An Introduction to Survival Analysis

Dr Barry Leventhal

Henry Stewart Briefing on Marketing Analytics
19th November 2010
Agenda

- Survival Analysis concepts
- Descriptive approach
- 1st Case Study – which types of customers lapse early
- Predicting survival times
- 2nd Case study – lifetimes of mobile phone customers
- Business applications of survival analysis
- Applications to different industries and problems
- Summary of business benefits
Tracking the Customer Lifecycle - Financial Services

Starting Out
- Financial Indicators
  - Mortgage
  - Loan
  - Protection
  - Joint Accounts

Forming a Family
- Life Insurance Loans
- Higher monthly debits

Moving up the Ladder
- Investments
- Increased monthly deposits
- Retirement Plans

Golden Years
- Income Change Retirement
- Annuity Move home

Transforming Data
- Financial Indicators
- Transforming Data
- Analytics
Tracking the Customer Lifecycle – Telco

Kids
- Funky Phone
- Pay as You Go
- Heavy texting

Young Adults
- Pay Monthly
- Smart Phone
- Data users

Middle Aged
- Good to talk
- Bluetooth
- Location-based services

Golden Years
- Simpler handset
- Skype to grandchildren
- Emergency services
What is Survival Analysis?
- Analysis of **TIME**

- To understand length of time before an event occurs
- To predict time till next event
- To analyse duration of time in a particular state

“Event” can be:

- Customer churn
- Take-up new product
- Default on credit
- Make next purchase
- ...

How does Survival Analysis differ from Churn Analysis?

<table>
<thead>
<tr>
<th><strong>Churn Analysis</strong></th>
<th><strong>Survival Analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Examines customer churn within a set time window e.g. next 3 or 6 months</td>
<td>- Examines how churn takes place over time</td>
</tr>
<tr>
<td>- Predicts likelihood of customer to churn during the defined window</td>
<td>- Describes or predicts retention likelihood over time</td>
</tr>
<tr>
<td>- No indication about subsequent risk of churn</td>
<td>- Identifies key points in customer lifecycle</td>
</tr>
<tr>
<td>- Does not provide information on customer lifetime value</td>
<td>- Informs customer lifetime value</td>
</tr>
</tbody>
</table>
The value of understanding both Churn and Survival Time

<table>
<thead>
<tr>
<th>Churn</th>
<th>Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Act on imminent event</td>
<td>• Plan the customer lifecycle</td>
</tr>
<tr>
<td>• Understand combination of factors that are causing the current high probability of churn</td>
<td>• Understand how to extend time as customer before churn is imminent</td>
</tr>
<tr>
<td>• Understand why some customers churn</td>
<td>• Understand why some customers are retained longer than others</td>
</tr>
<tr>
<td></td>
<td>• Act on predicted changes in survival time</td>
</tr>
</tbody>
</table>
Customer Survival – a Censored Data Problem

- You know most about the customers you’ve lost
- You want to predict the future retention of customers you haven’t yet lost

Diagram:
- Lapsed Case
- Censored Case

(time) now
Terminology used in Survival Analysis

• Hazard Function
  – the risk of churn in a time interval after time t, given that the customer has survived to time t
  – usually denoted as: \( h(t) \)

• Survival Function
  – the probability that a customer will have a survival time greater than or equal to t
  – usually denoted as: \( S(t) \)

• Hazard and Survival functions are mathematically linked - by modelling Hazard, you obtain Survival
Example Hazard Function – the classic “Bathtub” curve
Example Survival Curve

80% probability of surviving beyond 7 years

50% probability of surviving beyond 8 years

Area under curve = expected survival time
Descriptive Survival Analysis

• Compute the survival curve for your customer base
  – Understand ‘natural patterns’ in customer survival
  – Identify key points where survival rates fall
• Compare survival curves between
  – Demographic groups
  – Customer segments
  – Sales channels
  – Product plans, etc
• Identifies key factors influencing ‘time till churn’
• Enables you to predict monthly numbers of churners
  – but does not identify which customers will churn
• Most widely used method: Kaplan-Meier
1st Case Study
Which types of customers lapse early?

- Financial services company cross-selling Personal Accident insurance via telemarketing

- Company experienced an increase in monthly lapse rates and reduction in retention levels

- Wanted to understand which types of customers were lapsing early and identify optimal intervention point for reducing lapse rates
Descriptive Survival Analysis – by Age Bands

- Survival chances increase with Age
  - the older the customer, the longer they are likely to retain PA insurance

Results have been disguised
Predicting Survival Times

- **Hazards Model**
  - a model for predicting the hazard of an individual

- **Cox Proportional Hazards Model**
  - a particular form of hazards model, for predicting hazard as a combination of survival time and individual characteristics

\[ h(t,x,b) = h_0(t) \cdot e^{xb} \]

