CS & DA

Probability and Statistics

Sampling Theory & Distribution

DPP- 01

Discussion Notes

By- Dr. Puneet Sharma Sir

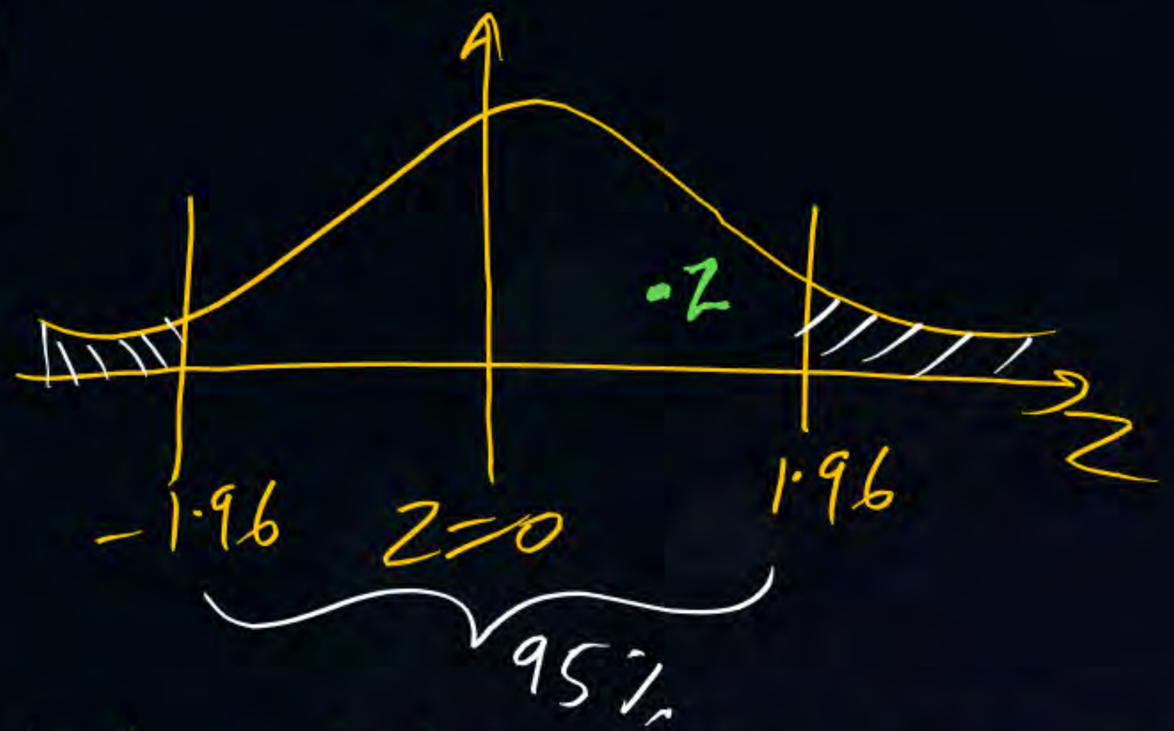
[NAT]

#Q. If a sample of 400 male workers has a mean height of 67.47 inches, is it reasonable to regard the sample as a sample from a large population with a mean height of 67.39 inches and a standard deviation of 1.30 inches at a 5% level of significance ? $Z_\alpha(0.05) = 1.96$

$n = 400,\ \bar{x} = 67.47,\ \mu = 67.39,\ \sigma = 1.30$

$H_0 : \boxed{\mu_0 = 67.39},\ H_1 : \mu_0 \neq 67.39$

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{67.47 - 67.39}{1.3/\sqrt{400}} = 1.231$$



$\because$ Z lies in Confidence Region so H_0 accepted

[NAT]

#Q. A quality engineer wants to check whether there is a difference between population and sample proportion for the rejection rate for parts manufactured on a production line. The rejected part proportion is 5% during production, whereas it was 8% when we selected 50 random samples is this difference in proportion statistically significant if the Level of Significance is 0.05 ? $Z_\alpha(0.05) = 1.96$

$n = 50$, $\tilde{p} = \frac{x}{n}$ = sample proportion = 0.08

$p_0 = \frac{X}{N}$ = Pop. " = 0.05

$H_0 : \boxed{p_0 = 0.05}$ $H_1 : p_0 \neq 0.05$

$q_0 = 0.95$

$$Z = \frac{\tilde{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.08 - 0.05}{\sqrt{\frac{0.05 \times 0.95}{50}}}$$

$= 0.974$

H_0 is Accepted.

