STEP#02: INTERPRET CO-EFFICIEN[T] OF CORELATION:

Scatter plot indicicates strong negative relationship between variables

For r:

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2) - (\Sigma x)^2][n(\Sigma y^2) - (\Sigma y)^2]}}$$

$$r = \frac{6(2682) - (37)(482)}{\sqrt{[6(337) - (37)^2][6(39526) - (482)^2]}}$$

$$r = \frac{-1742}{\sqrt{(653)(4832)}}$$

$$r = -0.9806$$

3-FOR SIMPLE LINEAR REGRESSION:

$Y' = a + bx$

Here, $a = \dfrac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$

$$a = \frac{(482)(337) - (37)(2682)}{6(337) - (37)^2}$$

$$a = 96.7840$$

Here, $b = \dfrac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$

$$b = \frac{6(2682) - (37)(482)}{6(337) - (37)^2}$$

$b = -2.6676$

Now, $Y' = a$ ...

$Y' = 9$ ...

$Y' =$ ...

4- PREDICT:

→ Final grade ... would be ...

Here $x = 7$

5- STANDAR[D] ...

Sest = ...

Sest ...

∴ Sest ...

Sest ...

$b = -2.6676$

Now, $Y' = a + bx$

$Y' = 96.78 + (-2.66)x$

$Y' = 96.78 - 2.66x$

4- PREDICT:

→ Final grade when number of absences would be 7

Here $x = 7$ so, $y' = 96.78 - 2.66(7)$

$y' = 78.16$

5- STANDARD ERROR ESTIMATION:

$$S_{est} = \sqrt{\frac{\Sigma y^2 - a\Sigma y - b\Sigma xy}{n-2}}$$

$$S_{est} = \sqrt{\frac{39526 - (96.78)(482) - (-2.66)(268}{6-2}}$$

$S_{est} = \sqrt{3.04}$

$S_{est} = 1.7435$

K = Number of categories

Chi Square for Qualitative data:

→ Which car do you like? Honda civic, Toyota

→ Non - Numeric

Topics

* Test of Goodness - of - Fit
* Test of Independance

TEST FOR GOODNESS - OF - FIT:

When you are testing to see whether a frequency distribution fits a specific pattern, you can use the chi - square goodness - of - fit test. For example, suppose as a market analyst you wished to see whether consumers have any preference among 5 flavours of a new fruit Soda. A sample of 100 people provided these data:

Cherry	Strawberry	Orange	Lime	Grape
32	28	16	14	10

- If there were no preference, you would expect each flavor to be selected with equal frequency
- In the goodness - of - fit test, the degrees of freedom are equal to the number of categories (K) minus 1
- For this example, there are 5 categories (Cherry, Strawberry, Orange, Lime & Grape) hence, the degree of Freedom are $5-1=4$

$\frac{32+28+16+14+10}{5} = 20$ ∵ K = Number of Categories

Preference → one is different among others

No Preference → all are equal → Equal Frequency → average

Formula for the Chi-Square Goodness-of-fit Test

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

with degrees of freedom equal to the number of categories minus 1, and where

O = observed frequency

E = expected frequency

Two assumptions are needed for the goodness-of-fit

Assumptions for Chi-Square Goodness-of-fit Test

1- The data are obtained from a random sample

2- The expected Frequency for each category must be 5 or more

This test is a right-tailed-test.

→ Fruit Soda Flavor Preference:

Is there enough evidence to reject the claim that there is no preferance in the selection of fruit soda flavors, using the data shown previously? Let $\alpha = 0.05$

<u>Solution:</u>

<u>Step 1:</u> State the Hypothesis & identify the claim.

H0: Consumer shows no preferance for flavors (claim)

H1: Consumer show a preferance

<u>Step 2:</u> Find the critical value. The degree of freedom are $5-1=4$, and $\alpha = 0.05$,

<u>Step 3:</u> Compute the test value by subtracting the expected value from the corresponding observed value, squaring the result & dividing the expected value, & finding the sum. The expected value for each category is 20, as shown previously.

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

$$= \frac{(32-20)^2}{20} + \frac{(28-20)^2}{20} + \frac{(16-20)^2}{20} + \frac{(14-20)^2}{2} + \frac{(10-20)^2}{20}$$

