TUTORIAL-9

1.

Solution:

Given:

- Expected ratio: 9:3:3:1
- Total beans: 1600
- Observed counts: G1 = 882, G2 = 313, G3 = 287, G4 = 118

Step 1: Calculate Expected Counts

Total ratio parts = $9 + 3 + 3 + 1 = 16$

- Expected for G1: $\frac{9}{16} \times 1600 = 900$
- Expected for G2: $\frac{3}{16} \times 1600 = 300$
- Expected for G3: $\frac{3}{16} \times 1600 = 300$
- Expected for G4: $\frac{1}{16} \times 1600 = 100$

Step 2: Chi-Square Test

The formula for the chi-square statistic is:

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

Where O_i is the observed count and E_i is the expected count.

- For G1: $\frac{(882-900)^2}{900} = \frac{(-18)^2}{900} = 0.36$
- For G2: $\frac{(313-300)^2}{300} = \frac{(13)^2}{300} = 0.5667$
- For G3: $\frac{(287-300)^2}{300} = \frac{(-13)^2}{300} = 0.5667$
- For G4: $\frac{(118-100)^2}{100} = \frac{(18)^2}{100} = 3.24$

$$\chi^2 = 0.36 + 0.5667 + 0.5667 + 3.24 = 4.7334$$

Step 3: Degrees of Freedom

Degrees of freedom = $k - 1 = 4 - 1 = 3$

Step 4: Compare with Critical Value

For $\alpha = 0.05$ and 3 degrees of freedom, the critical value from the chi-square table is approximately 7.815.

Since $\chi^2 = 4.7334$ is less than 7.815, we fail to reject the null hypothesis.

Conclusion:

The experimental result supports the theory, as the chi-square statistic is not significant.

2.

Solution:

Given the number of accidents across the days of the week:

- **Mon:** 14
- **Tue:** 18
- **Wed:** 12
- **Thu:** 11
- **Fri:** 15
- **Sat:** 14

Hypotheses:

- **Null hypothesis (H_0):** Accidents are uniformly distributed across the week (i.e., the number of accidents should be the same on each day).
- **Alternative hypothesis (H_1):** Accidents are not uniformly distributed across the week.

Step 1: Calculate the expected number of accidents for each day

Since there are 6 days, if the accidents are uniformly distributed, the expected number for each day is:

$$\text{Expected number of accidents} = \frac{\text{Total number of accidents}}{6}$$

Total number of accidents = $14 + 18 + 12 + 11 + 15 + 14 = 84$

$$\text{Expected for each day} = \frac{84}{6} = 14$$

Step 2: Calculate the chi-square statistic

The formula for the chi-square statistic is:

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

E_i

Where:

- O_i = observed number of accidents
- E_i = expected number of accidents (14)

Calculate for each day:

- For Mon: $\frac{(14-14)^2}{14} = 0$
- For Tue: $\frac{(18-14)^2}{14} = \frac{16}{14} = 1.1429$
- For Wed: $\frac{(12-14)^2}{14} = \frac{4}{14} = 0.2857$
- For Thu: $\frac{(11-14)^2}{14} = \frac{9}{14} = 0.6429$
- For Fri: $\frac{(15-14)^2}{14} = \frac{1}{14} = 0.0714$
- For Sat: $\frac{(14-14)^2}{14} = 0$

$$\chi^2 = 0 + 1.1429 + 0.2857 + 0.6429 + 0.0714 + 0 = 2.1429$$

Step 3: Degrees of Freedom

Degrees of freedom = $k - 1 = 6 - 1 = 5$

Step 4: Compare with the Critical Value

For $\alpha = 0.05$ and 5 degrees of freedom, the critical value from the chi-square table is approximately **11.070.**

Step 5: Conclusion

Since $\chi^2 = 2.1429$ is **less** than 11.070, we **fail to reject** the null hypothesis.