- **Baseline hazard**
- **Individual effect:** data value \( x \), regression coefficient \( b \)
Case Study Example: Survival Model for European Pre-pay Mobile Phone Operator

• Data from the Data Warehouse extracted for a sample of pre-pay mobile customers

• Both active customers and previous churners were represented

• Wide range of variables and attributes were extracted, that could help to explain length of customer relationship

Source of Case Study: Teradata Partners User Group Conference
Example data for Pre-pay Survival Analysis

- **Calling data**
  - Inbound / Outbound
  - Home / Roam
  - Voice / SMS (inbound and outbound)
  - Voice Mail usage
  - In-network / Out of network
  - Dropped calls
  - Customer care interactions
  - Product usage
  - Volatility of call patterns

- **Top-up data**
  - Frequency of top-ups
  - Time between top-ups
  - Value of top-ups

- **Customer data**
  - Age
  - Gender
  - Geodemographic data - postcodes
  - Handset information
  - Registered
Example Results: Key factors that influence lifetime of a pre-pay customer

- Prepayment top-up behaviour
  - High value prepayments
  - Medium value prepayments
  - Frequent prepayments made

- Calling behaviour in home calling area
  - Value of outbound voice calls
  - Number of inbound calls and text messages
  - Use of added-value services, such as voicemail
  - Out of network outbound voice calls

- Customer Demographics
  - Gender
  - Age
  - Geodemographic segments

- Quality issues
Example Results: How Factors Influence Survival - Customers making frequent pre-payments

![Graph showing overall survival probability over months of survival for different prepayment frequencies.]
Example Results: How Factors Influence Survival – Customers making high-value pre-payments
Outputs from Predictive Analysis

• Survival curve – all customers and sub-sets

• Key factors influencing “time till churn”

• Survival model – can apply to individual customers
  – Customers should be regularly rescored, and their scores saved and monitored
Business Applications of Survival Analysis
Customer Management

• Examine and act on predicted customer survival rates over time:

  – Identify customers whose predicted survival rates are low or rapidly falling

  – Examine implications if a key behaviour could be changed

  – Take the right marketing actions aimed at influencing behaviours with greatest impact on predicted survival rates

  – Address some behaviours by modifying service design or terms of use
What are the implications of changes in the customer’s behaviour on predicted survival?
What are the implications of changes in the customer’s behaviour on predicted survival?

<table>
<thead>
<tr>
<th>Frequent Prepayments</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 Euro Prepayment</td>
<td>0</td>
</tr>
<tr>
<td>30 Euro Prepayment</td>
<td>0</td>
</tr>
<tr>
<td>Recent Outbound Voice Calls</td>
<td>2</td>
</tr>
<tr>
<td>Outbound Voice Calls 2 Months ago</td>
<td>8</td>
</tr>
<tr>
<td>Recent Inbound Voice Calls</td>
<td>2</td>
</tr>
<tr>
<td>Recent Text Messages Sent</td>
<td>2</td>
</tr>
<tr>
<td>Text Messages Sent 2 Months ago</td>
<td>3</td>
</tr>
<tr>
<td>Recent Voicemail Use</td>
<td>5</td>
</tr>
<tr>
<td>Recent Out-of-network Voice Calls</td>
<td>5</td>
</tr>
</tbody>
</table>
What are the implications of changes in the customer’s behaviour on predicted survival?

| Frequent Prepayments | 0 |
| 20 Euro Prepayment | 0 |
| 30 Euro Prepayment | 1 |
| Recent Outbound Voice Calls | 2 |
| Outbound Voice Calls 2 Months ago | 8 |
| Recent Inbound Voice Calls | 2 |
| Recent Text Messages Sent | 2 |
| Text Messages Sent 2 Months ago | 3 |
| Recent Voicemail Use | 5 |
| Recent Out-of-network Voice Calls | 5 |
Further Business Applications

• Business Planning
  – Forecast monthly numbers of lapses and use to monitor current lapse rates

• Lifetime Value prediction
  – Derive LTV predictions by combining expected survival times with monthly revenues

• Active customers
  – Predict each customer’s time to next purchase, and use to identify “active” vs. “inactive” customers

• Campaign evaluation
  – Monitor effects of campaigns on survival rates
Applications to different industries and business problems

- Telco – customer lifetime and LTV
- Insurance – time to lapsing on policy
- Mortgages – time to mortgage redemption
- Mail Order Catalogue – time to next purchase
- Retail – time till food customer starts purchasing non-food
- Manufacturing - lifetime of a machine component
- Public Sector – time intervals to critical events
Business Benefits of Survival Analysis

• Improved planning and budgeting through better understanding of future events over time

• Ability to plan timing of churn-related customer communications

• Greater ability to manage customer lifecycles

• Better understanding of factors causing customers to stay for different lengths of time, enabling those factors to be influenced - either by improving service design or at customer level
Thank you!

Barry Leventhal

+44 (0)7803 231870
Barry@barryanalytics.com