$$= 18.0$$

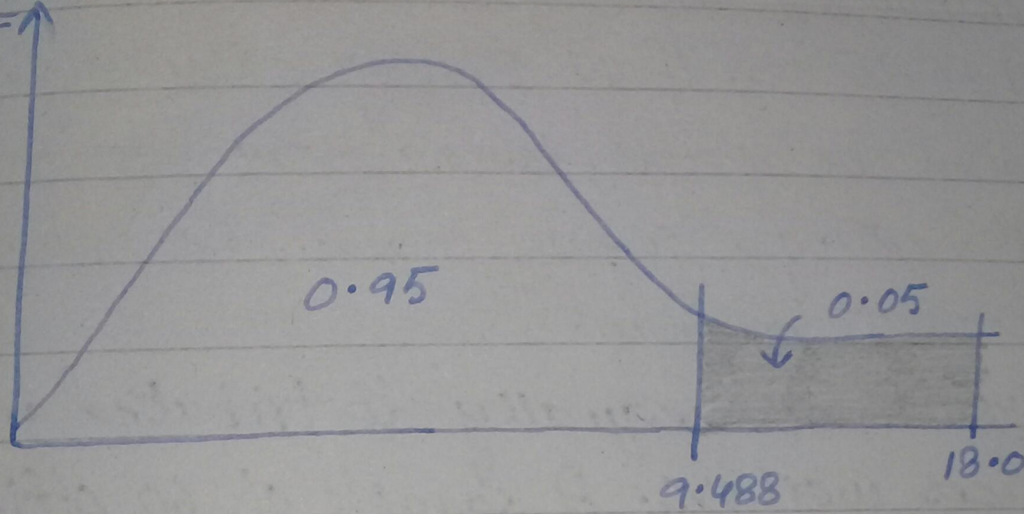
0 is on extreme left

rejection region = right side

Goodness of fitness → right tail

Table G

STEP #04:



Since chi square cal value lies in shaded region (Rejection region) therefore reject Ho.

Step 5:

Summarize the results. There is enough evidence to reject the claim that consumers show no preferance for the flavours.

6- Combatting Midday Drowsiness: A researcher wishes to see if the 5 ways (drinking decaffeinated bevarages, taking a nap, going for walk, eating a sugary snack, other) people to combat midday.
drowsiness are equally distributed among offices workers. A sample of 60 office workers is selected, & the following data are obtained. At $\alpha = 0.10$ can it be concluded that there is no preferance. Why would the result's be of interest to an employer?

Method	Bevarage	Nap	Walk	Snack	Other
Number	21	16	10	8	5

STEP 1: State the Hypotheses & identify the claim.

Ho: Office workers shown no preferance

H1: office workers shown preferance

STEP 2: Find the critical value The degree of freedom are $5-1=4$, & $\alpha=0.1$ Hence, the critical value from Table G in appendix C is 7.779

EV = 12

9442

STEP 3: Co[…] the expected […] observed […] by the Ex[…] The expecte[…]

$\chi^2 = \frac{(21-[…]}{12}$

$\chi^2 = \frac{81}{12}$ […]

$\chi^2 =$ […]

STEP 4: Since, chi[…] in Sha[…] Reject […]

EV = 12

9442

STEP 3: Compute the "test" value by subtracting the expected value from the corresponding observed value, squaring the result & dividing by the Expected value & finding the sum
The expected value for each category is 12

$$x^2 = \Sigma \frac{(O-E)^2}{E}$$

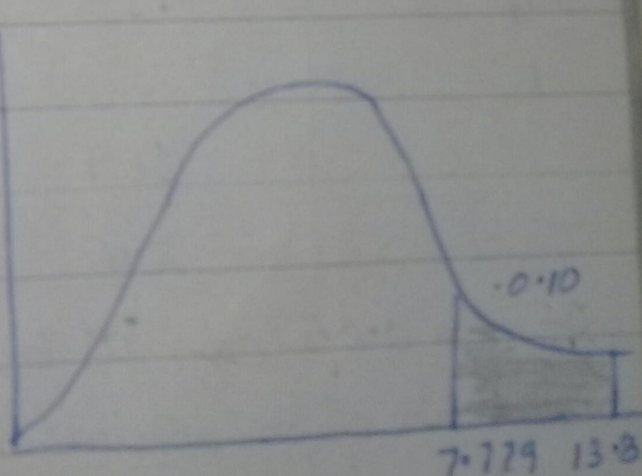
$$x^2 = \frac{(21-12)^2}{12} + \frac{(16-12)^2}{12} + \frac{(10-12)^2}{12} + \frac{(8-12)^2}{12} + \frac{(5-12)^2}{12}$$

$$x^2 = \frac{81}{12} + \frac{16}{12} + \frac{4}{12} + \frac{16}{12} + \frac{49}{12}$$

$$x^2 = 13.833$$

STEP 4:

Since, chi square value lies in shaded region, Reject Ho



TEST OF INDEPENDANCE:

* The chi-square independance test can b used to test the independance of tw variables.

Hospitals and Infections:

A researcher wishes to see if there is a relati between the Hospital & the number of patien infections. A Sample of 3 hospitals was sele & the number of infections for a specific year has been reported. The data are shown next

Hospital	Surgical Site infections	Pnemonia Infections	Bloodstream infections
A	41	27	51
B	36	3	40
C	169	106	109
Total	246	136	200

At $\alpha = 0.05$ can it be concluded that the number of infections is related to the Hospit where they occured?

