Appendix B: Chi-Square

The Chi-square probability value was used to indicate statistical significance and will be briefly discussed below

The Chi-square test (or $\chi 2$) is a test for assessing the statistical significance of cross-tabulated variables. This is based on a comparison between the observed cell frequencies of a cross-tabulation or joint contingency table with the frequencies that would be expected if the null hypothesis of no relationship were in fact true. The key to the calculation of the χ^2 statistic is to obtain these expected frequencies [87].

If two variables are statistically independent, within categories of the independent variable in a contingency table, there are identical proportions of the dependant variable. Similarly, within each category of the dependant variable the same proportion of the independent responses occurs.

If two variables, X and Y, are statistically independent, the formula for the expected frequency in row I and column j is [87]:

$$\hat{f}_{ij} = \frac{(f_{i.})(f_{.j})}{N} \qquad \textbf{(B1)}$$

Where:

$\hat{f}_{ij}$ = The expected frequency of the cell in row *i* and column *j*.

$f_{i.}$ = The total in the i^{th} row marginal.

$f_{.j}$ = The total in the j^{th} column marginal.

N = The grand total, or sample size of the entire table.

If $\hat{f}_{ij}$ is the expected frequency under the null hypothesis and f_{ij} is the observed frequency for the same cell, the value of the statistic for the table is found by the formula [87]:

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C}\frac{(\hat{f}_{ij} - f_{ij})^2}{\hat{f}_{ij}} \qquad \textbf{(B2)}$$

Where:
C = Number of Columns.
R = Number of rows.

The difference between the observed and expected frequencies in a cell is first squared (to remove the minus signs) and divided by the expected frequency for that cell. After this operation has been performed for all cells, the results are summed for all cells of the table.

The outcome of the chi-square test has a value that can be translated to a probability through the use of a Chi-square table. For the variables to be statistically dependant on one another the value of the probability should be less than 0.05 (χ^2<0.05). This is also the test value that was used through the data analysis for the survey done in Sekhukhuneland. For more information about the Chi-square test (χ^2) the reader is referred to Bohrnstedt, [87].