1. **Research Question:**

The research question was to determine some of the variables associated with a student being enrolled in an advanced placement program. In particular what role do socio-economic status and enrollment in a gifted class have on the odds of a student being in an advanced placement program. The data used came from the National Education Longitudinal Study of 1988 (NELS:88).

1. **Hypothesis**

The purpose of the research was to test the hypothesis that student enrolled in an advanced program were from higher socio economic class and typically had been enrolled as well in a gifted class. The null hypothesis, (H0 : β = 0 ) that student who have attended a special advanced program ensured that they promoted to next class. The alternative hypothesis ( H1: β ≠0 ) , there would be other variables that helps students to promote to next class.

1. **Selected variables and Analysis approach**

The dependent variable was F1S34E which as a binary variable had values of 1 = yes, 0 = no and indicated whether a student was in an Advanced Placement class. There were 2 independent variables: 1) BYS68A which indicated if a student had been enrolled in a class for gifted or talented students (1 = yes, 2 = no), and 2) socioeconomic status which was a composite developed by NELS made up of parents’ level of education, occupation, and household income from the survey which ranged from -2.97 to 2.56.

The questions asked in the survey were specifically the following

1. Have you ever been in any of the following kinds of courses or programs in high school? F1S34E Advanced placement program

Code Label

1 YES

2 NO

{blank} {NONR/NOT IN SAMPLE THIS WAVE}

8 {MISSING}

1. Are you enrolled in any of the following special programs/ services? (MARK ONE EACH). BYS68A Classes for gifted or talented students.

Code Label

1 YES

2 NO

{blank} {NONR/NOT IN SAMPLE THIS WAVE}

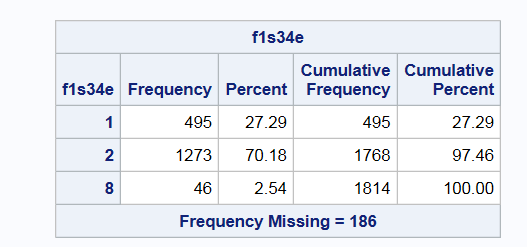
8 {MISSING}

The dependent variable fis34e (Advanced placement program) is a binary variable which means that there are typically 2 possible outcomes, either the event has occurred where the variable has a value or 1 (yes) or the event did not occur where the variable has a value of 0 or no.

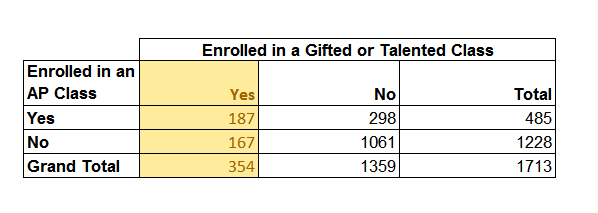
1. **Data Source and Measures:**

The data used came from the National Education Longitudinal Study of 1988 (NELS:88) which contained 2000 records based on sampling. This was a representative sample done in the USA of eighth-graders. The participants were then resurveyed in 1990, 1992, 1994, and 2000. The survey had a comprehensive list of topics including school, work, home, parent, support, self-concept, income and lifestyle. Students' teachers, parents, and school administrators were also surveyed in this project.

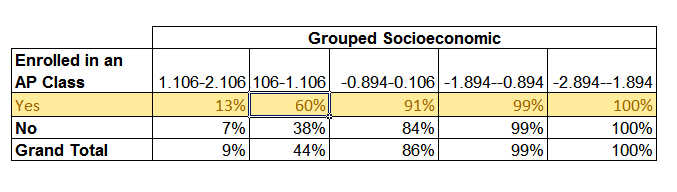
The frequency of enrollment in advanced courses appears as follow:



In our logistic regression model, we first investigated the variable to determine if a student had been enrolled in a gifted and talented class in eighth grade (recoded as 1 = yes, 0 = no). Running a basic pivot table in Excel indicated that more students who were in a gifted or talented class were also enrolled in an AP class than were not.



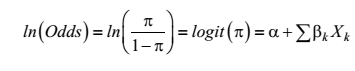
In addition when Socioeconomic categories were grouped the following results were returned a significant amount of students enrolled in an AP class were on the higher end of the socio-economic category.



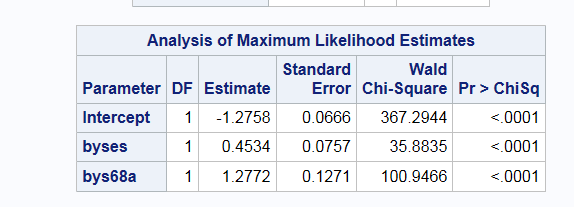
Because the dependent variable determined for the research question was binary, the SAS logistic regression model was used. Linear regression with least square method can result in unpredictable and useless predictor values when used to analyze binary dependent variables.

**The Logistic Regression Model:**

* In logistic regression, the dependent variable Y is binary
* It only contains data coded as 1 (True, Success, etc.) or 0 (False, Failure, etc.)
* By logistic regression the best fitting linear model that can describe the relationship between the binary characteristic of interest, the dependent variable and a set of independent variable is the following where π is the probability of presence of characteristic of interest



1. **The Results:**

****

The table above shows the coefficients (labeled Estimate), their standard errors, the Wald Chi-Square statistic, and associated p-values. The coefficients for byses (Socioeconomic), and bys68a (enrollment in gifted) are statistically significant as shown by their p value. The logistic regression coefficients suggests the change in the log odds of the outcome for a one unit increase in the predictor variable. For every one unit change in socio economic, the log odds of advanced enrollment versus non-enrollment increases by 0.4534. When a student is enrolled in an gifted program the enrollment odds increase by 1.2772 for an advanced course.

Coefficients are often interpreted using an odds ratio (OR), which, is a ratio of two odds. The OR is computed by exponentiating the regression coefficient (i.e., e0.4534), which results in an OR of 1.57 for socioeconomic and 3.59 for Gifted enrollment.

If the OR is less than one, then the odds associated with the coefficient are lower. If the OR is greater than one, then the odds are higher. If the OR is equal to one, then the likelihood is neither higher nor lower and the odds are the same (Huang, L, & Moon 2013). For example in our example, since the OR is greater than one, we can say that the odds of being enrolled in an AP class increase, on average, by a factor of 3.59 for students who were enrolled in a gifted program compared to students who were not enrolled in a gifted program.

1. **Conclusion :**

To make our work simple and easy understanding, we consider only one dependent variable and two independent variables, whereas to get more accurate result we have to consider other variables that has impacts on the hypothesis. While using logistic regression model, we first check to determine if the model tested fits the data well. There are three common tests that can be used to assess model fit in a simple logistic regression. The likelihood ratio (LR) test, the Wald test, and the Lagrange multiplier test (also known as the score test). In our observation we go for Wald test. Although all three tests address the same basic question, they go about answering the question in a different way.

1. **Reference**

Huang, L, & Moon, T. (2013). What Are the Odds of That? A Primer on Understanding Logistic Regression. *Gifted Child Quarterly*, 57(3) 197–204.