Final Conclusion:

There is **no significant evidence** to suggest that the number of accidents is not uniformly distributed over the days of the week. The accidents appear to be uniformly distributed.

3.

Solution:

Given the data, we can use the chi-square test to see if the distribution of satisfaction is consistent with the business owner's prediction of 80% satisfaction and 20% dissatisfaction.

Observed Data:

Department	Satisfied (O_1)	Dissatisfied (O_2)	Total
Finance	12	7	19
Sales	38	8	46
Human Resources	5	19	24
Technology	1	8	9
Total	56	42	98

Step 1: Expected Values

The business owner's predicted distribution is 80% satisfied and 20% dissatisfied. Calculate the expected values for each category:

- **Expected Satisfied:** $0.80 \times 98 = 78.4$
- **Expected Dissatisfied:** $0.20 \times 98 = 19.6$

For each department, calculate the expected values assuming the predicted proportions:

Department	Expected Satisfied (E_1)	Expected Dissatisfied (E_2)	Total
Finance	$0.80 \times 19 = 15.2$	$0.20 \times 19 = 3.8$	19
Sales	$0.80 \times 46 = 36.8$	$0.20 \times 46 = 9.2$	46
Human Resources	$0.80 \times 24 = 19.2$	$0.20 \times 24 = 4.8$	24
Technology	$0.80 \times 9 = 7.2$	$0.20 \times 9 = 1.8$	9
Total	78.4	19.6	98

Step 2: Chi-Square Calculation

The formula for the chi-square statistic is:

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

Where:

- O_i = observed frequency
- E_i = expected frequency

Calculate for each category:

- For Finance (Satisfied): $\frac{(12-15.2)^2}{15.2} = \frac{(-3.2)^2}{15.2} = 0.675$
- For Finance (Dissatisfied): $\frac{(7-3.8)^2}{3.8} = \frac{(3.2)^2}{3.8} = 2.6947$
- For Sales (Satisfied): $\frac{(38-36.8)^2}{36.8} = \frac{(1.2)^2}{36.8} = 0.0393$
- For Sales (Dissatisfied): $\frac{(8-9.2)^2}{9.2} = \frac{(-1.2)^2}{9.2} = 0.1578$
- For Human Resources (Satisfied): $\frac{(5-19.2)^2}{19.2} = \frac{(-14.2)^2}{19.2} = 10.5125$
- For Human Resources (Dissatisfied): $\frac{(19-4.8)^2}{4.8} = \frac{(14.2)^2}{4.8} = 41.4167$
- For Technology (Satisfied): $\frac{(1-7.2)^2}{7.2} = \frac{(-6.2)^2}{7.2} = 5.3889$
- For Technology (Dissatisfied): $\frac{(8-1.8)^2}{1.8} = \frac{(6.2)^2}{1.8} = 21.2222$

Sum the chi-square values:

$$\chi^2 = 0.675 + 2.6947 + 0.0393 + 0.1578 + 10.5125 + 41.4167 + 5.3889 + 21.2222 = 81.107$$

Step 3: Degrees of Freedom

Degrees of freedom = $(r-1)(c-1) = (2-1)(4-1) = 3$

Step 4: Critical Value

For $\alpha = 0.05$ and 3 degrees of freedom, the critical value from the chi-square distribution table is approximately **7.815**.

Step 5: Conclusion

Since $\chi^2 = 81.107$ is **greater** than 7.815, we **reject** the null hypothesis.

4.

i) How many blocks are involved in the design?

The degrees of freedom for blocks are given as 6 (since it is $6 = 7 - 1$). The number of blocks is equal to the degrees of freedom for blocks plus 1.

So, the number of blocks is:

$$\text{Number of blocks} = 6 + 1 = 7$$

ii) Fill in the blanks in the ANOVA table.

To complete the ANOVA table, we need to calculate the missing values.

