Print

Practice: Using the Binary Logistic Regression Task to Perform a Multiple Logistic Regression Analysis with Categorical Variables

The insurance company wants to model the relationship between three of a car's characteristics, weight, size, and region of manufacture, and its safety rating. The **safety** data set contains the data about vehicle safety.

1. Use the Binary Logistic Regression task to fit a multiple logistic regression model with **Unsafe** as the response variable and **Weight**, **Size**, and **Region** as the predictor variables.
    1. Use the EVENT= option to model the probability of Below Average safety scores.
    2. Specify **Size** and **Region** as classification variables and use reference cell coding. Specify *3* (large cars) as the reference level for **Size** and *Asia* as the reference level for **Region**.
    3. Request profile likelihood confidence limits, an odds ratio plot, and the effect plot.
    4. Run the task and view the results.

        1. In the Navigation pane, select **Tasks and Utilities**.
        2. Expand **Tasks**.
        3. Expand **Statistics** and open the **Binary Logistic Regression** task.
        4. Select the **stat1.safety** table.
        5. Assign **Unsafe** to the Response role, and use the Event of interest drop-down list to specify **1**.
        6. Assign **Size** and **Region** to the Classification variables role.
        7. Expand the **Parameterization of Effects** property and use the Coding drop-down list to select **Reference coding**.
        8. Assign **Weight** to the Continuous variables role.
        9. On the MODEL tab, verify that **Main effects model** is selected.
        10. On the OPTIONS tab, in the Select statistics to display drop-down list, select **Default and additional statistics**.
        11. Expand the **Parameter Estimates** property. In the Confidence intervals for odds ratios drop-down list, select **Based on profile likelihood**.
        12. Expand **PLOTS**, and in the Select plots to display drop-down list, select **Default and additional plots**.
        13. Select **Effect plot** and **Odds ratio plot**.
        14. Modify the code to specify specific levels of each class variable to use as reference levels. On the CODE tab, click the **Edit SAS code** icon.
        15. In the CLASS statement, add the options **(REF='3')** immediately after **Size** and **(REF='Asia')** immediately after **Region** .
        16. Click **Run**.

    Here are the results.

2. Do you reject or fail to reject the null hypothesis that all regression coefficients of the model are 0?

    The *p*-value for the Likelihood Ratio test is <.0001, and therefore, you reject the null hypothesis.

3. If you reject the global null hypothesis, then which predictors significantly predict safety outcome?

    Only **Size** is significantly predictive of **Unsafe**.

4. Interpret the odds ratio for significant predictors.

Only **Size** is significant. The design variables show that Size=1 (Small or Sports) cars have 14.560 times the odds of having a Below Average safety rating compared to the reference category 3 (Large or Sport/Utility). The 95% confidence interval (3.018, 110.732) does not contain 1, implying that the contrast is statistically significant at the 0.05 level.

The contrast from the second design variable is 1.931 (Medium versus Sport/Utility), implying a trend toward greater odds of low safety as size decreases. However, the 95% confidence interval (0.343, 15.182) contains 1, and therefore, the contrast is not statistically significant.

Hide Solution