Model Risk With Estimates of Probabilities of Default

Dirk Tasche
Imperial College, London

June 2015
The views expressed in the following material are the author’s and do not necessarily represent the views of the Global Association of Risk Professionals (GARP), its Membership or its Management.
Outline

Two forecasting problems

A taxonomy of dataset shift

Estimation under prior probability shift assumption

Estimation under 'invariant density ratio' assumption

An application to the mitigation of model risk for PD estimation

Concluding remarks

References
Two forecasting problems

Single borrower’s probability of default (PD)

▶ Moody’s corporate issuer and default counts in 2008\(^2\).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Caa-C</th>
<th>B</th>
<th>Ba</th>
<th>Baa</th>
<th>A</th>
<th>Aa</th>
<th>Aaa</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issuers</td>
<td>417</td>
<td>1151</td>
<td>528</td>
<td>1021</td>
<td>966</td>
<td>582</td>
<td>140</td>
<td>4805</td>
</tr>
<tr>
<td>Defaults</td>
<td>63</td>
<td>25</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>108</td>
</tr>
</tbody>
</table>

▶ January 1, 2009: What is a Baa-rated borrower’s probability to default in 2009?

▶ Natural (?) estimate: \(\frac{5}{1021}\) \(\approx\) 0.49%.

\(^2\)Source: Moody’s (2015)
Two forecasting problems

Rating profile known

Moody’s corporate issuer proportions and default rates in 2008 and issuer proportions in 2009\(^3\). All numbers in %.

<table>
<thead>
<tr>
<th>Grade</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Issuers</td>
<td>Default rate</td>
</tr>
<tr>
<td>Caa-C</td>
<td>8.7</td>
<td>15.1</td>
</tr>
<tr>
<td>B</td>
<td>24.0</td>
<td>2.2</td>
</tr>
<tr>
<td>Ba</td>
<td>11.0</td>
<td>1.1</td>
</tr>
<tr>
<td>Baa</td>
<td>21.2</td>
<td>0.5</td>
</tr>
<tr>
<td>A</td>
<td>20.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Aa</td>
<td>12.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Aaa</td>
<td>2.9</td>
<td>0.0</td>
</tr>
<tr>
<td>All</td>
<td>100.0</td>
<td>2.2</td>
</tr>
</tbody>
</table>

How to take account of the additional data?

\(^3\)Source: Moody’s (2015)
Some thoughts

- Compared to 2008, the rating profile in January 2009 has changed.
- Why should the grade-level or total default rates remain the same?
- Invariant grade-level default rates would imply almost invariant discriminatory power:
  - Observed accuracy ratio in 2008: 63.4%
  - Forecast accuracy ratio for 2009: 66.6%
- What other ways are there to reflect 'almost invariant' discriminatory power?
- Assuming an invariant accuracy ratio is not sufficient for inferring PDs if only the rating profile is known.
Two forecasting problems

Alternatives to ’invariant grade-level default rates’

- Geometric interpretations of accuracy ratio (for continuous score):
  - Based on area under Receiver Operating Characteristic (ROC)
  - Based on area between Cumulative Accuracy Profile (CAP) and diagonal

- Derivative of CAP is (essentially) the PD curve of the scores.
- Derivative of ROC is (essentially) the density ratio of the scores.
- Is ’invariant density ratio’ a viable alternative?
- We also look at ’invariant rating profiles of defaulters and non-defaulters’ as an alternative assumption on ’almost invariant’ discriminatory power.
The Machine Learning perspective

- Classification on datasets with changed distributions is a problem well-known in **Machine Learning**.
- **Moreno-Torres et al. (2012)** proposed a taxonomy for **dataset shifts**.

