

Today's Target

- *Basic terms of Hypothesis testing*

  (i) Population

  (ii) Sample

  (iii) Parameter and Statistics

  (iv) Hypothesis: Null Hypothesis and Alternative Hypothesis

  (v) Test of significance

  (vi) Level of significance

UNIT-5 : *Statistical Techniques – III*

Aktu Syllabus

Testing a Hypothesis

- *Introduction of sampling theory, Hypothesis, Null hypothesis, Alternative Hypothesis, Testing a Hypothesis, Level of Significance, Confidence limit,*
- *Test of significance of difference of means : (i) t – test (ii) Z – test (iii) Chi-square test* 14 Marks

Statistical Quality Control 7 Marks

- *Control charts*
- *Control charts for variables (X and R charts)*
- *Control charts for variables (p, np and C charts)*

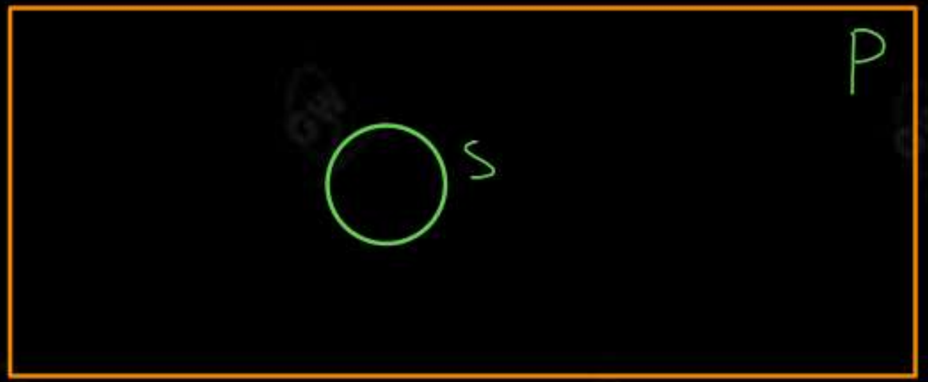
Basic terms of Hypothesis testing

(i) Population or Universe:

Any collection of individual | members | objects | set of all experimental data

Population

Finite (countable)



Infinite (Uncountable)

(ii) SAMPLE : A Finite subset of population

- *To study population, we select a random sample in such a way that each member of population has equal chance to include in sample*
- ***Sample Size*** : *The number of member | objects | individual in a sample*
- ***Sampling*** : *Process of selecting a sample*

(iii) Parameter and Statistics: Symbols for Population and Samples

	Parameter	*Statistic*
S.NO.	Measures used for Population	Measures used for Sample
1	Population Size (N)	*Sample Size* (n)
2	Population *Mean* (μ)	*Sample Mean* ($\overline{x}$)
3	Population *S.D* (σ)	*Sample S.*D (s)
4	*Population Variance* (σ^2)	*Sample Variance* (s^2)
5	*PopulationProportion* (P)	*Sample Proportion* (p)

(iv) Hypothesis / Statistical Hypotheses

Hypotheses → Assumptions required to make decisions about population on the basis of sample informatiion

Test of Hypotheses → Deciding a hypotheses is to be accepted or rejected

Null Hypothesis (H_0)

- It is a definite statement about population parameter
- Tested under the assumption that it is true

Alternative Hypotheses(H_1)

It is complementary to the null hypothesis

Note: (1) (H_0) : *Accept* then (H_1) : *Reject*

(2) (H_0) : *Reject* then (H_1) : *Accept*

Test of significance *(or Null Hypothesis testing)*

The tests which enable us, to decide whether the Null Hypothesis is eccepted or rejected on the basis of sample information

(i) t − test

(ii) Z − test

(iii) F − test

(iv) chi, −square test

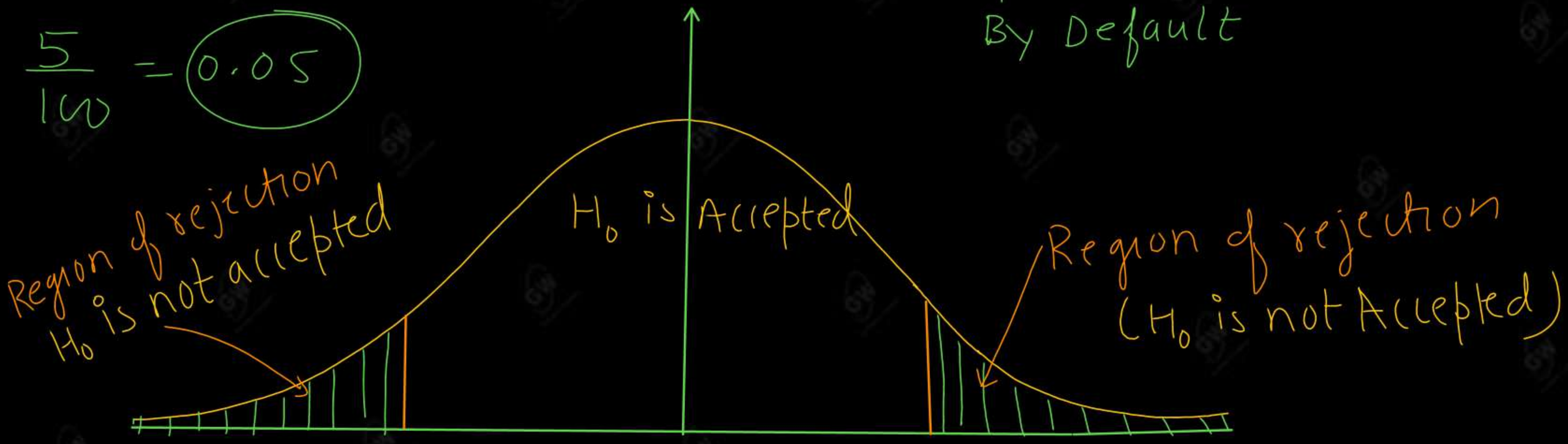
Level of significance ($\propto$)

- *The value of probability above which we do not reject the Null Hypothesis.*

- *Usualiy, level of significance in testing of hypothesis are 5 % , 1%*

↑ By Default

$$\frac{5}{100} = 0.05$$



AKTU Full Courses (Paid)
Download Gateway Classes Application
From Google Play store
All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-02

Today's Target

- $t - test\ for\ one\ sample$
- PYQ
- DPP

Parameter and Statistics: ***Symbols for Population and Samples***

	Parameter	***Statistic***
S.NO.	Measures used for Population known as Parameter	Measures used for Sample known as statistic
1	**Population Size** (N)	***Sample Size*** (n)
2	**Population** ***Mean*** (μ)	***Sample Mean*** $(\overline{x})$
3	**Population** ***S.D*** (σ)	***Sample S.*** **D** (s)
4	***Population Variance*** (σ^2)	***Sample Variance*** (s^2)
5	***PopulationProportion*** (P)	***Sample Proportion*** (p)

Hypothesis Testing

Test of significance

(i) t – test

(ii) z – test

(iii) F – test

(iv) chi – square test

7M

7M

t – test

OR

Students t – distribution

- It is a parametric test of hypothesis testing based on student's T distribution.
- To test the significance of the mean of a sample.
- t – test is applicable when

(i) sample size is small $(n \leq 30)$

(ii) Population standard deviation (σ) is not known

$t - test$

$\downarrow$

For one sample ✓ | For two sample — Next Lecture

$t - test\ for\ one\ sample$

➢ *To test the significant difference between the sample mean and population mean.* ($\bar{x}$, μ)

Steps to solve Questions

STEP − 1 : *Set Null Hypothesis* (H_0)

H_0 : *There is no significant difference between sample mean and population mean*

$H_0 : \mu = \mu_0$ (*given*)

STEP − 2 : *Alternative Hypothesis* (H_1)

$H_1 : \mu \neq \mu_0$ (Two tailed test)

Step – 3 : Calculate |t|

(i) When S.D is not given

$$\boxed{t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}}$$

$$\boxed{t = \left(\frac{\bar{x}-\mu}{s}\right)\sqrt{n}}$$

$\bar{x}$ = Sample mean

μ = Population mean

n = Sample size

$$\boxed{S = \text{Sample S.D} = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}}$$

(ii) When S.D is given

$$\boxed{t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n-1}}}}$$

Step − 4 : Write degree of freedom (n-1) and level of significance $\propto$ ($\propto = 5\%$ *OR* $\propto = 1\%$)

By default

Step − 5: *From the t − test table*

Find the value of $t_{\propto}$ *at the level of significance* $\propto$ *for the degree of freedom* $(n-1)$

Step − 6: Conclusion

(i) *If* $|t| < t_{\propto}$ *then* H_0 *is accepted*

- *There is no significant difference between sample mean and population mean*
- *Sample have been taken from correct population.*

(ii) *If* $|t| > t_{\propto}$ *then* H_0 *is rejected*

- *There is significant difference between sample mean and population mean*
- *Sample could not have come from this population.*

Fiducial limits or confidence limits

$$\boxed{\bar{x} - t_{\propto}\frac{s}{n} < \mu < \bar{x} + t_{\propto}\frac{s}{n}}$$

(i) 95% *confidence limits (level of significance is* 5%) *are* $\bar{x} \pm t_{0.05}\frac{s}{n}$

(i) 99% *confidence limits (level of significance is* 1%) *are* $\bar{x} \pm t_{0.01}\frac{s}{n}$

Q.1 *A sample of 18 items has a mean 24 units and standard deviation 3 units. Test the hypothesis that it is a random sample from a normal population with mean 27 units.*

Given

$n = 18$

$\bar{x} = 24$

$s = 3$

$\mu = 27$

Null Hypothesis (H_0)

$H_0 : \_\_\_$

$H_0 : \mu = 27$

Alternative Hypothesis (H_1)

$H_1 : \mu \neq 27$ (two tailed test

Calculation of $|t|$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

$$t = \frac{24 - 27}{\frac{3}{\sqrt{18-1}}}$$

$$t = -\frac{\cancel{3}}{\cancel{3}} \times \sqrt{17}$$

$$t = -\sqrt{17} = -4.123$$

$$\boxed{|t| = 4.123}$$

Degree of freedom $= n-1$

$$n-1 = 18-1$$
$$= 17$$

Level of significance

$$\alpha = 5\% = \frac{5}{100}$$

$$\alpha = 0.05$$

From the t-test table

$$\boxed{t_{0.05} = 2.11}$$

Conclusion

$|t| = 4.123$ and $t_{0.05} = 2.11$

$\because |t| > t_{0.05}$

$\therefore H_0$ Hypothesis is rejected

$t_{0.05} = 2.11$

Table 2 : SIGNIFICANT VALUES t_v (α) OF t-DISTRIBUTION (TWO TAIL AREAS) $[\,|t| > t_v(\alpha)] = \alpha$

d.f. (v)	Probability (Level of Significance)					
	0.50	0.10	0.05	0.02	0.01	0.001
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

Q. 2 *The 9 items of a sample have the following values:* $\overset{x}{45, 47, 50, 52, 48, 47, 49, 53, 51}$.

Does the mean of these values differ significantly from the assumed mean 47.5 ?

AKTU – 2021

Given

$n = 9$

$\mu = 47.5$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
45	−4.1	16.81
47	−2.1	4.41
50	0.9	0.81
52	2.9	8.41
48	−1.1	1.21
47	−2.1	4.41
49	−0.1	0.001
53	3.9	15.21
51	1.9	3.61
$\Sigma x = 442$		$\Sigma(x-\bar{x})^2 =$ 54.89

$$\bar{x} = \frac{\Sigma x}{n}$$

$$\bar{x} = \frac{442}{9}$$

$$\boxed{\bar{x} = 49.1}$$

$$\Sigma(x-\bar{x})^2 = 54.89$$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$s = \sqrt{\frac{54.89}{8}}$$

$$\boxed{s = 2.619}$$

Null Hypothesis (H_0)

H_0: ______

$H_0: \mu = 47.5$

Alternative Hypothesis (H_1)

$H_1: \mu \neq 47.5$ (Two tailed test)

calculation of $|t|$

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$$

$$= \frac{49.1-57.5}{\frac{2.619}{\sqrt{9}}}$$

$$t = -1.8327$$

$$\boxed{|t| = 1.8327}$$

Degree of freedom

$$n-1 = 9-1$$
$$= 8$$

Level of significance

$$\alpha = 5\% = \frac{5}{100}$$

$$\alpha = 0.05$$

From t-test table

$$\boxed{t_{0.05} = 2.31}$$

Conclusion

$$|t| = 1.8327 \text{ and } t_{0.05} = 2.31$$

$$\because |t| < t_{0.05}$$

$\therefore$ H_0 Hypothesis accepted

$t_{0.05} = 2.31$

Table 2 : SIGNIFICANT VALUES $t_v(\alpha)$ OF t-DISTRIBUTION (TWO TAIL AREAS) $[\,|t| > t_v(\alpha)] = \alpha$

d.f. (v)	Probability (Level of Significance) 0.50	0.10	0.05	0.02	0.01	0.001
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

Q.3 *The lifetime of electric bulbs for a random sample of 10 from a large consignment gave the following data:*

Item	1	2	3	4	5	6	7	8	9	10
Life in 000hrs.	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

Can we accept the hypothesis that the average lifetime of bulb is 4000 hrs? a

Given

$n = 10$

$\mu = 4000$

$\mu = 4$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
4.2	-0.2	0.04
4.6	0.2	0.04
3.9	-0.5	0.25
4.1	-0.3	0.09
5.2	0.8	0.64
3.8	-0.6	0.36
3.9	-0.5	0.25
4.3	-0.1	0.01
4.4	0	0
5.6	1.2	1.44
$\Sigma x = 44$		$\Sigma(x-\bar{x})^2 = 3.12$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{44}{10}$$

$$\boxed{\bar{x} = 4.4}$$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{3.12}{9}}$$

$$\boxed{S = 0.589}$$

Null Hypothesis (H_0)

H_0 : ______

$H_0 : \mu = 4000$

Alternative Hypothesis (H_1)

$H_1 : \mu \neq 4000$ (Two tailed test)

Calculation of $|t|$

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{4.4 - 4}{\frac{0.589}{10}}$$

$$t = 2.123$$

$$\boxed{|t| = 2.123}$$

Remaning Step → HW

$t_{0.05} = 2.26$

Table 2 : SIGNIFICANT VALUES t_v (α) OF t-DISTRIBUTION (TWO TAIL AREAS) [$|t| > t_v(\alpha)$] = α

d.f. (v)	Probability (Level of Significance) 0.50	0.10	0.05	0.02	0.01	0.001
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

GW GATEWAY CLASSES

Q.4 *A random sample of size 16 has 53 as mean. The sum of squares of the deviation from mean is 135. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.*

