

Estimación de Lifetime Value

Basada en Comportamiento Transaccional¹

Pablo Marín V.²

pamarin@analytics.cl

Luis Aburto L.

luaburto@analytics.cl

Máximo Bosch

mbosch@dii.uchile.cl

Marcel Goic

mgoic@dii.uchile.cl

Septiembre 2005

¹ Estudio financiado por proyecto FONDEF D03I1057

² Este paper resume el trabajo de título de Pablo Marín para optar a la mención de Ingeniería Civil Industrial en la Universidad de Chile. Luis Aburto, Máximo Bosch y Marcel Goic son los profesores que guiaron el desarrollo del trabajo, pertenecientes al Departamento de Ingeniería Industrial de la Universidad de Chile.

Abstract

Lifetime Value Estimation Based on Transactional Data

The *lifetime value* of a customer is the present value of all the future earnings that he will represent for the firm. The estimation of this indicator is central to the customer relationship management concept. It empowers managers with a tool they can use determine how much time, effort and money they can invest on their customers. In retail, Recency, Frequency and Monetary Value (RFM) attributes can be used to characterize a customer's transactional behavior. Models such as the Pareto/NBD (and other simpler ones) use this information to estimate the customer's expected future behavior. Specifically, the number of future purchases, their value, and the activity probability can be estimated to generate a customer's lifetime value. The proposed lifetime value estimation methodology is applied on a cohort of customers of a wholesale supermarket chain. The models prove to be useful in predicting the behavior of segments of customers, specifically on which groups are expected to change their behavior in the future.

Keywords: LTV, lifetime value, CLV, RFM, Pareto/NBD

Abstract

Estimación de Lifetime Value Basada en Comportamiento Transaccional

El *lifetime value* de un cliente es el valor presente de todos los beneficios futuros que éste generará para la empresa. La estimación de este indicador es central para la gestión de las relaciones con clientes: es útil para determinar cuánto tiempo, esfuerzo y dinero invertir en ellos, bajo una perspectiva de rentabilidad de largo plazo. En el ámbito del retail, atributos de Recency, Frequency y Monetary Value (RFM) pueden utilizarse para caracterizar el comportamiento transaccional de un cliente. Modelos como el Pareto/NBD (y otros más simples) se basan en esta información para estimar el comportamiento futuro de los clientes. Específicamente, puede estimarse para el número de compras futuras, el valor de éstas y la probabilidad de fuga, para finalmente obtener el lifetime value del cliente. La metodología propuesta se aplica sobre una base de clientes de un supermercado mayorista. Los modelos utilizados muestran ser útiles a nivel de segmentos de clientes, específicamente en la predicción de cambios en el comportamiento futuro.

Palabras clave: LTV, lifetime value, CLV, RFM, Pareto/NBD

Introducción

Uno de los temas más relevantes de la gestión de las relaciones con los clientes consiste en entender y diferenciarlos. No todos ellos son igualmente valiosos. El principio de Pareto (regla 80-20), postula que comúnmente se observa que el 80% del valor de una empresa es aportado por un 20% de sus clientes. Por otro lado, un buen cliente hoy no necesariamente lo será mañana. El valor de un cliente puede resumirse utilizando un único indicador, su lifetime value. Éste se define como el valor presente de todos los beneficios futuros que éste generará para la empresa. Su estimación es central para la gestión de las relaciones con clientes: es útil para determinar cuánto tiempo, esfuerzo y dinero invertir en ellos, bajo una perspectiva de rentabilidad de largo plazo.

La estimación del indicador mencionado no es trivial, ya que son muchos los factores influyentes. En esta investigación, la estimación del lifetime value se centrará únicamente sobre el comportamiento transaccional de los clientes. En el ámbito del retail, atributos de Recency, Frequency y Monetary Value (RFM) pueden utilizarse para caracterizar el comportamiento transaccional de un cliente. En base a esta información, se construirán modelos que estimen el comportamiento futuro de los clientes, permitiendo la obtención de su lifetime value.

