Logistic regression

**We use logistic regression to identify Demographic factors on marks**

**When we consider the reference category as : Second class**

**Table 2 Model fit Indices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Model Fitting Criteria | Likelihood Ratio Tests | | |
| -2 Log Likelihood | Chi-Square | df | Sig. |
| Intercept Only | 362.216 | 49.236 | 22 | .001 |
| Final | 312.980 |

Initially log likelihood model was developed. The initial log likelihood value obtained is 362.216, which is a measure of a model with no independent variable, that is only constant or intercept. The final log likelihood value obtained is 312.980 and this is measure of a model by considering all independent variables. The difference between these two measures is the model chi-square value, which is obtained as 49.236 and this has a significance as p <0.001, if we are setting the significance at 0.05 level, we can conclude that there is a significant relationship between dependent variable and the independent variables.

**Table 3 Likelihood Ratio Tests**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Effect | Model Fitting Criteria | Likelihood Ratio Tests | | |
| -2 Log Likelihood of Reduced Model | Chi-Square | df | Sig. |
| Intercept | 318.004 | 5.024 | 2 | 0.081 |
| School Distance | 316.102 | 3.122 | 2 | 0.210 |
| Residence | 313.429 | 0.449 | 2 | 0.799 |
| Income | 313.438 | 0.458 | 2 | 0.795 |
| Father - Occupation | 317.099 | 4.119 | 2 | 0.127 |
| Mother Occupation | 314.111 | 1.131 | 2 | 0.568 |
| Father Education | 317.389 | 4.410 | 2 | 0.110 |
| Mother Education | 313.556 | 0.577 | 2 | 0.749 |
| Caste | 318.204 | 5.224 | 2 | 0.073 |
| Education board | 322.476 | 9.497 | 2 | 0.009 |
| Aid | 314.273 | 1.293 | 2 | 0.524 |
| Missed Due to financial reasons | 328.475 | 15.495 | 2 | 0.000 |

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

**Table 4: Parameter Estimates**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Marklevela | | B | Std. Error | Wald | df | Sig. | Exp(B) |
|
| First class | Intercept | 2.826 | 1.306 | 4.687 | 1 | 0.030 |  |
| School Distance | 0.535 | 0.329 | 2.638 | 1 | 0.104 | 1.707 |
| Residence | -0.039 | 0.202 | 0.037 | 1 | 0.848 | 0.962 |
| Income | 0.143 | 0.243 | 0.344 | 1 | 0.557 | 1.153 |
| Father - Occupation | -0.122 | 0.179 | 0.466 | 1 | 0.495 | 0.885 |
| Mother Occupation | -0.124 | 0.176 | 0.497 | 1 | 0.481 | 0.883 |
| Father Education | -0.452 | 0.252 | 3.217 | 1 | 0.073 | 0.636 |
| Mother Education | 0.089 | 0.227 | 0.154 | 1 | 0.695 | 1.093 |
| Caste | -0.110 | 0.335 | 0.107 | 1 | 0.743 | 0.896 |
| Education board | -0.333 | 0.226 | 2.164 | 1 | 0.141 | 0.717 |
| Aid | 0.153 | 0.630 | 0.059 | 1 | 0.808 | 1.166 |
| Missed Due to financial reasons | -0.523 | 0.491 | 1.134 | 1 | 0.287 | 0.593 |
| Distinction | Intercept | 2.139 | 1.618 | 1.747 | 1 | 0.186 |  |
| School Distance | 0.358 | 0.386 | 0.862 | 1 | 0.353 | 1.431 |
| Residence | -0.154 | 0.244 | 0.401 | 1 | 0.527 | 0.857 |
| Income | 0.024 | 0.280 | 0.007 | 1 | 0.933 | 1.024 |
| Father - Occupation | -0.427 | 0.222 | 3.713 | 1 | 0.054 | 0.652 |
| Mother Occupation | -0.224 | 0.214 | 1.102 | 1 | 0.294 | 0.799 |
| Father Education | -0.584 | 0.314 | 3.459 | 1 | 0.063 | 0.558 |
| Mother Education | 0.219 | 0.290 | 0.568 | 1 | 0.451 | 1.244 |
| Caste | 0.580 | 0.384 | 2.281 | 1 | 0.131 | 1.787 |
| Education board | 0.252 | 0.230 | 1.200 | 1 | 0.273 | 1.287 |
| Aid | -0.607 | 0.861 | 0.496 | 1 | 0.481 | 0.545 |
| Missed Due to financial reasons | -3.181 | 1.113 | 8.163 | 1 | 0.004 | 0.042 |

a. The reference category is: Second class.

The multinomial logit model shown in the table corresponds to the following equations

Log (p(First class)/p(Second class)) = 2.826 + 0.535 School distance - 0.039 Residence + 0.143 Income -0.122 Father - Occupation – 0.124 Mother Occupation – 0.452 Father Education + 0.089 Mother Education + 0.896 Caste + 0.717 Education board + 1.166 Aid + 0.593 Missed Due to financial reasons

Log (p(Distinction)/p(Second class)) = 2.139+ 0.358 School distance -0.154Residence + 0.024Income -0.427Father - Occupation – 0.224Mother Occupation – 0.584Father Education + 0.219 Mother Education + 0.580 Caste + 0.252 Education board - 0.607Aid -3.181 Missed Due to financial reasons

The other way of interpreting the result is based on the Exponential beta. Exponential beta gives the odd ratio of the dependent variable.  We can find the probability of the dependent variable from this odd ratio.  When the exponential beta value is greater than one, then the probability of other category increases (High group), and if the probability of exponential beta is less than one, then the probability of reference (low) category increases.  Exponential beta value is interpreted with the reference category, where the probability of the dependent variable will increase or decrease.  In continuous variables, it is interpreted with one unit increase in the independent variable, corresponding to the increase or decrease of the units of the dependent variable.

Based on this one can conclude School distance, Income, Mother education and Aid are the factors positively influencing the first class group when compared to second class group for mark level as the Exp(B) is greater than one (See the column Exp(B)).

Based on this one can conclude School distance, Income, Mother education, Caste and Education board are the factors positively influencing the distinction group when compared to second class group for mark level as the Exp(B) is greater than one (See the column Exp(B)).