Statistical Methods in CRM

V. Kumar
Andrew Petersen

Instructor's Presentation Slides

©Dr. V. Kumar

Chapter Three

Customer Acquisition

©Dr. V. Kumar

Introduction

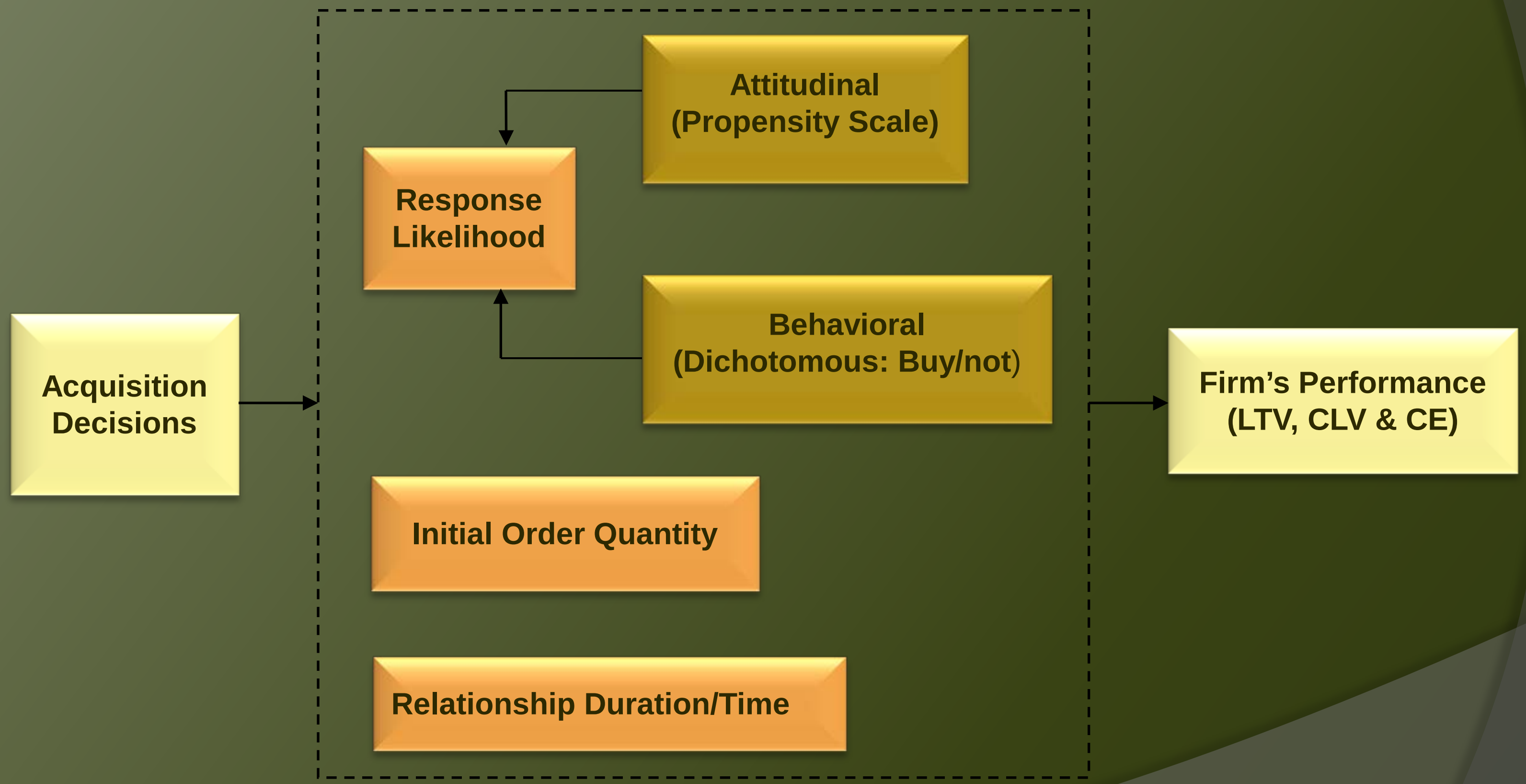
- This chapter provides guidelines on using a company's prospect & customer data to aid in successful and profitable customer acquisition.

- Type of statistical model to be used depends on customer purchase being;
  - Contractual *like Newspaper Subscriptions and telecommunications services*
  - Non-Contractual *like Retailing firms and Catalog Companies*

  - Researchers in the past have paid attention to key issues which occur in the Customer Acquisition process.