Modelos de lifetime value

En forma general, el lifetime value de un cliente puede definirse utilizando la siguiente fórmula que representa el valor presente de todas las transacciones futuras que realizará un cliente en la empresa.

$$LTV = \int_0^{\infty} v(t)S(t)d(t)dt \quad (1)$$

En ella, $v(t)$ es el valor del cliente en el instante t , $S(t)$ es la función de sobrevivencia (probabilidad de actividad) en t y $d(t)$ es el descuento aplicado en t . Para llevar la fórmula a una más concreta, se establecen los siguientes supuestos sobre el comportamiento de compra de los clientes:

1. Los indicadores de Recency, Frequency y Monetary Value (RFM) de un cliente son información suficiente para representar los procesos de compra y fuga de los clientes. En particular, los datos RFM que se utilizarán serán los siguientes:
 - x : número de transacciones realizadas.
 - t_x : período de tiempo transcurrido desde la primera hasta la última transacción.
 - T : período de tiempo transcurrido desde la primera transacción hasta el final del período.
 - M : monto promedio por transacción.
2. No existe relación entre la frecuencia de compra de un cliente y el monto de sus transacciones.
3. El gasto por transacción de un cliente es estable en el tiempo.

Si se cumplen todas las condiciones anteriores, la fórmula de lifetime value toma la siguiente forma:

$$LTV = (\hat{M} - \hat{C}) \cdot VP(r_{anual}, \hat{x}) \quad (2)$$

donde LTV es la estimación de lifetime value, \hat{M} es el valor promedio de monto por transacción pronosticado, \hat{C} es la estimación del costo asociado al cliente, $DET = VP(r_{anual}, \hat{x})$ es el valor presente del número de transacciones futuras esperadas (dada una tasa de descuento anual r_{anual}) y \hat{x} es el número de transacciones futuras pronosticadas (la estimación de este indicador incluye calcular la probabilidad de actividad del cliente).

Dado que se asumen constantes tanto el monto por transacción como los montos por cliente, el objetivo se reduce a pronosticar frecuencia de compra futura. Por ello, en lo que sigue, el trabajo se enfocará en la estimación del *DET*, a partir del cual finalmente se obtendrá el lifetime value.

Modelo Simple

Este modelo estima el comportamiento futuro del cliente directamente a partir de los datos RFM. Supone que el número de transacciones futuras será proporcional a la que se observaría si el cliente mantuviera su frecuencia de compra constante en el tiempo. Por otro lado, incorpora una estimación de la probabilidad de actividad de un cliente, basada en los indicadores de Recency t_x y T . No se incorpora el efecto de la frecuencia de compra en la probabilidad de actividad, por lo que todo cliente puede “demorarse de la misma forma”. De esta forma, los indicadores de interés se calculan de la siguiente manera:

- Probabilidad de actividad al finalizar el período de calibración

$$P(\text{activo}) = \frac{t_x + 1}{T + 1} \quad (3)$$

- Número de transacciones pronosticadas para un período futuro de largo t

$$\hat{x} = x \left(\frac{t}{T + 1} \right) P(\text{activo}) \quad (4)$$

- Valor presente del número de transacciones futuras esperadas

$$DET = VP(r_{\text{anual}}, \hat{x}) = \sum_{t=1}^{\infty} \frac{\hat{x}}{(1 + r_{\text{per}})^t} [P(\text{activo})]^{t-1} \quad (5)$$

donde r_{per} es la tasa de descuento para el período.

Este modelo, mediante la incorporación de una probabilidad de actividad, dirá que “los mejores clientes de hoy no necesariamente serán los mejores mañana”. De hecho, este modelo exige que los clientes hayan comprado recientemente para que puedan tener un alto lifetime value, independiente de su frecuencia de compra.

La probabilidad de fuga de un cliente tiene efecto directo sobre su *DET*. Los mejores clientes serán aquellos que compraron más frecuente y recientemente.