**Setting:**
- Each item in a dataset has a class $y$ and a covariates vector $x$.
- $p_{test}(x, y)$ and $p_{trai}(x, y)$ are the joint distributions of $(x, y)$ on the test and training sets respectively.
- $p_{trai}(x, y)$ is known from observation but for $p_{test}(x, y)$ only the marginal distribution $p_{test}(x)$ is observable now.
- How to determine unconditional class probabilities $p_{test}(y = c)$ and conditional class probabilities $p_{test}(y = c \mid x)$?

**Definition:** **Dataset shift** occurs if $p_{test}(x, y) \neq p_{trai}(x, y)$.

- On Slide 4, $y = $ default status, $x = $ rating grade.
The Moreno-Torres et al. taxonomy

- Four types of dataset shift $p_{test}(x, y) \neq p_{trai}(x, y)$:
  - **Covariate shift:**
    
    $$p_{test}(x) \neq p_{trai}(x), \text{ but } p_{test}(y \mid x) = p_{trai}(y \mid x).$$
  
  - **Prior probability shift:**
    
    $$p_{test}(y) \neq p_{trai}(y), \text{ but } p_{test}(x \mid y) = p_{trai}(x \mid y).$$
  
  - **Concept shift:**
    
    $$p_{test}(x) = p_{trai}(x), \text{ but } p_{test}(y \mid x) \neq p_{trai}(y \mid x), \text{ or }$$
    
    $$p_{test}(y) = p_{trai}(y), \text{ but } p_{test}(x \mid y) \neq p_{trai}(x \mid y).$$
  
  - **Other shifts.**

- Assuming 'invariant grade-level default rates' on Slide 4 is equivalent to an assumption of covariate shift.
Moody’s corporate rating profiles 2008

- Estimation under prior probability shift assumption

- Moody's corporate rating profiles 2008

- Defaulters
- All
- Non-defaulters

- Frequency

- Caa-C
- B
- Ba
- Baa
- A
- Aa
- Aaa

- Frequency

- Defaulters
- All
- Non-defaulters
The least-squares estimator

- Setting as on Slide 4:
  - \( y = \) default status (classes \( D \) default and \( N \) non-default), \( x = \) rating grade
  - \( p_{test}(x) \) known, but \( p_{test}(x) \neq p_{trai}(x) \)
  - Want to determine \( p_{test} = p_{test}(y = D) \) and \( p_{test}(y = D | x) \)
- Prior probability shift assumption:
  \[
p_{test}(x | y = c) = p_{trai}(x | y = c) \text{ for } c = D, N.
\]
- Hence, for all \( x \), the class probability \( p_{test} \) should satisfy
  \[
p_{test}(x) = p_{test} p_{trai}(x | y = D) + (1 - p_{test}) p_{trai}(x | y = N).
\]
- This is unlikely to be achievable. Therefore least squares approximation
  \[
  \widehat{p}_{test} = \frac{\int \left( p_{test}(x) - p_{trai}(x | y=N) \right) \left( p_{trai}(x | y=D) - p_{trai}(x | y=N) \right) dx}{\int \left( p_{trai}(x | y=D) - p_{trai}(x | y=N) \right)^2 dx}.
  \] (1)
Fitted and observed Moody’s corporate rating profiles

Estimation under prior probability shift assumption
The 'invariant density ratio' estimator

- Setting as on Slide 4:
  - $y =$ default status (classes $D$ default and $N$ non-default), $x =$ rating grade
  - $p_{\text{test}}(x)$ known, but $p_{\text{test}}(x) \neq p_{\text{trai}}(x)$
  - Want to determine $p_{\text{test}} = p_{\text{test}}(y = D)$ and $p_{\text{test}}(y = D \mid x)$

- Invariant density ratio assumption:
  \[
  \frac{p_{\text{test}}(x \mid y = D)}{p_{\text{test}}(x \mid y = N)} = \frac{p_{\text{trai}}(x \mid y = D)}{p_{\text{trai}}(x \mid y = N)} \overset{\text{def}}{=} \lambda(x).
  \]