©Dr. V. Kumar

Issues in Customer Acquisition Models



Issues Addressed in Customer Acquisition Models (LTV, Lifetime Value; CE, Customer Equity

©Dr. V. Kumar

Review of Customer Acquisition Models

Research Interest	Specification	Estimation	Representative Studies
Probability of being acquired	Logit	MLE	Hansotia and Wang (1997)
			Wangenheim and Bayon (2007)
	Probit	MLE	Reinartz, Thomas and Kumar (2005)
	Linear Regression	OLS	Lix, Berger and Magliozzi (1995)
	Log-linear	MLE	Lix, Berger and Magliozzi (1995)
No. of newly acquired customers	Systems of Linear Regression	3SLS	Lewis (2006b)
	Vector Autoregression		Villanueva, Yoo and Hanssens (2008)
Initial order quantity	Systems of Linear Regression	3SLS	Lewis (2006b)
Duration/Time	Hazard Function		Lewis (2006a)
			Schweidel, Fader and Bradlow (2008)
Firm's Performance (LTV, CLV and CE)	Tobit	MLE	Lewis (2006)
			Hansotia and Wang (1997)
			Reinartz, Thomas and Kumar (2005)
			Thomas, Reinartz and Kumar (2004)
	Vector Autoregression		Villanueva, Yoo and Hanssens (2008)
	Decision Calculus	Deterministic	Blattberg and Deighton (1996)
	Decision Calculus	Deterministic	Berger and Bechwati (2001)

MLE, Maximum Likelihood Estimation; OLS, Ordinary Least Squares; 3SLS, Three-Stage Least Squares

©Dr. V. Kumar

Data for Empirical Examples

A dataset titled: 'Customer Acquisition'. *A representative sample of 500 prospects from a typical B2B firm*

Variable	
Customer	Customer Number (from 1 to 500)
Acquisition	1 if the prospect was acquired, 0 otherwise
First_Purchase	Dollar value of the first purchase (0 if the customer was not acquired)
CLV	The predicted Customer Lifetime Value score. It is 0 if the prospect was not acquired or has already churned from the firm (000's)
Duration	The time in days that the acquired prospect has been or was a customer, right-censored at 730 days
Censor	1 if the customer was still a customer at the end of the observation window, 0 otherwise
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm
Ret_Expense	Dollars spent on marketing efforts to try and retain that customer
Ret_Expense_SQ	Square of dollars spent on marketing efforts to try and retain that customer
Crossbuy	The number of categories the customer has purchased
Frequency	The number of times the customer purchased during the observation window
Frequency_SQ	The square of the number of times the customer purchased during the observation window

Response Probability 1/6

- The first issue in customer acquisition is to model the probability of prospects being acquired.
- Hansotia and Wang (1997) adopted a logistic regression to model prospects' probability of response and used prospect profile variables as predictors.
- In logistic regression, because we can only observe whether the prospects respond or not, a latent response variable *y** indicating unobserved utility is assumed. Thus, *y* is usually defined that:

$$y_i^* = \beta x_i + \varepsilon_i \quad (1)$$

$$y_i = 1 \quad if\ y_i^* > 0$$

$$y_i = 0 \quad if\ y_i^* \leq 0$$

where y_i = the acquisition of customer i (1 = acquired, 0 = not acquired) and x_i = a vector of covariates affecting the acquisition of customer i.

©Dr. V. Kumar

Response Probability 2/6

The probability that the prospect responds is given by:

2) $$\Pr(y=1)=\Pr(y_i^*\ >0)=\Pr(\varepsilon_i\ >\ -\beta x_i)$$

For a logistic regression, ϵ_i has logistic distribution, with mean zero and variance equal to $\pi^2/3$. The cumulative distribution function of ϵ_i is expressed as:

3) $$F_\varepsilon(\beta x_i)=\frac{1}{1+e^{-\beta x_i}}$$

and

4) $$\Pr(y=1)=1-\Pr(\varepsilon_i\le-\beta x_i)=1-F_\varepsilon(-\beta x_i)=\frac{1}{1+e^{-\beta x_i}}$$

Setting the estimated value of Pr(y=1) to $\hat{p}$ when β is estimated by $\hat{\beta}$ we have:

