

The node's Scaling Options properties enable you to control the factor and offset values of the score equation. Scaling enables you to control the range of the scores as well as the rate of change in odds for a given increase in the score. The Points to Double Odds property determines the value of factor. The Scorecard Points and Odds properties jointly determine the value of offset. To understand how this is accomplished, consider what happens to the score when you double the odds. The score equation must satisfy the following: $\text{score} + \text{Points to Double Odds} = \ln(2 * \text{odds}) * \text{factor} + \text{offset}$.

Subtracting the original score equation and solving for Points to Double Odds yields this result: $\text{Points to Double Odds} = \text{factor} * \ln(2)$.

Solving for factor yields this result: $\text{factor} = \text{Points to Double Odds} / \ln(2)$.

Thus, when you set the value of the Points to Double property, you are, in fact, controlling the value of factor in the score equation. The default value for Points to Double Odds is 20. This value is interpreted to mean that a 20 point increase in an applicant's score means the odds of the applicant being a "good" risk is doubled.

If you substitute the result for factor into the original score equation and rearrange to isolate offset, you get the following result: $\text{offset} = \text{score} - (\text{Points to Double Odds} / \ln(2)) * \ln(\text{odds})$.

To solve for offset, you need one fixed pair of values for score and odds. You use the Scorecard Points and Odds properties to specify these two values. Rewriting the offset equation using the property names you get the following result: $\text{offset} = \text{Scorecard Points} - (\text{Points to Double Odds} / \ln(2)) * \ln(\text{Odds})$.

The values you choose for Scorecard Points and Odds are arbitrary. The default values are 200 and 50, respectively. These values are interpreted to mean that a score of 200 represents odds of 50 to 1 (that is, when $P(\text{"good"})/P(\text{"bad"}) = 50$, the score = 200).

Because the logistic regression models the $\ln(\text{odds})$ as a linear function of the characteristics, it is easy to see each characteristic's contribution to the $\ln(\text{odds})$, and thus, the score. The weight of evidence (WOE) variables enter the model as interval variables and the group variables enter the model as ordinal variables. Thus, when you use the WOE variables as inputs for the regression model, the score points for each characteristic i are calculated as follows:

$$- (\text{woe}_i * \beta_i + \frac{\alpha}{n}) * \text{factor} + \frac{\text{offset}}{n}$$

where α is the intercept from the logistic regression, the β_i is the parameter estimate associated with the i th characteristic, and n is the number of characteristics. An observation's total score is the sum of the score points across all n characteristics.

When group variables are used as inputs into the logistic regression model, the ordinal group variables are automatically replaced with a set of binary indicator variables by the DMREG procedure. Thus, when you use the group variables as inputs for the regression model, the score points for each attribute is calculated as follows:

$$- (\beta_{ij} + \frac{\alpha}{n}) * \text{factor} + \frac{\text{offset}}{n}$$

where α is the intercept from the logistic regression, β_{ij} is the parameter estimate associated with the j th attribute of the i th characteristic, and n is the number of characteristics. When a characteristic has k attributes and $k > 2$, the DMREG procedure generates $k-1$ binary indicator variables to represent the characteristic's ordinal group variable. One indicator variable must be omitted to prevent collinearity. Therefore, one attribute of such a characteristic does not have a regression coefficient associated with it. In such cases, the coefficient is assumed to be 0 in the preceding equation. An observation's total score is the sum of the score points across all n characteristics.