Given:

- Degrees of freedom for treatments = 4
- Degrees of freedom for blocks = 6
- Sum of squares for treatments (SS_Treatments) = 14.2
- Sum of squares for blocks (SS_Blocks) = 18.9
- Total degrees of freedom = 34 (i.e., the total number of observations minus 1)
- Total sum of squares = 41.9

Step 1: Calculate the Error Sum of Squares (SS_Error)

From the total sum of squares and the known sums of squares for treatments and blocks, we can calculate the error sum of squares:

$$SS_{\text{Error}} = SS_{\text{Total}} - SS_{\text{Treatments}} - SS_{\text{Blocks}} = 41.9 - 14.2 - 18.9 = 8.8$$

Step 2: Calculate Mean Sum of Squares (MS)

Mean sum of squares (MS) is calculated by dividing the sum of squares by the corresponding degrees of freedom.

$$MS_{\text{Treatments}} = \frac{SS_{\text{Treatments}}}{df_{\text{Treatments}}} = \frac{14.2}{4} = 3.55$$

$$MS_{\text{Blocks}} = \frac{SS_{\text{Blocks}}}{df_{\text{Blocks}}} = \frac{18.9}{6} = 3.15$$

$$MS_{\text{Error}} = \frac{SS_{\text{Error}}}{df_{\text{Error}}} = \frac{8.8}{24} = 0.3667$$

Step 3: Calculate F-statistics

The F-statistic is calculated by dividing the mean sum of squares for treatments or blocks by the mean sum of squares for error.

$$F_{\text{Treatments}} = \frac{MS_{\text{Treatments}}}{MS_{\text{Error}}} = \frac{3.55}{0.3667} = 9.7$$

$$F_{\text{Blocks}} = \frac{MS_{\text{Blocks}}}{MS_{\text{Error}}} = \frac{3.15}{0.3667} = 8.6$$

Completed ANOVA Table:

Sources of Variation	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F-Calculated
Treatments	4	14.2	3.55	9.7
Blocks	6	18.9	3.15	8.6
Error	24	8.8	0.3667	
Total	34	41.9		

iii) Do the data present sufficient evidence to indicate differences among the treatment means? Test at α=0.05.

To test if the treatment means differ significantly, we compare the F-statistic for treatments with the critical F-value from the F-distribution table.

- The degrees of freedom for treatments = 4, and for error = 24.
- The significance level $\alpha = 0.05$.

From the F-distribution table, the critical F-value for $df_1 = 4$ and $df_2 = 24$ at $\alpha = 0.05$ is approximately **2.76**.

Since the calculated F-value for treatments is **9.7**, which is greater than 2.76, we **reject the null hypothesis.**

Conclusion:

There is sufficient evidence to indicate differences among the treatment means at $\alpha = 0.05$.

iv) Do the data present sufficient evidence to indicate differences among the block means? Test at α=0.05.

To test if the block means differ significantly, we compare the F-statistic for blocks with the critical F-value from the F-distribution table.

- The degrees of freedom for blocks = 6, and for error = 24.
- The significance level $\alpha = 0.05$.

From the F-distribution table, the critical F-value for $df_1 = 6$ and $df_2 = 24$ at $\alpha = 0.05$ is approximately **2.53**.

Since the calculated F-value for blocks is **8.6**, which is greater than 2.53, we **reject the null hypothesis.**

Conclusion:

There is sufficient evidence to indicate differences among the block means at $\alpha = 0.05$.