5) $$\hat{p}=\Pr(y=1)=\frac{1}{1+e^{-\hat{\beta}x_i}}$$

©Dr. V. Kumar

Response Probability 3/6

Equation (5), the probability of response p is estimated by the log-odds function as the well-known logistic regression model:

6) $$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}X$$

The likelihood functions for the logit model is the product of the choice probabilities over the *i* individuals, that is:

7) $$L(\beta) = \prod_{i=1}^{N} \left(\Lambda(X_i\beta)\right)^{y_i} (1-\Lambda(X_i\beta))^{1-y_i}$$

where $\Lambda(X_i\beta)$ is the cumulative distribution function according to the standardized logistic distribution. And the log-likelihood is:

8) $$\boldsymbol{l(\beta) = \sum_{i=1}^{N} y_i \log \Lambda(X_i\beta) + \sum_{i=1}^{N} (1-y_i) \log\left(1-\Lambda(X_i\beta)\right)}$$

Response Probability 4/6

Owing to fact that:

9) $$\frac{\partial \Lambda(X_i\beta)}{\partial \beta} = \Lambda(X_i\beta)(1 - \Lambda(X_i\beta))X_i'$$

the gradient is given by:

10) $$G(\beta) = \frac{\partial l(\beta)}{\partial \beta} = -\sum_{i=1}^{N}(\Lambda(X_i\beta))X_i' + \sum_{i=1}^{N} X_i' y_i$$

and the Hessian matrix is given by:

11) $$H(\beta) = \frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} = -\sum_{i=1}^{N}(\Lambda(X_i\beta)(1 - \Lambda(X_i\beta))X_i' X_i$$

A binary logistic regression model to examine the effect the independent variable has on the dependent variable:

12) $$prob\ (y = 1) = \frac{e^{\beta \cdot x}}{1 + e^{\beta \cdot x}}$$

Response Probability 5/6

Conditioning on receiving WOM referral.

13) $$prob\ (y = 1|x = 1) = \frac{e^{\alpha \cdot z}}{1 + e^{\alpha \cdot z}}$$

where z indicated the two independent variables, source expertise and source similarity

In this case the cumulative distribution function of ϵ for a probit model is:

14) $$F(x_i\beta) = \Phi(x_i\beta) = \int_{-\infty}^{x_i\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

The probit model is also estimated by the maximum likelihood estimation (MLE) method. Following Franses and Paap (2001), the relevant likelihood function is given by:

15) $$L(\beta) = \prod_{i=1}^{N} \left(\Phi(X_i\beta)\right)^{y_i} (1 - \Phi(X_i\beta))^{1-y_i}$$

and the corresponding log-likelihood function is:

16) $$\boldsymbol{l(\beta) = \sum_{i=1}^{N} y_i \log \Phi(X_i\beta) + \sum_{i=1}^{N} (1 - y_i) \log\left(1 - \Phi(X_i\beta)\right)}$$

©Dr. V. Kumar

Response Probability 6/6

Differentiating *l(β)* with respect to *β* gives:

17) $$G(\beta) = \frac{\partial l(\beta)}{\partial \beta} = -\sum_{i=1}^{N} \frac{y_i - \Phi(X_i\beta)}{\Phi(X_i\beta)(1-\Phi(X_i\beta))} \phi(X_i\beta)X_i',$$

and the Hessian matrix:

18) $$H(\beta) = \frac{\partial^2 l(\beta)}{\partial\beta\partial\beta'} = \sum_{i=1}^{N} \frac{\phi(X_i\beta)^2}{\Phi(X_i\beta)(1-\Phi(X_i\beta))} X_i'X_i$$

- Similar to the logit model, the probit model is often used to model binary response variables, especially in cases where there is a desire to estimate a two-stage model.

- Whether it is the case that a logit or probit framework is used to model response probability, the output of the model is quite useful for determining which customer the firm is likely to acquire. In addition, the results of the binary choice model can provide the drivers of customer acquisition which can be useful for managers to make decisions in future customer acquisition campaigns.

©Dr. V. Kumar

Empirical Example – Response Problablity

- Whether we can determine which future prospects have the highest likelihood of adoption. At the end of this example we should be able to identify;
  - The drivers of customer acquisition likelihood
  - The parameter estimates from the Response Probability model

Dependent Variable	
Acquisition	1 if the prospect was acquired, 0 otherwise
Independent Variables	
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm

Empirical Example – Response Probablity (Contd.)