Given

$n = 16$

$\bar{x} = 53$

$\Sigma(x-\bar{x})^2 = 135$

$\mu = 56$

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$s = \sqrt{\frac{135}{16-1}}$$

$$s = \sqrt{9}$$

$$\boxed{s = 3}$$

Null Hypothesis (H_0)

H_0: ______

$H_0: \mu = 56$

Alternative Hypothesis (H_1)

$H_1: \mu \neq 56$ (Two tailed test)

calculation of |t|

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{53 - 56}{\frac{3}{\sqrt{16}}}$$

$$t = -\frac{\cancel{3}}{\cancel{3}}\sqrt{16}$$

$$t = -4$$

$$\boxed{|t| = 4}$$

Degree of freedom

$$n-1 = 16-1$$
$$= 15$$

Level of significance

$$\alpha = 5\% = \frac{5}{100}$$

$$\alpha = 0.05$$

From t-test table

$$\boxed{t_{0.05} = 2.13}$$

conclusion

$|t| = 4$ and $t_{0.05} = 2.13$

$\therefore |t| > t_{0.05}$

$\therefore H_0$ Hypothesis is rejected

When level of significance is 1%

$\alpha = 1\% = \frac{1}{100} = 0.01$

$\alpha = 0.01$

$\boxed{t_{0.01} = 2.95}$

confidence limit of mean at 95%

$= \bar{x} \pm t_{0.05} \frac{s}{n} = 53 \pm 2.13 \times \frac{3}{16}$

$\boxed{= 51.4025, 54.5975}$

confidence limit of mean at 99%

$= \bar{x} \pm t_{0.01} \frac{s}{n} = 53 \pm 2.95 \times \frac{3}{16}$

$\boxed{= 50.7875, 55.2125}$

$t_{0.05} = 2.13$

Table 2 : SIGNIFICANT VALUES t_v (α) OF t-DISTRIBUTION (TWO TAIL AREAS) $[\,|\,t\,| > t_v(\alpha)] = \alpha$

d.f.	*Probability (Level of Significance)*					
(v)	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*	*0.001*
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

UNIT-5 : DPP- 02

Topic : $t-test\ for\ one\ sample$

Q.1 *A sample of 20 items has mean 42 units and S.D. 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.*

Q.2 *Ten individuals are chosen at random from a normal population of students and their marks are found to be 63, 63, 66, 67, 68, 69, 70, 70, 71, 71. In the light of these data, discuss the that mean mark of the population of students is 66.*

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-03

Today's Target

- $\boxed{t - test\ for\ two\ sample}$
- PYQ
- DPP

$t-$ test for two sample

- To test the significant difference between mean of two independent samples.
- Let two independent samples are sample $-$ 1 and sample $-$ 2

Sample $-$ 1	Sample $-$ 2
(i) $x_1, x_2, x_3 \dots \dots x_n$	(i) $y_1, y_2, y_3 \dots \dots y_n$
(ii) Mean $\boxed{\bar{x} = \frac{\Sigma x}{n_1}}$	(ii) Mean $\boxed{\bar{y} = \frac{\Sigma y}{n_2}}$
(iii) $S_1 = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n_1-1}}$	(iii) $S_2 = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n_2-1}}$
(iv) Size $n_1 \leq 30$	(iv) Size $n_2 \leq 30$
(v) Population mean μ_1	(v) Population mean μ_2

Step to sovle Questions

STEP – 1: *Set Null Hypothesis* (H_0)

H_0: *The means of the two Population are same*

$H_0: \mu_1 = \mu_2$

STEP – 2: *Alternative Hypothesis* (H_1)

$H_1: \mu_1 \neq \mu_2$

STEP – 3: *Calculate* $|t|$

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

where $\bar{x}$ = Mean of Sample-1

$\bar{y}$ = Mean of Sample-2

n_1 = Size of Sample-1

n_1 = Size of sample-2

$S_1 = S.D\ of\ sample - 1$

$S_2 = S.D\ of\ sample - 2$

(i) If S.D S_1 and S_2 are given

$$\boxed{S = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}}$$

(ii) If S.D S_1 and S_2 are not given

$$\boxed{S = \sqrt{\frac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1 + n_2 - 2}}}$$

STEP – 4: Write degree of freedom $(n_1 + n_2 - 2)$ and level of significance $\propto$ ($\propto = 5\%$ OR $\propto = 1\%$)

By Default

STEP – 5 : From the t – test table

Find the value of $t_\propto$ at level of significance $\propto$ for degree of freedom $(n_1 + n_2 - 2)$

GW GATEWAY CLASSES

STEP – 6: *Conclusion*

(i) *If* $|t| < t_{\propto}$ *then* H_0 *is accepted*

There is no significant difference between ~~between~~ their means

(ii) *If* $|t| > t_{\propto}$ *then* H_0 *is rejected*

There is ~~no~~ significant difference between ~~between~~ their means

Q.1 *Samples of sizes 10 and 14 were taken from two normal populations with S.D. 3.5 and 5.2 The sample means were found to be 20.3 and 18.6. Test whether the means of the two populations are the same at 5% level.*

Given

$n_1 = 10$

$n_2 = 14$

$s_1 = 3.5$

$s_2 = 5.2$

$\bar{x} = 20.3$

$\bar{y} = 18.6$

We know that

$$S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$$S = \sqrt{\frac{10 \times (3.5)^2 + 14 \times (5.2)^2}{10 + 14 - 2}}$$

$$\boxed{S = 4.772}$$

Null Hypothesis (H_0)

H_0: ______

$H_0: \mu_1 = \mu_2$

Alternative Hypothesis (H_1)

$H_1: \mu_1 \neq \mu_2$

Calculation of $|t|$

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{20.3 - 18.6}{4.772\sqrt{\frac{1}{10} + \frac{1}{14}}}$$

$$t = 0.8604$$

$$\boxed{|t| = 0.8604}$$

Degree of freedom $= n_1 + n_2 - 2$

$= 10 + 14 - 2$

$= 22$

Level of significance (α)

$$\alpha = 5\% = \frac{5}{100} = 0.05$$

$$\alpha = 0.05$$

From t-test table

$$\boxed{t_{0.05} = 2.07}$$

conclusion

$|t| = 0.8604$ and $t_{0.05} = 2.07$

$\because |t| < t_{0.05}$

$\therefore$ H_0 Hypothesis is accepted

$t_{0.05} = 2.07$

Table 2 : SIGNIFICANT VALUES $t_v(\alpha)$ OF t-DISTRIBUTION (TWO TAIL AREAS) $[\,|t| > t_v(\alpha)] = \alpha$

d.f.	Probability (Level of Significance)					
(v)	0.50	0.10	0.05	0.02	0.01	0.001
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

Q.2 ***The height of 6 randomly chosen sailors in inches are 63, 65, 68, 69, 71 and 72. Those of 9 randomly chosen soldiers are 61, 62, 65, 66, 69, 70, 71, 72 and 73. Test whether the sailors are on the average taller than soldiers.***

Given

Sample-1 → Height of Sailors

Sample-2 → Height of Soldiers

$n_1 = 6$

$n_2 = 9$

x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$y-\bar{y}$	$(y-\bar{y})^2$
63	-5	25	61	-6.67	44.4889
65	-3	9	62	-5.67	32.1489
68	0	0	65	-2.67	7.1289
69	1	1	66	-1.67	2.7889
71	3	9	69	1.33	1.7689
72	4	16	70	2.33	5.4283
			71	3.33	11.0889
			72	4.33	18.7489
			73	5.33	28.4089
$\Sigma x = 408$		$\Sigma(x-\bar{x})^2 = 60$	$\Sigma y = 609$		$\Sigma(y-\bar{y})^2 = 151.9995$

Mean

$$\bar{x} = \frac{\sum x}{n_1} = \frac{408}{6}$$

$$\boxed{\bar{x} = 68}$$

$$\bar{y} = \frac{\sum y}{n_2}$$

$$\bar{y} = \frac{609}{9}$$

$$\boxed{\bar{y} = 67.67}$$

We know that

$$S = \sqrt{\frac{\sum(x-\bar{x})^2 + \sum(y-\bar{y})^2}{n_1 + n_2 - 2}}$$

$$S = \sqrt{\frac{60 + 151.9995}{6 + 9 - 2}}$$

$$\boxed{S = 4.038}$$

Null Hypothesis (H_0)

H_0: ______

$H_0: \mu_1 = \mu_2$

Calculation of $|t|$

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$t = \frac{68 - 67.67}{4.038\sqrt{\frac{1}{6} + \frac{1}{9}}}$$

$$\boxed{|t| = 0.1505}$$

Degree of freedom $= n_1 + n_2 - 2$

$= 6 + 9 - 2$

$= 13$

Level of significance α

$\alpha = 5\% = \frac{5}{100} = 0.05$

$\alpha = 0.05$

From the t-test table

$\boxed{t_{0.05} = 2.16}$

Conclusion

$|t| = 0.1505$ and $t_{0.05} = 2.16$

$\because |t| < t_{0.05}$

$\therefore$ H_0 Hypothesis is accepted

Means of the two population are same

Hence,

Average Height of Sailors is not taller than soldiers

$t_{0.05} = 2.16$

Table 2 : SIGNIFICANT VALUES $t_v(\alpha)$ OF t-DISTRIBUTION (TWO TAIL AREAS) $[\,|t| > t_v(\alpha)] = \alpha$

d.f. (v)	*Probability (Level of Significance)*					
	0.50	*0.10*	*0.05*	*0.02*	*0.01*	*0.001*
1	1.00	6.31	12.71	31.82	63.66	636.62
2	0.82	0.92	4.30	6.97	6.93	31.60
3	0.77	2.32	3.18	4.54	5.84	12.94
4	0.74	2.13	2.78	3.75	4.60	8.61
5	0.73	2.02	2.57	3.37	4.03	6.86
6	0.72	1.94	2.45	3.14	3.71	5.96
7	0.71	1.90	2.37	3.00	3.50	5.41
8	0.71	1.80	2.31	2.90	3.36	5.04
9	0.70	1.83	2.26	2.82	3.25	4.78
10	0.70	1.81	2.23	2.76	3.17	4.59
11	0.70	1.80	2.20	2.72	3.11	4.44
12	0.70	1.78	2.18	2.68	3.06	4.32
13	0.69	1.77	2.16	2.05	3.01	4.22
14	0.69	1.76	2.15	2.62	2.98	4.14
15	0.69	1.75	2.13	2.60	2.95	4.07
16	0.69	1.75	2.12	2.58	2.92	4.02
17	0.69	1.74	2.11	2.57	2.90	3.97
18	0.69	1.73	2.10	2.55	2.88	3.92
19	0.69	1.73	2.09	2.54	2.86	3.88
20	0.69	1.73	2.09	2.53	2.85	3.85
21	0.69	1.72	2.08	2.52	2.83	3.83
22	0.69	1.72	2.07	2.51	2.42	3.79
23	0.69	1.71	2.07	2.50	2.81	3.77
24	0.69	1.71	2.06	2.49	2.80	3.75
25	0.68	1.71	2.06	2.49	2.79	3.73
26	0.68	1.71	2.06	2.48	2.78	3.71
27	0.68	1.70	2.05	2.47	2.77	3.69
28	0.68	1.70	2.05	2.47	2.76	3.67
29	0.68	1.70	2.05	2.46	2.76	3.66
30	0.68	1.70	2.04	2.46	2.75	3.65
∞	0.67	1.65	1.96	2.33	2.58	3.29

UNIT-5 : DPP- 03

Topic : $t-$ test for two sample

Q.1 *Two samples of sodium vapour bulbs were tested for length of life and the following results were got:*

	Size	Sample Mean	Sample S.D.
Type I	8	1234 hrs	36 hrs
Type II	7	1036 hrs	40 hrs

Is the difference in the means significant to generalise that Type I is superior to Type II regarding length of life?

Q.2 *The marks obtained by a group of 9 regular course students and another group of 11 part time course students in a test are given below:*

Regular:	56	62	63	54	60	51	67	69	58		
Part Time:	62	70	71	62	60	56	75	64	72	68	66

Examine whether the marks obtained by regular students and part time students differ significantly at 5% and 1% level of significance.

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-04

Today's Target

- $Z\text{--}test$
- PYQ
- DPP

Hypothesis Testing

GATEWAY CLASSES

Test of significance

(i) t − test

(ii) z − test

(iii) F − test

(iv) chi − square test

Z−Test is applicable when

	Population S.D (σ) is Known	Population S.D (σ) is Unknown
$n \leq 3o$	z − test	t − test
$n > 3o$	z − test	z − test

Z – Test

(1) $Z-$ *test for single mean (one sample)*

(2) $Z-$ *test for Difference of mean (two sample)*

(3) $Z-$ *test for single proportion*

(4) $Z-$ *test for difference of proportion*

(5) $Z-$ *test for difference of standard deviation*

Z – Test Table

	Level of Significance		
	1%	5%	10%
Two Tailed Test	$\lvert Z \rvert = 2.58$	$\lvert Z \rvert = 1.96$	$\lvert Z \rvert = 1.645$
Right Tailed Test	$Z = 2.33$	$Z = 1.645$	$Z = 1.28$
Left Tailed Test	$Z = -2.33$	$Z = -1.645$	$Z = -1.28$

Z–test for one sample

To test the significant difference between the sample mean and population mean.

Steps to solve Questions

STEP – 1 : Set Null Hypothesis (H_0)

$$H_0 : \mu = \mu_0 \text{ (given)}$$

STEP – 2 : Alternative Hypothesis (H_1)

$$H_1 : \mu \neq \mu_0 \text{(Two Tail test)},\ \mu > \mu_0 \text{(Right Tail test)},\ \mu < \mu_0 \text{(Left Tail test)}$$

Step – 3 : Calculate Z

$$Z_{cal} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} \quad \text{OR} \quad Z_{cal} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

$\bar{x}$ = Sample mean

μ = Population mean

n = Sample size

σ = Population S.D

S = Sample S.D

$Step - 4 : From\ the\ Z - test\ table$

Find the tabulated value Z_{tab} at the given level of significance $\propto$

$Step - 5 : Conclusion$

(i) If $Z_{cal} < Z_{tab}$ then H_0 is accepted

(ii) If $Z_{cal} > Z_{tab}$ then H_0 is rejected

Confidence limits or fiducial limits

$$\bar{x} - z_{\propto}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\propto}\frac{\sigma}{\sqrt{n}}$$

(i) 95% confidence limits (level of significance is 5%) are $\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$

(i) 99% confidence limits (level of significance is 1%) are $\bar{x} - 2.58\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}$

Q.1 *A random sample of 900 members has a mean 3.4 cms. Can it be reasonably regarded as a sample from a large population of mean 3.2 cms and S.D. 2.3 cms?*

Given

$n = 900$

$\bar{x} = 3.4$

$\mu = 3.2$

$\sigma = 2.3$

Null Hypothesis (H_0)

Sample is drawn from large population with mean 3.2

$H_0 : \mu = 3.2$

Alternative Hypothesis (H_1)

$H_1 : \mu \neq 3.2$ (two tailed test)

Calculation of $|Z|$

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

$$Z = \frac{3.4 - 3.2}{\frac{2.3}{\sqrt{900}}}$$

$$Z = \frac{0.2 \times 30}{2.3}$$

$$\boxed{Z_{cal} = 2.61}$$

From Z-test table

At the level of significance 5%

$$\boxed{Z_{tab} = 1.96}$$

conclusion

$Z_{cal} = 2.61$

$Z_{tab} = 1.96$

$\because |Z_{cal}| > |Z_{tab}|$

$\therefore H_0$ is rejected

Hence

H_1 is accepted

$Z - Test\ for\ two\ sample$

To test the significant difference between mean of two large independent samples.

