**NON-PARAMETRIC TEST**

**Cochran Q-Test**

Cochran Q-test is a more than two related samples non-parametric test. It is used to test whether three or more matched sets of frequency differ significantly among themselves when the observation can be dichotomized i.e. pass (P) or fail (F), accept (A) or reject (R), increase (I) or decrease (D).

|  |  |
| --- | --- |
| Sample | 1 2 3 ------------------ i ------------------------- n |
| T1 | P P F ------------------ P ------------------------ P |
| T2 | P P P ------------------ P ----------------------- F |
| . | ------------------------------------------------------------------------------------ |
| Tj | F F F ------------------- P ------------------------ P |
| . | ------------------------------------------------------------------------------------ |
| Tk | P P F ------------------ P ------------------------ P |

**Assumptions:**

(a) The distribution of P’s is independent of (I, j) classification.

(b) The data are at least in a nominal scale of are dichotomous form.

**Problem: T o test**

H0: Performance on all the k-samples is similar.

H1: At least one performance of sample is different.

**Mechanism:**

(a) Compute the total number of P’s along row wise and denoted by R I and column wise C j.

|  |  |  |  |
| --- | --- | --- | --- |
| Sample | 1 2 3 ---------- i --------- n | Total () |  |
| T1 | P P F --------- P -------- P |  |  |
| T2 | P P P --------- P ------- F |  |  |
| . | ----------------------------------------------- | . | . |
| T j | F F F -----------P ---------- P |  |  |
| . | ----------------------------------------------- | . | . |
| T k | P P F ---------- P ---------- P |  |  |
| Total | -------- -------- | n = |  |
|  | ---------- -------- |  |  |

(b) Compute square of the row totals and column totals and also computes and.

**Test statistics:** Under H0 test statistics is given by

Q =

=

**Critical region:** Next for a pre-assigned level of significance we obtain from the chi-square table, the critical value of chi-square is or probability value is

P0 = Pr

**Decision:** If Q; we reject H0. Otherwise accept H0.

**Numerical problem:**

**Example (1):** Three puzzles are given to 8 students for solving. Performance of students in each puzzles is as follows:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Puzzles | Students No. | | | | | | | |
| A | B | C | D | E | F | G | H |
| P1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| P2 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| P3 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |

Use Cochran’s Q-test to know if the degree of difficulty of each puzzle is the same. Where 1 = success and 0 = failure.

**Solution:** Given,

No. of row (k) = 3

No. of column (n) = 8

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Puzzles | Students No. | | | | | | | | Row total (R i) |  |
| A | B | C | D | E | F | G | H |
| P1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 4 | 16 |
| P2 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 4 | 16 |
| P3 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 5 | 25 |
| C j | 2 | 1 | 0 | 2 | 2 | 1 | 3 | 2 | =13 | =57 |
|  | 4 | 1 | 0 | 4 | 4 | 1 | 9 | 4 | = 27 |  |

Now,

**Problem: T o test**

H0: Performance on all the three Puzzles is similar.

H1: At least one performance of Puzzle is different.

**Test statistics:** Under H0 test statistics is given by

Q = = = = = 0.3333

**Critical region:** The tabulated value of chi-square at = 0.5 and k-1 = 3- 1 = 2 df. is given by

= = = 5.991

**Decision:** If Q; we accept H0.

**Conclusion:** Therefore performance of all the three puzzles is same.

**Example (2):** Zigzag leaf-hopper that damages the crop plant is havoc to rural peasants of Western Nepal. To get rid from this problem the state government performed an experiment to study the effect of four killing agents (insecticides) on the zigzag leaf-hopper. 32 zigzag leaf-hoppers were cased randomly into bottles grouped into 8 of 4 (matched set) each and certain constant level of each of the insecticide was injected to one bottle of a group and exposed for a specified period of time. The results of the experiment were as follows:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Insecticide | Zigzag leaf-hopper dead or alive | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| W | D | D | D | A | D | D | D | D |
| X | D | A | D | D | D | D | D | D |
| Y | A | A | A | A | A | D | A | D |
| Z | D | A | D | D | A | D | D | A |

Do the insecticides differ significantly so far as killing of zigzag leaf-hopper is concerned? Where D = Dead and A = Alive leaf-hopper.

