Performance measures of LGD models

Katarzyna Bijak and Lyn Thomas

Credit Scoring and Credit Control XIV
26th-28th August 2015, Edinburgh
Outline

• Introduction
• LGD
• LGD models
• Performance measures of LGD models
• Example
• Conclusions
LGD

• Loss Given Default (LGD)
  – The lender’s loss on a loan due to the customer’s default, i.e. failure to meet the credit commitment

• Recovery Rate (RR)
  – $LGD = 1 - RR$

• Basel II and III
  – Under the Advanced Internal Ratings-Based (AIRB) approach, lenders are allowed to use their own predictions of risk parameters, including LGD
LGD distribution example
LGD models

- Unsecured loans
  - One-stage models
  - Multi-stage approaches
    - Separation of 0s (+ Separation of 1s) + Prediction

- Mortgage loans
  - One-stage models
  - Two-stage approaches
    - Repossession model + Haircut model
LGD models

- Separation stage(s)
  - Logistic regression
  - Decision trees

- Prediction stage/one-stage models
  - Regression models
  - Tobit models
  - Survival analysis
  - Classification and Regression Trees (CART)
  - Other nonlinear models
Performance measures

• PD
  – Gini coefficient
  – KS statistic

• LGD
  – ???
MSE

- Mean Square Error (MSE):

\[ MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

- Sensitive to extreme values of the residuals
- E.g. Bellotti and Crook (2008)
SSE

• Sum of Squared Errors (SSE) a.k.a. residual sum of squares:

\[ SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]

• E.g. Qi and Zhao (2011)

• Depends on the sample size → the results are incomparable between samples of different sizes

• Useful when comparing a number of LGD models developed on the same sample
RMSE

- Root Mean Square Error (RMSE):

\[ RMSE = \sqrt{MSE} \]

- Expressed in the same units as LGD
- Bastos (2010)
MAE

• Mean Absolute Error (MAE) a.k.a. Mean Absolute Deviation (MAD):

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \]

• Expressed in the same units as LGD
• Compare with RMSE
• E.g. Bellotti and Crook (2008)
MAPE

- Mean Absolute Percentage Error (MAPE) a.k.a. Mean Absolute Percentage Deviation (MAPD):

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}
\]

- Zero observations \(\rightarrow\) the division-by-zero problem

- SAS (2012) unconventionally uses the predicted LGD as the denominator \(\rightarrow\) low errors when the predicted values are e.g. ten or more times larger than the observed values

- We do not recommend using MAPE for LGD
RSE

• Relative Square Error (RSE):

\[ RSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \]

• Ratio of MSE of the model and MSE of a simple predictor

• E.g. Bastos (2014)
RAE

- Relative Absolute Error (RAE):

\[ RAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{\sum_{i=1}^{n} |y_i - \bar{y}|} \]

- Ratio of MAE of the model and MAE of a simple predictor

- E.g. Bastos (2010)
AOC

• Regression Error Characteristic (REC) curve estimates the CDF of the squared or absolute residual

• Area Over the REC Curve (AOC) estimates the expected regression error (Bi and Bennett, 2003)

• If the REC curve is derived using the squared residuals, then $\text{AOC} \to \text{MSE}$ as the sample size $\to \infty$

• If the REC curve is derived using the absolute residuals, then $\text{AOC} \to \text{MAE}$ as the sample size $\to \infty$
AOC

- Loterman et al. (2012) calculated both RMSE and AOC (based on the squared residuals)
  - LGD models: 24 various techniques and six datasets
  - Differences between AOC and the squared RMSE:
    - < 0.001 for five larger datasets
    - < 0.01 for the smallest dataset (test: ca. 1100 loans)

- We recommend applying either AOC or MSE/MAE in order to avoid information redundancy
R-squared

• Coefficient of determination (R-squared):

\[ R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \]

• E.g. Loterman et al. (2012)

• In an OLS regression model with a constant term, R-squared can be interpreted as the proportion of variation in LGD that is explained by variation in the regressors

• We only recommend using R-squared in OLS models
Out-of-sample R-squared

- To assess the out-of-sample performance, the out-of-sample mean is normally used.
- Out-of-sample R-squared is calculated using the in-sample mean instead of the out-of-sample mean.
- In-sample and out-of-sample means are often similar → no difference between out-of-sample and “normal” R-squared.
- Useful?
Adjusted R-squared