$E_{11} = \frac{R_1 C_1}{G}$ → Grand Total
↗ Total of Column 1

SOLUTION:

STEP 1: State the Hypothesis & identify the claim

H_0: The number of infections is independent of the Hospital.

H_1: The number of infections is dependant on the Hospital (claim)

STEP 2: Find the critical value at $\alpha = 0.05$ with $(3-1)(3-1) = (2)(2) = 4$ degrees of freedom is 9.488

STEP 3: Compute the test value. First find the expected value

$$E_{1,1} = \frac{(119)(246)}{582} = 50.298$$

$$E_{1,2} = \frac{(119)(136)}{582} = 27.807$$

$$E_{1,3} = \frac{(119)(200)}{582} = 40.893$$

$$E_{2,2} = \frac{(79)(136)}{582} = 18.460$$

$$E_{2,3} = \frac{(79)(200)}{582} = 27.147$$

$$E_{2,1} = \frac{(79)(246)}{582} = 33.391$$

$$E_{3,1} = \frac{(384)(246)}{582} = 162.309$$

$$E_{3,2} = \frac{(384)(136)}{582}$$

$$E_{3,2} = 89.731$$

$$E_{3,3} = \frac{(384)(200)}{582}$$

$$E_{3,3} = 131.958$$

$$x^2 = \Sigma \frac{(O-E)^2}{E}$$

$$x^2 = \frac{(41-50.30)^2}{50.30} + \frac{(27-27.81)^2}{27.81} + \frac{(51-40.89)^2}{40.89} +$$

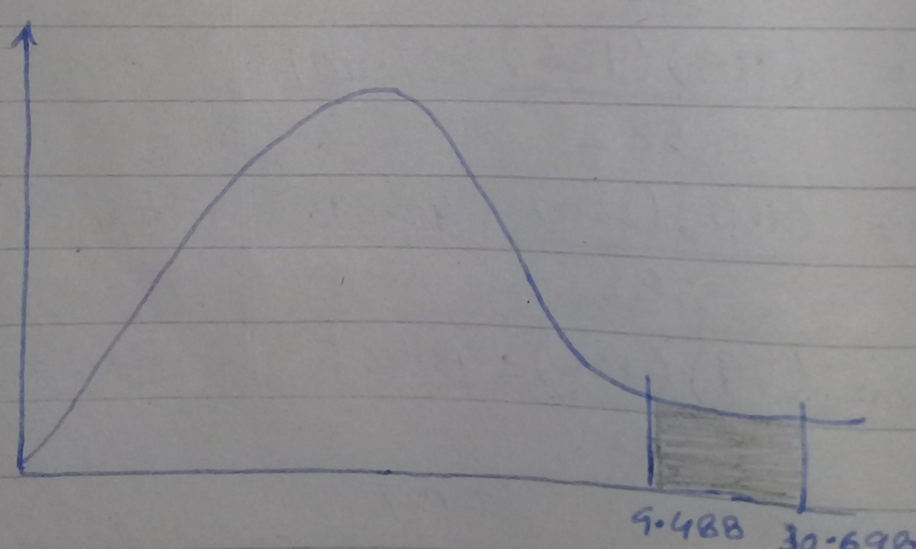
$$\frac{(36-33.39)^2}{33.39} + \frac{(3-18.46)^2}{18.46} + \frac{(40-27.15)^2}{27.15} +$$

$$\frac{(169-\cancel{1}62.31)^2}{162.31} + \frac{(106-89.73)^2}{89.73} + \frac{(109-131.96)^2}{131.96}$$

$$= 1.719 + 0.024 + 2.500 + 0.204 +$$
$$12.498 + 6.082 + 0.276 + 2.950 + 3.995$$

$$= 30.698$$

Step # 4:



Step # 5:

Summarize the results. There is enough evidence to support the claim that the number of infections is related to the Hospital where they occured.

17- Weekend Furniture Sales: A large furniture retailer with stores in three cities had the following results from a special weekend sale. At $\alpha = 0.05$ is there suffucient evidence that the type of furniture sold was dependent upon the store?

