Level Importance of Risk Factors in PD Modeling

WoE Regression and Scorecard Scaling

Andrija Djurovic

www.linkedin.com/in/andrija-djurovic

Importance of Risk Factors in PD Modeling

According to Achen (1982), in Interpreting and Using Regression, the concept of variable importance is ambiguous without clarifying: important for what purpose?
He classifies importance into three groups, each answering a different question:

1. *Theoretical Importance*: This measures the potential effect of a risk factor, holding all else constant. The unstandardized coefficient in the regression captures it and indicates how much the dependent variable changes per unit increase in the risk factor.
2. *Level Importance*: This measures a risk factor's actual contribution to the mean level of the dependent variable in a given sample. It is computed as the product of the regression coefficient and the mean of the predictor.
3. *Dispersion Importance*: This measures how much variation in the dependent variable is explained by a risk factor. The standardized coefficient captures it.

This presentation focuses on the level importance of risk factors in logistic regression with Weights of Evidence (WoE) encoding of categorical risk factors, which is the most commonly used approach in practice for Probability of Default (PD) modeling.

Too often, practitioners first scale the logistic regression scores and then evaluate a risk factor's contribution to the average score, both in absolute and relative terms. However, this approach raises a few important questions:

- Is this always appropriate?
- What factors influence the conclusion on risk factor importance?
- How does this approach align with the interpretation of risk factor contribution when using WoE encoding?

The following slides first introduce the concept of scorecard scaling in WoE regressions (logistic regression with WoE encoding of categorical risk factors) and then present a simulation study with several examples demonstrating how the selection of scaling inputs and the treatment of the regression intercept affect the assessment of risk factors level importance.

Scorecard Scaling

Scorecard scaling refers to transforming logistic regression outputs into a user-friendly scoring format. The objective is to map probabilities or odds into a numerical score range that is easy to interpret and use for decision-making. A standard industry approach uses a logarithmic scale, where the odds double every fixed number of points - known as "points to double the odds" (PDO).
The following equation gives the general relationship between score and odds:

$$\text{Score} = \text{Offset} + \text{Factor} \cdot \ln(\text{odds})$$

where the offset and factor are determined based on reference points chosen by the user.
For instance, if a score of 600 corresponds to odds of 50:1, and the odds double every 20 points, then the factor and offset can be calculated as:

$$\text{Factor} = \frac{\text{PDO}}{\ln(2)} = \frac{20}{\ln(2)} \approx 28.85$$

$$\text{Offset} = 600 - (28.85 \cdot \ln(50)) \approx 487.12$$

Each score in the scorecard is then derived using this formula, ensuring consistency and interpretability across different scorecards.

The example above demonstrates how to transform the model output into scores. However, practitioners often include an additional step in which each modality of a risk factor (i.e., each characteristic attribute) is assigned a specific score. The total sum of the individual modality scores should match the overall model score produced by the method described earlier. The following slides illustrate score allocation to risk factor modalities for WoE encoding.

Scorecard Scaling - WoE Encoding

Since the scorecard is typically developed using WoE encoding, the final score for an applicant is determined by summing the points assigned to each modality of the risk factors in the scorecard.

By applying the Factor and Offset, points are allocated to the each modality. The final credit score is obtained by summing the points assigned to all relevant modalities. The point allocation follows the trends established in the WoE analysis. Expressed in terms of formulas, the steps below demonstrate the process of point allocation to the modalities of each risk factor:

$$\mathsf{Score} = \mathsf{Offset} + \ln(\mathsf{odds}) \cdot \mathsf{Factor} =$$

$$-\left(\sum_{j=1}^{k}\sum_{i=1}^{n}(\mathsf{WoE}_j \cdot \beta_i) + \alpha\right) \cdot \mathsf{Factor} + \mathsf{Offset} =$$

$$-\left(\sum_{j=1}^{k}\sum_{i=1}^{n}(\mathsf{WoE}_j \cdot \beta_i + \frac{\alpha}{n})\right) \cdot \mathsf{Factor} + \mathsf{Offset} =$$

$$\sum_{j=1}^{k}\sum_{i=1}^{n}\left(-\mathsf{WoE}_j \cdot \beta_i + \frac{\alpha}{n}\right) \cdot \mathsf{Factor} + \frac{\mathsf{Offset}}{n}$$

Scorecard Scaling - WoE Encoding cont.

where:

- WoE_j is the Weight of Evidence for each modality;
- β_i is the regression coefficient for each risk factor;
- α denotes the intercept term from logistic regression;
- n is the number of risk factors in the model;
- k denotes the number of modalities in each risk factor;
- Factor is a scaling parameter used to adjust the scores of each modality within a given risk factor;
- Offset is the constant used to shift the score into a desired range.