- Adjusted coefficient of determination (adjusted $R^2$):

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - k - 1}$$

- Corrected for the number of regressors ($k$)

- E.g. Caselli et al. (2008)

- Useful when comparing a number of linear LGD models
Pseudo R-squared

• For example, McFadden’s pseudo R-squared a.k.a. likelihood ratio index
  – for models estimated using the ML method

  \[ Pseudo \ R^2 = 1 - \frac{lnL(M_1)}{lnL(M_0)} \]

• E.g. Dermine and Neto de Carvalho (2006)

• Useful when applying non-linear transformations of LGD
Correlation coefficients

- Measure correlation between the observed and predicted LGD (Loterman et al., 2012)

- Pearson’s correlation coefficient:

\[
r = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \hat{y})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (\hat{y}_i - \hat{y})^2}}
\]

- Measures the strength of the linear relationship between the observed and predicted LGD (\(r^2 = R^2\) in OLS models)

- Spearman’s and Kendall’s correlation coefficients
AUC

• Loans need to be classified into two groups based on the observed LGD, e.g. below-the-mean and over-the-mean
  – Analogy with PD: the predicted LGD – the estimated probability; the predicted RR – a score

• CDFs of the predicted RR are computed for the groups

• Receiver Operating Characteristic (ROC) curve is drawn by plotting the CDFs against each other

• Area Under the ROC Curve (AUC) measures how well the model separates loans belonging to the two groups

• E.g. Gupton and Stein (2005)
AR

• Accuracy Ratio (AR) a.k.a. Gini coefficient

\[ AR = 2AUC - 1 \]

• E.g. Chalupka and Kopecsni (2008)

• The probability that a randomly selected “Good” will have a lower predicted LGD value than a randomly selected “Bad”
Proposed measure: MAUC

- AUC has a drawback when applied to LGD as it requires an arbitrary classification of the dependent variable
- \(m\) – the number of unique values of the observed LGD
- Mean AUC (MAUC) is calculated as the average of AUC for all possible divisions into two groups:

\[
MAUC = \frac{1}{m-1} \sum_{j=1}^{m-1} AUC_j
\]

- MAUC takes values from the interval \([0.5, 1]\) like AUC
Example

- Two-stage model applied to the data on personal loans granted by a large UK bank

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.143</td>
<td>AOC (s. r.)</td>
<td>0.143</td>
</tr>
<tr>
<td>MAE</td>
<td>0.329</td>
<td>AOC (a. r.)</td>
<td>0.329</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.378</td>
<td>RSE</td>
<td>0.928</td>
</tr>
<tr>
<td>MAPE</td>
<td>N/A</td>
<td>RAE</td>
<td>0.940</td>
</tr>
</tbody>
</table>
Example -- continued

- Two-stage model applied to the data on personal loans granted by a large UK bank

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.072</td>
<td>Pearson</td>
<td>0.268</td>
</tr>
<tr>
<td>Adjust. R-sq.</td>
<td>N/A</td>
<td>Spearman</td>
<td>0.255</td>
</tr>
<tr>
<td>Pseudo R-sq.</td>
<td>N/A</td>
<td>Kendall</td>
<td>0.179</td>
</tr>
<tr>
<td>AUC</td>
<td>0.637</td>
<td>AR</td>
<td>0.272</td>
</tr>
</tbody>
</table>

MAUC = 0.616
Example -- continued

AUC for different divisions into two groups
New ideas

• Loterman *et al.* (2014)
  – Discussed statistical tests suitable for the central tendency and dispersion of the errors
  – Proposed four test-statistics-based measures

• Fischer and Pfeuffer (2014)
  – Focused on measures related to discriminatory power
  – Proposed new measures
Conclusions

• Recommendations for LGD model developers/users
  – Do not use MAPE
  – Apply either AOC or MSE/MAE
  – Only use R-squared in OLS models
  – Look for alternatives to AUC
- Do you prefer any of the discussed performance measures? Why?

- Do you use any other performance measures of LGD models?

Thank you!
References


References


