**Chapter 14: Logistic Regression**

**Binomial Logistic Regression**

Admission.csv contains the data of 400 students’ GRE and GPA grades, and whether or not they were admitted to a university. Perform a binomial logistic regression using Jamovi to analyze if GRE and GPE grades are significant predictors of being admitted to a university.

Select the appropriate levels, and output all assumption checks and information needed to interpret the results and analyze the accuracy of the model.







Set the baseline level to 0 since we are more interested in a student being admitted than not.





| Model Fit Measures |
| --- |
|  |  |  |  |  |  |  |  |
| **Model** | **Deviance** | **AIC** | **R²McF** |
| 1 |  | 480 |  | 486 |  | 0.0393 |  |
|  |

The deviance and AIC values assess the model fit. AIC is more useful when comparing two or more models with different predictors. In terms of magnitude, the smaller the AIC, the better. Using fewer variables (removing the insignificant ones) often leads to a larger AIC as a trade-off. Thus, it is necessary to observe if the difference in AIC is large or negligible.

In this case, we are only considering one model since both predictors turned out to be significant (see model coefficients table).

| Omnibus Likelihood Ratio Tests |
| --- |
|  |  |  |  |  |  |  |  |
| **Predictor** | **χ²** | **df** | **p** |
| GRE |  | 6.62 |  | 1 |  | 0.010 |  |
| GPA |  | 5.71 |  | 1 |  | 0.017 |  |
|  |

The omnibus test looks at the whole model significance and the contribution of each predictor. The results show that both GRE and GPA are significant in predicting the success of a student in being admitted to a university.

| Model Coefficients - Admission |
| --- |
|  | **95% Confidence Interval** |  |
| **Predictor** | **Estimate** | **Lower** | **Upper** | **SE** | **Z** | **p** | **Odds ratio** |
| Intercept |  | -4.94938 |  | -7.057 |  | -2.84223 |  | 1.07509 |  | -4.60 |  | < .001 |  | 0.00709 |  |
| GRE |  | 0.00269 |  | 6.18e-4 |  | 0.00476 |  | 0.00106 |  | 2.54 |  | 0.011 |  | 1.00269 |  |
| GPA |  | 0.75469 |  | 0.128 |  | 1.38106 |  | 0.31959 |  | 2.36 |  | 0.018 |  | 2.12695 |  |
| Note. Estimates represent the log odds of "Admission = 1" vs. "Admission = 0" |
|  |
|  |

For the model coefficients, it is more beneficial and easier to interpret the odds ratio.

For a one-unit increase in GRE, the odds of a student being admitted is increased by a factor of 1.003.

For a one-unit increase in GPA, the odds of a student being admitted is increased by a factor of 2.127.

 **Assumption Check**

| Collinearity Statistics |
| --- |
|  |  |  |  |  |  |
|  | **VIF** | **Tolerance** |
| GRE |  | 1.13 |  | 0.885 |  |
| GPA |  | 1.13 |  | 0.885 |  |
|  |

 **Prediction**

| Classification Table – Admission |
| --- |
|  | **Predicted** |  |
| **Observed** | **0** | **1** | **% Correct** |
| 0 |  | 263 |  | 10 |  | 96.3 |  |
| 1 |  | 118 |  | 9 |  | 7.09 |  |
| Note. The cut-off value is set to 0.5 |
|  |

 The prediction table shows that the model is better at predicting those who will not be admitted than the admitted students. The cut-off was set to 0.5 which is the default. You can experiment with the cut off and see which produces the better prediction table. In this case, 0.5 is working fine in terms of overall accuracy. However, you can still change this to prioritize correct classification for either admitted students or non-admitted students.

| Predictive Measures |
| --- |
|  |  |  |  |  |  |  |  |
| **Accuracy** | **Specificity** | **Sensitivity** | **AUC** |
| 0.680 |  | 0.963 |  | 0.0709 |  | 0.635 |  |
| Note. The cut-off value is set to 0.5 |
|  |

The overall accuracy of the model is 68%. It is also highly specific, meaning that it is better at predicting those who will not be admitted than those who will get admitted. This is still useful in a sense, in that, if a student is predicted to be admitted, realistically, there is a high chance that the student really will be admitted to university.

**ROC Curve**



The area under the ROC curve is 0.63 which falls under the 'poor' category.

Note that there are only 2 predictors in this model. The amount of variance that cannot be accounted by these two predictors is still large, causing low predictive power.