모형 변별력 평가 지표
- AR & AUC

모델 변별력

리스크 업계에서 많이 사용되는 지표 : KS, GINI 계수 or AR

BASEL 위원회에서 권장하는 지표 : GINI 계수 or AR, AUC

Table 3: Measures of Model Discrimination

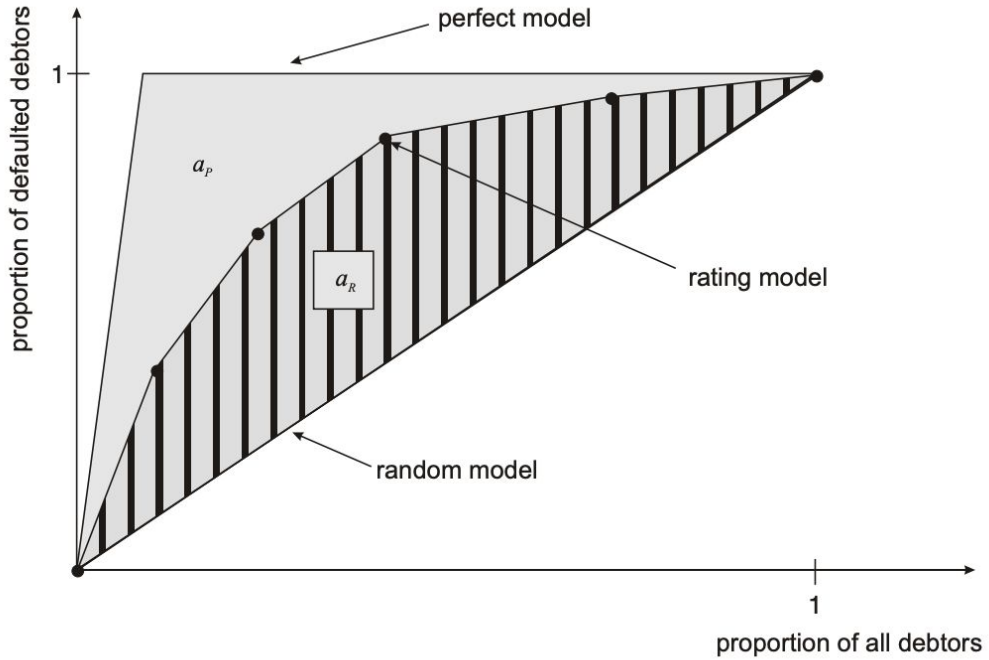
Component	Description
Gini Coefficient or Accuracy Ratio (AR)	• Area under Gini Curve/Lorenz Curve of model as compared to the 'perfect' model, or one that would capture 100% of events in the first score bucket/decile. • Values ranging from 0-1. • This measure enables a direct comparison across models.
Kolmogorov–Smirnov (KS) Statistic	• The maximum separation between the percentage of events captured and percentage of non-events captured by the model in cumulative distributions of events and non-events. • Values ranging from 0-1. • This measure enables a direct comparison across models.
Receiver Operating Characteristic Curve/ Area Under Receiver Operating Characteristic (ROC) Curve	• Enables a comparison of two models on their ability to identify a true positive (that is, an event) as opposed to a false positive. • The area under the ROC curve (AUC) is a measure of the probability that the model will rank a randomly chosen event higher than a randomly chosen non-event and is related to the Gini Coefficient (G) by the formula G = 2(AUC) - 1.
Pietra Index	• The maximum vertical distance between the model Lorenz Curve and the line representing a random decision rule. This distance may be interpreted as the maximum 'lift' over random provided by the model.
Change in Gini Coefficient	• A comparison of the model's most recent Gini calculation with that observed for the previous tracking period. A large decrease or increase would indicate a need for further investigation.