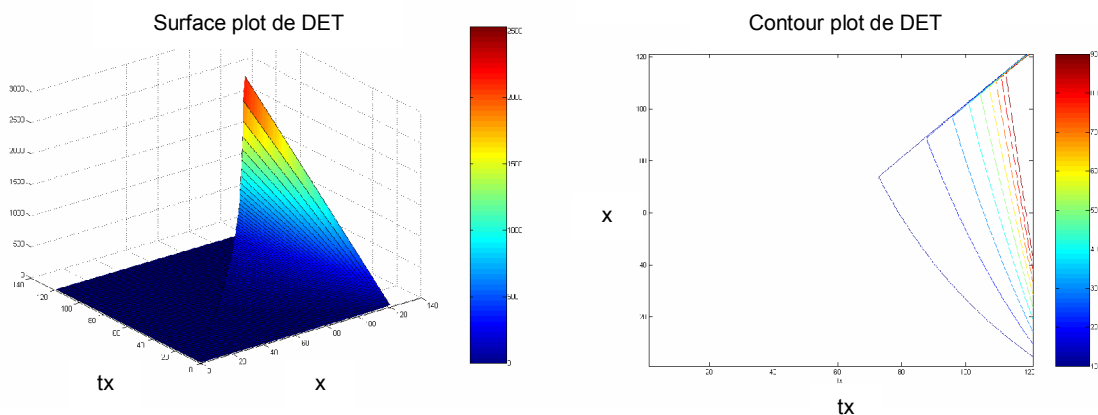


Figura 1: DET según Modelo Simple

Pareto/NBD

El modelo principal de lifetime value que se utilizará es el Pareto/NBD de Schmittlein, Morrison y Colombo (1987), adaptado por Fader, Hardie y Lee (2005). Este modelo incorpora procesos estocásticos para representar el comportamiento de compra de los clientes. Los supuestos básicos que lo conforman son los siguientes:

- Un cliente puede encontrarse en uno de dos estados: activo o inactivo. Clientes activos son aquellos que realizan compras en la empresa, los inactivos son aquellos que se han “fugado” y no volverán a comprar. Para el modelo, un cliente que ha pasado al estado inactivo no puede volver a activarse.
- Clientes activos realizan compras que puede representarse mediante un proceso de Poisson de tasa λ . Esto es equivalente a suponer que el tiempo entre compras se distribuye según una exponencial de media $1/\lambda$.
- Un cliente permanece activo por un período no observable τ , que se distribuye según una exponencial de tasa μ . Transcurrido este período el cliente pasa al estado de inactividad.
- Las tasas de compra (λ de los clientes) se distribuyen en la población según una gamma de parámetros r (forma), α (escala).
- Las tasas de fuga (μ de los clientes) se distribuyen en la población según una gamma de parámetros s (forma), β (escala).
- Las tasas de compra (λ de los clientes) y de fuga (μ de los clientes) varían independientemente entre los clientes.

Bajo los supuestos anteriores, los indicadores de interés para el Pareto/NBD se obtienen de la siguiente manera:

- Probabilidad de actividad al finalizar el período de calibración

$$P(\text{activo}) = \left\{ 1 + \left(\frac{s}{r+s+x} \right) (\alpha + T)^{r+x} (\beta + T)^s A_0 \right\}^{-1} \quad (6)$$

- Número de transacciones pronosticadas para un período futuro de largo t

$$\hat{x} = \frac{(r+x)(\beta+T)}{(\alpha+T)(s-1)} \left[1 - \left(\frac{\beta+T}{\beta+T+t} \right)^{s-1} \right] \cdot P(\text{activo}) \quad (7)$$

- Valor presente del número de transacciones futuras esperadas

$$DET = VP(\delta, x^{\text{futuro}}) = \frac{\alpha^r \beta^s \delta^{s-1} \Gamma(r+x+1) \Psi(s; s; \delta(\beta+T))}{\Gamma(r)(\alpha+T)^{r+x+1} L(r, \alpha, s, \beta / X = x, t_x, T)} \quad (8)$$

donde δ es la tasa de interés compuesta continuamente y $\Psi(\cdot)$ es la función hipergeométrica confluyente de segundo orden.