- When we run the logistic regression we get the following result;

Variable	Estimate	Standard Error	p-value
Intercept	-26.206	3.537	< 0.0001
Acq_Expense	0.067	0.010	< 0.0001
Acq_Expense_SQ	-0.00004	0.000008	< 0.0001
Industry	0.033	0.389	0.9326
Revenue	0.032	0.012	0.0070
Employees	0.005	0.001	< 0.0001

- 4 of the 5 independent variables are significant at a p-value of 5% or better. *Industry is the only one that is non significant.*
- Acquisition expense has a positive, but diminishing effect on acquisition likelihood
- A prospect who is B2B (vs. B2C) will not matter will not matter in terms of acquisition likelihood all else being equal
- The higher the *Revenue* the prospect has, the more likely the prospect will be acquired
- The more *Employees* the prospect has, the more likely the prospect will be acquired

Log-odds ratio

- Shows how changes in the drivers of acquisition likelihood are likely to lead to either increases or decreases in acquisition likelihood (when dealing with a logistic regression)

- When we compute the log odds ratio for each of the statistically significant variables we get the following results for an increase in 1 unit of the independent variable.

Variable	**Log Odds Ratio**
Acq_Expense	exp(0.06696 – 0.00008*Acq_Expense)
Revenue	1.033
Employees	1.005

- With regard to *Acq_Expense*, the Odds Ratio is dependent on the level of *Acq_Expense ($500->$501, then 2.7% increase)*

- For each increase in *Revenue* by $1 million the acquisition likelihood should increase by 3.3%.

- For each increase in *Employees* by 1 person the acquisition likelihood should increase by 0.5%.

©Dr. V. Kumar

Number of Newly Acquired Customers & Initial Order Quantity

The number of newly acquired customers and the average order size for new customers are modelled in the following two equations:

19) $$CA_{cq} = \beta_1 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_3 + \beta_5 x_4 + \beta_6 x_5 + \varepsilon_1$$

where CA_{cq} denoted the number of newly acquired customers, x_1 ~ x_5 denoted the explanatory variables, including shipping, pricing and customer base terms, and

20) $$Amount(new) = \beta_1 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_3 + \beta_5 x_4 + \beta_6 x_5 + \varepsilon_2$$

where *Amount (new)* the average order size for new customers, x_1 ~ x_5 denoted the shipping variables and pricing variable.

As the estimated variables are treated as potentially endogenous, the authors used a three-variable vector autoregression (VAR) modeling technique:

21) $$\begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \sum_{l=1}^{p} \begin{pmatrix} a_{11}^l & a_{12}^l & a_{13}^l \\ a_{21}^l & a_{22}^l & a_{23}^l \\ a_{31}^l & a_{32}^l & a_{33}^l \end{pmatrix} \begin{pmatrix} x_{t-l} \\ y_{t-l} \\ z_{t-l} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix}$$

Where x denoted the number of customers acquired through the firm's marketing actions, y denoted the number of customers acquired through word of mouth, and z denoted the firm's performance. The subscript t denoted for time, p denoted the lag order of the model and (e_{1t}, e_{2t}, e_{3t})' are white-noise disturbances distributed as N(0, Σ). The direct effects are captured by a_{31}, a_{32} ; the feedback effects are captured by a_{13}, a_{23}; the cross-effects are captured *by a_{12}, a_{21}* ; and the reinforcement effects are captured *by a_{11}, a_{22}, a_{33}.*

Empirical Example: Number of Newly Acquired Customers

- Besides being able to predict, it is interesting to know how well our Response Probability model accurately predicts the total number of customers we are likely to acquire and specifically which prospects we are most likely to acquire. By the end of this chapter, we should be able to;
  - Predict the number of prospects we are likely to acquire
  - Determine the accuracy of our prediction

- The information we need for this prediction includes;

Dependent Variable	
Acquisition	1 if the prospect was acquired, 0 otherwise
Independent Variables	
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm

The predictive accuracy of the model

- For a logistic regression we must apply the proper probability function as noted earlier in the chapter (see Equation 5).

$$P(Acquisition = 1|X\beta) = \frac{1}{1 + exp(-X\beta)}$$

		Actual Acquisition		
		0	1	Total
Predicted Acquisition	0	183	19	**202**
	1	25	273	**298**
	Total	**208**	**292**	**500**

