Survival Analysis Methods to Predict Loss Rates in Credit Card Portfolios

Landon Thompson
Goal: Determine an alternative method to model losses that occur in a credit card portfolio

Business Question: Is there a more statistical way to measure losses over a period of time?

• Introduction
• Survival Analysis
• Competing Risks
• Application
Loss Rates

- Two types of ‘losses’
  - Closure: Balance is paid off and card is deactivated
    - Reasons include: didn’t understand terms, not happy with credit line
  - Default: 6+ months delinquent, balance written off as bad debt and attempted to be collected
    - Reasons include: spent too much on the card, lost a job

$Y = 1$ if account has either closed or defaulted in time
Survival Analysis

• Define Random Variable $T$ such that $T =$ Time until event of interest has occurred

• Survival Function:
  $$S(t) = P(T > t)$$

Survival Function Estimator:

$$\hat{S}_{KM}(t) = \prod_{j < t} \left[ 1 - \frac{d_j}{n_j} \right]$$

$d_j = \# \text{ of events in time period } j$

$n_j = \# \text{ of obs at risk in time period } j$
Data

• How data is laid out depends on question being asked

"Vintage"

<table>
<thead>
<tr>
<th>ID</th>
<th>Year</th>
<th>Age</th>
<th>Event</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1900</td>
<td>1</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>1</td>
<td>1901</td>
<td>2</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>1</td>
<td>1902</td>
<td>3</td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td>1</td>
<td>1903</td>
<td>4</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>1</td>
<td>1904</td>
<td>5</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>1900</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1901</td>
<td>2</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>2</td>
<td>1902</td>
<td>3</td>
<td>0</td>
<td>81</td>
</tr>
<tr>
<td>2</td>
<td>1903</td>
<td>4</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>1904</td>
<td>5</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>1905</td>
<td>6</td>
<td>0</td>
<td>34</td>
</tr>
</tbody>
</table>

"Population"

<table>
<thead>
<tr>
<th>ID</th>
<th>Year</th>
<th>Age</th>
<th>Event</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1900</td>
<td>25</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>1</td>
<td>1901</td>
<td>26</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>1</td>
<td>1902</td>
<td>27</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>1900</td>
<td>49</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>2</td>
<td>1901</td>
<td>50</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>1900</td>
<td>18</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>1901</td>
<td>19</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>1902</td>
<td>20</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>1900</td>
<td>21</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>4</td>
<td>1901</td>
<td>22</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>1902</td>
<td>23</td>
<td>0</td>
<td>97</td>
</tr>
</tbody>
</table>
Cox Proportional Hazards Regression

- **Motivation:** Use covariates to estimate survival curve

\[ h_i(t) = h_0(t)e^{X_i(t)\beta} \]
Proportional Hazards

Smoothed Hazard for Closure

Hazard Ratios Over Time
Multiple Events

- Instead of looking at overall loss rates, what if we look at Closure and Defaults separately?
Solution?

- Could look at each event independently
- Over Represented “Non-Event” groups
Competing Risks

- Another solution that addresses mutually exclusive events
  - If an account is closed, it cannot be defaulted or current
Competing Risks Proportional Hazard Regression

• Like before, we can use covariates to estimate survival curves with competing risks

Model Specification:
\[ \lambda_j(t \mid X) = \lambda_{0j}(t) \exp(X\beta) \]

• Results in \( j \) sets of coefficient estimates
• Can obtain predicted cumulative event curves for each event \( j \)

\[ \widehat{F}_{ij}(t \mid X) = \exp(-\widehat{H}_{0j}(t) \exp(X_i\beta)) \]

*Implementation of this in R can be done using Terry Therneau’s “Survival” package*
Competing Risks Regression

Competing Risks Forecast

Independent Forecast

variable
- Closure
- Chargeoff
- Attrition

F(x)

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

0 12 24 36 48 60 72 84 96 108 120

Time - In Months
Application

- Risk underwriting determines what accounts are given a credit card solicitation
- According to stakeholder’s risk appetite, we can set cutoffs accordingly
Predicted Curves

Strategy Change - Default

Strategy Change - Closure
Questions?