Uno de los aspectos más interesantes del modelo es que tanto los indicadores Recency como Frequency afectan en la probabilidad de actividad de un cliente. En particular, el Pareto/NBD exigirá que los clientes que hayan comprado más frecuentemente hayan comprado más recientemente. Asimismo, exigirá menos de aquellos que compraron con menor frecuencia. El efecto de la incorporación del la frecuencia de compra en la probabilidad de fuga sobre el *DET* puede observarse en la siguiente figura.

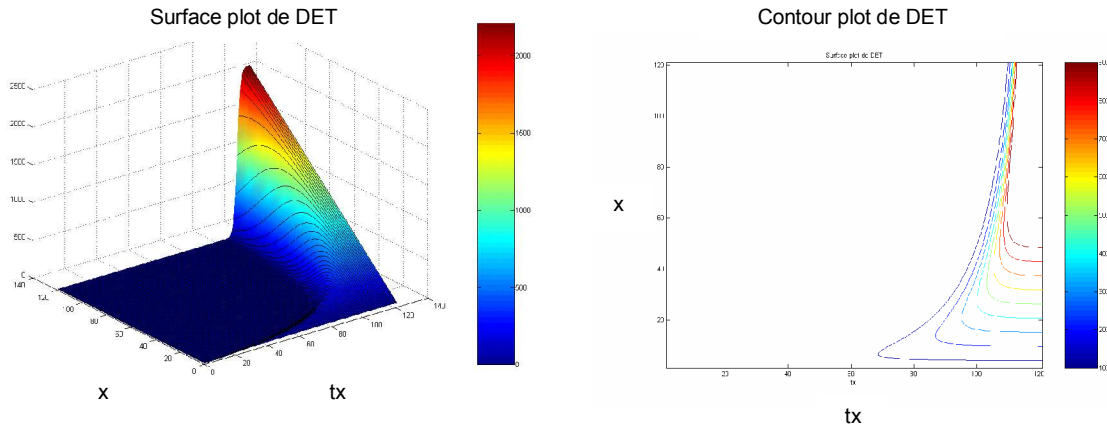


Figura 2: DET según Pareto/NBD

Las curvas de isovalor cambian con respecto a las observadas anteriormente. A medida que aumenta la frecuencia de compra las curvas de isovalor se van haciendo cada vez menos “permisivas” con t_x : clientes más frecuentes deben haber comprado recientemente. Por otro lado, las curvas de isovalor son más “optimistas” al disminuir la frecuencia de compra, pues la probabilidad de actividad no será tan baja como la que pronosticará el modelo simple.

Aplicación de los modelos

Los modelos de lifetime value son aplicados sobre una base de clientes de un supermercado mayorista. Se poseen datos de 24.378 clientes que realizan un total de 372.378 transacciones con la tarjeta de la empresa en un período de ocho meses, gastando un total de \$12.915 millones de pesos. Para la calibración de los modelos se utilizan los primeros cuatro meses de datos. Una vez calibrados los modelos, se utilizan para predecir lo que ocurrirá en los siguientes cuatro meses de datos, de modo de validarlos. El comportamiento de los clientes se resume en la siguiente tabla, que muestra los valores promedio de los quintiles de cada variable.

