ЖАРАТЫЛЫСТАНУ ЖӘНЕ ТЕХНИКАЛЫҚ ҒЫЛЫМДАР

NATURAL AND TECHNICAL SCIENCES

IRSTI 06.58.45

N. Sultanova[1], A. Tulegenova[2], B. Suleimen[3]
[1,2,3]Suleyman Demirel University, Kaskelen, Kazakhstan

USING THE GINI COEFFICIENT TO BUILD A CREDIT SCORING MODEL

Abstract. The credit scoring model is widely used to predict the likelihood of a customer default. To measure the quality of such scoring models, you can use quantitative indicators such as the GINI index, Kolmogorov-Smirnov (KS) statistics, Lift, Mahalanobis distance, information statistics. This article discusses and illustrates the practical use of the GINI index.

Keywords: Credit scoring, Fair Isaac Corporation (FICO), GINI, Probability of default (PD), Capital assistance program (CAP), Receiver operating characteristic (ROC), Area under the ROC curve (AUC), Confusion matrix, True Positive Ratio, True False Positive Ratio.

***

Аңдатпа. Несиелік скоринг моделі клиенттің дефолт ықтималдығын болжау үшін кеңінен қолданылады. Мұндай балдық модельдердің сапасын өлшеу үшін сіз GINI индексі, Колмогоров-Смирнов (KS) статистикасы, Лифт, Махаланобис арақашықтық, ақпараттық статистика сияқты сандық көрсеткіштерді пайдалана аласыз. Бұл мақалада GINI индексінің практикалық қолданылуы талқыланады және суреттеледі.

Түйін сөздер: Несиелік скоринг, Fair Isaac Corporation (FICO), GINI, дефолт ықтималдығы (PD), капиталға көмек бағдарламасы (CAP), алушының жұмыс сипаттамасы (ROC), ROC қисығы астындағы аймақ (AUC), шатасу матрицасы, шынайы қатынас коэффициенті, шын Жалған оң қатынас.

***

Аннотация. Модель кредитного рейтинга широко используется для прогнозирования вероятности дефолта клиента. Для измерения качества таких скоринговых моделей можно использовать количественные показатели, такие как индекс GINI, статистика Колмогорова-Смирнова (KS), Lift, расстояние Махаланобиса, информационная статистика. В этой статье обсуждается и иллюстрируется практическое использование индекса GINI.

Ключевые слова: Кредитный скоринг, Fair Isaac Corporation (FICO), GINI, Вероятность дефолта (PD), Программа капитальной помощи (CAP), Операционная характеристика получателя (ROC), Площадь под кривой ROC (AUC), Матрица неточностей, Истинный положительный коэффициент, Истина Ложноположительный коэффициент.

Introduction

Banks and other financial institutions receive thousands of loan applications every day (in the case of a consumer loan, this can be tens or hundreds of thousands every day). Since they cannot be processed manually, these institutions make extensive use of automatic systems to assess the creditworthiness of loan applicants. Credit risk assessment relies on one of the most successful applications of statistics and transactions research: credit scoring.

Credit scoring is a collection of decision-making models and underlying techniques that assist lenders in providing consumer loans. Credit scoring is a method mainly used in consumer lending to help loan providers make a loan decision. Credit scoring is a decision-making aid used by lenders in providing consumer credit. The main idea of credit scoring is to differentiate and identify a specific sample of population groups [1]. Credit scoring is used to assess the risk of providing a loan to an individual. Whether an individual will be assessed for creditworthiness or not. This method was used by the bank to help make decisions related to granting loans to borrowers. The goal is to construct a classification that can distinguish between good and bad customers based on a specific standard. Credit scoring encourages lenders to create a credit card in which each characteristic has its own weight, and the overall score across all characteristics will determine the individual creditworthiness.

