1. A commonly discussed issue is why people are successful (for example, <http://freakonomics.com/podcast/sports-ep-3/>. One perspective is that success depends on effort/hard work, whereas another perspective is that luck has a major part in determining success. What can you add to the conversation, based on the topic discussed in this module?
   1. Looking at this discussion statistically, we can define success as the dependent variable Y. Effort/hard work is the independent variable X. The moderator variable M is luck. All variables seem to be categorical. A moderation analysis needs to be preformed to determine whether or not luck impacts the relationship between effort/hard work and success. If the p-value for the product of effort/hard work and luck is less than alpha, we can determine the interaction effect between the effort/hard work is significant.
2. Use the Cars dataset from SASHELP directory to explore if Drive Train impacts the relationship between MSRP(y) and Horsepower (x). Conduct a thorough analysis and state your conclusions.
   1. As shown in the appendix under question 2 we can see the moderation linear regression preformed. The model F value of 222.70 and resulting p-value of less the 0.0001 display that the model is significant. The parameter estimates table contains the interaction effect results for the independent variable horsepower and the moderator variable drive train. Since drive train is a categorical variable with three variables, there are two p-values and a referent group. The referent group is drive train rear/horsepower which doesn’t have any values. The p-values for both drive train all/horsepower and drive train front/horsepower are less than 0.0001 which is less than a nominal alpha of 0.05. We can then determine that the interaction between drive train and horsepower is significant. In conclusion, drive train does impact the relationship between MSRP and Horsepower.
3. Use the dataset named “FinalQ2” posted in the datasets folder to solve this question. Examine if Grade category and total amount interact with each other to impact VAR7 Project Approval (0 = not approved, 1 = approved).
   1. As shown in the appendix under question 3 we can see the binary moderation logistic regression model. The Likelihood Ratio, Score, and Wald tests each have a p-value of less than 0.0001 so we can conclude the model to be significant. The analysis of effects tables displays an interaction effect for total amount and grade category of 0.0432 which is less than a nominal alpha value of 0.05. Concluding that grade category and total amount is significant and do interact with each other to impact VAR7. However, when looking at the Analysis of Maximum Likelihood Estimates table, we can see the p-values for total amount/grade category for Elementary and High are 0.0647 and 0.7563 respectively. Concluding that the interaction between amount/grade category for Elementary and High are not significant. This would lead us to believe grade category middle is significant, however this can not be proven since middle grade category is the referent group. This would explain the drop in the interaction p-value for grade category and total amount.

Appendix

Question 2:

|  |  |
| --- | --- |
| **Data Set** | SASHELP.CARS |
| **Dependent Variable** | MSRP |
| **Selection Method** | None |

|  |  |
| --- | --- |
| **Number of Observations Read** | 428 |
| **Number of Observations Used** | 428 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **DriveTrain** | 3 | All Front Rear |

| **Dimensions** | |
| --- | --- |
| **Number of Effects** | 4 |
| **Number of Parameters** | 8 |

| **Least Squares Summary** | | | | |
| --- | --- | --- | --- | --- |
| **Step** | **Effect Entered** | **Number Effects In** | **Number Parms In** | **SBC** |
| **\* Optimal Value of Criterion** | | | | |
| **0** | **Intercept** | 1 | 1 | 8457.7685 |
| **1** | **Horsepower** | 2 | 2 | 7970.9854 |
| **2** | **DriveTrain** | 3 | 4 | 7970.1467 |
| **3** | **Horsepowe\*DriveTrain** | 4 | 6 | 7935.2525\* |

**Least Squares Model (No Selection)**

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 5 | 1.169209E11 | 23384186059 | 222.70 | <.0001 |
| **Error** | 422 | 44310688407 | 105001631 |  |  |
| **Corrected Total** | 427 | 1.612316E11 |  |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 10247 |
| **Dependent Mean** | 32775 |
| **R-Square** | 0.7252 |
| **Adj R-Sq** | 0.7219 |
| **AIC** | 8340.89774 |
| **AICC** | 8341.16441 |
| **SBC** | 7935.25248 |