Tabla 1: Valores Promedio por Quintiles RFM

| Quintil | Datos en Calibración | | | | | Datos en Validación | | |
|---------|----------------------|-------|-----|----------|-------------|---------------------|----------|-------------|
| | x | t_x | T | M | $Monto$ | x | M | $Monto$ |
| 1 | 1 | 1 | 43 | \$6.585 | \$15.879 | 0 | 0 | 0 |
| 2 | 3 | 47 | 96 | \$16.074 | \$70.877 | 2 | \$8.478 | \$39.442 |
| 3 | 8 | 96 | 114 | \$25.867 | \$206.560 | 8 | \$20.746 | \$196.626 |
| 4 | 18 | 113 | 119 | \$39.665 | \$563.169 | 20 | \$35.228 | \$602.028 |
| 5 | 45 | 119 | 121 | \$71.070 | \$1.792.684 | 51 | \$66.752 | \$2.008.947 |

Un 20% de los clientes no vuelven a realizar transacciones en el período de calibración. Por otro lado, otro 40% aumenta (en promedio) su frecuencia de compra.

Se han presentado dos modelos predictivos, el modelo simple y el Pareto/NBD. La aplicación de los modelos para predecir el número de transacciones en el período de validación se resume en el siguiente gráfico. La media incondicional (predecir que se repite el comportamiento observado en el período de calibración) se incluye como referencia. En el período de calibración existen clientes que realizan hasta 121 transacciones, sin embargo, la gran mayoría realiza menos de 20 compras, por lo que el gráfico se centra en el comportamiento de los modelos en estos casos.

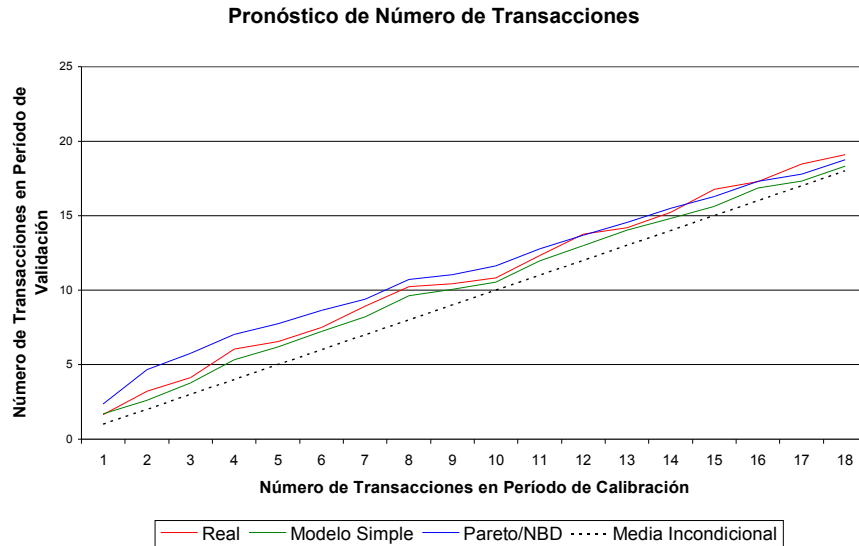


Figura 3: Pronóstico de Número de Transacciones

El Modelo Simple tiende a ajustarse mejor que el Pareto/NBD para los clientes de menor frecuencia. Sin embargo, a medida que aumenta la frecuencia de compra, el Pareto/NBD mejora progresivamente. Como se mencionó anteriormente, el objetivo se había reducido a pronosticar de manera adecuada la frecuencia de compra futura. En la siguiente tabla se resume el error promedio, la desviación y el promedio de los errores absolutos (MAD) cometidos por los modelos, tanto para el número de transacciones (frecuencia de compra futura), como para el monto por transacción y el gasto total en el período de validación.

Tabla 2: Resumen de Errores

| | Error | Media Incondicional | Modelo Simple | Pareto/NBD |
|------------------------------|--------------|----------------------------|----------------------|-------------------|
| Transacciones | Promedio | -0,96 | -0,23 | 0,16 |
| | (Desv. Est). | (10,9) | (11) | (10) |
| | MAD | 6,31 | 6,04 | 6,13 |
| Monto por transacción | Promedio | | -\$295 | |
| | (Desv. Est). | | (\$15.196) | |
| | MAD | | \$9.541 | |
| Gasto Total | Promedio | -\$39.570 | -\$18.449 | -\$6.154 |
| | (Desv. Est). | (\$415.690) | (\$437.840) | (\$414.250) |
| | MAD | \$216.320 | \$211.140 | \$218.690 |