[NAT]

#Q. Let's say you're testing two flu drugs A and B. Drug A works on 41 people out of a sample of 195. Drug B works on 351 people in a sample of 605. Are the two drugs comparable? Use a 5% alpha level. i.e $Z_\alpha(0.05) = 1.96$

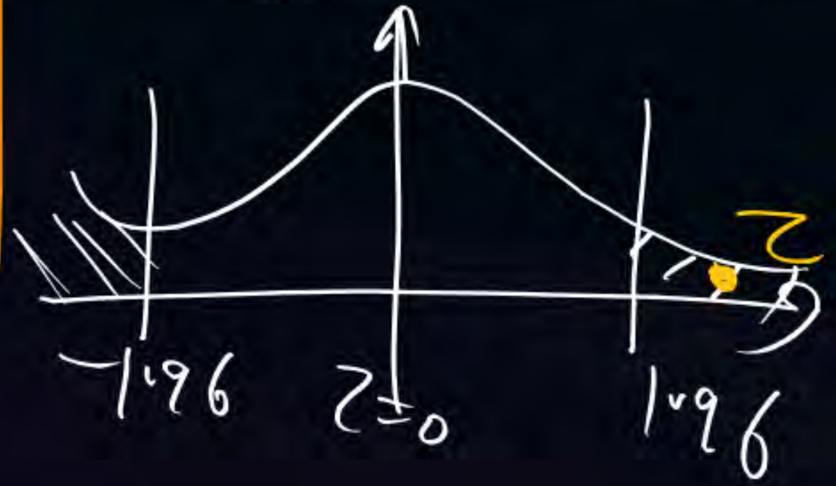
$\tilde{p}_1 = \frac{x_1}{n_1} = \frac{41}{195}$, $H_0 : \boxed{p_1 = p_2}$, $H_1 : p_1 \neq p_2$

$\tilde{p}_2 = \frac{x_2}{n_2} = \frac{351}{605}$

$\tilde{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{41+351}{195+605} = ?$

$\tilde{q} = 1 - \tilde{p} = ?$

$$Z = \frac{\tilde{p}_1 - \tilde{p}_2}{\sqrt{\frac{\tilde{p}\tilde{q}}{n_1} + \frac{\tilde{p}\tilde{q}}{n_2}}} = -8.99$$



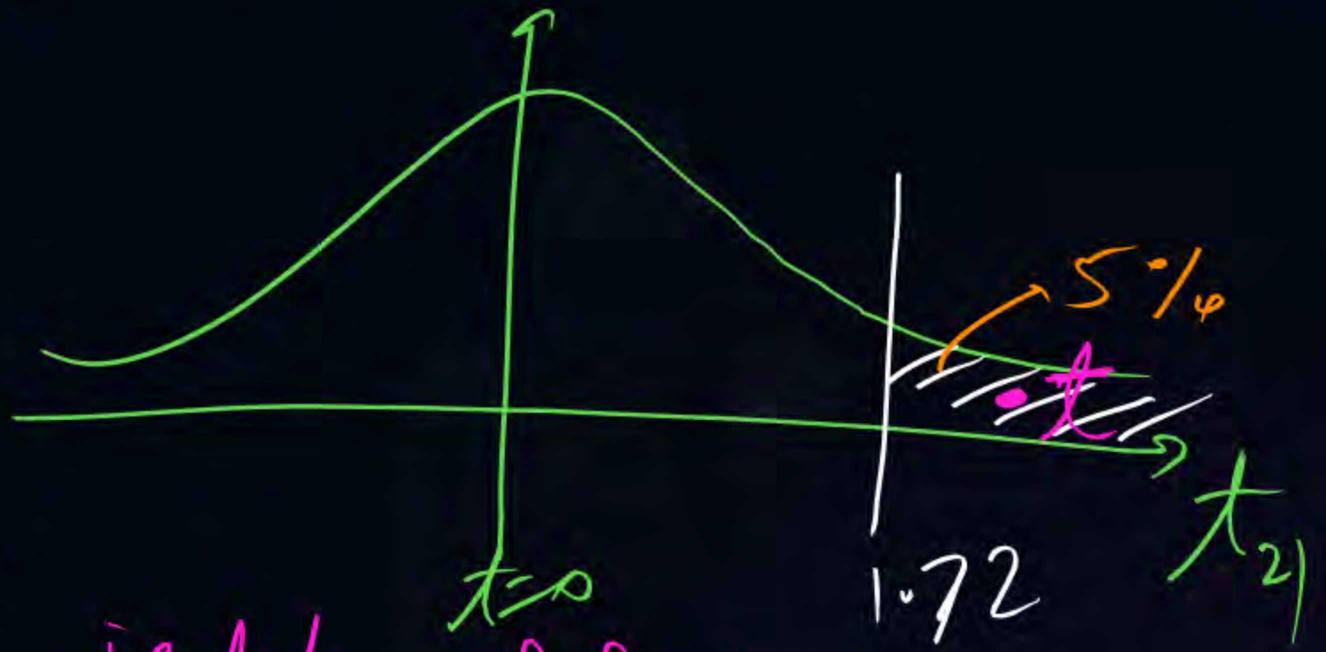
So H_0 is Rejected & H_1 is Accepted

#Q. The mean weekly sales of soap bars in departmental stores was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful ? $t_{21}(0.05) = 1.72$

$\mu = 146.3$, $\bar{x} = 153.7$, $n = 22$ so d.o.f $= n-1 = 22-1 = 21$

$s = 17.2$

$H_0 : \boxed{\mu = 146.53}$, $H_1 : \mu > 146.53$



$$t = \frac{\bar{x} - \mu}{\sqrt{s^2/n}} = \frac{153.7 - 146.53}{\sqrt{\frac{(17.2)^2}{22}}} = 2.017$$

i.e. t lies in R. Region so H_0 is Rejected & H_1 is Accepted

YES Adv Campaign was successful.

