ASNA Case Round 0

Team 4

2019-09-30

# Summary

In order to calculate the probability that a bridge needs to be repaired or replaced, we cleaned up the given data set. We did this by replacing missing data with the mean or mode of the subset. We used the bridge information to develop a logistic regression model. As a result, by plugging in a bridge’s parameters, our model is able to predict the probability that a bridge will need to be repaired.

# Method

After excluding the irrelevant variables and doing a correlation matrix test between each of the two numeric varibles, we decided to use RoadType, NumberofLanes, AvgDailyTraffic, OperationalStatus, StructureMaterial, DeckMaterial, BridgeDesign, NumberofSpans, BridgeLength, NumberofSpans, BridgeLength, DeckWidth, and Age as our explanatory variable. We made the needtorepair indicator as our response variable.

Being that we seeked to model probabilities, it was a natural choice to use the logistic regression model. We considered the following model for each of the aforementioned analysis:

is the probability of a bridge needs repair or replacement.

# Result

Given a bridges’ characteristic, we are able to use our logistic regression model and predict the probability of repair or replacement.

Exponential estimated parameters of logistic model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Exp. Estimate | Std. Error | t value | P-Value |
| Intercept | 0.268 | 0.148 | -8.915 | 0.000 |
| RoadTypeThoroughfare | 1.046 | 0.028 | 1.595 | 0.111 |
| RoadTypeSecondary | 1.159 | 0.035 | 4.203 | 0.000 |
| RoadTypeCityStreet | 1.025 | 0.034 | 0.728 | 0.467 |
| RoadTypeOther | 0.509 | 0.046 | -14.827 | 0.000 |
| NumberofLanes | 1.106 | 0.016 | 6.218 | 0.000 |
| AvgDailyTraffic | 1.000 | 0.000 | 4.786 | 0.000 |
| OperationalStatusTemporaryFix | 4.123 | 0.137 | 10.351 | 0.000 |
| OperationalStatusRestriction | 3.336 | 0.041 | 29.451 | 0.000 |
| OperationalStatusClosed | 2.687 | 0.116 | 8.521 | 0.000 |
| OperationalStatusNewNotOpen | 0.142 | 0.822 | -2.375 | 0.018 |
| StructureMaterialSteel | 1.341 | 0.022 | 13.085 | 0.000 |
| StructureMaterialWood | 0.637 | 0.066 | -6.814 | 0.000 |
| DeckMaterialSteel | 1.607 | 0.053 | 9.025 | 0.000 |
| DeckMaterialWood | 1.326 | 0.039 | 7.261 | 0.000 |
| BridgeDesign | 1.005 | 0.002 | 2.975 | 0.003 |
| NumberofSpans | 0.997 | 0.003 | -0.882 | 0.378 |
| BridgeLength | 1.000 | 0.000 | 2.565 | 0.010 |
| DeckWidth | 0.961 | 0.003 | -14.907 | 0.000 |
| Age | 1.030 | 0.000 | 68.968 | 0.000 |

Then we subsitute the data into the model again with the weights we just calculated. We predict the probability of repair or replacement and we plot probability of repair or replacement with respect to age in a graph shown below.