Step to sovle Questions

STEP – 1: *Set Null Hypothesis* (H_0)

$$H_0: \mu_1 = \mu_2$$

STEP – 2: *Alternative Hypothesis* (H_1)

$$H_1 : \mu \neq \mu_0 (\text{Two Tail test}),\ \mu > \mu_0 (\text{Right Tail test}),\ \mu < \mu_0 (\text{Left Tail test})$$

STEP – 3: *Calculate* $|z|$

$$Z_{cal} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad OR \quad Z_{cal} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where $\bar{x}$ = Mean of Sample-1

$\bar{y}$ = Mean of Sample-2

$$\sigma^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}$$

STEP − 4 : From the $Z -$ test table

Find the tabulated value Z_{tab} at the level of significance $\propto$

STEP − 5: Conclusion

(i) If $Z_{\text{cal}} < Z_{tab}$ then H_0 is accepted

(ii) If $Z_{\text{cal}} > Z_{tab}$ then H_0 is rejected

Q.2 *The average income of persons was Rs. 210 with a S.D. of Rs. 10 in sample of 100 people of a city. For another sample of 150 persons, the average income was Rs. 220 with S.D. of Rs. 12. The S.D. of incomes of the people of the city was Rs. 11. Test whether there is any significant difference between the average incomes of the localities.*

Given

Sample-1

$n_1 = 100$

$\bar{x} = 210$

$s_1 = 10$

Sample-2

$n_2 = 150$

$\bar{y} = 220$

$s_2 = 12$

Null Hypothesis (H_0)

H_0: There is no significant difference between the average income of the localities

$H_0: \bar{x} = \bar{y}$

Alternative Hypothesis (H_1)

$H_1: \bar{x} \neq \bar{y}$

Calculation of $|Z_{cal}|$

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$z = \frac{210-220}{\sqrt{\frac{10^2}{100} + \frac{12^2}{150}}}$$

$$z = -7.1428$$

$$\boxed{|z_{cal}| = 7.1428}$$

From z-test table

At the level of significance 5%

$$\boxed{z_{tab} = 1.96}$$

Conclusion

$\because |z_{cal}| > |z_{tab}|$

$\therefore$ H_0 is rejected

Hence

H_1 is accepted

$Z-Test\ for\ one\ single\ proportion$

To test the significant difference between Proportion of the sample and the population.

Let X be the number of successes in n independent trials with constant probability P of success for each trial.

Steps to solve Questions

STEP − 1 : Set Null Hypothesis (H_0)

$$H_0 : \mu = \mu_0 \ (given)$$

STEP − 2 : Alternative Hypothesis (H_1)

$$H_1 : \mu \neq \mu_0$$

Step − 3 : Calculate Z

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$$

Where

Observed Proportion of Success $p = \frac{x}{n}$

$x =$ number of success

$n =$ Sample Size

$P =$ Probability of success

$Q = 1 - P$ (Probability of failure)

Step -5: *From the* $Z-$ *test table*

Find the tabulated value Z_{tab} *at the level of significance* $\propto$

Step -6: *Conclusion*

(i) $If Zca_l < Z_{tab} then\ H_0\ is\ accepted.$

(ii) $If Zca_l > Z_{tab}\ then\ H_0\ is\ rejected$

95% *confidence limits (level of significance is* 5%) *are* $P \pm Z_{tab}\sqrt{\frac{PQ}{n}}$

99% *confidence limits (level of significance is* 1%) *are* $P \pm Z_{tab}\sqrt{\frac{PQ}{n}}$

$Q.3$ *A coin was tossed 400 times and the head turned up 216 times. Test the hypothesis that the coin is unbiased.*

Given

$n = 400$

$P = 0.5$

$Q = 1 - 0.5$

$Q = 0.5$

$x = 216$

$p = \frac{x}{n}$

$p = \frac{216}{400}$

$p = 0.54$

Null Hypothesis (H_0)

H_0: The coin is unbiased

$H_0: P = 0.5$

Alternative Hypothesis (H_1)

$H_1: P \neq 0.5$ (two tailed test)

Calculation of $|Z|$

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{400}}}$$

$Z = 1.6$

$$\boxed{|Z_{cal}| = 1.6}$$

From z-test table

At the level of singnificance 5%

$$|Z_{tab}| = 1.96$$

conclusion

As $|Z_{cal}| < |Z_{tab}|$

$\therefore$ H_0 is accepted

$Z-$test for difference of proportion

To test the significant difference between the sample proportions p_1 and p_2

Step to sovle Questions

STEP – 1: Set Null Hypothesis (H_0)

$$H_0: \mu_1 = \mu_2$$

STEP – 2: Alternative Hypothesis(H_1)

$$H_1: \mu_1 \neq \mu_2$$

STEP – 3: Calculation of $|z|$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Where $p_1 = \frac{x_1}{n_1}$

$p_2 = \frac{x_2}{n_2}$

$$\text{Where } P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$Q = 1 - P$$

$STEP - 5:$ From the $Z -$ test table

Find the tabulated value Z_{tab} at the level of significance $\propto$

$STEP - 6:$ Conclusion

(i) If $Z_{cal} < Z_{tab}$ then H_0 is accepted

(ii) If $Z_{cal} > Z_{tab}$ then H_0 is rejected

Q. 4 *Before an increase in excise duty on tea, 800 people out a sample of 100 persons were found to be tea drinkers. After an increase in the duty, 800 persons were known to be tea drinkers in a sample of 1200 people. Do you think that there has been a significant decrease in the consumption of tea after the increase in the exise duty?*

Given

$n_1 = 1000$

$x_1 = 800$

$n_2 = 1200$

$x_2 = 800$

$p_1 = \frac{x_1}{n_1} = \frac{8\cancel{00}}{10\cancel{00}}$

$p_1 = 0.8$

$p_2 = \frac{x_2}{n_2} = \frac{8\cancel{00}}{12\cancel{00}}$

$p_2 = 2/3$

$p_2 = 0.67$

$P = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{800+800}{1000+1200}$

$P = \frac{8}{11}$

$Q = 1 - P = 1 - \frac{8}{11}$

$Q = \frac{3}{11}$

Null Hypothesis (H_0)

H_0: There is no significant decrease in the consumption of tea after increase in the excise duty

$H_0 : p_1 = p_2$

Alternative Hypothesis (H_1)

$H_1 : p_1 > p_2$ (one tailed test)

calculation of $|Z|$

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$Z = \frac{0.8 - 0.67}{\sqrt{\frac{8}{11} \times \frac{3}{11}\left(\frac{1}{1000} + \frac{1}{1200}\right)}}$$

$$Z = 6.817$$

$$\boxed{|Z_{cal}| = 6.817}$$

From t-test table

At the level of significance 5%

$|Z_{tab}| = 1.645$

Conclusion

As $|Z_{cal}| > |Z_{tab}|$

$Z - test\ for\ difference\ of\ standard\ deviations$

To test the significant difference between the standard deviation of two independent samples.

Step to sovle Questions

$STEP-1:$ *Set Null Hypothesis* (H_0)

$$H_0: \mu_1 = \mu_2$$

$STEP-2:$ *Alternative Hypothesis* (H_1)

$$H_1: \mu_1 \neq \mu_2$$

$STEP-3:$ *Calculation of* $|z|$

$$Z = \frac{s_1 - s_2}{\sqrt{\frac{{\sigma_1}^2}{2n_1} + \frac{{\sigma_2}^2}{2n_2}}} \qquad OR \qquad Z = \frac{s_1 - s_2}{\sqrt{\frac{{s_1}^2}{2n_1} + \frac{{s_2}^2}{2n_2}}}$$

$STEP-5$: *From the* $Z-$ *test table*

Find the tabulated value Z_{tab} *at the level of significance* $\propto$

$STEP-6$: *Conclusion*

(i) $If Z_{\text{cal}} < Z_{tab}$ *then* H_0 *is accepted*

(ii) $If Z_{\text{cal}} > Z_{tab}$ *then* H_0 *is rejected*

Q.5 *Random samples drawn from two countries gave the following data relating to the heights of adult males :*

	country A	country B
Mean height(in inche	67.42	67.25
Standard deviation	2.58	2.50
number in samples	1000	1200

Is the difference between the standard deviation significant?

Given

$n_1 = 1000$

$n_2 = 1200$

$s_1 = 2.58$

$s_2 = 2.50$

Null Hypothesis (H_0)

H_0: There is no significant diffeunce between the standard deviation

H_0: $s_1 \neq s_2$ (Two tailed test)

Alternative Hypothesis(H_1)

$H_1 : s_1 \neq s_2$

Calculation of Z

$$Z = \frac{s_1 - s_2}{\sqrt{\frac{s_1^2}{2n_1} + \frac{s_2^2}{2n_1}}}$$

$$Z = \frac{2.58 - 2.50}{\sqrt{\frac{2.58^2}{2\times1000} + \frac{2.50^2}{2\times1200}}}$$

$$Z = 1.0387$$

$$\boxed{|Z_{cal}| = 1.0387}$$

From z-test table

At the level of significance 5%

$$|Z_{tab}| = 1.96$$

Conclusion

As $|Z_{cal}| < |Z_{tab}|$

$\therefore$ H_0 is accepted

DPP-4

Topic : Z - Test

Q.1 – *A machine is producing bolts of which a certain fraction is defective. A random sample of 400 is taken from a large batch and is found to contain 30 defective bolts. Does this indicate that the proportion of defectives is larger than that claimed by the manufacturer where the manufacturer claims that only 5% of his product are defective. Find 95% confifence limits of the proportion of defective bolts in batch.*

Q.2 – *A machine produced 16 defective articles in a batch of 500. After overhauling it produced 3 defectives in a batch of 100. Has the machine improved?*

Q.3 – *The mean weight obtained from aa random sample of size 100 is 64 gms. The S.D. of the weight distribution of the population is 3 gms. Test the statement that the mean weight of the pupulation is 67 gms at 5% level of significance. Also set up 99% confidence limits of the mean weight of the pupulation.*

Q.4 – *Intelligence tests were given to two group of boys and girls.*

	Mean	*S.D.*	*Size*
Girls	75	8	60
Boys	73	10	100

Examine if the difference mean score is significant

Q .4 – *Intelligence tests were given to two group of boys and girls.*

	Mean	S.D.	Size
Girls	75	8	60
Boys	73	10	100

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-05

Today's Target

- $\chi^2 - test\ of\ goodness\ of\ FIT\ (Part - 1)$
- PYQ
- DPP

Hypothesis Testing

Test of significance

(i) t − test

(ii) z − test

(iii) F − test

(iv) chi − square test (χ^2 − test) v.imp

Hpothesis Testing

Parametric test

(Sample drawn from normal population)

(i) t − test

(ii) z − test

(iii) F − test

Non − Parametric test

(No information about population)

(i) χ^2 − test

Chi − Square test (χ^2 − test)

χ^2 − test

χ^2 − test of goodness of FIT

L-5, L-6

χ^2 − test as a test of independence

L-7

χ^2 − test of goodness of Fit : Given Data can be fitted in

(i) Poisson Distribution

(ii) Binomial Distribution

(iii) Normal Distribution

Fitting is good or not is test by χ^2 − test.

Given in Question

Find

In χ^2 − test $o_1, o_2 \dots . o_n$ is a set of observed frequencies and $E_1, E_2, E_3 \dots E_n$ are the corresponding set of expected frequencies such that $\sum o_i = \sum E_i$

Conditions for applying χ^2 – test

(i) Sum of frequencies must be greater than equal to 50.

(ii) The expected frequency E_i of any item must be greater than 5.

Note : If any item has frequency less than 5 then it should be combined with the next or preceding item until the total frequency exceed 5.

Steps to solve questions

Step – 1 : Null Hypothesis (H_o)

(H_o): There is no significant difference between observed and expected frequency.

Step – 2 : Calculate χ^2 under H_o

$$\boxed{\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i}\right]}$$

Where O_i = observed frequency

E_i = Expected frequency

Step – 3 : Degree of freedom = Number of variables – number of constraints

$$\vartheta = n - k = n-1$$

Level of significance $\propto = 5\% = 0.05$

Step – 4 : From χ^2 – table

Calculate ${\chi^2}_{0.05}$ at the given level of significance for the given degree of freedom

Step – 5 : conclusion

(i) If $\chi^2 < {\chi^2}_{0.05}$, then H_0 accepted

- There is no significants difference between observed frequency and expected frequency.
- Fit of the data is considered to be good.

(ii) $\chi^2 > {\chi^2}_{0.05}$, then H_0 is rejected.

- There is significants difference between observed frequency and expected frequency.
- Fit of the data is not good.