[NAT]

$\Rightarrow n \Rightarrow df = 10-1 = 9$

#Q. A random sample of 10 boys had the following I.Q.'s: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of a population mean I.Q. of 100? Find a reasonable range in which most of the mean I.Q. values of samples of 10 boys lie.

$t_9(0.05) = 2.262$

$\mu = 100, \quad n = 10,$

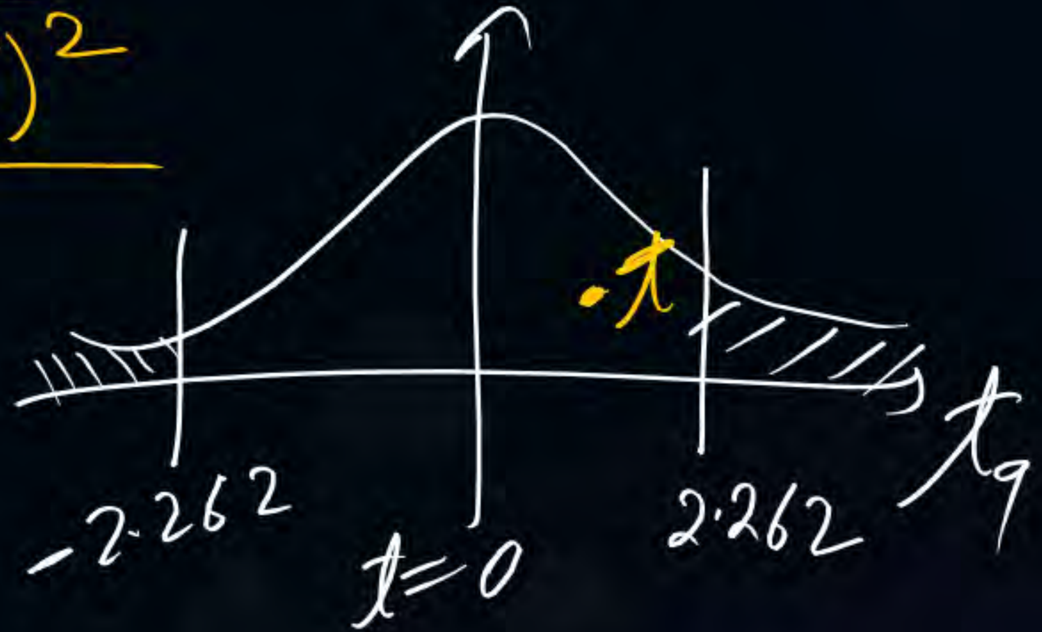
$H_0: \boxed{\mu_0 = 100}, \quad H_1: \mu_0 \neq 100$

$\bar{x} = \frac{\sum x}{n} = \frac{972}{10} = 97.2$

$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{1833.6}{9} = 203.73$

$t = \frac{\bar{x}-\mu_0}{\sqrt{s^2/n}} = \frac{97.2-100}{\sqrt{\frac{203.73}{10}}} = 0.62$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
70		
120		
110		
101		
88		
83		
95		
98		
107		
100		
$\Sigma = 972$		$\Sigma = 1833.60$



i.e. H_0 is Accepted i.e. Pop Av IQ = 100

[NAT]

#Q. Samples of two types of electric light bulbs were tested for length of life and following data were obtained:

	Type I	Type II
Sample No.	$n_1 = 8$	$n_2 = 7$
Sample Means	$\bar{x}_1 = 1{,}234 hrs. = \bar{x}$	$\bar{x}_2 = 1{,}036 hrs. = \bar{y}$
Sample S.D.'s	$s_1 = 36$ hrs $= s_x$	$s_2 = 40$ hrs $= s_y$

Is the difference in the means sufficient to warrant that type 1 is superior to type II regarding length of life? $t_{13}(0.05) = 1.77$

$Df = (n_1 + n_2) - 2$

$\nu = 13$

$H_0 : \mu_x = \mu_y$

$H_1 : \mu_x > \mu_y$

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

where s^2 = Common S. Var.