Source: Accenture, November 2014

CAP (Cumulative Accuracy Profile)

신용 리스크 모델링에 주로 사용되는 평가 지표

AR : CAP의 통계 요약 지표



CAP 계산법

Ex) 차주 신용등급 예측 모형

모형 예측 결과 1~10

Label : 0, 1 (정상차주, 부도차주)

$$p_T^i = \pi p_D^i + (1-\pi) p_{ND}^i.$$

$$CD_D^i = \sum_{j=1}^{i} p_D^j,\ i = 1, \ldots, k,$$

$$CD_{ND}^i = \sum_{j=1}^{i} p_{ND}^j,\ i = 1, \ldots, k,$$

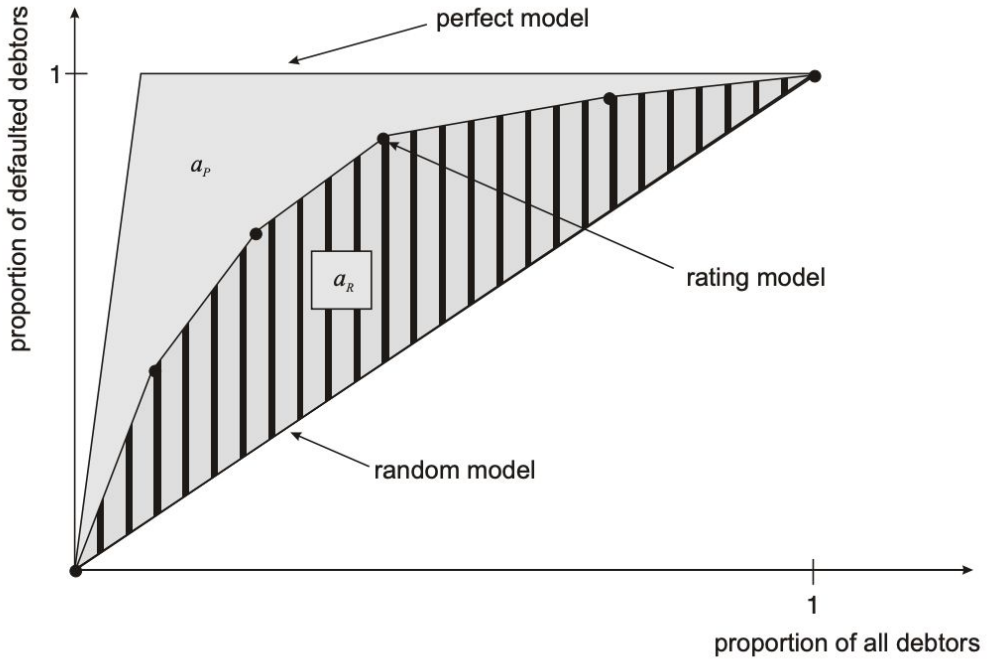
$$CD_T^i = \sum_{j=1}^{i} p_T^j,\ i = 1, \ldots, k,$$