Q.1 ***A die is thrown 276 times and the results of these throws are given below:***

No. appeared on the die	1	2	3	4	5	6
Frequency	40	32	29	59	57	59

Test whether the die is biased or not. **(*A.K.T.U. 2019*)**

Null Hypothesis (H_0)

H_0: The die is an unbiased die

calculation of χ^2, under H_0

$$E_i = \frac{\sum O_i}{n} = \frac{276}{6}$$

$$E_i = 46$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$
40	46	−6	36
32	46	−14	196
29	46	−17	289
59	46	13	169
57	46	11	121
59	46	13	169
$\sum O_i = 276$	$\sum E_i = 276$		$\sum (O_i - E_i)^2 = 980$

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i}\right]$$

$$\chi^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$

$$= \frac{980}{46}$$

$$\boxed{\chi^2 = 21.30}$$

Degree of Freedom $= n-1$
$= 6-1$
$= 5$

Level of significance α

$\alpha = 5\% = 0.05$

From χ^2-test table

$$\chi^2_{0.05} = 11.070$$

Conclusion

$\chi^2 = 21.30$, $\chi^2 = 11.070$

$\because \chi^2 > \chi^2_{0.05}$

$\therefore H_0$ is rejected

Hence,

Given die is a biased die

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r\,(\chi^2 > \chi^2(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q. 2 *The following table gives the number of accidents that took place in an industry during various days of the week. Test if accidents are uniformly distributed over the week.*

Day	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	14	18	12	11	15	14

Null Hypothesis (H_0)

H_0 : Accidents are uniformly distributed over the week

calculation of χ^2, under H_0

$$E_i = \frac{\sum O_i}{n} = \frac{84}{6} = 14$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$
14	14	0	0
18	14	4	16
12	14	−2	4
11	14	−3	9
15	14	1	1
14	14	0	0
$\sum O_i =$ 84			$\sum (O_i - E_i)^2 =$ 30

$$\chi^2 = \sum\left[\frac{(O_i - E_i)^2}{E_i}\right]$$

$$= \frac{\sum (O_i - E_i)^2}{E_i}$$

$$= \frac{30}{14}$$

$$\boxed{\chi^2 = 2.1428}$$

Degree of freedom $= n-1$

$= 6-1$

$= 5$

Level of Significance (α)

$\alpha = 5\% = 0.05$

From χ^2-test table

$$\boxed{\chi^2_{0.05} = 11.070}$$

Conclusion

$\chi^2 = 2.1428$, $\chi^2_{0.05} = 11.070$

$\because \chi^2 < \chi^2_{0.05}$

$\therefore H_0$ is accepted

Hence,

Accidents are uniformly distributed over the week

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r\,(\chi^2 > \chi^2\,(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	Probability (Level of Significance)						
	0.99	0.95	0.50	0.10	0.05	0.02	0.01
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q. 3 ***In experiments on pea breeding, the following frequencies of seeds were obtained:***

Round and yellow	*Wrinkled and yellow*	*Round and green*	*Wrinkled and green*	*Total*
315	101	108	32	556

Theory predicts that the frequencies should be in proportions 9 : 3 :: 3 : 1.

Examine the correspondence between theory and experiment.

Null Hypothesis (H_0)

H_0: Experimental Results support theory

Calulation of χ^2 under H_0

$$E_1 = \frac{9}{16} \times 556 = 312.75$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
315	312.75	2.25	5.0625	0.016187
101	104.25	−3.25	10.5625	0.101319
108	104.25	3.75	14.0625	0.134892
32	34.75	−2.75	7.5625	0.217626
$\sum O_i = 556$	$\sum E_i = 556$			$\sum \left[\frac{(O_i - E_i)^2}{E_i}\right] = ?$

$E_2 = \frac{3}{16} \times 556 = 104.25$

$E_3 = \frac{3}{16} \times 556 = 104.25$

$E_4 = \frac{1}{16} \times 556 = 34.75$

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

$$\boxed{\chi^2 = 0.470024}$$

Degree of freedom $= n - 1$

$= 4 - 1$

$= 3$

Level of significance (α)

$\alpha = 5\% = 0.05$

From χ^2-test table

$\chi^2_{0.05} = 7.185$

Conclusion

$\chi_{cal} = 0.470024$, $\chi_{tab} = 7.185$

$\because \chi_{cal} < \chi_{tab}$

$\therefore H_0$ is Accepted

Hence,

Experimental Results support theoritical results

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r\,(\chi^2 > \chi^2(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

UNIT-5 : DPP- 05

Topic : χ^2 − test of goodness of FIT (Part − 1)

Q.1 What is $\chi^2 -$ test?

A die is thrown 90 times with the following results: [AKTU 2010]

Face	:	1	2	3	4	5	6	Total
Frequency	:	10	12	16	14	18	20	90

Use $\chi^2 -$ test to test whether these data are consistent with the hypothesis that die is unbiased.

Q.2 The demand for a particular spare part in a factory was found to vary from day − to − day. In a sample study, the following information was obtained:

Days :	Mon	Tue	Wed	Thurs	Fri	Sat
No. of parts demanded :	1124	1125	1110	1120	1126	1115

Test the hypothesis that the number of parts demanded does not depend on the day of the week.

Q.3 The theory predicts the proportion of beans in the four groups, G_1, G_2, G_3, G_4 should be in the ratio 9: 3: 3: 1. In an experiment with 1600 beans the numbers in the four groups were 882, 313, 287 and 118. Does the experimental result support the theory?

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-06

Today's Target

- $\boxed{\chi^2 - test\ of\ goodness\ of\ FIT\ (Part-2)}$
- PYQ
- DPP

Hpothesis Testing

Parametric test

(Sample drawn from normal population)

(i) t – test

(ii) z – test

(iii) F – test

Non – Parametric test

(No information about population)

(i) χ^2 – test

χ^2 – test

χ^2 – test of goodness of FIT

L-5, L-6

χ^2 – test as a test of independence

L-7

χ^2 – test of goodness of Fit :

- Given Data can be fitted in

(i) Poisson Distribution

(ii) Binomial Distribution

~~(iii) Normal Distribution~~

- Fitting is good or not is test by χ^2 – test.
- In χ^2 – test $o_1, o_2 \ldots . o_n$ is a set of observed frequencies and $E_1, E_2, E_3 \ldots E_n$ are the corresponding set of expected frequencies such that $\sum o_i = \sum E_i$

Conditions for applying χ^2 – test

(i) Sum of frequencies must be greater than equal to 50.

(ii) The expected frequency E_i of any item must be greater than 5.

Note : If any item has frequency less than 5 then it should be combined with the next or preceding item until the total frequency exceed 5.

Steps to solve questions

Step – 1 : Null Hypothesis (H_o)

(H_o): There is no significant difference between observed and expected frequency.

Step – 2 : Calculate χ^2 under H_o

$$\chi^2 = \sum\left[\frac{(O_i - E_i)^2}{E_i}\right]$$

Where O_i = observed frequency

E_i = Expected frequency

Step – 3 : Degree of freedom = Number of variables – number of constraints

$$\vartheta = n - k$$

Level of significance $\propto = 5\% = 0.05$

Step – 4 : From χ^2 – table

Calculate ${\chi^2}_{0.05}$ at the given level of significance for the given degree of freedom

Step – 5 :

(i) If $\chi^2 < {\chi^2}_{0.05}$, then H_0 accepted

- There is no significants difference between observed frequency and expected frequency.
- Fit of the data is considered to be good.

(ii) $\chi^2 > {\chi^2}_{0.05}$, then H_0 is rejected.

- There is significants difference between observed frequency and expected frequency.
- Fit of the data is not good.

Q.1 *A survey of 320 families with 5 children shows the following distribution:*

No of boys & girls:	*5 boys & 0 girl*	*4 boys & 1 girl*	*3 boys & 2 girl*	*2 boys & 3 girl*	*1 boys & 4 girl*	*0 boys & 5 girl*	*Total*
No. of families:	18	56	110	88	40	8	320

Test the hypothesis that data are binomially distributed and male and female births are equal probability. **(G.B.T.U. 2010)**

Null Hypothesis (H_0)

H_0: Data are binomially distributed and probability of Male and Female birth are equal

Let p = Prob. of Girl (success)

q = Prob. of boy (Failure)

$p = q = \frac{1}{2}$

$N = 320$

$n = 5$

$r = 0, 1, 2, 3, 4, 5$

Binomial Distribution

$$P(X=r) = {}^{n}C_{r}\, p^{r} q^{n-r}$$

Expected Frequencies

$$E_i = N\, {}^{n}C_r\, p^r q^{n-r}$$

$$\boxed{E_i = 320 \times {}^{5}C_r \left(\tfrac{1}{2}\right)^r \left(\tfrac{1}{2}\right)^{5-r}} \quad \text{— (1)}$$

$$E_1 = 320 \times {}^{5}C_0 \times \left(\tfrac{1}{2}\right)^0 \times \left(\tfrac{1}{2}\right)^5$$

$$= \cancel{320}^{10} \times 1 \times 1 \times \frac{1}{\cancel{32}}$$

$$\boxed{E_1 = 10}$$

$$E_2 = 320 \times {}^{5}C_1 \times \left(\tfrac{1}{2}\right)^1 \times \left(\tfrac{1}{4}\right)^4 = \cancel{320}^{10} \times 5 \times \frac{1}{\cancel{32}} = 50$$

$$\boxed{E_2 = 50}$$

$$E_3 = 320 \times {}^{5}C_2 \times \left(\tfrac{1}{2}\right)^2 \times \left(\tfrac{1}{2}\right)^3 = \cancel{320}^{10} \times \frac{5 \times \cancel{4}^{2}}{\cancel{2}} \times \frac{1}{\cancel{32}} = 100$$

$$\boxed{E_3 = 100}$$

$$E_4 = 320 \times {}^{5}C_3 \times \left(\tfrac{1}{2}\right)^3 \times \left(\tfrac{1}{2}\right)^2 = \cancel{320}^{10} \times \frac{5 \times \cancel{4}^{2}}{\cancel{2}} \times \frac{1}{\cancel{32}}$$

$$\boxed{E_4 = 100}$$

$$E_5 = 320 \times {}^{5}C_4 \times \left(\tfrac{1}{2}\right)^4 \times \left(\tfrac{1}{2}\right)^1 = \cancel{320}^{10} \times 5 \times \frac{1}{\cancel{32}} = 50$$

$$\boxed{E_5 = 50}$$

$$E_6 = 320 \times {}^5C_5 \times \left(\frac{1}{2}\right)^5 \times \left(\frac{1}{2}\right)^0$$

$$= \cancel{320}^{10} \times 1 \times \frac{1}{\cancel{32}}$$

$$\boxed{E_6 = 10}$$

calulation of χ^2, under H_0

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
18	10	8	64	6.4
56	50	6	36	0.72
110	100	10	100	1
88	100	−12	144	1.44
40	50	−10	100	2
8	10	−2	4	0.4
				$\sum \left[\frac{(O_i - E_i)^2}{E_i}\right] = 11.96$

$$\boxed{\chi^2_{cal} = 11.96}$$

Degree of freedom $= n - k$

$= 6 - 1$

$= 5$

Level of significance (α)

$\alpha = 5\% = 0.05$

From χ^2-table

$\chi_{tab} =$

$$\boxed{\chi_{tab} = 11.070}$$

Conclusion

$\chi^2_{cal} = 11.96$ and $\chi_{tab} = 11.070$

$\because \chi^2_{cal} > \chi^2_{tab}$

$\therefore H_0$ is rejected

Hence,

Data are not Binomially Distributed and Probability of birth of boy and girl are not equal

$\Rightarrow p \neq q$

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q.1 *Fit a Poisson distribution to the following data and test the goodness of fit:*

x	:	0	1	2	3	4
		O_1	O_2	O_3	O_4	O_5
f	:	109	65	22	3	1

Null Hypothesis(H_0)

H_0: Fitting of data in poisson Distribution is good

x	f	fx
0	109	0
1	65	65
2	22	44
3	3	9
4	1	4
	$\Sigma f = 200$	$\Sigma fx = 122$

$$\text{Mean} = \frac{\Sigma fx}{\Sigma f}$$

$$\boxed{\lambda = \frac{122}{200} = 0.61}$$

By poisson Distribution

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$P(x) = \frac{e^{-0.61} \times (0.61)^{x}}{x!}$$

Expected Frequencies

$$E_i = \frac{N e^{-0.61} \times (0.61)^x}{x!}$$

$$\boxed{E_i = 200 \times \frac{e^{-0.61} \times (0.61)^x}{x!}}$$

(1)

$$E_1 = \frac{200 \times e^{-0.61} \times (0.61)^0}{0!}$$

$$E_1 = \frac{200 \times e^{-0.61}}{1}$$

$E_1 = 108.67 \simeq 109$

$E_2 = 66.29 \simeq 66$

$E_3 = 20.22 \simeq 20$

$E_4 = 4.11 \simeq 4$

$E_5 = 0.63 \simeq 1$

calculation of χ^2, under H_0

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
109	109	0	0	0
65	66	-1	1	0.01515
22	20	2	4	0.2
$\left.\begin{matrix}3\\1\end{matrix}\right\}4$	$\left.\begin{matrix}4\\1\end{matrix}\right\}5$	-1	1	0.2
				$\chi^2 = 0.41515$

$$\boxed{\chi_{(a)} = 0.41515}$$

Degree of freedom $= n - K$

$$d.f = 5 - 1 - 1 - 1$$

$$d.f = 2$$

Level of significance (α)

$$\alpha = 5\% = \frac{5}{100} = 0.05$$

From χ^2-table

$$\chi^2_{tab} = 5.991$$

Conlusion

$\chi^2_{cal} = 0.41515$ and $\chi^2_{tab} = 5.991$

$\because \chi^2_{cal} < \chi^2_{tab}$

$\therefore$ H_0 is accepted

Hence, Fitting of data in poisson Distribution is good

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r\,(\chi^2 > \chi^2(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
~~2~~	~~.0201~~	~~.103~~	~~1.386~~	~~4.605~~	~~5.991~~	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q.2 *When the first proof of 392 pages of a book of 1200 pages were read, the distribution of printing mistakes were found to be as follows:*

No. of mistakes in a page (x)	:	0	1	2	3	4	5	6
No. of pages (f)	:	275	72	30	7	5	2	1

Fit a poisson distribution to the above data and test the goodness of fit.

Null Hypothesis (H_0)

H_0 : Poisson Distribution is a good fit to the data

$$\text{Mean} = \frac{\sum fx}{\sum f} = \frac{189}{392}$$

$$\boxed{\lambda = 0.4821}$$

x	f	fx
0	275	0
1	72	72
2	30	60
3	7	21
4	5	20
5	2	10
6	1	6
	$\sum f = 392$	$\sum fx = 189$

By Poisson Distribution

$$P(x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$$

Expected Frequency

$$E_i = \frac{N \times e^{-\lambda} \lambda^{x}}{x!}$$

$$\boxed{E_i = 392 \times \frac{e^{-0.4821} \times (0.4821)^{x}}{x!}} \quad (1)$$

$$E_1 = 392 \times \frac{e^{-0.4821} \times (0.4821)^{0}}{0!}$$

$$E_1 = 242.05 \approx 242.1$$

$$E_2 = 116.69 \approx 116.7$$

$$E_3 = 28.13 \approx 28.1$$

$$E_4 = 4.52 \approx 4.5$$

$$E_5 = 0.5 \approx 0.5$$

$$E_6 = 0.052 \approx 0.1$$

$$E_7 = 0.0042 \approx 0$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
275	242.1	32.9	1082.41	4.471
72	116.7	-44.7	1998.09	17.121
30	28.1	1.9	3.61	0.128
7	4.5			
5 = 15	0.5 = 5.1	9.9	98.01	19.217
2	0.1			
1	0			
				$\chi^2 = 40.937$

$$\chi^2_{cal} = \sum \left[\frac{(O_i - E_i)^2}{E_i}\right]$$

$$\boxed{\chi^2_{cal} = 40.937}$$

Degree of Freedom $= n - K$

$$d.f = 7 - 1 - 1 - 3$$

$$\boxed{d.f = 2}$$

Level of significance (α)

$$\alpha = 5\% = 0.05$$

From χ^2- table

$$\boxed{\chi^2_{tab} = 5.991}$$

conclusion

$$\chi^2_{cal} = 40.937$$

$$\chi^2_{tab} = 5.991$$

$\because \chi^2_{cal} > \chi^2_{tab}$

$\therefore H_0$ is rejected

Hence

Poisson Distribution is not a good fit to the given data

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r(\chi^2 > \chi^2(\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

UNIT-5 : DPP- 06

Topic : χ^2 – test of goodness of FIT (Part – 2)

Q.1 *Test for goodness of fit of a poisson distribution at 5% level of significance to the following frequency distribution:*

x :	0	1	2	3	4	5	6	7	8
f :	52	151	130	102	45	12	5	1	2

Q.2 *Test for goodness of fit of a poisson distribution at 5% level of significance to the following frequency distribution:*

x :	0	1	2	3	4	5
f :	275	138	75	7	4	1

Q.3 *Records taken of the number of male and female births in 800 families having four children are as follows:*

No. of male births	0	1	2	3	4
No. of female births	4	3	2	1	0
No. of families	32	178	290	236	64

Test whether the data are consistent with the hypothesis that the Binomial law holds and the chance of male birth is equal to that of female birth, namely $p = q = 1/2$. [U.P.T.U. 2009]

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-07

Today's Target

- $\boxed{\chi^2 - test\ of\ as\ a\ test\ of\ independence}$
- PYQ
- DPP

$\chi^2 - test$

$\chi^2 - test\ of\ goodness\ of\ FIT$

L-5, L-6

$\chi^2 - test\ as\ a\ test\ of\ independence$

L-7

$\chi^2 - test\ as\ a\ test\ of\ independence$

Two attributes A and B are given with size r and S respectively.

