Modelling LGD using Bayesian methods

Katarzyna Bijak
Lyn C. Thomas

Credit Scoring and Credit Control XII
24-26 August 2011, Edinburgh
Outline

• Introduction
• Modelling LGD: traditional approach
• Modelling LGD: Bayesian approach
• Empirical results
• Conclusions
Loss Given Default (LGD)

- Loss borne by a bank when a customer defaults on a loan
- Often used in the form of LGD rate = 1 – Recovery Rate (RR)
- LGD estimates:
  - Expected LGD that helps calculate the expected loss
  - Downturn LGD, i.e. LGD in an economic downturn, under the new Basel Accord
- In this research: unsecured loans (no collateral)
- LGD is often found difficult to model, especially using the one-step approach, e.g. Bellotti and Crook (2008)
LGD distribution example

- LGD distribution usually has a high peak at zero: many customers default but finally pay in full
Modelling LGD: traditional approach

- Two-step approach
  - Matuszyk, Mues and Thomas (2010)
  - Loterman, Brown, Martens, Mues and Baesens (2009)

- There are two separate models estimated independently:
  - The first model (logistic regression) separates positive values from zeroes (and negative values, if any)
  - The second model (e.g. linear regression) allows for the estimation of the positive values

- The independent estimation can be considered problematic from the methodological point of view
Traditional approach: suggested solutions

• In order to apply this approach, one needs:
  – either to set a cut-off for the first model and take zero if \( \text{Prob}(\text{LGD} > 0) < \text{cut-off} \) and the estimated value otherwise \textbf{(cut-off approach)}
  – or to calculate a product of \( \text{Prob}(\text{LGD} > 0) \) and the estimated value \textbf{(probability times value approach)}

• In particular, one can draw a cut-off from \( U(0, 1) \) or draw a number from a Bernoulli distribution with \( \text{Prob}(\text{LGD} > 0) \) \textbf{(random cut-off approach)}

• The result is a point estimate of LGD for each customer
Modelling LGD: Bayesian approach

• Bayesian graphical models
  – Single, hierarchical model instead of two separate ones

• **Random cut-off approach**
  – $p_i \sim \text{Logistic}(x_i, \theta_1)$ – this is $\text{Prob} \!(\text{LGD} > 0)$
  – $b_i \sim \text{Bernoulli}(p_i)$ – this is equal to either 0 or 1
  – If $b_i = 0$ then $y_i \sim N(0, \sigma_1^2)$; else $y_i \sim N(\text{Linear}(z_i, \theta_2), \sigma_2^2)$
  – this is the estimated LGD value

• **Probability times value approach**
Random cut-off approach: graph

\[ y_i = \begin{cases} 1 & \text{if } b_i = 1 \\ 0 & \text{otherwise} \end{cases} \]

\[ b_i \sim \text{Bernoulli}(p_i) \]

- Deterministic node
- Stochastic node

\[ i = 1, \ldots, N \]
MCMC methods

• Markov chain Monte Carlo (MCMC) methods

  – Stochastic simulation methods that are used in the Bayesian inference to generate samples from posterior distributions

  – Based on the construction of a Markov chain that converges to the posterior distribution

  – Under some assumptions, as $t \to \infty$, $\theta^{(t)}$ converges to its equilibrium distribution (posterior distribution) that is independent from its initial state $\theta^{(0)}$

  – Results: distributions of model parameters and outcomes as well as model performance measures
**MCMC algorithm**

1) Selection of the initial values $\theta^{(0)}$

2) Generating $T$ values until the equilibrium is reached

3) Convergence monitoring

4) Discarding the first $B$ values (burnin period)

5) Treating $\{\theta^{(B+1)}, \ldots, \theta^{(T)}\}$ as the sample (MCMC output)

6) Analysis of the posterior distributions: calculating posterior summary statistics, plotting densities etc.

Bayesian framework

- Software: OpenBUGS 3.0.3
- Method: Markov chain Monte Carlo (MCMC)
- Burnin period: 10K iterations
- MCMC output: 100K iterations
- Sampling lag (thinning interval) $L = 5$
  - MCMC output is not independent: there are correlations between $\theta^{(t)}$ and $\theta^{(t+k)}$, autocorrelations of lag $k = 1, 2, \ldots$
  - Keeping the first value from each batch of $L$ iterations $\rightarrow$ independent sample
Empirical research

• Research design
  – Traditional approach
    • Random cut-off
    • Probability times value
  – Bayesian approach
    • Random cut-off
    • Probability times value
• Data on ca 48K personal loans granted by a large UK bank
• Random training and validation samples (10K records each)
## Traditional approach: separate models

<table>
<thead>
<tr>
<th>Logistic regression</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAE</strong></td>
<td>0.367</td>
<td>0.370</td>
</tr>
<tr>
<td><strong>MSE</strong></td>
<td>0.183</td>
<td>0.186</td>
</tr>
<tr>
<td><strong>Gini</strong></td>
<td>0.420</td>
<td>0.421</td>
</tr>
<tr>
<td><strong>KS</strong></td>
<td>0.309</td>
<td>0.317</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Linear regression</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAE</strong></td>
<td>0.182</td>
<td>0.181</td>
</tr>
<tr>
<td><strong>MSE</strong></td>
<td>0.057</td>
<td>0.056</td>
</tr>
<tr>
<td><strong>Pearson correlation</strong></td>
<td>0.398</td>
<td>0.405</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.159</td>
<td>0.164</td>
</tr>
</tbody>
</table>
Traditional approach: entire model