For more details on scorecard scaling using WoE encoding, practitioners can refer to Siddiqi, N. (2012). Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring. John Wiley & Sons, Inc.

In addition, the presentation available at this link builds on the work presented in Siddiqi (2012) and further extends the example of the dummy encoding method.

From Scorecard Scaling to Risk Factor Level Importance

- After assigning scaled points to each modality of the risk factors, practitioners often assess the level importance of each factor.
- In the context of scorecard scaling, level importance measures the average contribution each risk factor makes to the overall model score.
- A common approach is to rank factors based on their average score contribution, though practitioners are often more interested in their relative contributions.
- Practitioners often measure relative contribution as the proportion of each risk factor's average score to the model's overall average score.
- However, is calculating a risk factor's relative contribution to the average score in this way always meaningful - and does it truly reflect what practitioners intend to assess?

Simulation Study

Simulation Design

Assume that the modeling dataset consists of a binary target variable and three risk factors. The simulation dataset is available here with the column `target` representing the binary dependent variable, and columns `rf1`, `rf2`, and `rf3` representing categorical risk factors.

After estimating the logistic regression of the form `target ~ rf1 + rf2 + rf3` with WoE encoding of the risk factors, the following simulation steps involve defining scaling inputs and producing scores for each risk factor modality, following the approach from the previous slide. The different scaling inputs in the box on the right represent those used in various examples throughout this simulation study.

The goal of this simulation is to examine how different scaling inputs affect conclusions on risk factor level importance, measured as the average relative contribution of each risk factor to the overall model score.

Scaling Inputs

Example 1:

Factor: 1
Offset: 0

Example 2:

Factor: 2
Offset: 6

Example 3:

Factor: 30
Offset: 450

Simulation Results

Regression Results and Scores Overview

Columns score.1, score.2, and score.3 represent score allocations for different scaling inputs.
Columns intercept and beta show the logistic regression coefficients with WoE encoding of the risk factors.

```
##    rf bin  no  ng  nb     woe intercept    beta score.1 score.2 score.3
##   rf1   1 274 139 135 -0.8181   -0.8461 -0.9274 -0.4767    1.05  135.70
##   rf1   2 269 164 105 -0.4014   -0.8461 -0.9274 -0.0902    1.82  147.29
##   rf1   3  63  49  14  0.4055   -0.8461 -0.9274  0.6581    3.32  169.74
##   rf1   4 394 348  46  1.1763   -0.8461 -0.9274  1.3729    4.75  191.19
##   rf2   0  40  15  25 -1.3581   -0.8461 -0.7980 -0.8017    0.40  125.95
##   rf2   1  49  21  28 -1.1350   -0.8461 -0.7980 -0.6237    0.75  131.29
##   rf2   2 530 361 169 -0.0883   -0.8461 -0.7980  0.2116    2.42  156.35
##   rf2   3  88  60  28 -0.0852   -0.8461 -0.7980  0.2141    2.43  156.42
##   rf2   4 293 243  50  0.7337   -0.8461 -0.7980  0.8676    3.74  176.03
##   rf3   1 282 222  60  0.4610   -0.8461 -0.8768  0.6863    3.37  170.59
##   rf3   2 232 161  71 -0.0286   -0.8461 -0.8768  0.2570    2.51  157.71
##   rf3   3 332 230 102 -0.0342   -0.8461 -0.8768  0.2521    2.50  157.56
##   rf3   4 154  87  67 -0.5861   -0.8461 -0.8768 -0.2318    1.54  143.04
```

Risk Factor Relative Level Importance

The following table shows the relative level importance for the different scaling inputs.

```
##  Example    rf1    rf2    rf3
##        1 40.67% 30.68% 28.65%
##        2 35.24% 32.65% 32.12%
##        3 33.81% 33.16% 33.03%
```

Conclusions

- Scorecard scaling refers to transforming logistic regression outputs into a user-friendly scoring format and is one of the most commonly performed steps when developing PD models.
- After assigning scaled points to each modality of the risk factors, practitioners often assess level importance by measuring the average contribution each risk factor makes to the overall model score.
- As the simulation study showed, the scaling inputs affect the final results of relative importance. This raises the question of whether calculating relative importance in this way always provides sufficient insights.
- Moreover, practitioners often overlook the fact that the intercept is allocated proportionally to each risk factor and incorporated into the score calculation. This calls for attention to the meaning of the intercept in WoE regressions and a precise formulation of the importance concept.
- Finally, when performing this exercise, practitioners are strongly encouraged to carefully investigate and understand each process step to ensure correct conclusions for the intended purpose.