Steps to solve questions

Step – 1 : Null Hypothesis (H_o)

(H_o) : *Two attributes are independent*

Step – 2: Contingency table : *Simple data is given is tabular form, called contingency table*

a o_1	b o_2
c o_3	d o_4

The expected frequency for any cell in a contingency table can be calculated by using formula

$$\boxed{\text{Expected frequency} = \frac{\text{Row total } X \text{ column total}}{\text{grand total}}}$$

Observed frequency

a o_1	b o_2	$a+b$
c o_3	d o_4	$c+d$
$a+c$	$b+d$	$N = a+b+c+d$

↑
Grand total

Expected frequency

$E_1 = \frac{(a+b)(a+c)}{N}$	$E_2 = \frac{(a+b)(b+d)}{N}$	$a+b$
$E_3 = \frac{(c+d)(a+c)}{N}$	$E_4 = \frac{(c+d)(b+d)}{N}$	$c+d$
$a+c$	$b+d$	$N = a+b+c+d$

Step – 3 : *Calculate χ^2 under H_0*

$$\boxed{\chi^2 = \sum\left[\frac{(O_i - E_i)^2}{E_i}\right]}$$

Step – 4 : *Degree of freedom* $= (r-1)(S-1)$

Level of significance $(\propto)$

$\propto = 5\% = 0.05$

From χ^2-table $\chi^2_{cal} = ?$

Step – 5 : *Conclusion*

(ii) *If* $\chi^2 < {\chi^2}_{0.05}$, *then* H_0 *accepted*

Two attributes A and B are independent

(ii) *If* $\chi^2 > {\chi^2}_{0.05}$, *then* H_0 *is rejected.*

Two attributes A and B are not *independent*

Q. 1 ***From the following table regarding the colour of eyes of father and son, test if the colour of son's eye is associated with that of the father.*** **(A. K. T. U. 2021)**

		Eye colour of son	
		Light	Not light
Eye colour of father	Light	471	51
	Not light	148	230

Null Hypothesis(H_0)

H_0: The colour of son's eye is not associated with that of father

Contingency table

Observed Frequency

471 O_1	51 O_2	522
148 O_3	230 O_4	378
619	281	900

Grand total

Expected Frequency

$E_1 = \frac{522 \times 619}{900} = 359.02$	$E_2 = \frac{522 \times 281}{900} = 162.98$	522
$E_3 = \frac{378 \times 619}{900} = 259.98$	$E_4 = \frac{378 \times 281}{900} = 118.02$	378
619	281	900

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

$$\chi^2 = \frac{(O_1-E_1)^2}{E_1} + \frac{(O_2-E_2)^2}{E_2} + \frac{(O_3-E_3)^2}{E_3} + \frac{(O_4-E_4)^2}{E_4}$$

$$= \frac{(471-359.02)^2}{359.02} + \frac{51-162.98}{162.98} + \frac{(148-259.98)^2}{259.98}$$

$$+ \frac{(230-118.02)^2}{118.02}$$

$$= 34.9290 + 76.9390 + 48.2326$$

$$+ 106.3497$$

$$\boxed{\chi^2_{(a)} = 266.3497}$$

Degree of freedom

$$d.f = (r-1)(s-1)$$
$$= (2-1)(2-1)$$
$$= 1$$

Level of significance

$$\alpha = 5\% = 0.05$$

From χ^2-table

$$\boxed{\chi^2_{tab} = 3.841}$$

Conclusion

$$\chi^2_{cal} = 266.3497, \quad \chi^2_{tab} = 3.841$$

$$\because \chi^2_{cal} > \chi^2_{tab}$$

$\therefore H_0$ is rejected

Hence

The colour of son's eye is associated with that of father's eye

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$$P = P_r (\chi^2 > \chi^2 (\alpha)) = \alpha$$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q. 2 *To test the effectiveness of inoculation against cholera, the following table was obtained*

	Attacked	*Not attacked*	*Total*
Inoculated	30 O_1	160 O_2	190
Not inoculated	140 O_3	460 O_4	600
Total	170	620	790

(The figures represent the number of persons.) ***(A. K. T. U. 2009, 2018)***

Use χ^2 − test to defend or refute the statement that the inoculation prevents attack from cholera.

Null Hypothesis (H_0)

H_0: The inoculation does not prevent from cholera attack

Contingency table

Observed Frequency

30 O_1	160 O_2	190
140 O_3	460 O_4	600
170	620	790

$E_1 = \frac{190 \times 170}{790} = 40.89$	$E_2 = \frac{190 \times 620}{790} = 149.11$	190
$E_3 = \frac{600 \times 170}{790} = 129.11$	$E_4 = \frac{600 \times 620}{790} = 470.89$	600
170	620	790

$$\chi^2 = \sum \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^3}{E_3} + \frac{(O_4 - E_4)^2}{E_4}$$

$$\chi^2 = \frac{(30 - 40.89)^2}{40.89} + \frac{(160 - 149.11)^2}{149.11} + \frac{(140 - 129.11)^2}{129.11} + \frac{(460 - 470.89)^2}{470.89}$$

$$\chi^2 = 2.90 + 0.795 + 0.918 + 0.252$$

$$\boxed{\chi^2_{cal} = 4.865}$$

Degree of freedom

$$d.f = (r-1)(s-1)$$

$$= (2-1)(2-1)$$

$$= 1$$

Level of significance

$$\alpha = 5\% = 0.05$$

From χ^2-table

$$\chi^2_{tab} = 3.841$$

Conclusion

$$\chi_{cal} = 4.865$$

$$\chi^2_{tab} = 3.841$$

$$\because \chi^2_{cal} > \chi^2_{tab}$$

$\therefore H_0$ is rejected

Hence

The inoculation prevent from cholera Attack

Table 4 : CHI-SQUARE

Significant Values χ^2 (α) of Chi-Square Distribution Right Tail Areas for Given Probability α,

$P = P_r (\chi^2 > \chi^2(\alpha)) = \alpha$

And v is Degrees of Freedom (d.f.)

Degree of freedom (v)	*Probability (Level of Significance)*						
	0.99	*0.95*	*0.50*	*0.10*	*0.05*	*0.02*	*0.01*
1	.000157	.00393	.455	2.706	3.841	5.214	6.635
2	.0201	.103	1.386	4.605	5.991	7.824	9.210
3	.115	.352	2.366	6.251	7.815	9.837	11.341
4	.297	.711	3.357	7.779	9.488	11.668	13.277
5	.554	1.145	4.351	9.236	11.070	13.388	15.086
6	.872	2.635	5.348	10.645	12.592	15.033	16.812
7	1.239	2.167	6.346	12.017	14.067	16.622	18.475
8	1.646	2.733	7.344	13.362	15.507	18.168	20.090
9	2.088	3.325	8.343	14.684	16.919	19.679	21.669
10	2.558	3.940	9.340	15.987	18.307	21.161	23.209
11	3.053	4.575	10.341	17.275	19.675	22.618	24.725
12	3.571	5.226	11.340	18.549	21.026	24.054	26.217
13	4.107	5.892	12.340	19.812	22.362	25.472	27.688
14	4.660	6.571	13.339	21.064	23.685	26.873	29.141
15	4.229	7.261	14.339	22.307	24.996	28.259	30.578
16	5.812	7.962	15.338	23.542	26.296	29.633	32.000
17	6.408	8.672	15.338	24.769	27.587	30.995	33.409
18	7.015	9.390	17.338	25.989	28.869	32.346	34.805
19	7.633	10.117	18.338	27.204	30.144	33.687	36.191
20	8.260	10.851	19.337	28.412	31.410	35.020	37.566
21	8.897	11.591	20.337	29.615	32.671	36.343	38.932
22	9.542	12.338	21.337	30.813	33.924	37.659	40.289
23	10.196	13.091	22.337	32.007	35.172	38.968	41.638
24	10.856	13.848	23.337	32.196	36.415	40.270	42.980
25	11.524	14.611	24.337	34.382	37.65	41.566	44.314
26	12.198	15.379	25.336	35.363	38.885	41.856	45.642
27	12.879	16.151	26.336	36.741	40.113	41.140	46.963
28	13.565	16.928	27.336	37.916	41.337	45.419	48.278
29	14.256	17.708	28.336	39.087	42.557	46.693	49.588
30	14.933	18.493	29.336	40.256	43.773	47.962	50.892

Q.3 *In an experiment on the immunisation of cattle from a disease, the following results are obtained. Derive your inferences on the efficiency of the vaccine.* **(A.K.T.U. 2021)**

	Affected	*Unaffected*
Inoculated with vaccine	12 O_1	28 O_2
Not incoulated	13 O_3	7 O_4

$$\chi^2_{cal} = 6.73$$

Engineering Mathematics by Gulshan Sir

UNIT-5 : DPP- 07

Topic : $\chi^2 -$ *test of as a test of independence*

Q.1 *In an experiment on the immunisation of goats from anthrox, the following results were obtained. Derive your inferences on the efficiency of the vaccine.*

	Died anthrox	*Survived*
Inoculated with vaccine	2	10
Not incoulated	6	6

Q.2 *By using $\chi^2 -$ test, find out whether there is any association between income level and type of schooling:*

Income	*Public School*	*Govt. school*
Low	200	400
High	1000	400

Q.3 *The following data is collected on two characters:*

	Smokers	*Non smokers*
Literate	83	57
Illiterate	45	68

Based on this information can you say that there is no relation between habit of smoking and literacy.

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-08

Today's Target

- $Control\ Charts\ for\ variables\ (\bar{X}-chart\ and\ R-chart)$
- PYQ
- DPP

Statistical Quality control (SQC)

- A quality control system performs inspection, testing and analysis to ensure that the quality of the products produced is as per the quality standards.
- It is called SQC ,When statistical techniques are used to control, improve and maintain quality or to solve quality problems.

Techniques of SQC

To control the quality characteristics of the product there are two main techniques.

(i) Process Quality Control

Control the quality of product during production process.

(ii) Product Quality control

Control the quality of product after production.

Control Chart

Control chart is a graphical representation of quality characteristics, which indicates whether the process is under control or not.

Types of control charts

There are many types of control charts. Most commonly used control charts are

(i) Control Charts for variables

They are used to measure quality characteristics

(1) Control chart for sample mean ($\bar{X}-chart$)

(2) Control chart for sample range ($R-chart$)

L-8

(ii) Control Charts for attributes

They are used to maintain and achieve an acceptable quality level

(1) Control chart for fraction defective ($p-chart$)

(2) Control chart for number of defective ($np-chart$)

(3) Control chart for number of defects ($c-chart$)

L-9

Control Limits

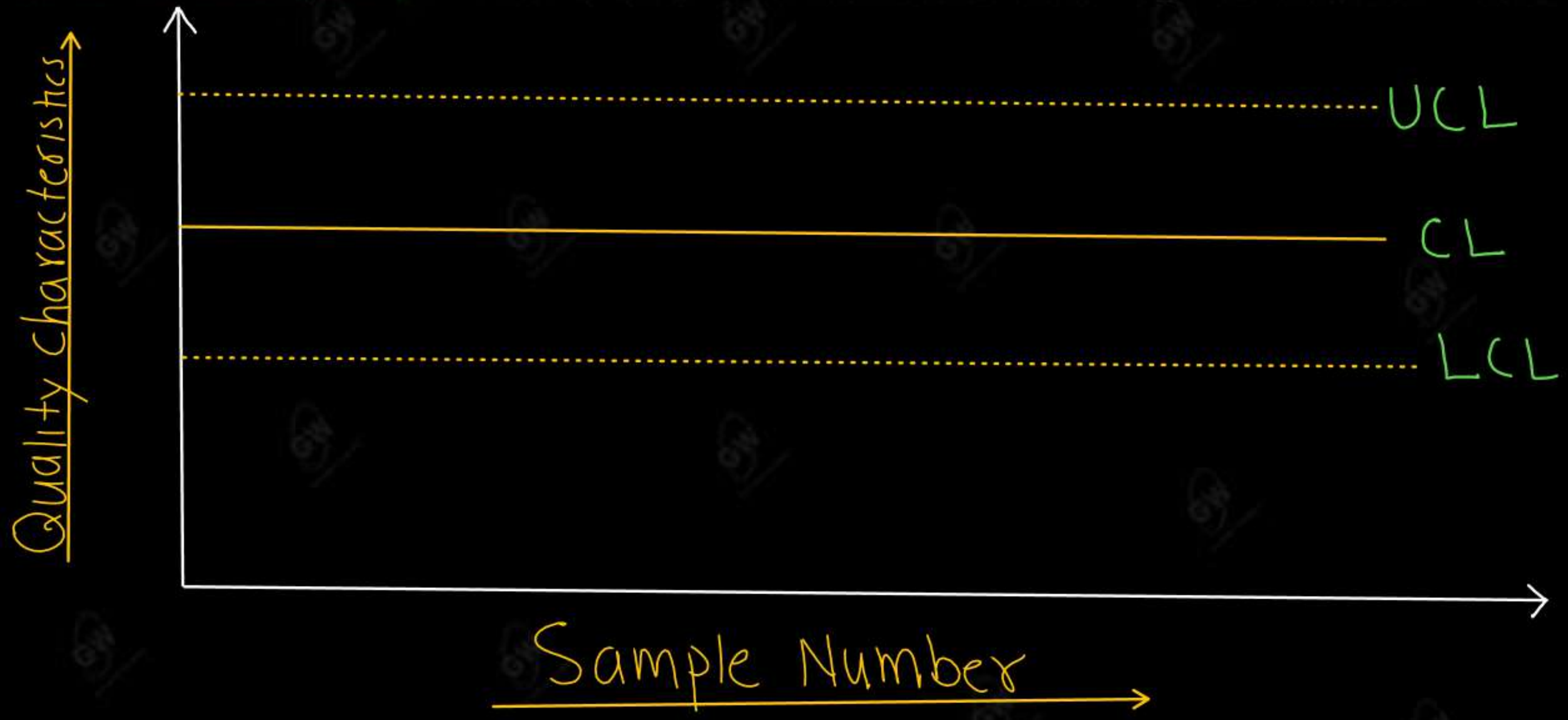
Control limits are the limits within which variations are acceptable

A Control chart consist of 3 Horizontal lines

(i) Control line (C L) : It is a solid line and represent desired control level of porcess.