| **Parameter Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | 1 | -27020 | 3413.380075 | -7.92 | <.0001 |
| **Horsepower** | 1 | 278.450623 | 12.455849 | 22.36 | <.0001 |
| **DriveTrain All** | 1 | 21090 | 5664.556531 | 3.72 | 0.0002 |
| **DriveTrain Front** | 1 | 21787 | 4120.815561 | 5.29 | <.0001 |
| **DriveTrain Rear** | 0 | 0 | . | . | . |
| **Horsepowe\*DriveTrain All** | 1 | -98.044295 | 22.455315 | -4.37 | <.0001 |
| **Horsepowe\*DriveTrain Front** | 1 | -116.505251 | 17.227412 | -6.76 | <.0001 |
| **Horsepowe\*DriveTrain Rear** | 0 | 0 | . | . | . |

Question 3:

| **Model Information** | | |
| --- | --- | --- |
| **Data Set** | \_TEMP0.FINALQ2 |  |
| **Response Variable** | VAR7 | Approved (0= Not approved, 1 = Approved) |
| **Number of Response Levels** | 2 |  |
| **Model** | binary logit |  |
| **Optimization Technique** | Fisher's scoring |  |

|  |  |
| --- | --- |
| **Number of Observations Read** | 54624 |
| **Number of Observations Used** | 54624 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **VAR7** | **Total Frequency** |
| **1** | 0 | 8227 |
| **2** | 1 | 46397 |

**Probability modeled is VAR7='1'.**

| **Class Level Information** | | | | |
| --- | --- | --- | --- | --- |
| **Class** | **Value** | **Design Variables** | | |
| **Grade\_Category** | **Elementary** | 1 | 0 | 0 |
|  | **High** | 0 | 1 | 0 |
|  | **Middle** | 0 | 0 | 1 |

| **Model Convergence Status** |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Model Fit Statistics** | | |
| --- | --- | --- |
| **Criterion** | **Intercept Only** | **Intercept and Covariates** |
| **AIC** | 46297.834 | 46224.992 |
| **SC** | 46306.742 | 46278.442 |
| **-2 Log L** | 46295.834 | 46212.992 |

| **Testing Global Null Hypothesis: BETA=0** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > ChiSq** |
| **Likelihood Ratio** | 82.8416 | 5 | <.0001 |
| **Score** | 90.9145 | 5 | <.0001 |
| **Wald** | 87.8670 | 5 | <.0001 |

| **Type 3 Analysis of Effects** | | | |
| --- | --- | --- | --- |
| **Effect** | **DF** | **Wald Chi-Square** | **Pr > ChiSq** |
| **TotalAmount** | 1 | 38.2570 | <.0001 |
| **Grade\_Category** | 2 | 14.6466 | 0.0007 |
| **TotalAmou\*Grade\_Cate** | 2 | 6.2827 | 0.0432 |

| **Analysis of Maximum Likelihood Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameter** |  | **DF** | **Estimate** | **Standard Error** | **Wald Chi-Square** | **Pr > ChiSq** |
| **Intercept** |  | 1 | 1.7774 | 0.0403 | 1943.6384 | <.0001 |
| **TotalAmount** |  | 1 | -0.00011 | 0.000041 | 7.3194 | 0.0068 |
| **Grade\_Category** | **Elementary** | 1 | 0.0796 | 0.0448 | 3.1511 | 0.0759 |
| **Grade\_Category** | **High** | 1 | -0.1099 | 0.0629 | 3.0539 | 0.0805 |
| **Grade\_Category** | **Middle** | 0 | 0 | . | . | . |
| **TotalAmou\*Grade\_Cate** | **Elementary** | 1 | -0.00009 | 0.000048 | 3.4125 | 0.0647 |
| **TotalAmou\*Grade\_Cate** | **High** | 1 | 0.000018 | 0.000060 | 0.0963 | 0.7563 |
| **TotalAmou\*Grade\_Cate** | **Middle** | 0 | 0 | . | . | . |

| **Association of Predicted Probabilities and Observed Responses** | | | |
| --- | --- | --- | --- |
| **Percent Concordant** | 54.2 | **Somers' D** | 0.084 |
| **Percent Discordant** | 45.8 | **Gamma** | 0.084 |
| **Percent Tied** | 0.0 | **Tau-a** | 0.022 |
| **Pairs** | 381708119 | **c** | 0.542 |