**Solution:** Given,

k = 4, n = 8

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Insecticide | Zigzag leaf-hopper dead or alive | | | | | | | | Row total |  |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| W | D | D | D | A | D | D | D | D | 1 | 1 |
| X | D | A | D | D | D | D | D | D | 1 | 1 |
| Y | A | A | A | A | A | D | A | D | 6 | 36 |
| Z | D | A | D | D | A | D | D | A | 3 | 9 |
|  | 1 | 3 | 1 | 2 | 2 | 0 | 1 | 1 | =11 | =47 |
|  | 1 | 9 | 1 | 4 | 4 | 0 | 1 | 1 | = 21 |  |

Now,

**Problem: T o test**

H0: Performance on all the four insecticides is similar.

H1: At least one performance of insecticides is different.

**Test statistics:** Under H0 test statistics is given by

Q = = = = = 8.739

**Critical region:** The tabulated value of chi-square at = 0.05 and k-1 = 4- 1 = 3 df. is given by

= = = 7.82

**Decision:** If Q; we reject H0.

**Conclusion:** Therefore performance of at least one insecticide is different.

**Kruskal Wallis H Test**

Kruskal Wallis test is called one way ANOVA test. It is used to test significance difference between three or more than three samples independent or not.

**Problem: T o test**

H0: Performance on all the k-samples is similar.

H1: At least one performance of sample is different.

**Test statistic:** Under H0, test statistics is

H =

In case of tied observations, then test statistic is given by

H =

Where T =

= No. of tied ranks in the tied group or no. of times of tied scores.

= Sum of ranks of sample.

N = No. of all observations

k = No. of samples

**Critical region:** (i) For k = 3 and , the critical value of P­ 0 is obtained from kruskal wallis probability table

P 0 = P (H H calculated)

(ii) Other than, the critical value of H is chi-square value is

**Decision:** (i) If P 0 , we reject H 0. Otherwise accept H 0.

(ii) If H , we accept H 0. Otherwise reject H 0.

**Numerical problem**

**Example (1):** A bacteriologist was interested to study the number of plankton organism inhabiting the lake water. He made hauls of water from three lakes and the following results were found.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Lake | Number of plankton organism | | | | |
| Phewa | 12 | 19 | 16 |  |  |
| Rara | 4 | 8 | 5 | 2 | 3 |
| Taudaha | 14 | 15 | 20 | 11 |  |

Using Kruskal Wallis test. Do the data provide substantial evidence to infer significant variation between lake water?

**Solution:** Given,

k = 3, = 3, = 5, = 4

N = + + = 3 + 5 + 4 = 12

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Lake | Ranks of plankton organism | | | | |  |  |  |
| Phewa | 7 | 11 | 10 |  |  | 28 | 784 | 261.3333 |
| Rara | 3 | 5 | 4 | 1 | 2 | 15 | 225 | 45 |
| Chaudaha | 8 | 9 | 12 | 6 |  | 35 | 1225 | 306.25 |
|  |  |  |  |  |  |  |  | = 612.5833 |

Now,

**Problem: T o test**

H0: Performance on all the three lakes is similar.

H1: At least one performance of lakes is different.

**Test statistic:** Under H0, test statistics is

H =

=

= 18

**Critical region:** The critical value of P­ 0 at = 0.05 obtained from kruskal wallis is

P 0 = P (H H calculated) = P (H 8.1218) = 0.010

**Decision:** (i) If < , we reject H 0.

**Conclusion:** There is significantly variation between water of three lakes.

**Example (2):** An I.Q. test was given to fifteen students taken at random from different faculties of Tribhuvan University. The individual scores (out of 30) were recorded as below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Faculty | Scores | | | | | |
| Science | 22 | 29 | 11 | 17 | - | - |
| Arts | 28 | 15 | 26 | 28 | 11 | 23 |
| Management | 25 | 11 | 20 | 18 | 25 | - |

Carry out Kruskal Wallis ANOVA tests and states your conclusion.

**Solution:** Given,

k = 3, = 4, = 6, = 5

N = + + = 4 + 6 + 5 = 15

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Faculty | Scores | | | | | |  |  |  |
| Science | 8 | 15 | 2 | 5 |  |  | 30 | 900 | 225 |
| Arts | 13.5 | 4 | 12 | 13.5 | 2 | 9 | 54 | 2916 | 486 |
| Management | 10.5 | 2 | 7 | 6 | 10.5 |  | 36 | 1296 | 259.2 |
|  |  |  |  |  |  |  |  |  | = 970.2 |

Rank 2 repeats 3 times, = 3 = = 33 – 3 = 27 – 3 = 24

Rank 10.5 repeats 2 times, = 2 = = 23 – 2 = 8 – 2 = 6

Rank 13.5 repeats 2 times, = 2 = = 23 – 2 = 8 – 2 = 6

Ʃ T = + + = 24 + 6 + 6 = 36

**Problem: To test**

H0: Performance of all three faculties is similar.