(ii) Upper control line (UCL) : It is a dotted line and represents upper tolerance limit.

(iii) Lower control line (LCL) : It is a also a dotted line and repreent lower tolerance limit.



Control Charts for variables

Consider a random sample of size n is drawn during manufacturing process

Sample mean

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots\ldots + x_n}{n} = \frac{\sum x}{n}$$

Sample range

$$R = x_{max} - x_{min}$$

Let K consecutive samples are selected with mean $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots \ldots \bar{x}_K$ and range $R_1, R_2, R_3 \; \ldots . \; R_K$

Mean of sample mean ($\bar{\bar{x}}$)

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \ldots\ldots \bar{x}_K}{K} = \frac{\sum \bar{x}}{K}$$

Mean of sample Range

$$\bar{R} = \frac{R_1 + R_2 + R_3 + \ldots . + r_n}{K}$$

3σ Control limits for $\bar{X}$ – chart

$$CL = \begin{cases} \bar{\bar{x}}, & \text{When tolerance limits are not given} \\ \mu, & \text{When tolerance limits are given} \end{cases}$$

When σ is not given

(i) $ULC = \bar{\bar{x}} + A_2\bar{R}$

(ii) $LCL = \bar{\bar{x}} - A_2\bar{R}$

Where A_2 depend on sample size and found from table or given in the question

When σ is given

(i) $ULC = \bar{\bar{x}} + 3\,\sigma_x$

(ii) $LCL = \bar{\bar{x}} - 3\,\sigma_x$

$$\text{Where, } \sigma_x = \frac{\bar{\sigma}}{\sqrt{n}}$$

Control limits for $R-chart$

(i) $CL = \bar{R}$

(ii) $LCL = \bar{R}\, D_3$

(iii) $UCL = \bar{R}\, D_4$

Where D_3, D_4 depend on sample size and found form table or given in the question

Process capability (6σ)

$$6\sigma = 6\frac{\bar{R}}{d_2}$$

Where d_2 depend on sample size and found from table or given in the questions.

Q.1 *The given table shows the values of sample mean $\bar{X}$ and the range (R) for 10 samples each of size 5 each.* AKTU-2022-23

Draw mean and range charts and also comment on the state of control of the process.

Sample No.	1	2	3	4	5	6	7	8	9	10
$\bar{X}$	45	46	48	52	53	37	51	46	47	38
R	4	5	6	7	4	5	7	6	6	4

Here $n = 5$, $A_2 = 0.58$, $D_4 = 2.115$ *and* $D_3 = 0$.

$$\bar{\bar{X}} = \frac{\sum \bar{X}}{n} = \frac{45+46+48+52+53+37+51+46+47+38}{10} = \frac{463}{10}$$

$$\boxed{\bar{\bar{X}} = 46.3}$$

$$\bar{R} = \frac{\sum R}{n} = \frac{4+5+6+7+4+5+7+6+6+4}{10} = \frac{54}{10}$$

$$\boxed{\bar{R} = 5.4}$$

control limits for $\bar{X}$-chart

$CL_{\bar{X}} = 46.3$

$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$

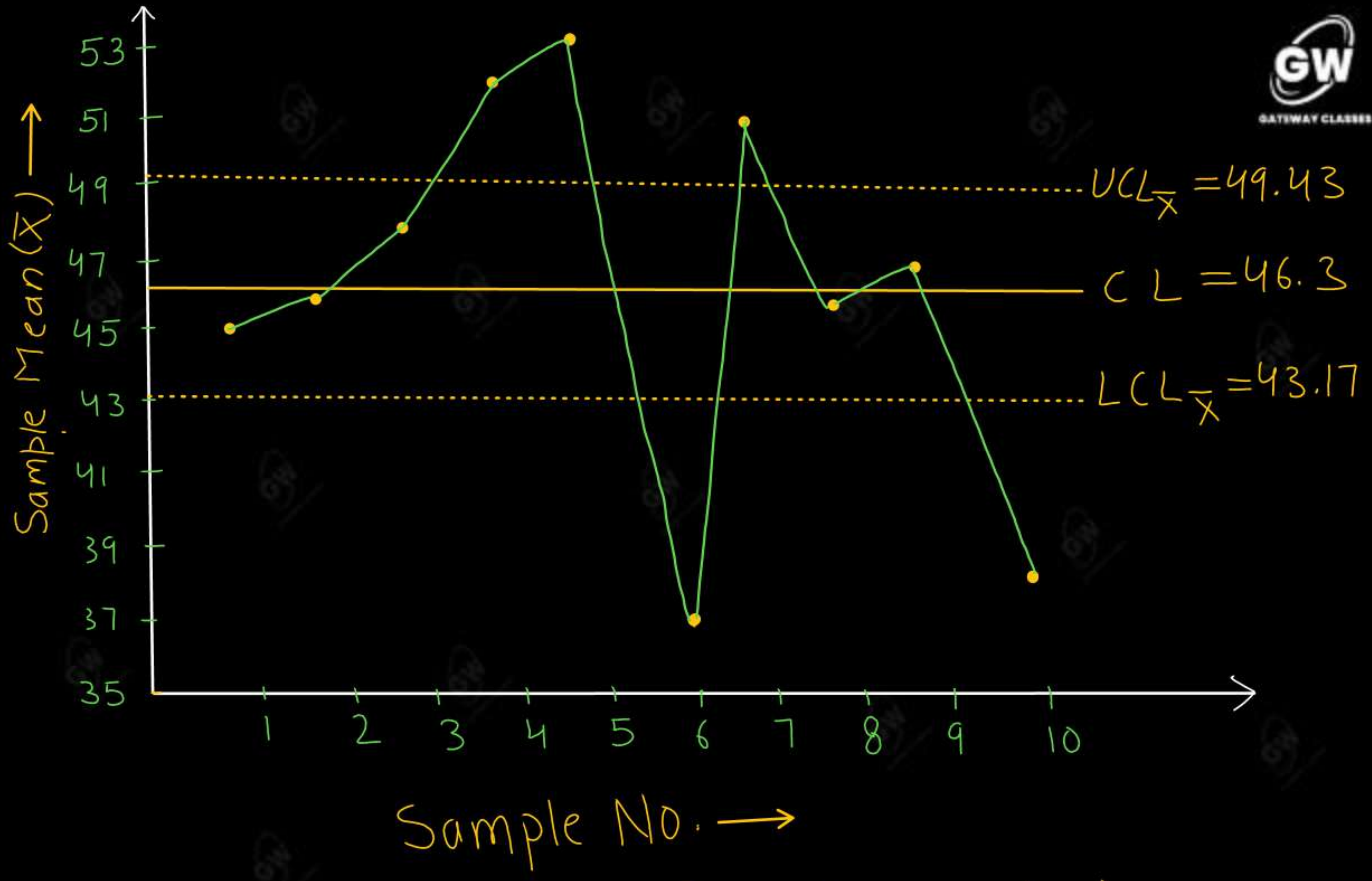
$= 46.3 + 0.58 \times 5.4$

$= 49.43$

$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$

$= 46.3 - 0.58 \times 5.4$

$= 43.17$



Hence, process is not under control

control limits for R-chart

$CL_R = \overline{R}$

$$\boxed{CL_R = 5.4}$$

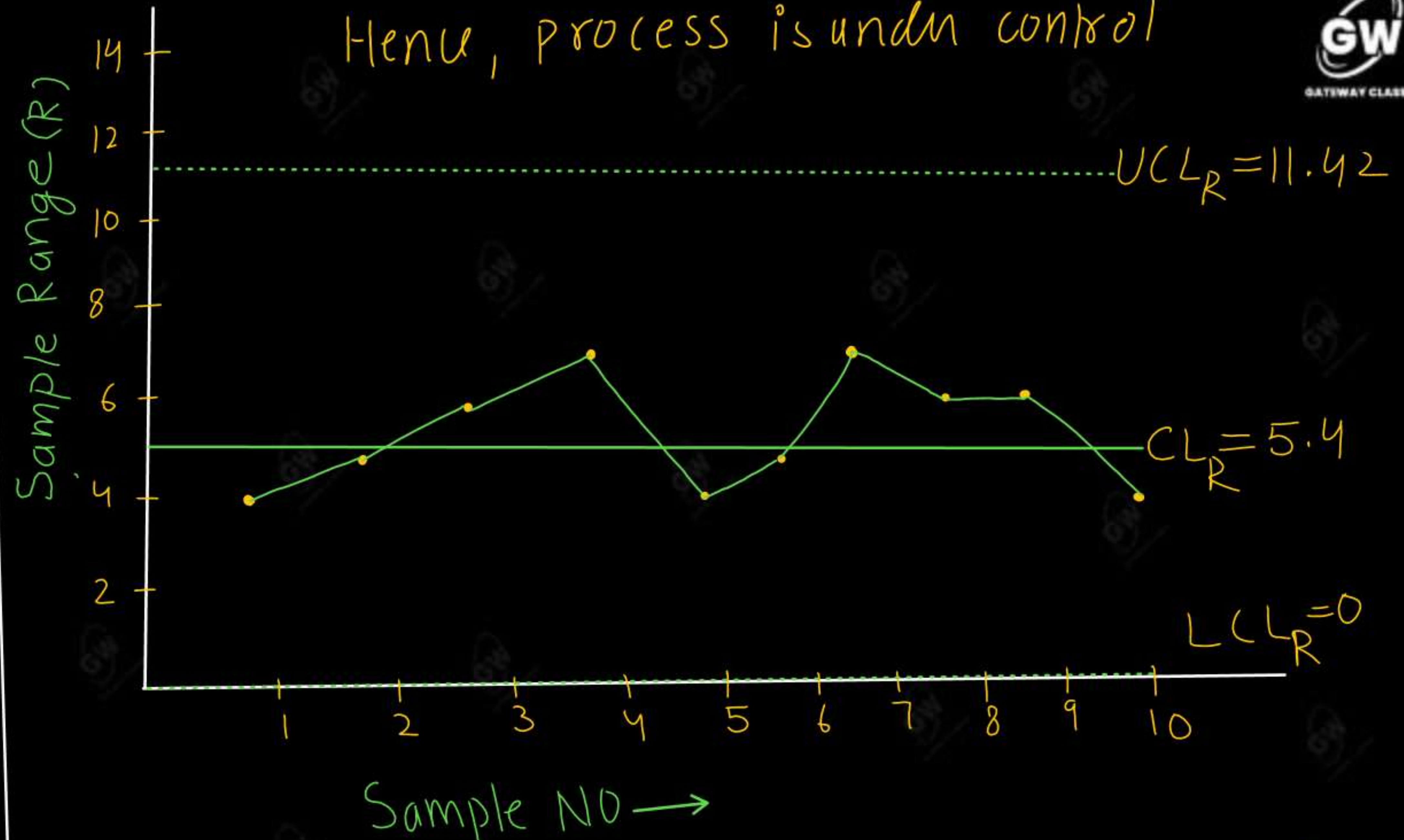
$UCL_R = D_4 \overline{R} = 2.115 \times 5.4$

$$\boxed{UCL_R = 11.42}$$

$LCL_R = D_3 \overline{R} = 0$

$$\boxed{LCL_R = 0}$$

Hence, process is under control



Conclusion: Process is not under Statistical control

Q.2 *The following are the mean lengths and ranges of lengths of a finished product from 10 samples each of size 5. The specification limits for length are 200 ± 5 cm. Construct $\bar{X}$ − chart and R − chart and examine whether then process is undercontrol and state your recommendations.*

Sample No.	1	2	3	4	5	6	7	8	9	10
Mean ($\bar{X}$)	201	198	202	200	203	204	199	196	199	201
Range (R)	5	0	7	3	3	7	2	8	5	6

Assume for $n = 5$, $A_2 = 0.58$, $D_4 = 2.11$ *and* $D_3 = 0$.