- Then the class probability $p_{\text{test}}$ must satisfy
  \[
  0 = \int \frac{\lambda(x) - 1}{1 + (\lambda(x) - 1) p_{\text{test}}} d p_{\text{test}}(x). \tag{2a}
  \]

- There is a unique solution to (2a) if and only if
  \[
  \int \lambda(x) d p_{\text{test}}(x) > 1 \quad \text{and} \quad \int \lambda(x)^{-1} d p_{\text{test}}(x) > 1. \tag{2b}
  \]
Properties

- If condition (2b) is not satisfied then the profiles $p_{\text{test}}(x)$ and $p_{\text{trai}}(x)$ are so different that any 'inheritance' of discriminatory power seems questionable.

- **Exact fit:** With the 'invariant density ratio' estimate $\tilde{p}_{\text{test}}$ come estimates of the conditional densities $\tilde{p}_{\text{test}}(x \mid y = c)$, $c = D, N$, such that

  $$
  \lambda(x) = \frac{\tilde{p}_{\text{test}}(x \mid y = D)}{\tilde{p}_{\text{test}}(x \mid y = N)}, 	ext{ and}
  $$

  $$
  p_{\text{test}}(x) = \tilde{p}_{\text{test}} \tilde{p}_{\text{test}}(x \mid y = D) + (1 - \tilde{p}_{\text{test}}) \tilde{p}_{\text{test}}(x \mid y = N).
  $$

- Call $\bar{p}_{\text{test}} = \int p_{\text{trai}}(y = D \mid x) \, d \, p_{\text{test}}(x)$ the 'covariate shift' estimator of $p_{\text{test}}$. Then it holds that

  $$
  \bar{p}_{\text{test}} = (1 - \pi) p_{\text{trai}}(y = D) + \pi \tilde{p}_{\text{test}}.
  $$

  Where $0 \leq \pi \leq 1$ and $\pi$ is the closer to 1 the more discriminatory power the scores $x$ have on the training set.
Conditional profiles 2009: Invariant from 2008 vs. fitted

Defaulter frequencies:
- Caa-C
- B
- Ba
- Baa
- A
- Aa
- Aaa

Non-defaulter frequencies:
- Caa-C
- B
- Ba
- Baa
- A
- Aa
- Aaa

Comparison between observed and fitted distributions for different credit ratings in 2008 and 2009.
Different forecast methods

- Three methods of forecasting portfolio-wide default rate $p_{test}$:
  - Covariate shift estimator $\bar{p}_{test}$ (Slide 13)
  - Prior probability shift estimator $\hat{p}_{test}$, (1)
  - Invariant density ratio estimator $\tilde{p}_{test}$, (2a)

- $\hat{p}_{test}$ and $\tilde{p}_{test}$ give the same estimate under a prior probability shift.

- Estimates $\hat{p}_{test}$ and $\tilde{p}_{test}$ of $p_{test}$ provide estimates of the conditional default rates $p_{test}(y = D | x)$ by

$$
p_{test}(y = D | x) = \frac{p_{test} \lambda(x)}{1 + \left(\lambda(x) - 1\right) p_{test}}.
$$

- (3) suggests that $\max(\bar{p}_{test}, \tilde{p}_{test})$ could be an upper bound for next year’s portfolio-wide default rate.
An application to the mitigation of model risk for PD estimation

Observed vs. forecast corporate default rates

Source of observed rates: Moody’s (2015).
Concluding remarks

- Straightforward ‘covariate shift’ (or ‘invariant conditional default rate’) PD estimates sometimes may seriously underestimate future default rates.

- The ‘invariant density ratio’ approach often provides very different estimates that may be used for model risk mitigation.

- The ‘invariant density ratio’ approach can also be applied for the estimation of loss rates.

- Rating agency data like Moody’s (2015) possibly are ‘subjective’, making results of the approach conservative.

- Further reading:
  - Background and more details: Tasche (2013), Tasche (2014)


