**Practical No: 04**

**OBJECT:**A software company conducted a survey to determine whether there is a relationship between the programming language preference (Python, Java, ) and the level of difficulty experienced by programmers (Easy, Moderate, Difficult). The results are as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Programmers |  | Difficulty Level | l |
| Easy | Moderate | Difficult |
| Python | 25 | 15 | 10 |
| Java | 10 | 20 | 30 |
| C++ | 20 | 25 | 15 |

Perform Chi-Square test to determine if there is a significant association between the programming language preference and the level of difficulty experienced by programmers.

**WORKING EXPRESSION:**

Chi - square is a non-parametric test and it is used to measure data in nominal and ordinal scale.

**Chi -square Test for goodness of fit:**

It is used to test the significant difference between observed frequency and expected frequency.

Let us consider n observation of random variables. X is classified into k classes with their respective frequency.

Different steps in the test are;

**Test statistic:** Under Ho

**χ² = ~** χ2(r-1) (c-1), r = number of rows, c=number of columns

Where, O = Observed frequency

E = Expected frequency

=

**Hypothesis Setting:**

Null hypothesis (Hₒ): There is no significant association between two factors.

Alternative hypothesis (H₁): There is significant association between two factors.

**Level of significance(α):** 5% = 0.05

Practical No: 04

Name: Aakash Shrestha

Roll No.: 02

Subject: Statistics

Date: 2080/04/08

Faculty: BSc. CSIT 3rd Semester

**OUTPUT:**

WEIGHT BY Frequency.

CROSSTABS

/TABLES=Programmers BY Difficulty\_Level

/FORMAT=AVALUE TABLES

/STATISTICS=CHISQ

/CELLS=COUNT EXPECTED

/COUNT ROUND CELL.

**Crosstabs**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Case Processing Summary** | | | | | | | |
|  | Cases | | | | | | |
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| There are three Programmers \* There are three Difficulty level | 170 | 100.0% | 0 | 0.0% | 170 | 100.0% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **There are three Programmers \* There are three Difficulty level Crosstabulation** | | | | | | |
|  | | | There are three Difficulty level | | | Total |
| Easy | Moderate | Difficult |
| There are three Programmers | Python | Count | 25 | 15 | 10 | 50 |
| Expected Count | 16.2 | 17.6 | 16.2 | 50.0 |
| Java | Count | 10 | 20 | 30 | 60 |
| Expected Count | 19.4 | 21.2 | 19.4 | 60.0 |
| C++ | Count | 20 | 25 | 15 | 60 |
| Expected Count | 19.4 | 21.2 | 19.4 | 60.0 |
| Total | | Count | 55 | 60 | 55 | 170 |
| Expected Count | 55.0 | 60.0 | 55.0 | 170.0 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  |  |  |  |
|  |  |  |  |
|  | | | |
|  | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **Chi-Square Tests** | | | |
|  | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 19.683a | 4 | .001 |
| Likelihood Ratio | 19.596 | 4 | .001 |
| Linear-by-Linear Association | 1.404 | 1 | .236 |
| N of Valid Cases | 170 |  |  |
| a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 16.18. | | | |

**RESULTS:**

From the chi-square Tests Table, we obtain the value of X2cal = 19.683.

**Tabulated value:** At 5% level of significance and (r-1) (c-1) = (3-1) (3-1) = 4 d.f. The tabulated value is 9.488.

**Decision: χ²cal =**19.683 > **χ² tab**= 9.488. Hence, we reject Ho. Hence, we can conclude that there is a significant association between the programming language preference and the level of difficulty experienced by programmers.

**CONCLUSION:**

Hence, we have calculated the value of chi-square. There are three programming language preference and three level of difficulty, and the total number of observations was 170. i.e., N=170. As from the result we have rejected Null Hypothesis and the tabulated value was written from chi-square table.