En promedio, el Pareto/NBD es el mejor modelo. Sin embargo, a nivel de clientes no lo es. De hecho, los modelos predictivos no superan a la media incondicional. Esto significa que los modelos no son confiables a nivel de clientes individuales, pero sí lo son a niveles más agregados. En particular, los modelos predictivos son capaces de predecir cambios en el comportamiento futuro de segmentos de clientes, cosa que es imposible hacer utilizando la media incondicional. En la figura, el primer gráfico representa el cambio real en el gasto total observado. En el segundo (a la derecha), se muestra la predicción del Pareto/NBD. El tercer gráfico corresponde a la predicción del modelo simple y el cuarto a la de la media incondicional.

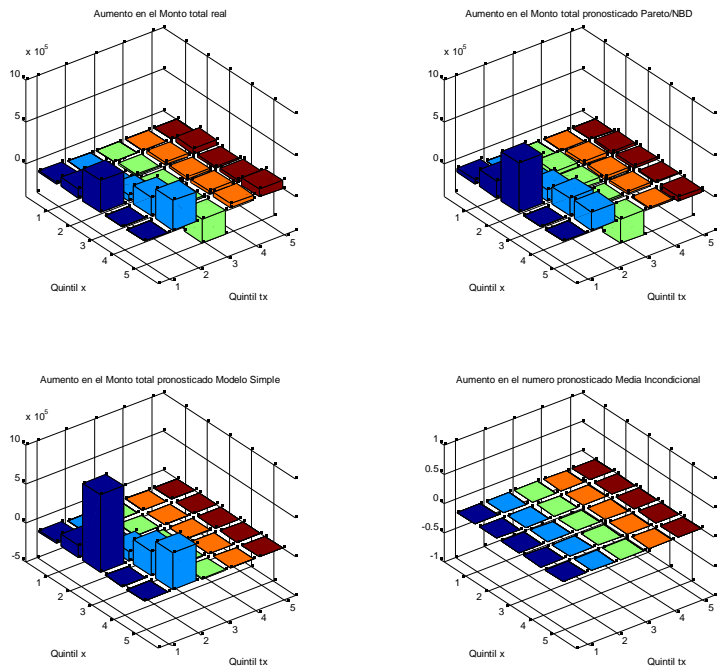


Figura 4: Pronósticos de Cambios de Comportamiento

El Pareto/NBD genera una predicción muy adecuada de los cambios futuros. Supera claramente al modelo simple (y a la media incondicional), por lo que se infiere que es el modelo más confiable a nivel de segmentos de clientes.

Conclusiones

El lifetime value de un cliente sirve para determinar cuánto puede invertirse en un cliente, en base a una perspectiva de rentabilidad de largo plazo. Los modelos de lifetime value aplicados son capaces de predecir adecuadamente cambios de comportamiento a nivel de segmentos de clientes. En particular, el Pareto/NBD muestra ser especialmente útil en este sentido. Basado en sus predicciones, pueden generarse nuevas segmentaciones basadas en base a las estimaciones del comportamiento futuro esperado, de modo de agrupar clientes según lo que se espera harán en el futuro en vez de en función de lo que han hecho.

Por otro lado, se concluye que los atributos de Recency, Frequency y Monetary Value no contienen toda la información necesaria para generar una predicción precisa del comportamiento futuro a nivel de clientes individuales. Se cree que la incorporación de otras variables (por ejemplo variables sociodemográficas de los clientes) complemente las utilizadas y permitan mejorar las predicciones del comportamiento futuro a nivel de clientes individuales.

Referencia Principal

P. Fader, B. Hardie, K. Lee. RFM and CLV: Using Iso-value Curves for Customer Base Analysis. 2005. brucehardie.com/abstracts/abstract-fhl_rfm_clv_2005-02.html