$$S_1^2 = \frac{\sum(x-\bar{x})^2}{n_1-1} \Rightarrow \sum(x-\bar{x})^2 = 7\times S_1^2$$

$$S_2^2 = \frac{\sum(y-\bar{y})^2}{n_2-1} \Rightarrow \sum(y-\bar{y})^2 = 6\times S_2^2$$

$$S^2 = \frac{\sum(x-\bar{x})^2+\sum(y-\bar{y})^2}{(n_1-1)+(n_2-1)}$$

$$= \frac{7\times(36)^2 + 6\times(40)^2}{8+7-2} = ? = 1659.08$$

$\bar{x} = 1234$, $\bar{y} = 1036$, $n_1 = 8$, $n_2 = 7$

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{S^2}{n_1}+\frac{S^2}{n_2}}} = 9.39$$

H_0 is Rejected & H_1 is accepted

i.e. $\mu_x > \mu_y$

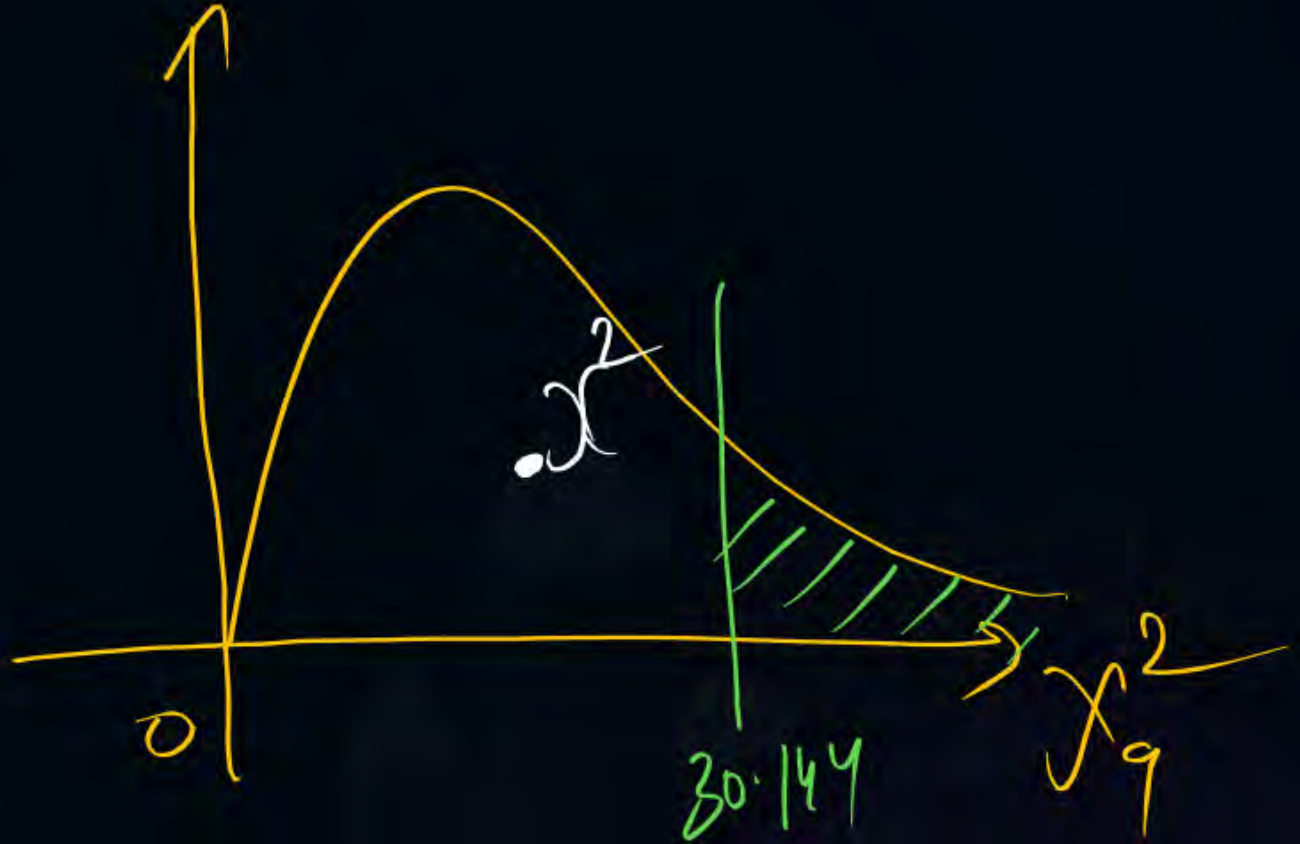
[NAT]

#Q. A random sample of size 20 $=n$ from a normal population gives the sample standard deviation of 6. Test the hypothesis that the population standard deviation is 9. $\chi^2_{19}(0.05) = 30.144$

$s=6,\quad n=20,\quad \nu = n-1 = 19,\quad \sigma = 9$

$H_0 : \boxed{\sigma = 9},\quad H_1 : \sigma \neq 9$

$$\chi^2 = \frac{ns^2}{\sigma^2} = \frac{20\times 36}{81} = 8.89$$



Hence H_0 is Accepted i.e $\boxed{\sigma = 9}$

[NAT]