Predicted versus actual acquisition

→ In this case the sum of the diagonal is 456 and it is accurate 91.2% (456/500) (> the best alternative: 58.4%)

Empirical Example: Initial Order Quantity

- How much value the customer is likely to provide
  - Determine the drivers of initial order quantity (value)
  - Predict the expected initial order quantity for each prospect
  - Determine the predictive accuracy of the model

Dependent Variables	
Acquisition	1 if the prospect was acquired, 0 otherwise
First_Purchase	Dollar value of the first purchase (0 if the customer was not acquired)
Independent Variables	
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm

E(*Initial Order Quantity*) = P(*Acquisition* = 1)* E(*First Purchase* | *Acquisition* = 1)

Inverse Mills ratio

- Once we estimate the probit model we need to create a new variable, λ, representing the correlation in the error structure across the two equations. This variable, known as the sample selection correction variable, will then be used as an independent variable in the *First Purchase* model to remove the sample selection bias in the estimates. To compute λ we use the following equation, also known as the **inverse Mills ratio**:

$$\lambda = \frac{\phi(X'\beta)}{\Phi(X'\beta)}$$

ϕ: the normal probability density function, Φ: the normal cumulative density function, X: the value of the variables in the Acquisition model, & β: the coefficients derived from the estimation of the Acquisition model

Empirical Example: Initial Order Quantity

Acquisition Model	Estimate
Intercept	-15.134
Acq_Expense	0.039
Acq_Expense_SQ	-0.00002
Industry	0.051*
Revenue	0.018
Employees	0.003

First_Purchase Model	Estimate
Intercept	-62.261
Acq_Expense	0.705
Acq_Expense_SQ	-0.0008
Industry	-2.764*
Revenue	3.026
Employees	0.254
Lambda (λ)	19.101

** denotes not significant at $p < 0.05$*

- A potential selection bias problem since the error term of our selection equation is correlated positively with the error term of our regression equation (λ is positive and significant)
- All other variables of the *First Purchase* model are significant with the exception of *Industry*

The mean absolute deviation (MAD)

$$E(\textit{Initial Order Quantity}) =$$

$$P(\textit{Acquisition} = 1) * E(\textit{First\_Purchase} | \textit{Acquisition} = 1)$$

$$= \Phi(X'\beta) * (\gamma'\alpha + \mu\lambda)$$

- In this case Φ is the normal CDF distribution, *X* is the matrix of independent variable values from the *Acquisition* equation, *β* is the vector of parameter estimates from the *Acquisition* equation, *γ* is matrix of independent variables form the *First_Purchase* equation, *α* is the vector of parameter estimates from the *First_Purchase* equation, *μ* is the parameter estimate for the inverse Mills ratio, and *λ* is the inverse Mills ratio.

- The mean absolute deviation (MAD) and mean absolute percent error (MAPE) equations are;

  MAD = Mean(Absolute Value[E(*Initial Order Quantity*) – *First_Purchase*])
  MAPE = Mean(Absolute Value[(E(*Initial Order Quantity*) – *First_Purchase*)/*First_Purchase*])

- We would find that MAD = 127.17 and a MAPE = 135.48%.

Ours	Benchmark
$51.96 (18.69%)	$127.17 (135.48%)

Duration/Time

Accelerated failure time models (Kalbfleish and Prentice 1980) estimated as shown in the following specification:

22)

$$\log(t_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \sigma\varepsilon_i,$$

where ϵ_i is a random disturbance term, x_i and x_i^2 are discount and discount squared, and β_s and σ are parameters needed to be estimated.

The hazard function for *T* can be written in terms of this base-line hazard $\lambda_0(\cdot)$ according to:

23)

$$\lambda(t; z) = \lambda_0(te^{-z\beta})e^{-z\beta}$$

The survival function is:

24)

$$F(t; z) = \exp\left[-\int_0^{te^{-z\beta}} \lambda_0(u)\, du\right]$$

and the density function is the product of (15) and (16).

©Dr. V. Kumar

Duration/Time

The parametric distribution for the probability of acquiring service at time :

25) $$f(t_A|\Theta) = S_A(t_A - 1|\Theta) - S_A(t_A|\Theta)$$

where $S_A(t_A|\Theta)$ is the survival function of the parametric distribution, $t_A = 1,2,\ldots,T$, and T is the length of the observation period.