Given

$\mu = \bar{\bar{X}} = 200$

$$\bar{R} = \frac{\Sigma R}{n} = \frac{5+0+7+3+3+7+2+8+5+6}{10} = \frac{46}{10}$$

$\bar{R} = 4.6$

Control limits for $\bar{X}$-Chart

$CL_{\bar{X}} = 200$

$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$

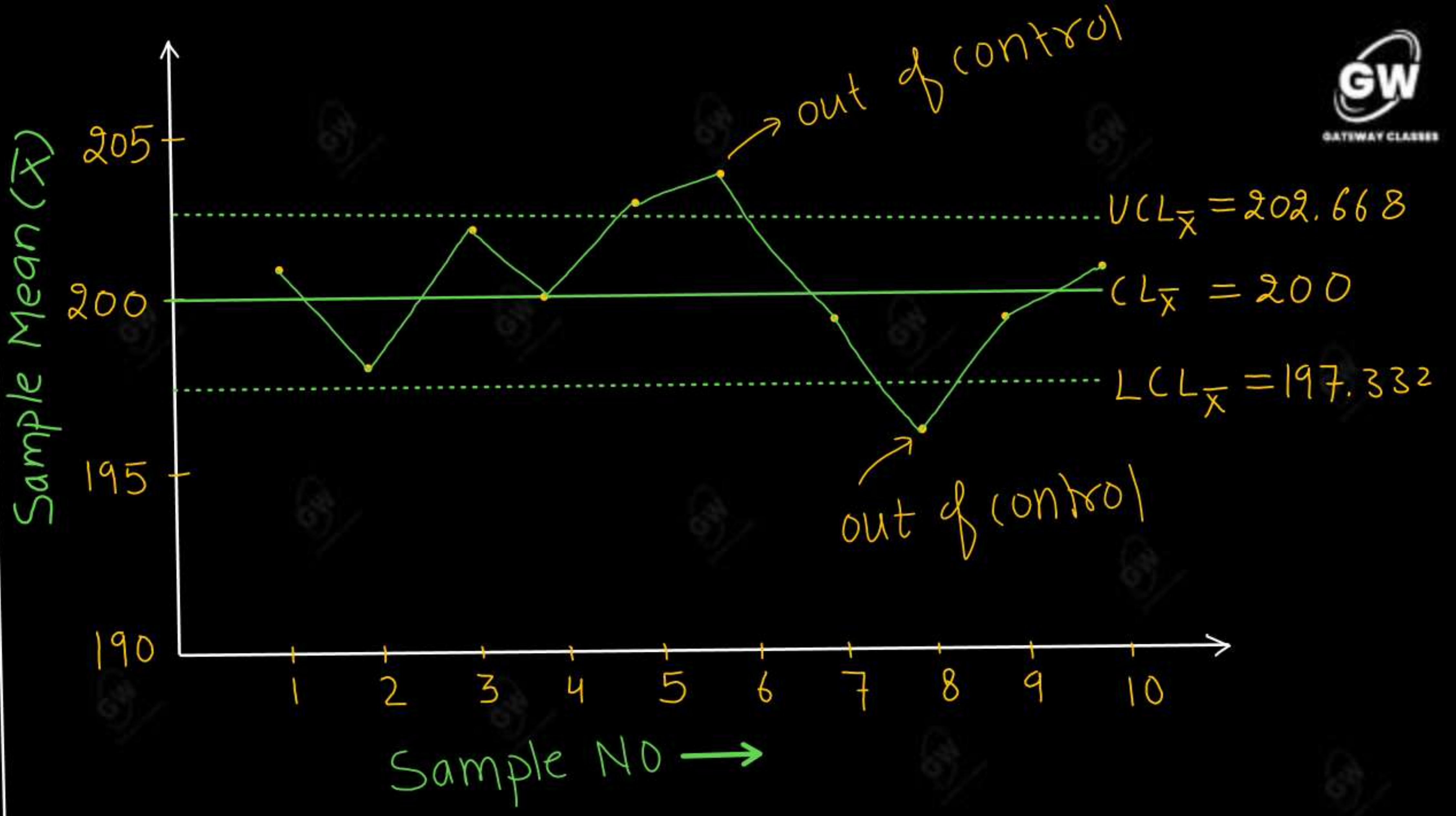
$= 200 + 0.58 \times 4.6$

$= 202.668$

$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$

$= 200 - 0.58 \times 4.6$

$= 197.332$



All points do not lie within control limits, Hence process is not under control

Control limits for R-chart

$$CL = \bar{R}$$

$$CL = 4.6$$

$$UCL_R = D_4 \bar{R}$$

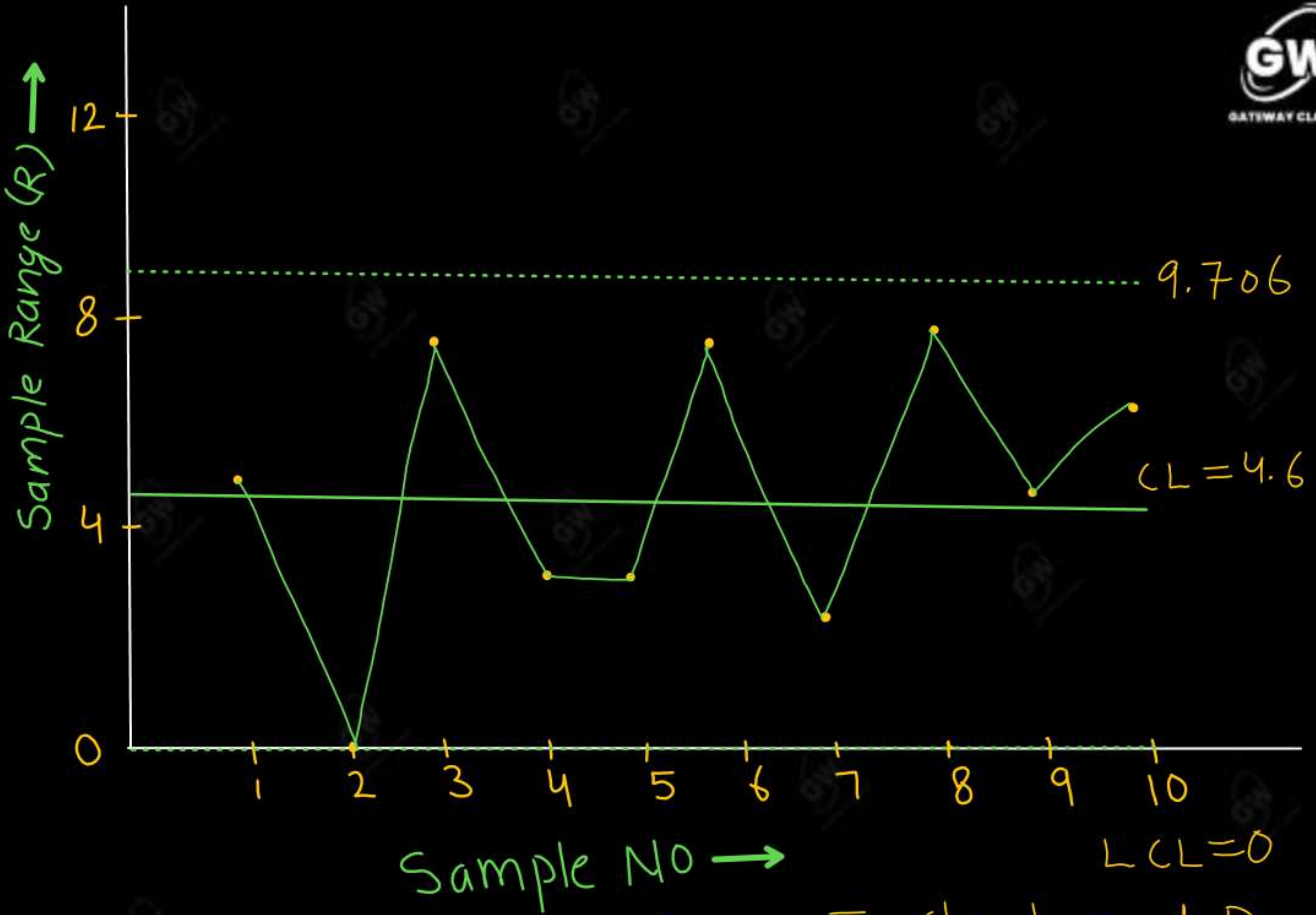
$$= 2.11 \times 4.6$$

$$= 9.706$$

$$LCL_R = D_3 \bar{R}$$

$$= 0 \times 4.6$$

$$= 0$$



Conclusion: It is clear from $\bar{X}$-Chart and R-Chart that process is not under statistical control

Q.3 *A drilling machine bores holes with a mean diameter of* 0.5230 *cm and a standard deviation of* 0.0032 *cm. Calculate the* 2 − *sigma and* 3 − *sigma upper and lower control limits for means of sample of* 4. (*GBTU* − 2010)

Given

$\bar{\bar{X}} = 0.5230$

$\sigma = 0.0032$

$n = 4$

2 − Sigma limit

$UCL = \bar{\bar{X}} + 2\frac{\sigma}{\sqrt{n}}$

$= 0.5230 + \frac{2 \times 0.0032}{\sqrt{4}}$

$\boxed{UCL = 0.5262 \text{ cm}}$

$LCL = \bar{\bar{X}} - 2\frac{\sigma}{\sqrt{n}}$

$= 0.5230 - \frac{2 \times 0.0032}{\sqrt{4}}$

$\boxed{LCL = 0.5198 \text{ cm}}$

3 − Sigma Limit

$UCL = \bar{\bar{X}} + 3\frac{\sigma}{\sqrt{n}}$

$\boxed{UCL = 0.5278 \text{ cm}}$

$LCL = \bar{\bar{X}} - 3\frac{\sigma}{\sqrt{n}}$

$\boxed{LCL = 0.5182}$

UNIT-5 : DPP- 08

Topic :*Control Charts for variables* ($\bar{X}$ − *chart and* R − *chart*)

Q. 1 *The following data shows the value of sample mean* $\bar{X}$ *and arange* R *for* 10 *samples of size* 5 *each. Calculate the values for central line and control limits for* $\bar{X}$ − *chart and* R *chart and determine whether the process is under control.*

Sample no.:	1	2	3	4	5	6	7	8	9	10
Mean $\bar{X}$:	11.2	11.8	10.8	11.6	11	9.6	10.4	9.6	10.6	10
Range R:	7	4	8	5	7	4	8	4	7	9

Assume for $n = 5,\ A_2 = 0.577,\ D_3 = 0$ *and* $D_4 = 2.115.$

Q. 2 *A company manufactures screws to a nominal diameter* 0.500 ± 0.030 *cm. Five samples were taken randomly from the manufactured lots and* 3 *measurements were taken on each sample at different lengths. Following are the readings:*

Sample no.	*Measurement per sample x(in cm)*		
	1	2	3
1	0.488	0.489	0.505
2	0.494	0.495	0.499
3	0.498	0.515	0.487
4	0.492	0.509	0.514
5	0.490	0.508	0.499

Calculate the control limits of $\bar{X}$ *and* R *charts. Draw* $\bar{X}$ *and* R *charts and examine whether the process s in statistical control?*

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

ENGINEERING MATHEMATICS

UNIT-5 : *Statistical Techniques – III*

(Testing a Hypothesis and Statistical Quality Control)

Lec-09

Today's Target

- $Control\ Charts\ for\ Attributes\ (p - chart\ ,\ np - chart\ and\ C - chart)$
- PYQ
- DPP

Types of control charts

There are many types of control charts. Most commonly used control charts are

(i) Control Charts for variables

They are used to measure quality characteristics

(1) Control chart for sample mean ($\bar{X}$ – chart)

(2) Control chart for sample range (R – chart)

L- 8

(ii) Control Charts for attributes

They are used to maintain and achieve an acceptable quality level

(1) Control chart for fraction defective (p – chart)

(2) Control chart for number of defective (np – chart)

(3) Control chart for number of defects (c – chart)

L - 9

Control Charts for attributes

(i) Control charts for Fraction Defective ($P - chart$)

Fraction Defective (p)

$$p = \frac{\text{Number of Defective item in a sample } (d)}{\text{Sample Size } (n)}$$

$$\boxed{p = \frac{d}{n}}$$

$$d = np$$

Total number of defective in all samples $= \sum d = \sum np$

Total number of items in all samples $= \sum n = N$

Mean Fraction Defective ($\bar{p}$)

$$\boxed{\bar{p} = \frac{\text{Total number of defective in all samples}}{\text{Total number of items in all samples}} = \frac{\sum np}{\sum n} = \frac{\sum d}{N}}$$

Control limit on P – Chart are based on Binomial Distribution

By Binomial Distribution

$$\boxed{\sigma_p = \sqrt{\frac{\bar{P}(1-\bar{P})}{n}}}$$

$3\,\sigma$ control limits for P – chart

(i) $CL_p = \bar{P}$

(ii) $UCL_p = \bar{p} + 3\sigma_P$

(iii) $LCL_p = \bar{p} - 3\sigma_P$

Note : If LCL is negative then it is considered as Zero

Q.1 *Following is the data of defective of 10 samples of size 100 each.* (AKTU – 2021 22)

Sample No.	1	2	3	4	5	6	7	8	9	10
Number of defective(d)	15	11	9	6	5	4	3	2	7	1

construct p – chart and state whether the process is in statistical control

Given

$K = 10$

$n = 100$

$\Sigma d = 63$

$N = 10 \times 100 = 1000$

$\bar{P} = \frac{\Sigma d}{N} = \frac{63}{1000} = 0.063$

Sample NO	d	$P = \frac{d}{100}$
1	15	0.15
2	11	0.11
3	9	0.09
4	6	0.06
5	5	0.05
6	4	0.04
7	3	0.03
8	2	0.02
9	7	0.07
10	1	0.01

$$\boxed{\bar{P} = 0.063}$$

$$\sigma_p = \sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$$

$$3\sigma_p = 3\sqrt{\frac{0.063(1-0.063)}{100}}$$

$$\boxed{3\sigma_p = 0.073}$$

3σ control limits

$$CL_p = \bar{P}$$

$$\boxed{CL_p = 0.063}$$

$$UCL_p = \bar{P} + 3\sigma_p$$

$$= 0.063 + 0.073$$

$$\boxed{UCL_p = 0.136}$$

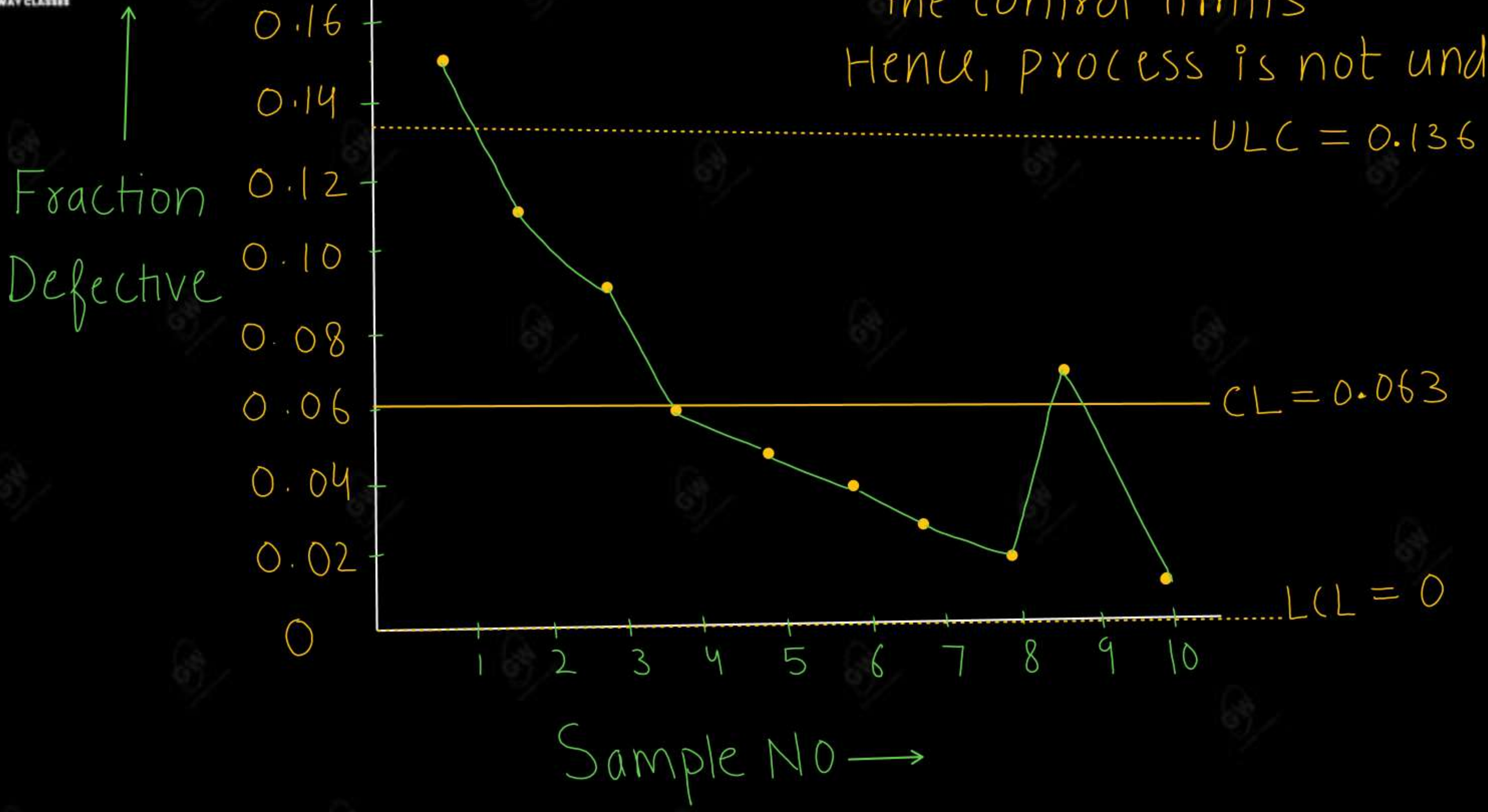
$$LCL_p = \bar{P} - 3\sigma_p$$

$$= 0.063 - 0.073$$

$$= -0.01$$

$$\boxed{LCL_p \simeq 0}$$

conclusion : since one sample point lie outside the control limits
Hence, process is not under statistical control