#Q. A sample of 20 observation gave a standard deviation 3.72. Is this compatible with the hypothesis that the sample is from a normal population with variance 4.35 ? $\chi^2_{19}(0.05) = 30.144$

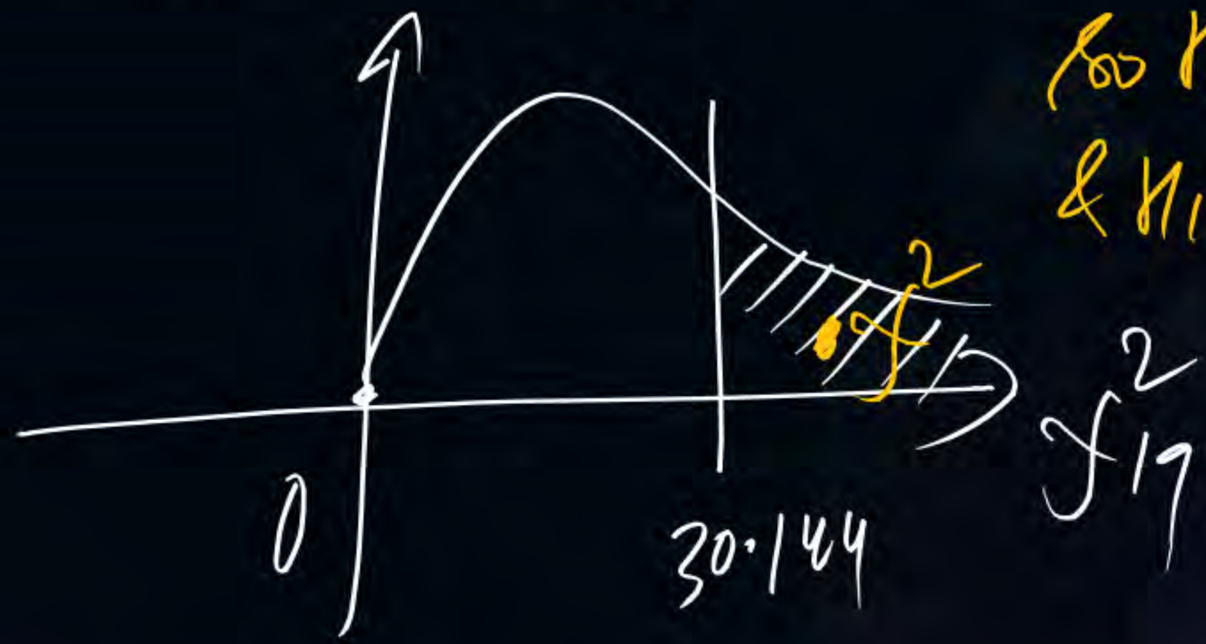
$n = 20,\ \nu = 19$

$s = 3.72,\ \sigma^2 = 4.35$

$H_0 : \sigma^2 = 4.35,$

$H_1 : \sigma^2 \neq 4.35$

$$\chi^2 = \frac{ns^2}{\sigma^2} = \frac{20 \times (3.72)^2}{4.35} = 63.62$$



So H_0 is Rejected & H_1 is Accepted.

[NAT]

#Q. Weights in kg. of 10 students are given 38, 40, 45, 53, 47, 43, 55, 48, 52, 49.can you say that variance of distribution of weights of all students from which the above sample of 10 students was drawn is equal to 20 square kg ?