CAP

$$(CD_T^i, CD_D^i)_{i=0,\ldots,k}$$

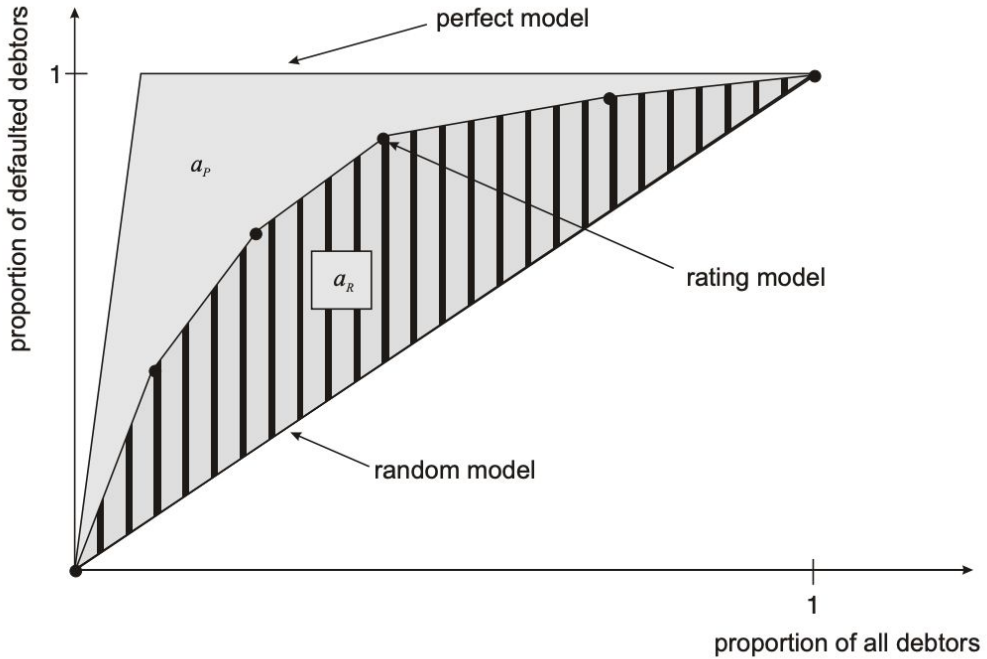
X축 : i등급 아래 데이터일 확률

Y축 : 누적 부도 확률



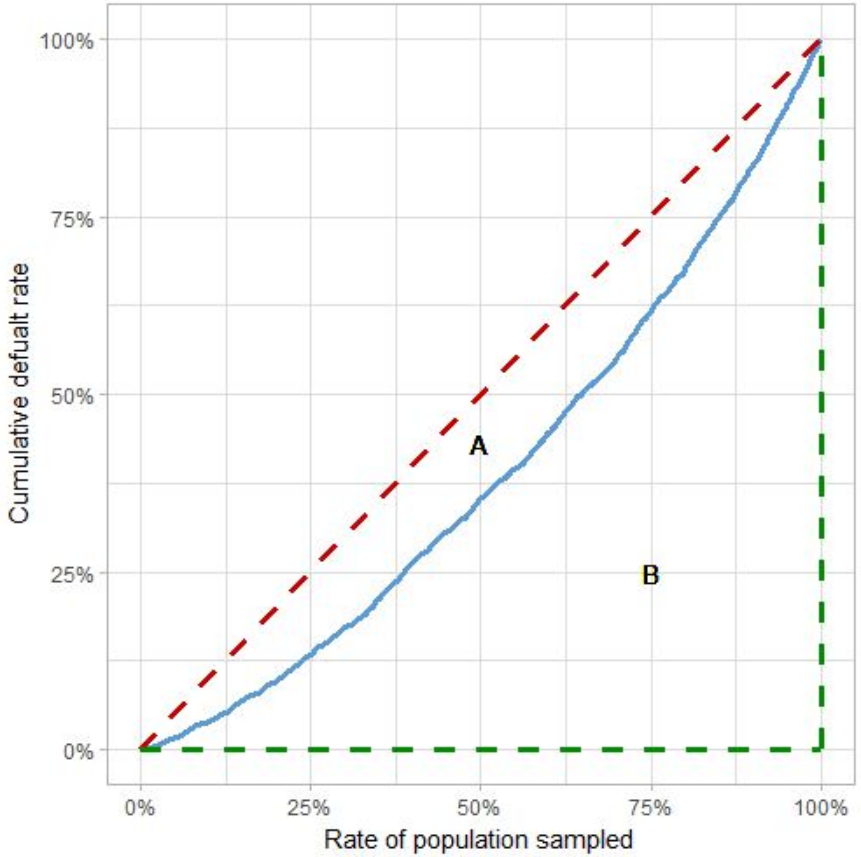
AR

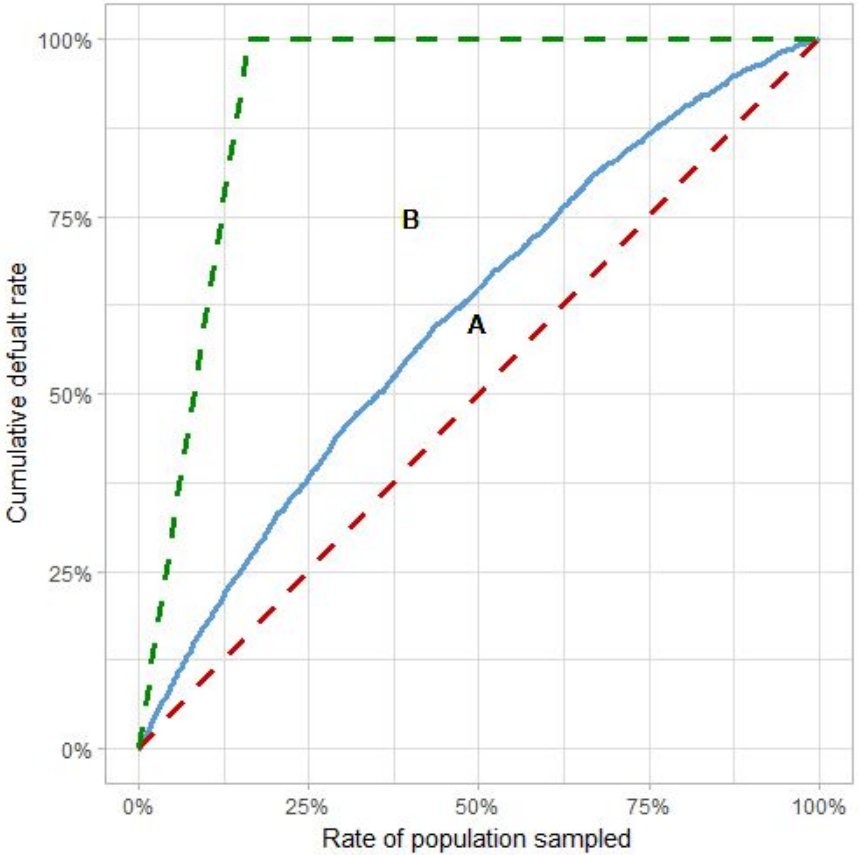
$$AR = \frac{a_R}{a_P}.$$

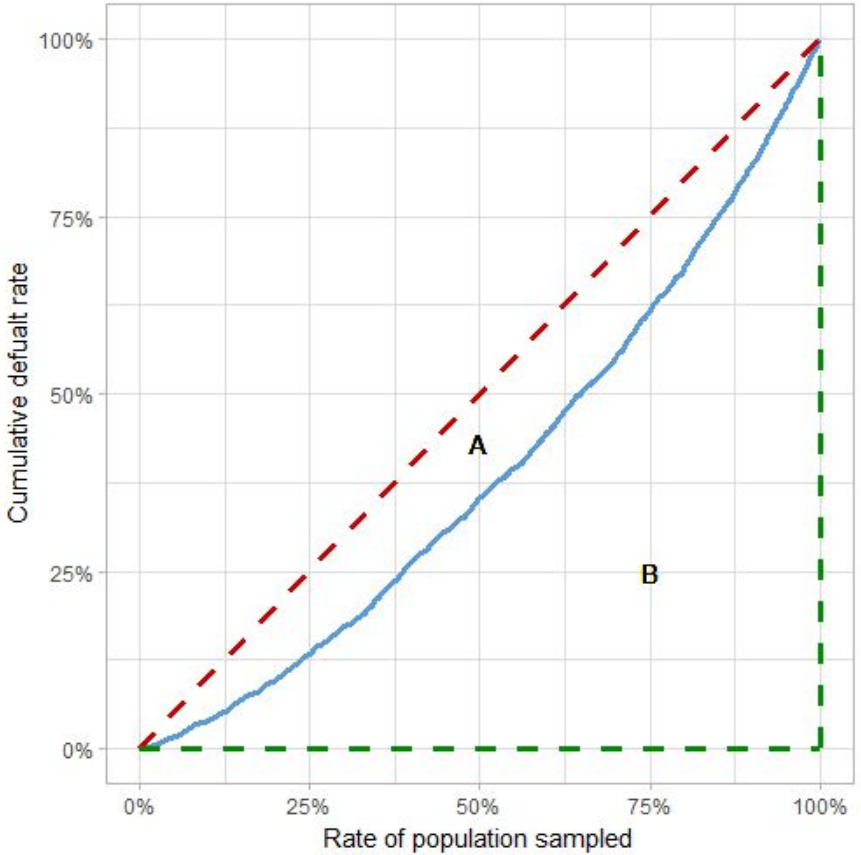


GINI 계수

Lorenz 커브와 Random model
사이의 면적




100%
75%
50%
25%
0%
Cumulative defualt rate
B
A
0%
25%
50%
75%
100%
Rate of population sampled


100%
75%
50%
25%
0%
Cumulative defualt rate
A
B
0%
25%
50%
75%
100%
Rate of population sampled

AR = GINI

ROC Curve & AUC

X축 : FPR (False Positive Rate)

Y축 : TPR (True Positive Rate)

AUC = Area Under the ROC Curve

$$(CD^i_{ND}, CD^i_D)_{i=0,\ldots,k}.$$

$$AR = 2\,AUC - 1.$$