(ii) Control chart for number of defectives $(np - chart)$

Fraction Defective (p)

$$p = \frac{Number\ of\ Defective\ item\ in\ a\ sample\ (d)}{Sample\ Size\ (n)}$$

$$p = \frac{d}{n}$$

$$d = np$$

Total number of defective in all samples $= \sum d = \sum np$

Total number of items in all samples $= \sum n = N$

Mean Fraction Defective $(\bar{p})$

$$\boxed{\bar{p} = \frac{Total\ number\ of\ defective\ in\ all\ samples}{Total\ number\ of\ items\ in\ all\ samples} = \frac{\sum np}{\sum n} = \frac{\sum d}{N}}$$

Control limits on np – Chart are based on Binomial Distribution

By binomial distribution

$$\sigma_{np} = n\sigma_p = \sqrt{\frac{n^2\bar{P}(1-\bar{P})}{n}}$$

$$\boxed{\sigma_{np} = \sqrt{n\,\bar{P}(1-\bar{P})}}$$

3 σ control limits for np – chart

(i) $CL_{np} = n\bar{p}$

(ii) $UCL_{np} = n\bar{p} + 3\sigma_{np}$

(iii) $LCL_{np} = n\bar{p} - 3\sigma_{np}$

Note : If LCL is negative then it is considered as Zero

Q.2 *Distinguish between the np – chart and p – chart. Following is the data of defectives of 10 samples of size 100 each. Construct np – chart and give your comments.*

Sample no.	:	1	2	3	4	5	6	7	8	9	10
No. of difectives	:	6	9	12	5	12	8	8	16	13	7

(AKTU – 2020 21)

Given

$K = 10$

$n = 100$

$\Sigma d = 96$

$N = 1000$

$\bar{p} = \frac{\Sigma d}{N}$

$\bar{p} = \frac{96}{1000}$

$\boxed{\bar{p} = 0.096}$

$\sigma_{np} = \sqrt{n\bar{p}(1-\bar{p})}$

$3\sigma_{np} = 3\sqrt{100 \times 0.096(1-0.096)}$

$\boxed{3\sigma_{np} = 8.838}$

3σ control limits

$CL = n\bar{p} = 100 \times 0.096$

$\boxed{CL = 9.6}$

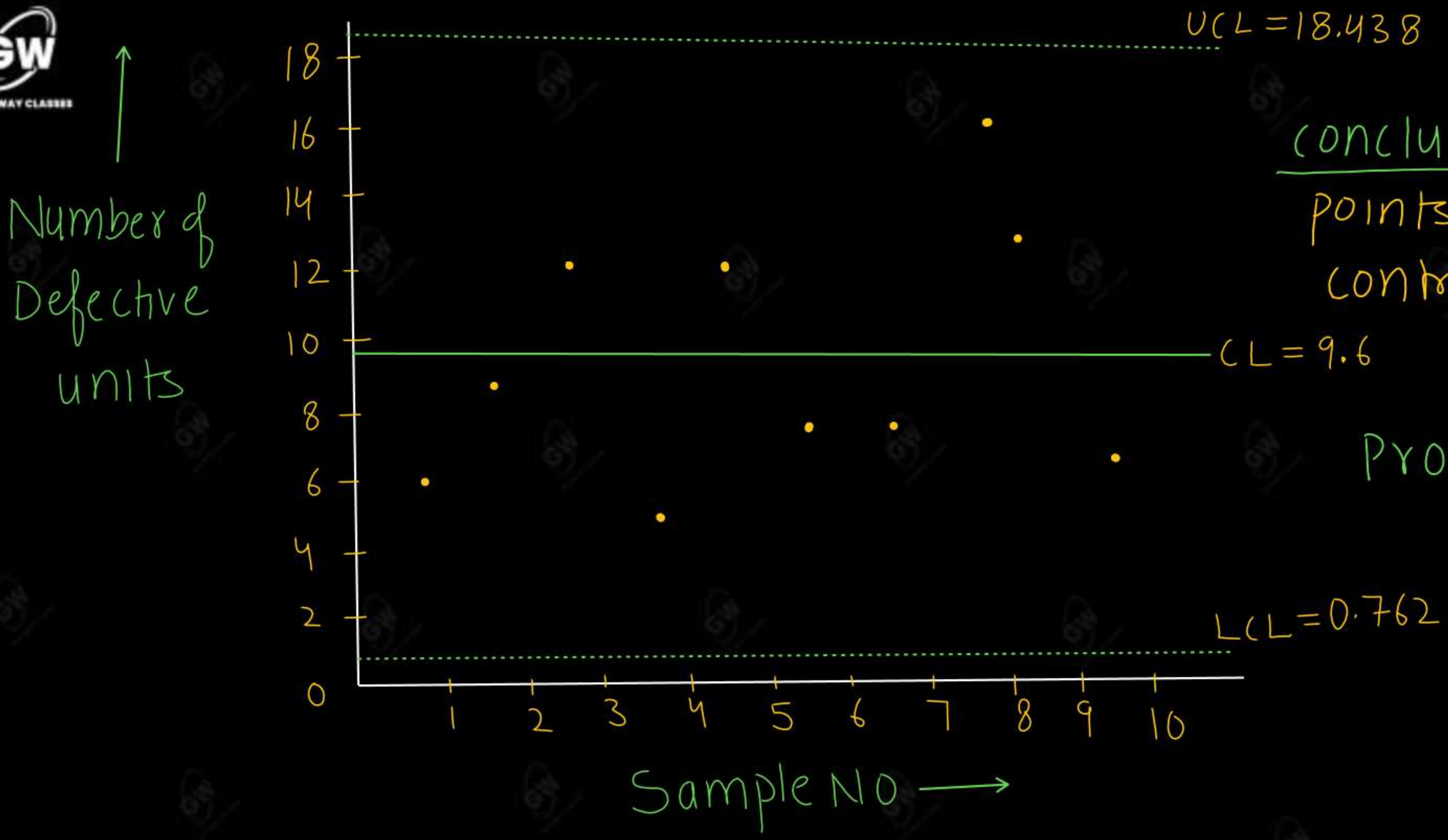
$UCL = n\bar{p} + 3\sigma_{np}$

$= 9.6 + 8.838$

$\boxed{UCL = 18.438}$

$LCL = n\bar{p} - 3\sigma_{np} = 9.6 - 838$

$\boxed{LCL = 0.762}$



Conclusion: Since all points lie within the control limits

Hence Process is under statistical control

(iii) Control chart for number of defects ($C - chart$)

$C -$ Chart is used when number of defects per unit are counted

Number of defects per unit $= C$

Mean of C

$$\boxed{\bar{C} = \frac{\text{Number of defective in all samples}}{\text{Total number of samples}} = \frac{\Sigma C}{K}}$$

Control limit on $C -$ Chart are based on poisson Distribution

Standard deviation

$$\boxed{\sigma_c = \sqrt{\bar{c}}}$$

$3\,\sigma$ control limits for $C - chart$

(i) $CL = \bar{C}$

(ii) $UCL_c = \bar{C} + 3\sigma_c$

(iii) $LCL_c = \bar{C} - 3\sigma_c$

Note : If LCL is negative then it is considered as Zero

Q.3 *The number of defects observed in 10 samples of sheet of equal dimension are given below*

Sample No.	1	2	3	4	5	6	7	8	9	10
Number of defects	2	5	3	5	4	2	3	6	7	4

Draw a control chart for the number of defects and check whether the process is under control or not

Given

$K = 10$

$\Sigma c = 41$

$\bar{c} = \frac{\Sigma c}{K}$

$\bar{c} = \frac{41}{10}$

$\boxed{\bar{c} = 4.1}$

$\sigma_c = \sqrt{\bar{c}}$

$3\sigma_c = 3\sqrt{4.1}$

$\boxed{3\sigma_c = 6.075}$

3σ control limits for c-chart

$CL = \bar{c}$

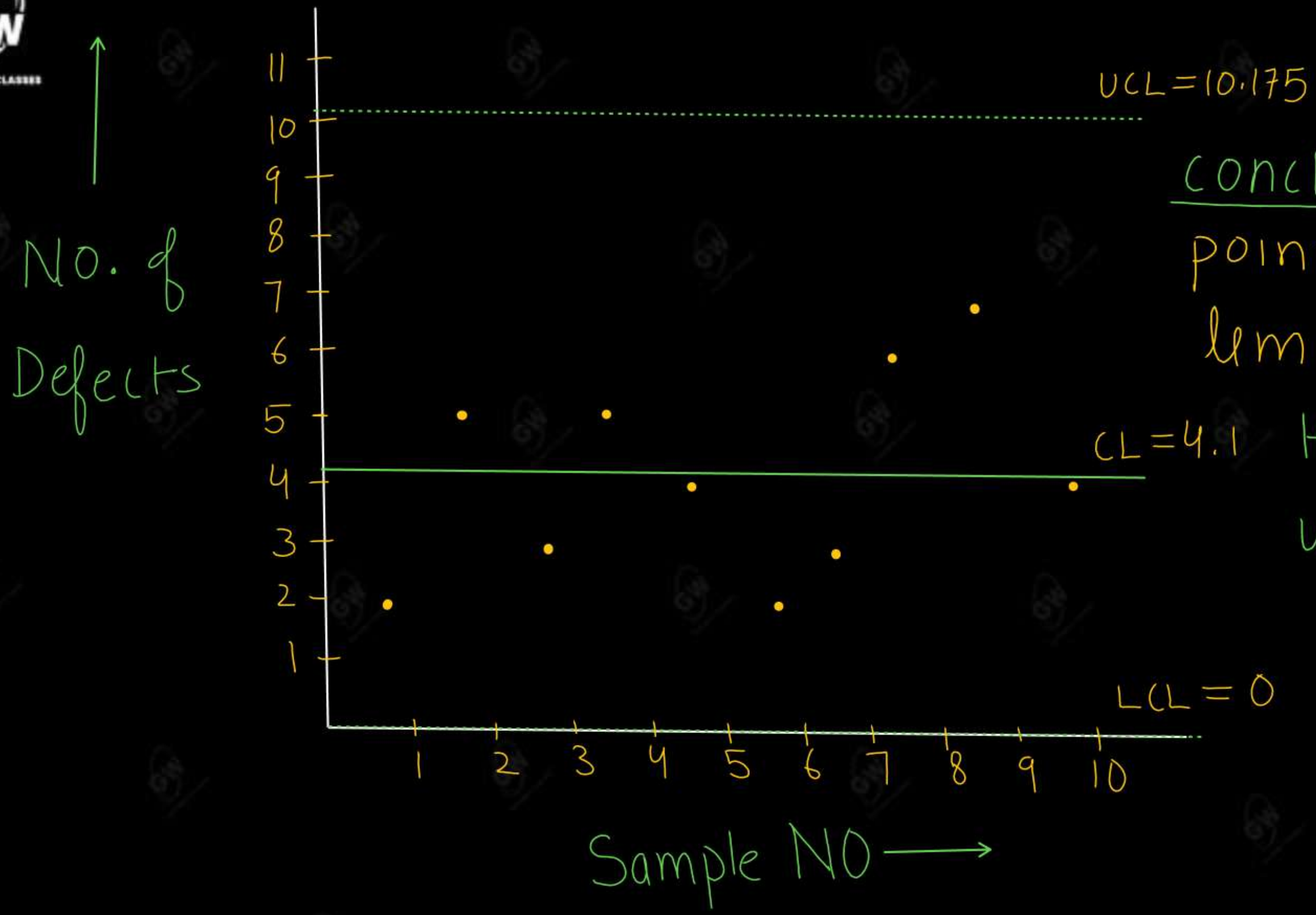
$\boxed{CL = 4.1}$

$UCL = \bar{c} + 3\sigma_c = 4.1 + 6.075$

$\boxed{UCL = 10.175}$

$LCL = \bar{c} - 3\sigma_c = 4.1 - 6.075 = -1.975$

$\boxed{LCL \simeq 0}$



Conclusion: All sample points lie within control limits

Hence, process is under statistical control

Topic : *Control Charts for Attributes* ($p-chart$, $np-chart$ and $C-chart$)

Q.1 *In a factory producing spark plug, the number of defectives found in inspection of 20 lots of 100 each is given below.*

Lot No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No. of defectives	5	10	12	8	6	4	6	3	3	5	4	7	8	3	3	4	5	8	6	10

Construct $p-$ chart and satte whether the process is in statistical control. (AKTU – 2011)

Q.2 *Discuss how control charts can be used in quality control of industrial products. The averge percentage of defectives in 27 samples of size 1500 each was found to be 13.7%. construct $P-$ chart for this situation. Explain how the control chart can be used to control quality.* (AKTU – 2014)

Q.3 *An inspection of 10 samples of size 400 each from 10 lots revealed the following number of defective units. 17,15,14,26,9,4,19,12,9,15. Calculate the control limits for the number of defective units and whether the process is under control or not.* (AKTU – 2021 22)

Q.4 *In a blade manufacturing factory, 1000 blades are examined daily. Draw the np chart for the following table and examine whether the process is under control ?* (AKTU – 2010)

Date	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No. of defective blades	9	10	12	8	7	15	10	12	10	8	7	13	14	15	16

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You