$n = 10$, $\sigma^2 = 20\ (kg)^2$

$H_0 : \sigma^2 = 20$ ✓

$H_1 : \sigma^2 \neq 20$ ✗

$\nu = n - 1 = 9$

$\bar{x} = \frac{470}{10} = 47$

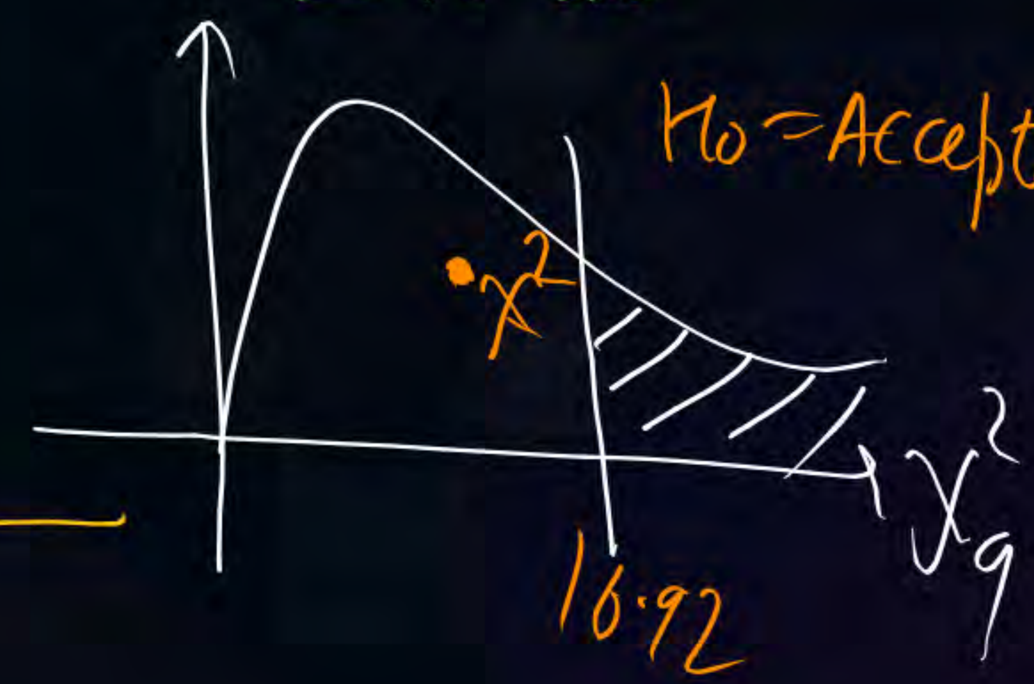
$S^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{280}{9} =$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
38	−9	81
40	−7	49
45	−2	4
53	6	36
47	0	0
43	−4	16
55	8	64
48	1	1
52	5	25
49	2	4
$\Sigma = 470$		$\Sigma = 280$

$\chi^2_9(0.05) = 16.92$

$\chi^2 = \frac{nS^2}{\sigma^2} = \frac{10 \times \frac{280}{9}}{20} = \frac{140}{9} = 15.56$

Ho = Accepted.



[NAT]

#Q. A random sample of size 10 drawn from normal population gave the following values: 65, 72, 68, 74, 77, 61, 63, 69, 73, 71. Test the hypothesis that the population variance is 32.

$\chi_9^2(0.05) = 16.92$

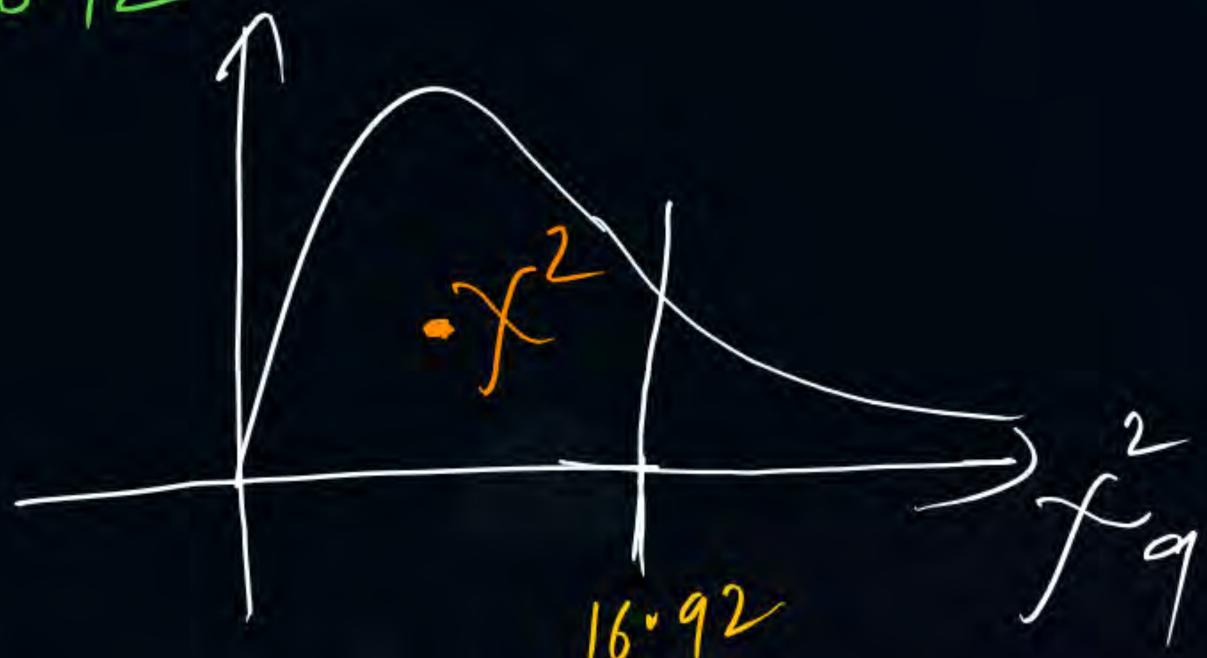
x	$(x-\bar{x})^2$
65	–
72	–
68	–
74	–
77	–
61	–
63	–
69	–
73	–
71	–
$\Sigma = 693$	$\Sigma = 234.1$