Cut-off and model performance measures (training sample)
Traditional approach: entire model

Cut-off and model performance measures (validation sample)
## Random cut-off approach: performance

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE (Training)</th>
<th>MAE (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - mean</td>
<td>0.364</td>
<td>0.365</td>
</tr>
<tr>
<td>Bayesian - post. mean</td>
<td>0.364</td>
<td>0.365</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE (Training)</th>
<th>MSE (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - mean</td>
<td>0.244</td>
<td>0.245</td>
</tr>
<tr>
<td>Bayesian - post. mean</td>
<td>0.243</td>
<td>0.245</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Corr (Training)</th>
<th>Corr (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - mean</td>
<td>0.081</td>
<td>0.085</td>
</tr>
<tr>
<td>Bayesian - post. mean</td>
<td>0.082</td>
<td>0.085</td>
</tr>
</tbody>
</table>
Random cut-off approach: performance

<table>
<thead>
<tr>
<th></th>
<th>MAE (Training)</th>
<th>MAE (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - sd</td>
<td>0.0029</td>
<td>0.0030</td>
</tr>
<tr>
<td>Bayesian - post. sd</td>
<td>0.0031</td>
<td>0.0031</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MSE (Training)</th>
<th>MSE (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - sd</td>
<td>0.0028</td>
<td>0.0028</td>
</tr>
<tr>
<td>Bayesian - post. sd</td>
<td>0.0030</td>
<td>0.0031</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Corr (Training)</th>
<th>Corr (Validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional - sd</td>
<td>0.0101</td>
<td>0.0103</td>
</tr>
<tr>
<td>Bayesian - post. sd</td>
<td>0.0104</td>
<td>0.0105</td>
</tr>
</tbody>
</table>
### Random cut-off approach: estimates

<table>
<thead>
<tr>
<th>Variable (logistic regression)</th>
<th>Traditional</th>
<th>Bayesian (p. mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.084</td>
<td>1.094</td>
</tr>
<tr>
<td>Age of exposure (months)</td>
<td>-0.545</td>
<td>-0.547</td>
</tr>
<tr>
<td>Amount of loan at opening</td>
<td>0.338</td>
<td>0.342</td>
</tr>
<tr>
<td>Total number of advances/ arrears within the whole life of the loan</td>
<td>-1.478</td>
<td>-1.510</td>
</tr>
<tr>
<td>Number of months with arrears &gt;0 within the life of the loan</td>
<td>0.073</td>
<td>0.069</td>
</tr>
<tr>
<td>Number of months with arrears &gt;1 within the last 12 months</td>
<td>-0.529</td>
<td>-0.539</td>
</tr>
</tbody>
</table>
## Random cut-off approach: estimates

<table>
<thead>
<tr>
<th>Variable (linear regression)</th>
<th>Traditional</th>
<th>Bayesian (p. mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.719</td>
<td>0.718</td>
</tr>
<tr>
<td>Joint applicant present</td>
<td>-0.012</td>
<td>-0.012</td>
</tr>
<tr>
<td>Total number of advances/ arrears within the whole life of the loan</td>
<td>-0.143</td>
<td>-0.141</td>
</tr>
<tr>
<td>Term of loan (months)</td>
<td>-0.037</td>
<td>-0.037</td>
</tr>
<tr>
<td>Worst arrears within the life of the loan</td>
<td>0.178</td>
<td>0.177</td>
</tr>
<tr>
<td>Number of months with arrears &gt;2 within the last 12 months</td>
<td>-0.053</td>
<td>-0.052</td>
</tr>
</tbody>
</table>
Bayesian approach: outcomes

- In the traditional approach: **a point estimate** of LGD (a single number) for each customer

- In the Bayesian approach: **an individual distribution** of LGD for each customer $\rightarrow$ various characteristics, e.g.:
  - Mean
  - Median
  - Other quantiles
Individual distributions: applications

- The predictive mean can be treated as **the expected LGD** → expected loss $EL = E(LGD) \times E(PD) \times E(EAD)$

- The individual predictive distributions of LGD can be used to estimate **the downturn LGD**:
  - One could choose e.g. the 0.999th quantile
  - The choice of the quantile would depend on the user’s perception of the severity of downturns (Kim, 2006)
  - This could be useful in case of lacking downturn data
Individual distributions: example 1

Compare this with the traditional approach:

\[ \text{Prob}(\text{LGD} > 0) = 0.53 \]
\[ \text{Estimated LGD} = 0.72 \]

Observed value (validation sample)
Individual distributions: example 2

Compare this with the traditional approach:

\[
\text{Prob}(\text{LGD} > 0) = 0.96 \\
\text{Estimated LGD} = 0.58
\]
Individual distributions: example 3

Compare this with the traditional approach:

\[
\text{Prob}(\text{LGD} > 0) = 0.75 \\
\text{Estimated LGD} = 0.86
\]
Conclusions

• The (posterior) means of model performance measures and parameter estimates are very similar in both approaches: Bayesian and traditional (using a random cut-off)

• Advantages of the Bayesian approach:
  – More coherent approach: a single hierarchical model instead of two separate ones
  – Much better description of uncertainty (distributions!) than in the traditional approach
Conclusions

• In the Bayesian approach, there is an individual distribution of LGD for each customer (rather than just a point estimate as in the traditional approach)

• The individual distributions of LGD can be used to estimate:
  – the expected LGD → expected loss
  – the downturn LGD

• What does it mean for the user?
  – More information
  – Need to make choices and decisions
References


Thank you!