5,

i) Analysis of Variance (ANOVA) for Treatments

Step 1: Organize the Data

Block	A	B	C	D	Total
1	6	4	8	9	33
2	9	3	12	5	25
3	10	2	4	7	23
Total	25	9	24	21	113

Step 2: Calculate the Grand Mean (GM)

$$GM = \frac{Total}{\text{Number of observations}} = \frac{113}{12} = 9.42$$

Step 3: Calculate Sum of Squares for Treatments (SS_Treatments)

- Mean of A: $\frac{6+9+10}{3} = 8.33$
- Mean of B: $\frac{4+3+2}{3} = 3$
- Mean of C: $\frac{8+12+4}{3} = 8$
- Mean of D: $\frac{9+5+7}{3} = 7$

$$SS_{\text{Treatments}} = \sum \left(\frac{(Total_{A,B,C,D} - GM)^2}{\text{Number of blocks}} \right)$$

$SS_{\text{Treatments}} = 3[(8.33 - 9.42)^2 + (3 - 9.42)^2 + (8 - 9.42)^2 + (7 - 9.42)^2] = 3[1.1889 + 40.9764 + 1.7936 + 5.5936] = 156.000$

Step 4: Calculate Sum of Squares for Blocks (SS_Blocks)

- Mean of Block 1: $\frac{6+4+8+9}{4} = 6.75$
- Mean of Block 2: $\frac{9+3+12+5}{4} = 7.25$
- Mean of Block 3: $\frac{10+2+4+7}{4} = 5.75$

$SS_{\text{Blocks}} = 4[(6.75 - GM)^2 + (7.25 - GM)^2 + (5.75 - GM)^2] = 4[0.0278 + 0.6721 + 13.6721] = 56.000$

Step 5: Calculate Error Sum of Squares (SS_Error)

$$SS_{\text{Error}} = SS_{\text{Total}} - SS_{\text{Treatments}} - SS_{\text{Blocks}} = 113 - 156 - 56 = 0$$

Step 6: Degrees of Freedom

- $df_{\text{Treatments}} = 4 - 1 = 3$
- $df_{\text{Blocks}} = 3 - 1 = 2$
- $df_{\text{Error}} = (3 - 1) \times (4 - 1) = 6$

Step 7: Calculate Mean Squares

$$MS_{\text{Treatments}} = \frac{SS_{\text{Treatments}}}{df_{\text{Treatments}}} = \frac{156}{3} = 52$$

$$MS_{\text{Blocks}} = \frac{SS_{\text{Blocks}}}{df_{\text{Blocks}}} = \frac{56}{2} = 28$$

$$MS_{\text{Error}} = \frac{SS_{\text{Error}}}{df_{\text{Error}}} = \frac{0}{6} = 0$$

Step 8: F-Statistic

$$F_{\text{Treatments}} = \frac{MS_{\text{Treatments}}}{MS_{\text{Error}}} = \frac{52}{0} \quad \text{(undefined)}$$

Conclusion

The data provides strong evidence that treatments differ significantly at α = 0.05. Since the error term is zero, the differences among the treatments are highly significant.

ii) Test for Differences Among Block Means

The null hypothesis is that there is no difference among block means, and the alternative is that there is a difference.

Using the same F-statistic procedure:

$$F_{\text{Blocks}} = \frac{MS_{\text{Blocks}}}{MS_{\text{Error}}} = \frac{28}{0} \quad \text{(undefined)}$$

Since the error term is zero, the blocks differ significantly as well.

Conclusion:

The data present sufficient evidence to indicate differences among block means at α = 0.05.

VIVA:

1. **Chi-Square Test**: It is a statistical test used to determine if there is a significant association between categorical variables or if the observed frequencies differ from the expected frequencies. It is used for goodness of fit, test of independence, and test of homogeneity.
2. **Null and Alternative Hypotheses for Chi-Square Test**:
   - Null hypothesis (H0H_0): There is no significant difference between the observed and expected frequencies.
   - Alternative hypothesis (H1H_1): There is a significant difference between the observed and expected frequencies.
3. **Conditions for the Validity of Chi-Square Test**:
   - The data should be in the form of counts or frequencies.
   - Each expected frequency should be at least 5.
   - The samples must be independent.
4. **ANOVA**: Analysis of Variance (ANOVA) is a statistical method used to test differences between the means of three or more groups.

Assumptions:

- Independence of observations.
- Normality of the data within each group.
- Homogeneity of variances (equal variances across groups).