We will use logistic regression to model the binary outcome of whether or not a loan is classified as risky. On the Logistic Regression Template sheet, you will find columns for a single predictor (Column B) and the outcome (Column C). You will also find spaces in which we will calculate the probability that the loan is classified as risky (Column D) or on-time (Column E), and the log-likelihood (Column F) that we will use to estimate the logistic regression.

For illustration purposes, we will relate home ownership to whether or not the loan is considered risky. In cell B2, reference the recoded value for home ownership on the Calibration Data worksheet in cell R2:

='Calibration Data'!R2

Copy this formula down the worksheet. In cell C2, reference the recoded value for the loan status on the Calibration Data worksheet in cell U2:

='Calibration Data'!U2

Copy this formula down the worksheet.

In cell D2, we will enter the probability (based on logistic regression) that the loan is classified as risky:

=EXP($I$1+$I$2*B2)/(1+EXP($I$1+$I$2*B2))

Copy this formula down the worksheet.

In cell E2, we can calculate the probability that the loan is classified as on-time as:

=1-D2

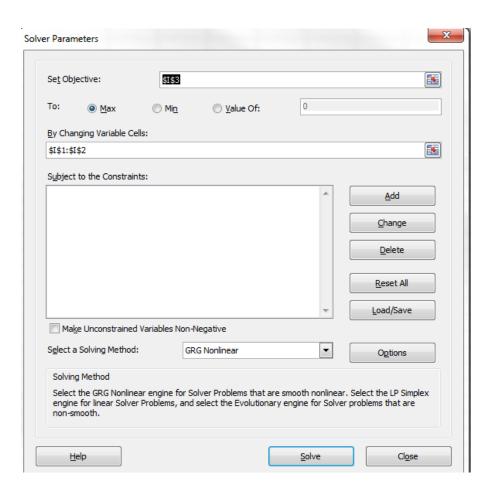Copy this formula down the worksheet.

In cell F2, we will calculate the log-likelihood of the observed outcome. If the outcome is equal to 1, this is the ln(P(Risky)). If the outcome is equal to 0, this is ln(P(On-time)):

=IF(C2=1,LN(D2),LN(E2))

Copy this formula down the worksheet. In Cell I3, calculate the log-likelihood of the data sample as:

=SUM(F2:F50001)

Having populated the worksheet, our goal is to now estimate the intercept and slope using Solver. Be sure that you have installed the Solver add-on on the Data tab. Using Solver, set the objective cell to I3. We will maximize this value by changing the values in cells I1 and I2. Be sure to uncheck the box that constrains the coefficients. The Solver dialog box should appear as depicted below:

Click on Solve to estimate the intercept and slope corresponding to this regression.