∴ conclude that the number of spare parts demanded does not depend on the day of the week.

Chi-square test as a test of Independence.

OR Chi-square test for testing independence of attributes:

In many business situations, a market researcher might be interested in understanding the relationship between two variables or to check whether they are independent of each other.

Suppose there are two attributes A and B. A is divided in r classes $A_1, A_2, \ldots A_r$ and B is divided in s classes $B_1, B_2, \ldots B_s$. The various frequencies in different cells of the $r \times s$ contingency table are $(A_i B_j)$, $i = 1, 2, \ldots r$ and $j = 1, 2, \ldots s$.

$A_i = \sum_{j=1}^{s} (A_i B_j)$ = Total number of persons possessing attribute A_i

$B_j = \sum_{i=1}^{r} (A_i B_j)$ = Total number of persons possessing attribute B_j

$\sum_{i=1}^{r} A_i = \sum_{j=1}^{s} = N$ = Total frequency.

We wish to test H_0: Attributes A and B are independent. When H_0 is true, expected frequency corresponding to $(A_i B_j)$ is

Scanned with CamScanner

$$(A_iB_j)_0 = \frac{(A_i)(B_j)}{N}$$

$$\therefore \boxed{\chi^2 = \sum \frac{(O-E)^2}{E} = \sum_{i=1}^{r}\sum_{j=1}^{s} \left[\frac{\left((A_iB_j) - \frac{(A_i)(B_j)}{N}\right)^2}{\frac{(A_i)(B_j)}{N}}\right]}$$

OR

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}}$$

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

We find $\chi^2_{(r-1)(s-1),\,0.05}$ value from the χ^2 table

If $\chi^2 < \chi^2_{(r-1)(s-1),\,0.05}$, we accept H_0 at 5% level of significance and conclude that the attributes A and B are independent.

Ex 1: A certain drug is claimed to be effective in curing colds. In an experiment on 164 people with colds, half of them were given the drug and half of them given sugar pills. The patients reactions to the treatment are recorded in the following table. Test the hypothesis that the drug is no better than sugar pills for curing colds. (Given $\chi^2_{0.05}$ for 2 d.f. = 5.99)

	Helped	Harmed	No effect.
Drug	104	20	40
Sugar Pills	88	24	52

Solⁿ: Let us take the hypothesis that drug is not better than sugar pills for curring colds i.e the two

attributes are independent.
The above information can be arranged in the form of a 2×3 contingency table as follows:

observed frequencies.(O).

	Helped	Harmed	No. effect	Total
Drug	104	20	40	164
sugar pills.	88	24	52	164
Total.	192.	44	92	328

Expected frequency for each cell has been calculated by using the formula.

$$= \frac{\text{Row total} \times \text{Column total}}{\text{Grand total.}}$$

$E_{11} = \frac{164 \times 192}{328} = 96$, $E_{12} = \frac{164 \times 44}{328} = 22$

$E_{13} = \frac{164 \times 92}{328} = 46$, $E_{21} = \frac{192 \times 164}{328} = 96$

$E_{22} = \frac{192 \times 44}{328} = 22$, $E_{23} = \frac{164 \times 92}{328} = 46$

Expected frequencies (E)

	Helped	Harmed	No. effect	Total
Drug	96	22	46	164
sugar Pills	96	22	46	164
Total	192.	44	92	328

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$; O_i = observed frequency, E_i = Expected frequency

$$= \frac{(104-96)^2}{96} + \frac{(20-22)^2}{22} + \frac{(40-46)^2}{46} + \frac{(88-96)^2}{96} + \frac{(24-22)^2}{22}$$

Scanned with CamScanner

$\therefore \chi^2 = 0.667 + 0.182 + 0.783 + 0.667 + 0.182 + 0.783$

$= 3.264$

$d.f = (r-1)(c-1) = (2-1)(3-1) = 2$

As the calculated value of $\chi^2 = 3.264 < \chi^2_{2, 0.05}$ which is 5.99.

$\therefore$ H_0 is accepted.

i.e we conclude that the result of the experiment does not provide any evidence against the hypothesis.

$\therefore$ Drug is no better than sugar pills in curing colds.

Ex-2 In an experiment to study the dependence of hypertension on smoking habit, the following data were obtained on 180 individuals.

	Non - smokers	Moderate smokers	Heavy smokers.
Hypertension	21	36	30
No. hypertension	48	26	19

Test the hypothesis that the presence or absence of hypertension is independent of smoking habit.

Solⁿ: H_0: Hypertension is independent of Smoking habit.

	Non smoker	Moderate smoker	Heavy smoker	Total
Hypertension	21 (33.35)	36 (29.97)	30 (23.68)	87
No Hypertension	48 (35.65)	26 (32.03)	19 (25.32)	93
Total	69	62	49	180

Expected frequencies, written in brackets, are calculated as follow.

$E_{11} = \frac{87 \times 69}{180} = 33.35$ $E_{21} = \frac{93 \times 69}{180} = 35.65$

$E_{12} = \frac{87 \times 62}{180} = 29.97$ $E_{22} = \frac{93 \times 62}{180} = 32.03$

$E_{13} = \frac{87 \times 49}{180} = 23.68$ $E_{23} = \frac{93 \times 49}{180} = 25.32$

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$; O_i = Observed frequency
E_i = Expected frequency

$$= \frac{(21-33.35)^2}{33.35} + \frac{(36-29.97)^2}{29.97} + \frac{(30-23.68)^2}{23.68} + \frac{(48-35.65)^2}{35.65} + \frac{(26-32.03)^2}{32.03} + \frac{(19-25.32)^2}{25.32}$$

$$= 4.57 + 1.21 + 1.69 + 4.28 + 1.14 + 1.58$$

$\chi^2 = 14.47$.

d.f. $= (r-1)(c-1) = (2-1)(3-1) = 2$.

from the table we obtain $\chi^2_{2,0.05} = 5.991$

$\therefore \chi^2 > \chi^2_{2,0.05}$

We reject H_0 at 5% level of significance and conclude that hypertension is dependent on smoking habit.