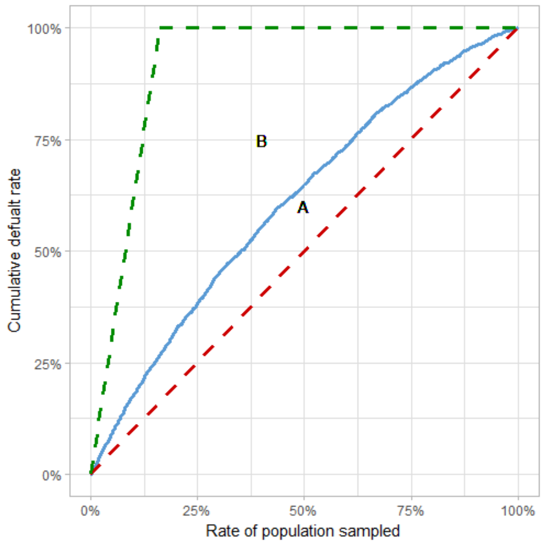
Main part

Finding the GINI coefficient using the CAP curve

The CAP curve in our context is intended to determine the ordinal relationship between the estimate (PD) and the default indicator. If our model does a good job of distinguishing between good and bad borrowers, we would expect to find more defaults in low-rated borrowers than high-rated borrowers. The CAP curve reflects this concept by aggregating the aggregate default rate for a sample of borrowers from the lowest to the highest score [2].

To plot the CAP curve, the entire population of the model must be ordered according to the predicted probability of default. Namely, the observation with the lowest score is the first, and the observation with the highest score is the last. We then sample from the first to the last, and after each sample, we calculate the cumulative default ratio. The X-axis of the CAP curve represents a fraction of the population sample, and the Y-axis represents the corresponding default cumulative factor. To plot the CAP curve, the entire population of the model must be ordered according to the predicted probability of default. If our model has perfect discriminating power, we expect to reach 100% of the cumulative default frequency after sampling a portion of the observations that is equal to the default frequency in our data (green line in the diagram below).

For example, if the default frequency in our data is 16%, after sampling 16% of the observations, we will capture all the default values in our data [5]. In contrast, if we use a random model, that is, a model that randomly assigns scores in an equal distribution, the cumulative coefficient will always default to a fraction of the selected observations (red line in the diagram below).



Picture 1. Predictive probability of a conventional model

The GINI coefficient is characterized as the ratio between the area inside the fair curve and the free fair line (A) and the area between the demonstration curve and the free fair line (A + B). In other words, the GINI coefficient can be a proportion that tells how close our demonstration is to the “ideal model” and how far from it is from the “random model”. Thus, the "ideal model" will receive a GINI of 1, and the “random model” will receive a GINI of 0.

Our model got a low GINI of 0.26:

$$\frac{A}{A+B}=\frac{0.109}{0.419}=0.26 \qquad (1)$$

Plotting the Lorentz curve, extracting the Corrado GINI measure and deriving the GINI coefficient.

The Lorentz curve is the inverse of the CAP curve; it is constructed using the same mechanism for sampling observations and aggregating the default cumulative ratio, but sampling is performed in reverse order (highest to lowest score). Lorentz also has a diagonal line that is equivalent to the "random CAP pattern" line and is called the "equal line" (red line in the diagram below). Another contrast between the two curves is the "ideal model" line. Since the Lorenz curve was designed to reflect the spread of wealth, the main distinguishing result may be the case where all the wealth of the population is concentrated in one perception [4]. However, a problem arises when using the Lorenz curve to assess the ability to recognize a credit rating model and assign the cumulative default rate to its y-axis. Since the y-axis describes the aggregation of a binary outcome (1 or 0), there is no case where the entire aggregate default rate is concentrated in one case. In other words, when evaluating a credit rating model using the Lorenz curve, it is impossible to achieve the "Perfect Model" line. Therefore, the corresponding "perfect model" line for this kind of estimate should be adjusted to the default level in the population, as in the CAP curve.

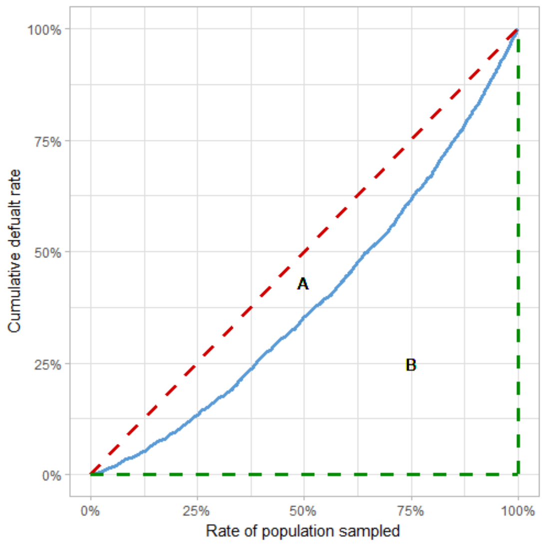


Figure 2. Credit rating models using the Lorenz curve.

