SAS® GLOBAL FORUM 2016

IMAGINE. CREATE. INNOVATE.

Customer Lifetime Value Modeling

Vadim Pliner, Verizon Wireless

#SASGF

What is Customer Lifetime Value (LTV)?

- LTV is a prediction of the net profit attributed to the entire future relationship with a customer.

$$LTV = \sum_{t=1}^{N} p_t * M_t,$$

where p_t is the estimated probability to remain the company's customer for the next t months/weeks/… and M_t is the profit margin during period t.

"Survival" Component of LTV

- Survival analysis is used for data on the length of time to a specific event (e.g. customer churn) in the scenario when some subjects (e.g. customers) have not experienced the event during the observation period.

- Survival analysis techniques for LTV modeling: parametric regression models, Cox proportional hazards regression, and discrete-time logistic hazard models.

Discrete-time Logistic Hazard Models

- Discrete-time logistic hazard models are better suited for events that are observed in discrete units and are able to deal with various complexities of real world applications, including time-dependent covariates and nonlinear hazards.

- Special case:

$$\text{Log}\,[h(t)/(1-h(t))] = \psi(t,\alpha) + \sum_{j=1}^{m} \beta_j x_j \;,$$

where $h(t)$ is the hazard function, x_j are covariates, α, β are vectors of parameters to be estimated, and …

Discrete-time Logistic Hazard Models

$\psi(t,\alpha)$ is a cubic spline: $\psi(t,\alpha) = \sum\limits_{j=1}^{J} \alpha_j\, b(t,k_j),$

where J is the number of knots k_j and

$b(t, k_j) = I\{t> k_j\}(t- k_j)^3 - t^3 + 3\, k_j\, t^2 - 3\, k_j^2 t.$

Such a selection of $\psi(t,\alpha)$ makes the hazard function constant when

$t > k_J$ (assuming the knots are sorted in the ascending order).

Discrete-time Logistic Hazard Models

- The likelihood function for the discrete-time logistic hazard models is equivalent to the likelihood function of the logistic regression model for expanded data:

t	event	covariates
0	0	$\mathbf{x}_i$
1	0	$\mathbf{x}_i$
……………………..		
T_{i-1}	0	$\mathbf{x}_i$
T_i	C_i	$\mathbf{x}_i$,

T_i is the time of the last observation; $C_i = 0$ means censoring and $C_i = 1$ indicates that the event of interest occurred at time T_i.

Discrete-time Logistic Hazard Models

- Macro %LOGHAZ (Pliner 2005) does data expansion, sampling non-events to reduce the computational burden caused by the amount of data (if needed), and applies PROC LOGISTIC to get maximum likelihood estimates of parameters α and β.

- The above approach of survival analysis has been also implemented in the SAS Enterprise Miner starting from version 7.1 (Schubert et al 2012).

Predicting Profit Margins

- Linear model:

$$M_t = \sum_{j=1}^{k} a_j M_{t-j} + \sum_{i=1}^{m} b_i z_i, + ct + d,$$

  using as predictors latest historic margins ($M_{t-1,\ldots,}$ M_{t-k}) in addition to other variables (z_i)

- The parameters above can be easily estimated with either PROC GLM or PROC REG, but one may encounter a problem if outliers are present in data. In fact profit margins can have extremely high positive and even negative (in case of refunds) values especially if we deal with business customers.

Predicting Profit Margins: Robust Regression

- Four methods of robust regression are implemented in PROC ROBUSTREG: the Huber M estimation, Least Trimmed Squares (LTS), S estimation, and MM estimation.

- The last three (LTS, S, and MM) are suited for scenarios with outliers in both dependent and independent variables, which is the case when we deal with outliers in profit margins. The method is specified on the PROC ROBUSTREG statement (METHOD=LTS or METHOD=S or METHOD=MM).

References

- Pliner V 2005. "Predictive Modeling Using Survival Analysis." *Proceedings of NESUG 2005 (18th Annual NorthEast SAS Users Group Conference)*; Portland, ME.

- Potts, W. 2003. "Survival data mining: predictive hazard modeling for customer history data." Will Potts, Data Miners Inc.

- Schubert, S, Haller, S, and Lee T. 2012. "It's About Time: Discrete Time Survival Analysis Using SAS® Enterprise Miner™." *Proceedings of the SAS Global Forum 2012 Conference.* Orlando, FL.

SAS® GLOBAL FORUM 2016
IMAGINE. CREATE. INNOVATE.

Contact information:
vadim.pliner@verizonwireless.com

Can you contact me now? ☺