$n = 10$, $\bar{x} = \frac{693}{10} = 69.3$

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{23.41}{9}$$

$H_0 : \boxed{\sigma^2 = 32}$, $H_1 : \sigma^2 \neq 32$

$$\chi^2 = \frac{ns^2}{\sigma^2} = \frac{10 \times 234.1}{32 \times 9} \simeq 7.33$$



∴ H_0 is accepted

i.e. $\sigma^2 = 32$

[NAT]

#Q. A dog trainer wants to know if golden retrievers and French bulldogs are equally good at learning how to skateboard. She tries to train 40 golden retrievers and 60 French bulldogs to skateboard and finds the following :

	Skateboards	Can't skateboard	
Golden retrievers	20	20	= 40
French bulldogs	50	10	= 60
	70	30	100

Should she reject the null hypothesis that the dog's breed is unrelated to their skateboarding ability? $\chi^2_1(0.05) = 3.84$

(a) She should reject the null hypothesis.

(b) She should fail to reject the null hypothesis

H_0: Dog's Breed is Ind from ability

H_1: " " Dep " "

$$E(20) = \frac{40 \times 70}{100} = 28$$

$$E(20) = \frac{40 \times 30}{100} = 12$$

$$E(50) = \frac{70 \times 60}{100} = 42$$

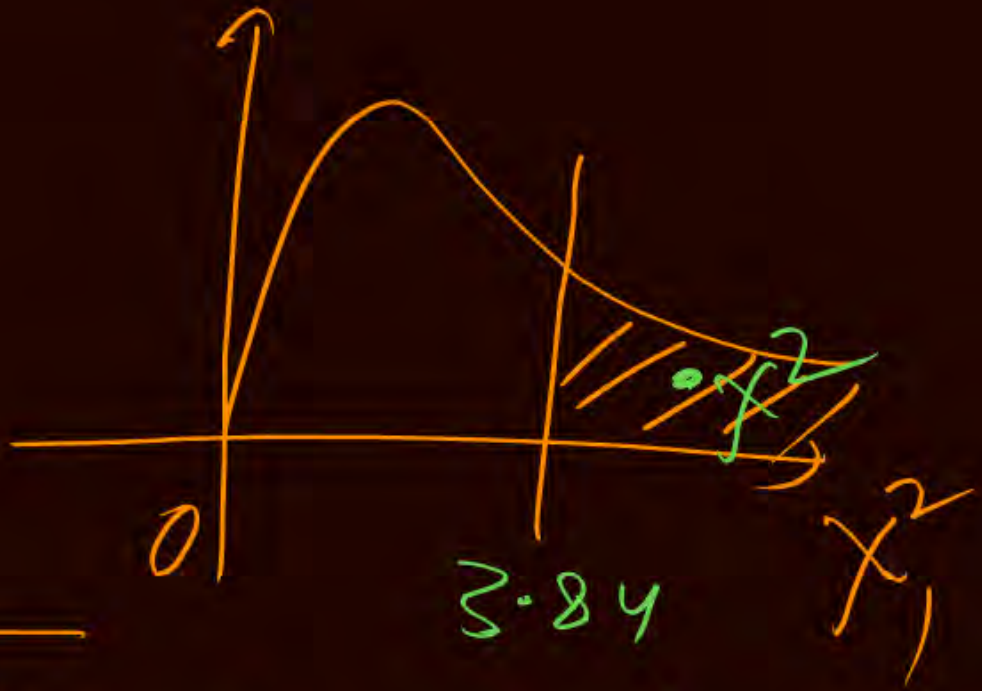
$$E(10) = \frac{30 \times 60}{100} = 18$$

$$\boxed{E_i = \frac{R_i \times C_j}{Total}}$$

$$\nu = (R-1)(C-1) = 1$$

O_i	E_i	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
20	28	64	—
20	12	64	—
50	42	144	—
10	18	64	—
			$\Sigma = 12.7$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 12.7$$



Ho is Rejected

& H_1 is accepted.

[NAT]

#Q. A restaurant reviewer wants to know if three popular burger restaurants are equally recommended by their customers. At each of the three restaurants, he asks 25 random customers whether they would recommend the restaurant to a friend. He finds the following :

	Would recommend	Would not recommend	
Tasty Burgers	~~10~~ 20	5	= 25
Burger Prince	22	3	= 25
Burger Town	18	7	= 25
	= 60	15	$N = 75$

Should he reject the null hypothesis that the proportion of customers recommending the restaurant is the same for the three restaurants? $\chi^2_2(0.05) = 5.99$

(a) He should reject the null hypothesis.