Therefore, in order to adjust the Corrado GINI measure to estimate the credit rating model, we need to subtract the inaccessible area from its denominator. Finally, to get the GINI factor from the Corrado GINI measure, we can use the following formula:

$$Gini = \frac{Corrado\ Gini}{1 - default\ rate} \qquad (2)$$

Our model received a Corrado GINI of 0.22:

$$\frac{A}{A+B}=\frac{0.109}{0.5}=0.22 \tag{3}$$

The default coefficient in my sample is 16%, so the GINI coefficient of the model can be calculated as follows:

$$Gini=\frac{0.22}{(1-0.16)}=0.26 \tag{4}$$

Plotting the ROC curve, finding the AUC and obtaining the GINI coefficient

The third strategy for calculating the GINI ratio is to use another common curve: the ROC curve. The area under the ROC curve, more commonly referred to as the AUC, is also the dominant metric for assessing and comparing the performance of credit rating models. The ROC curve sums up two proportions from the confusion network [6]: True Positive Ratio (TPR) and False Positive Ration (FPR).The confusion matrix summarizes, for a given threshold, the number of cases in which:

- The model predicted default and default of the borrower - True Positive.
- The model predicted default, but the borrower did not default - False Positive.
- The model did not predict default and the borrower did not default - True Negative.
- The model did not predict a default, and the borrower defaulted - False Negative.

The True Positive Return Ratio (TPR) is defined as the number of default borrowers for which the model overlapped the total number of default borrowers in our data. The false positive rate (FPR) is calculated as the number of cases in which the model incorrectly predicted the default based on the total number of instances other than the default.

$$TPR=\frac{TP}{TP+FN}=\frac{439}{439+1184}=27\% \tag{5}$$

$$FPR=\frac{FP}{FP+TN}=\frac{3628}{3628+4743}=43\% \tag{6}$$

The ROC curve is constructed using confusion matrices that originate from thresholds from 1 to 1000 and drive their TPR and FPR. The y-axis of the ROC curve represents TPR values and the x-axis represents FPR values. AUC is the area between the curve and the abscissa.

Advantages and Disadvantages of the GINI Ratio

Despite its advantages, the GINI ratio has several disadvantages that should be considered when using it to evaluate and compare credit rating models.

For the purposes of moderation in this area, we will depict a typical trap when trying to derive the GINI coefficient and its main drawback. The main disadvantage of the GINI coefficient stems from the fact that it is an ordinal indicator, i.e. it fixes a set of values, ignoring the separation between them. This characteristic of the GINI ratio can sometimes overshadow disadvantaged demonstration results. Thus, the failure of the GINI coefficient to capture the adequacy of the model for separating different levels of probability can be a serious disadvantage. To overcome this disadvantage, we recommend visually inspecting the distribution of the model's predictions and using the Precision-Recall curve to estimate the model's relationship between the outreach of “bad” borrowers and the erroneous prediction of default by “good” borrowers.

Conclusion

The Gini coefficient, which is used in the financial industry to assess the quality of a credit rating model, is actually a “Somers’ D, ”not Corrado's GINI inequality score. There are three common methods for calculating the GINI ratio:

- Extract the GINI factor from the CAP curve.
- Construct a Lorentz curve, extract the Corrado measure GINI, then output the GINI coefficient.
- Plot the ROC curve to extract the AUC, then output the GINI factor.

A common mistake when calculating the GINI coefficient is the same scores. The main disadvantage of the GINI coefficient is that it does not reflect the sensitivity of the model to different levels of risk.

References

1 Thomas, L.C., Edelman, D.B., Crook, J.N., Credit Scoring and Its Applications. *Philadelphia: Society for Industrial and Applied Mathematics*, 2006, 204 p.

2 Sabato, G., Assessing the Quality of Retail Customers: Credit Risk Scoring Models. *The IUP Journal of Financial Risk Management*, (2010): pp. 35 - 43.

3 Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. How do machine learning and non-traditional data affect credit scoring? (2016): pp. 204 -208.

4 Schechtman, E., & Schechtman, G. The Relationship between Gini Methodology and the ROC curve. Rochester, NY: Social Science Research Network, (2016): pp. 11-15.

5 Somers, R. A New Asymmetric Measure of Association for Ordinal Variables. American Sociological Review, (2011): pp. 799-811.

6 Gini, C. On the measurement of concentration and variability of characters, 2005, - 338 p.