Table **3.4** **Hazard Function and Survival Function of Three Baseline Distributions**

Baseline Distribution	**Hazard Function**	**Survival Function**
Weibull	$h(t\|\gamma,\alpha) = \gamma\alpha t^{\alpha-1}$	$S(t\|\gamma,\alpha) = e^{-\gamma t^{\alpha}}$
Log-logistic	$h(t\|\gamma,\alpha) = \frac{\gamma\alpha(\gamma t)^{\alpha-1}}{1+(\gamma t)^{\alpha}}$	$S(t\|\gamma,\alpha) = \frac{1}{1+(\gamma t)^{\alpha}}$
Expo-power	$h(t\|\gamma,\alpha,\theta) = \gamma\alpha t^{\alpha-1}e^{\theta t^{\alpha}}$	$S(t\|\gamma,\alpha,\theta) = e^{(\gamma/\theta)(1-e^{\theta t^{\alpha}})}$

Source: Schweidel, Fader and Bradlow (2008)

Empirical Example - Duration/Time

- How long a newly acquired customer will still be a customer
  - Determine the drivers of new customer duration
  - Predict the duration of each new customer

Dependent Variables	
Duration	The time in days that the acquired prospect has been or was a customer, right-censored at 730 days
Independent Variables	
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Ret_Expense	Dollars spent on marketing efforts to try and retain that customer
Ret_Expense_SQ	Square of dollars spent on marketing efforts to try and retain that customer
Crossbuy	The number of categories the customer has purchased
Frequency	The number of times the customer purchased during the observation window
Frequency_SQ	The square of the number of times the customer purchased during the observation window
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm
Censor	1 if the customer was still a customer at the end of the observation window, 0 otherwise

Empirical Example Duration/Time (Contd.)

Variable	Estimate	p-value
Intercept	2.837	< 0.0001
Acq_Expense	0.007	< 0.0001
Acq_Expense_SQ	-0.00001	< 0.0001
Ret_Expense	0.001	< 0.0001
Ret_Expense	-0.00000004	0.017
Crossbuy	0.098	< 0.0001
Frequency	0.111	< 0.0001
Frequency_SQ	-0.001*	0.173
Industry	0.524	< 0.0001
Revenue	0.012	< 0.0001
Employees	0.0001	< 0.0001
Scale	0.138	
Shape	7.252	

* Denotes not significant at p<0.05

Each of the variables in the model is significant to at least a level of 1% with the exception of *Frequency_SQ*.

The predictive accuracy of the model

$$\ln(\mathit{Duration}) = X'\beta + \sigma\epsilon$$

- We just need to recognize that σ is the scale parameter (0.138) and ϵ is derived from the 50% percentile of the Weibull distribution, or ln(-ln(1-*p*) = ln(-ln(1-0.5) = -0.367. Then to compute Duration we just need take the anti-log of the right hand side. We get the following.

$$\mathit{Duration} = \exp(X'\beta + 0.138*-0.367) = \exp(X'\beta - 0.051)$$

- Now we need to compare the actual *Duration* values for the 157 customers who churned during the observation window with the predicted values of *Duration*. We find a mean absolute deviation (MAD) of 45.97 days and a mean absolute percent error (MAPE) of 13.88%.

Ours	Benchmark
45.97 days (13.88%)	170.77 days (171.92%)

©Dr. V. Kumar

Firm's Performance (LTV, CLV & CE)

For the right-censored situation in which censoring occurs above a value , the Tobit model is:

26)

$$y_i = \begin{cases} y_i^* & if\ y_i^* < y_L \\ y_L & if\ y_i^* \geq y_L \end{cases}$$

$$y^* = \beta x_i + u_i, u_i \sim N(0, \sigma^2)$$

The likelihood function, which is the probability of observing the sample values, is:

27)

$$L = \prod_{i=1}^{M} \{Prob[PVR = observed\ value]\}^S \{Prob[PVR \geq observed\ value]\}^{1-S}$$

where S=1 if observation i is uncensored; 0 otherwise.

Blattberg and Deighton (1996) drew a curve of the actual acquisition probability by the equation:

28)

$$a = ceiling\ rate\ \times \left[1 - e^{(-k_1 \times \$A)}\right]$$

Where k is a constant that controls the steepness of the curve.