	Recliner	Sofa	Love seat	Total
Store 22A	15	12	18	45
Store 22B	20	10	12	42
Store 22C	10	10	10	30
	45	32	40	117

Step 1: State the Hypothesis & identify the claim

Ho: The number of furniture weekend sales is independant of store

H1: The number of furniture dependant on the store

Step 2: Find the critical value at $\alpha = 0.05$ with $(3-1)(3-1) = (2)(2) = 4$ degrees of freedom is

Step 3: Compute the test value: First find the expected values

$E_{1,1} = \frac{(45)(45)}{117} = 17.307$	$E_{1,3} = \frac{(45)(40)}{117}$
$E_{1,2} = \frac{(45)(32)}{117} = 12.30$	$E_{1,3} = 15.384$

$$E_{2,1} = \frac{(42)(45)}{117} = 16.153$$

$$E_{2,2} = \frac{(42)(32)}{117} = 11.487$$

$$E_{2,3} = \frac{(42)(40)}{117} = 14.358$$

$$E_{3,1} = \frac{(30)(45)}{117} = 11.538$$

$$E_{3,2} = \frac{(30)(32)}{117} = 8.205$$

$$E_{3,3} = \frac{(30)(40)}{117} = 10.256$$

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

$$x^2 = \sum \frac{(O-E)^2}{E}$$

$$x^2 = \frac{(15-17.307)^2}{17.307} + \frac{(12-12.30)^2}{12.30} + \frac{(18-15.384)^2}{15.384}$$

$$+ \frac{(20-16.153)^2}{16.153} + \frac{(10-11.487)^2}{11.487} + \frac{(12-14.358)^2}{14.358}$$

$$+ \frac{(10-11.538)^2}{11.538} + \frac{(10-8.205)^2}{8.205} + \frac{(10-10.256)^2}{10.256}$$

$$x^2 = 0.3705 + 0.007 + 0.444 + 0.916 +$$

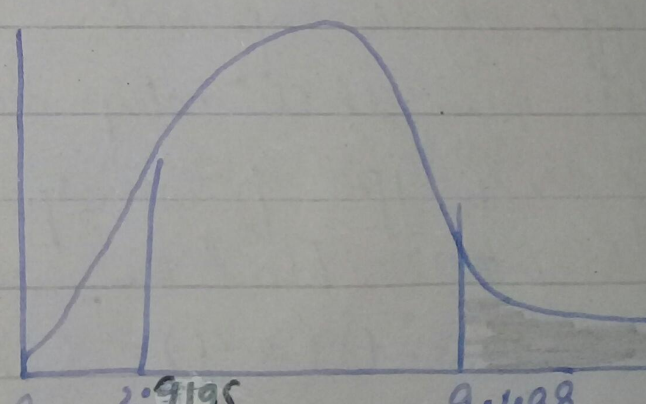
$$0.192 + 0.387 + 0.205 + 0.392 +$$

$$0.006$$

$$x^2 = 2.9195$$

STEP#04:

Since, the value does not lie in the shaded region so, we will accept the claim H_0



16- Organ Transplantation: Listed below is information regarding organ transplantation for 3 different years. Based on these data, is there sufficient evidence at $\alpha = 0.01$ to conclude that a relationship exists between year and type of transplant?

Year	Heart	Kidney/Pancreas	Lung	Total
2003	2056	870	1085	3984
2004	2016	880	1173	4069
2005	2127	903	1408	4438
Total	6199	2653	3639	12491

Step#01: State the Hypothesis & identify the claim

H_0: The relationship exists between year & type of transplant.

H_1: The relationship does not exist between year & type of transplant

Step#02: Find the critical value at $\alpha = 0.01$ with $(3-1)(3-1) = (2)(2) = 4$ degree of freedom

Step#03: Compute the test value. First find the expected values

$$E_{1,1} = \frac{(3984)(6199)}{12491} = 1977.16$$

$$E_{1,2} = \frac{(3984)(2653)}{12491} = 846.173$$

$$E_{1,3} = \frac{(3984)(3639)}{12491} = 1160.65$$

$$E_{2,1} = \frac{(4069)(6199)}{12491} = 2019.352$$

$$E_{2,2} = \frac{(4069)(2653)}{12491} = 864.226$$

$$E_{2,3} = \frac{(4069)(3639)}{12491} = 1185.420$$

$$E_{3,1} = \frac{(4438)(6199)}{12491} = 2202.478$$

$$E_{3,2} = \frac{(4438)(2653)}{12491} = 942.599$$

$$E_{3,3} = \frac{(4438)(3639)}{12491} = 1292.92$$

$$x^2 = \frac{\Sigma (O-E)^2}{E}$$

$$x^2 = \frac{(2056 - 1977.16)^2}{1977.16} + \frac{(870 - 846.173)^2}{846.173} + \frac{(1085 - 1160.65)^2}{1160.65}$$

$$+ \frac{(2016 - 2019.352)^2}{2019.352} + \frac{(880 - 864.226)^2}{864.226} + \frac{(1173 - 1185.420)^2}{1185.420}$$

$$+ \frac{(2127 - 2202.478)^2}{2202.478)} + \frac{(903 - 942.599)^2}{942.599} + \frac{(1408 - 1292.92)^2}{1292.92}$$

$$x^2 = 3.143 + 0.670 + 4.930 + 0.005 + 0.287 + 0.130 +$$
$$2.586 + 1.663 + 10.243$$

$$x^2 = 23.657$$