(b) He should fail to reject the null hypothesis,

H_0 : Ind : Prop of Recommendation by Customers for 3 Restaurants would be same

H_1 : Dep : " " " " " " not be same

$$E = \frac{R_i \times C_j}{Total} \qquad df = (R-1)(C-1) = (3-1)(2-1) = 2$$

$$E(20) = \frac{25 \times 60}{75} = 20$$

$$E(22) = \frac{25 \times 60}{75} = 20$$

$$E(18) = \frac{25 \times 60}{75} = 20$$

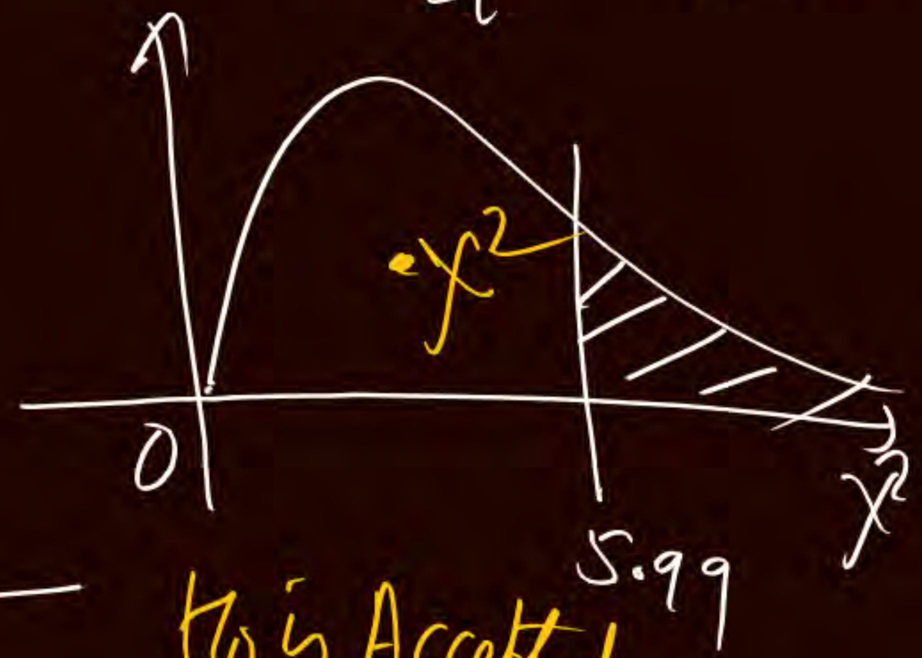
$$E(5) = \frac{25 \times 15}{75} = 5$$

$$E(3) = \frac{25 \times 15}{75} = 5$$

$$E(7) = \frac{25 \times 15}{75} = 5$$

O_i	E_i	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
20	20	0	0
22	20	4	0.2
18	20	4	0.2
5	5	0	0
3	5	4	0.8
7	5	4	0.8
			$\Sigma = 2.0$

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



H_0 is Accepted & H_1 is Rejected.

[NAT]

#Q. You work at a nut factory and you're in charge of quality control. The nut factory produces a nut mix that's supposed to be 50% peanuts, 30% cashews, and 20% almonds . To check that the nut mix proportions are acceptable, you randomly sample 1000 nuts and find the following frequencies :

H_0 : Nut mix has desired Prop of Nuts

H_1 : Nut mix does not have desired Prop of Nut

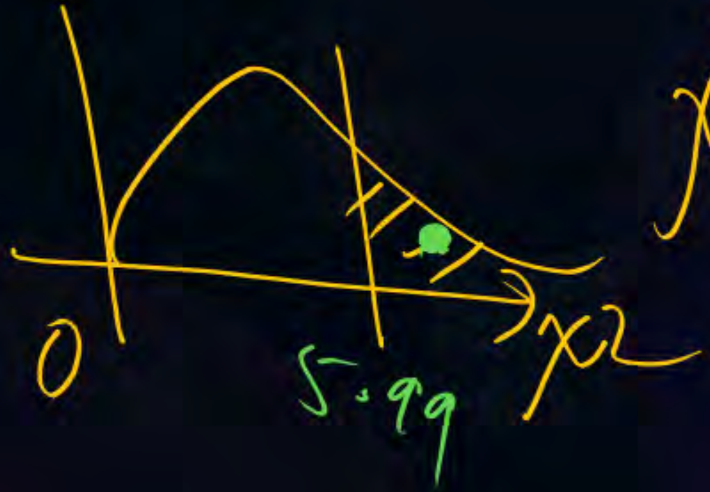
Nut	Frequency O_i	E_i	$(O_i-E_i)^2$	$\frac{(O_i-E_i)^2}{E_i}$
Peanuts	621	500	$(121)^2$	29.28
Cashew	189	300	$(-111)^2$	41.07
Almonds	190	200	$(-10)^2$	0.5
	$N=1000$	$N=1000$		

Should you reject the null hypothesis that the nut mix has the desired proportions of nuts?

$\chi^2_2(0.05) = 5.99$

(a) I should reject the null hypothesis.

(b) I should fail to reject the null hypothesis.



$\chi^2 = \Sigma = ? = 70.85$

H_1 = Accepted

THANK - YOU