H1: At least one performance of faculty is different.

**Test statistic:** Under H0, test statistics is

H =

=

=

=

= 0.5155

**Critical region:** The critical value of H is chi-square value at = 0.05 and k – 1 = 3 – 1 is

= = 5.991

**Decision:** Since H , we accept H 0.

**Conclusion:** There is no significance difference between scores of three faculties.

**Friedman F Test**

The Friedman two-way ANOVA by ranks test is used to test three or more matched samples have been drawn from same population or not.

**Problem:** **To test**

H0: Performance of all the k-samples is similar.

H1: At least one performance of sample is different.

**Test statistic:** Under H0, test statistic is given by

=

In case of tied observations, then test statistic is given by

=

Where,

T = ; no. of times ith rank is repeated.

k = No. of samples.

n = No. blocks or columns or observations.

= Sum of ranks of samples.

**Critical region:** (i) For k = 3 and or k = 4 and 2 the critical value of P­ 0 is obtained from Friedman probability table

P 0 = P (H H calculated)

(ii) Other than, the critical value of H is chi-square value is

**Decision:** (i) If P 0 , we reject H 0. Otherwise accept H 0.

(ii) If , we accept H 0. Otherwise reject H 0.

**Numerical problem**

**Example (1):** A researcher wants to compare the teaching of the standard of three English median schools on the basis of performance of the student final examination scores. The percentage of passers in I to IV grade the schools are presented in the following diagram:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| School | Grade | | | |
| I | II | II | IV |
| Alpha | 89 | 98 | 70 | 80 |
| Sigma | 45 | 76 | 40 | 55 |
| Gamma | 20 | 58 | 35 | 67 |

Using Friedman test, the performances of the schools with respect to pass percentage? Use = 5%.

**Solution:** Given,

k = 3, n = 4

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| School | Grade | | | |  |  |
| I | II | II | IV |
| Alpha | 3 | 3 | 3 | 3 | 12 | 144 |
| Sigma | 2 | 2 | 2 | 1 | 7 | 49 |
| Gamma | 1 | 1 | 1 | 2 | 5 | 25 |
|  |  |  |  |  |  | = 218 |

Now,

**Problem:** **To test**

H0: Performance of all the three samples is similar.

H1: At least one performance of sample is different.

**Test statistic:** Under H0, test statistic is given by

=

=

=

= 54.5 – 48

= 6.5

**Critical region:** The critical value of P­ 0 at = 0.05 level of significance and k = 3, n = 4 is obtained from Friedman probability table

P 0 = P ( F r calculated) = P ( 6.5) = 0.042

**Decision:** Since P0 < , we reject H 0.

**Conclusion:** Therefore at least one performance of all three schools is different.

**Example (2):** In a certain village of South Africa pain patients are duly treated with color visual aids. In other to assess the efficacy of the color visual aids (treatments) sixteen pain patients four of each kind were selected at random. Four visual aids of color are given to the patients of each group. The plan and recovery time of the patients being shown in the diagram appended:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Visual color aid | Pain site | | | | |
| A | B | C | D | E |
| Red | 6 | 8 | 15 | 10 | 8 |
| Yellow | 13 | 21 | 25 | 20 | 28 |
| Green | 5 | 14 | 12 | 10 | 6 |
| Blue | 12 | 10 | 9 | 15 | 8 |

Using Friedman test on these data and interpret your results?

**Solution:** Given,

k = 4, n = 5

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Visual color aid | Pain site | | | | |  |  |
| A | B | C | D | E |
| Red | 2 | 1 | 3 | 1.5 | 2.5 | 10 | 100 |
| Yellow | 4 | 4 | 4 | 4 | 4 | 20 | 400 |
| Green | 1 | 3 | 2 | 1.5 | 1 | 8.5 | 72.25 |
| Blue | 3 | 2 | 1 | 3 | 2.5 | 11.5 | 132.25 |
|  |  |  |  |  |  |  | = 704.5 |

Rank 1.5 repeats 2 time in column D, = 2, = = 23 – 2 = 6

Rank 2.5 repeats 2 time in column E, = 2, = = 23 – 2 = 6

Ʃ T = + = 6 + 6 = 12

Now,

**Problem:** **To test**

H0: Performance of all the four colors is similar.

H1: At least one performance of colors is different.

**Test statistic:** Under H0, test statistic is given by

=

=

=

=

=

= 9.9375

**Critical region:** The critical value of Chi-square at = 0.05 level of significance is obtained from Friedman probability table = = = 7.82

**Decision:** Since, >, we reject H 0.

**Conclusion:** Therefore at least one performance of all four colors is different.