©Dr. V. Kumar

Empirical Example – Firm's Performance

- Determine whether the customers that were acquired are profitable
  - Determine the drivers of customer profitability
  - Determine the predictive accuracy of the customer profitability model

Dependent Variables	
Censor	1 if the customer was still a customer at the end of the observation window, 0 otherwise
CLV	The predicted Customer Lifetime Value score. It is 0 if the prospect was not acquired or has already churned from the firm (000's).
Independent Variables	
Acq_Expense	Dollars spent on marketing efforts to try and acquire that prospect
Acq_Expense_SQ	Square of dollars spent on marketing efforts to try and acquire that prospect
Ret_Expense	Dollars spent on marketing efforts to try and retain that customer
Ret_Expense_SQ	Square of dollars spent on marketing efforts to try and retain that customer
First_Purchase	Dollar value of the first purchase (0 if the customer was not acquired)
Crossbuy	The number of categories the customer has purchased
Frequency	The number of times the customer purchased during the observation window
Frequency_SQ	The square of the number of times the customer purchased during the observation window
Industry	1 if the prospect is in the B2B industry, 0 otherwise
Revenue	Annual sales revenue of the prospect's firm (in millions of dollars)
Employees	Number of employees in the prospect's firm

Empirical Example – Firm's Performance (Contd.)

Censor Model	**Estimate**
Intercept	-32.151
Acq_Expense	0.070*
Acq_Expense_SQ	-0.0001
Ret_Expense	0.006
Ret_Expense_SQ	0.0000005*
First_Purchase	-0.002*
Crossbuy	0.758
Frequency	0.852
Frequency_SQ	-0.004*
Industry	4.588
Revenue	0.072
Employees	0.002*

CLV Model	**Estimate**
Intercept	-2.613
Acq_Expense	0.009
Acq_Expense_SQ	-0.00001
Ret_Expense	0.004
Ret_Expense_SQ	-0.0000007
First_Purchase	0.003
Crossbuy	0.203
Frequency	0.155
Frequency_SQ	-0.005
Industry	0.566
Revenue	0.008
Employees	0.0002*
Lambda (λ)	0.174*

** denotes not significant at $p < 0.05$*

- We cannot interpret this to mean that there is not selection problem, only that it is unlikely given the error term of our selection equation is not correlated with the error term of our regression equation. (λ is positive, but not significant)
- We also see that with the exception of *Employees*, the remaining variables in the *CLV* model are all significant at $p < 0.05$.

The mean absolute deviation (MAD)

- Our next step is to predict *CLV* for each of the customers and see if our predictions are accurate. We do this by starting with the equation for expected future profitability at the beginning of this example:

$$E(\textit{Future Profitability}) = P(\textit{Censor} = 1)*E(\textit{CLV}|\textit{Censor} = 1)$$

$$= \Phi(X'\beta)*(\gamma'\alpha + \mu\lambda)$$

- We do the comparison for the *CLV* values of the 135 customers where *Censor* = 1, i.e. the customers that are still active with the firm. We find for this data that the mean absolute deviation (MAD) is 0.29 (or $290) and the mean absolute percent error is 4.5%. For our benchmark model we use the mean of 6.58 as our predicted value for each customer. In this case we get a MAD of 0.71 (or $710) and a MAPE of 11.05%. We see that our *CLV* model provides a significant better prediction of *CLV* than the benchmark model. This shows that identifying the drivers of CLV can significantly help a firm understand which customers are most likely to be profitable in the future.

Ours	Benchmark
$290 (4.5%)	$710 (11.05%)

Implementation & Summary

1-1) The PROC Logistic feature in SAS to implement the logistic regression

2-1) A SAS Data step and the Freq procedure

3-1) PROC logistic with a probit link function to estimate the model of customer acquisition

3-2) A SAS Datastep to compute the inverse Mills Ratio using the output of the probit model.

3-3) Run an OLS regression using PROC Reg and added the inverse Mills ratio as an additional variable

4-1) PROC Lifereg procedure in SAS where the dependent variable was right-censored

5-1) ROC Logistic and a probit link function to model the probability that a customer has already quit.

5-2) A SAS Datastep to compute the inverse Mills ratio

5-3) PROC Reg in SAS to determine the drivers of CLV

- When firms engage in optimal prospect selection by understanding the drivers of customer acquisition and the ‘right’ amount of customer acquisition effort by understanding the relationship between marketing spending and customer value, the result can generate significant customer and firm profitability.

End of Chapter Three

©Dr. V. Kumar