Implementing the Pareto/NBD Model Given Interval-Censored Data

Peter S. Fader
www.petefader.com

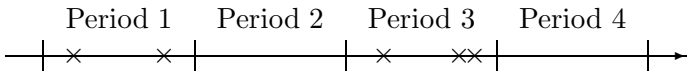
Bruce G. S. Hardie
www.brucehardie.com[†]

November 2005
Revised August 2010

1 Introduction

The Pareto/NBD model (Schmittlein et al. 1987) is a benchmark model for customer-base analysis in a noncontractual setting. The formulation of the model likelihood function developed by Fader and Hardie (2005) assumes we know when each of a customer's x transactions occurred during the period $(0, T]$, although it turns out that we only need to know the exact time of the last transaction (denoted by t_x).

In some model application settings, the data reporting procedures are such that we do not know when each of a customer's x transactions occurred, and therefore t_x is unknown. Rather, we know how many transaction occurred in each of a series of discrete time-intervals. For example, the following transaction history



would be recorded as two transactions occurring in Period 1, no transactions in Period 2, three transactions in Period 3, and no transactions in Period 4. We know that the last observed transaction (the fifth purchase) occurred in Period 3 but we do not know the exact time of this purchase (t_5). As a result, we cannot use the expression for the Pareto/NBD likelihood function given in Fader and Hardie (2005).

When the transaction history is reported in terms of the transaction counts for each of a series of discrete time-intervals, we say that the data are *interval censored*. The purpose of this note is to derive the Pareto/NBD likelihood function for the case of interval-censored data.

2 Model Assumptions

Before deriving the model likelihood function for the case of interval-censored data, let us review the underlying assumptions of the Pareto/NBD model and the key results for the "full information" case.

The Pareto/NBD model is based on the following assumptions:

i. Customers go through two stages in their "lifetime" with a specific firm: they are "alive" for some period of time, then become permanently "inactive".

ii. While alive, the number of transactions made by a customer follows a Poisson process with transaction rate λ. This is equivalent to assuming that the time between transactions is distributed exponential with transaction rate λ,

$$f(t_j \,|\, \lambda, t_{j-1}) = \lambda e^{-\lambda(t_j - t_{j-1})}\,, \quad t_j > t_{j-1} > 0\,,$$

where t_j is the time of the jth purchase.

iii. A customer's unobserved "lifetime" of length τ (after which he is viewed as being inactive) is exponentially distributed with dropout rate μ:

$$f(\tau \,|\, \mu) = \mu e^{-\mu\tau}\,.$$

iv. Heterogeneity in transaction rates across customers follows a gamma distribution with shape parameter r and scale parameter α.

v. Heterogeneity in dropout rates across customers follows a gamma distribution with shape parameter s and scale parameter β.

vi. The transaction rate λ and the dropout rate μ vary independently across customers.

For a customer with transaction history (x, t_x, T) (i.e., the individual made x purchases in the time interval $(0, T]$ with the last transaction occurring at t_x), the individual-level likelihood function is

$$L(\lambda, \mu \,|\, x, t_x, T) = \frac{\lambda^x \mu}{\lambda+\mu} e^{-(\lambda+\mu)t_x} + \frac{\lambda^{x+1}}{\lambda+\mu} e^{-(\lambda+\mu)T}\,. \tag{1}$$

Taking the expectation of this over the distributions of λ and μ yields the following expressions for the likelihood function for a randomly-chosen individual with transaction history (x, t_x, T):

- For $\alpha \geq \beta$,

$$L(r,\alpha,s,\beta\,|\,x,t_x,T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \times \left\{ \left(\frac{s}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,s+1;r+s+x+1;\frac{\alpha-\beta}{\alpha+t_x}\big)}{(\alpha+t_x)^{r+s+x}} + \left(\frac{r+x}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,s;r+s+x+1;\frac{\alpha-\beta}{\alpha+T}\big)}{(\alpha+T)^{r+s+x}} \right\}. \quad (2)$$

- For $\alpha \leq \beta$,

$$L(r,\alpha,s,\beta\,|\,x,t_x,T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \times \left\{ \left(\frac{s}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,r+x;r+s+x+1;\frac{\beta-\alpha}{\beta+t_x}\big)}{(\beta+t_x)^{r+s+x}} + \left(\frac{r+x}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,r+x+1;r+s+x+1;\frac{\beta-\alpha}{\beta+T}\big)}{(\beta+T)^{r+s+x}} \right\}. \quad (3)$$

The probability that a customer with purchase history (x,t_x,T) is alive at time T is given by

$$P(\text{alive}\,|\,r,\alpha,s,\beta,x,t_x,T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s} \Big/ L(r,\alpha,s,\beta\,|\,x,t_x,T)\,. \quad (4)$$

The random variable $Y(t)$ denotes the number of purchases made in the period $(T,T+t]$. The expected number of purchases in the period $(T,T+t]$ for a customer with purchase history (x,t_x,T), the so-called *conditional expectation*, is given by

$$E[Y(t)\,|\,r,\alpha,s,\beta,x,t_x,T] = \frac{\Gamma(r+x+1)}{\Gamma(r)(s-1)} \frac{\alpha^r\beta^s}{(\alpha+t)^{r+x+1}} \times \left[\frac{1}{(\beta+T)^{s-1}} - \frac{1}{(\beta+T+t)^{s-1}}\right] \Big/ L(r,\alpha,s,\beta\,|\,x,t_x,T)\,. \quad (5)$$

3 Derivation of the Likelihood Function

We observe an individual for n discrete time-periods.

- Let s_i be the length of the ith time period $(i=1,\ldots,n)$ and x_i the number of transactions observed in this period.

- Let
$$y_j = \sum_{i=1}^{j} x_i \text{ and } T_j = \sum_{i=1}^{j} s_i \,.$$
That is, y_n is the total number of transactions made by the individual in the time interval $(0, T_n]$, which is divided into n discrete periods:
$$(0, T_1],\ (T_1, T_2],\ \ldots,\ (T_{n-1}, T_n]$$
We typically observe $s_i = s_j\ \forall\, i, j$. However, if the recording periods are months (as opposed to quad weeks), $s_i \neq s_j$ for some i, j.

- Let m $(\leq n)$ be the last period in which at least one purchase took place. If no purchases are observed in $(0, T_n]$, $m = 0$; therefore $0 \leq m \leq n$.

As we seek to derive a general expression for the model likelihood function, let us consider the following three cases:

Case 1: $y_n = 0$

Suppose no purchases are observed in the time interval $(0, T_n]$ (i.e., $y_n = 0$). Assuming that the customer was alive at the start of the observation period, there are two ways this could have occurred:

i. The customer is still alive at the end of the observation period (i.e., $\tau > T_n$), in which case the individual-level likelihood function is simply the exponential survivor function evaluated at T_n:
$$L(\lambda \,|\, y_n = 0, \tau > T_n) = e^{-\lambda T_n} \,.$$

ii. The customer became inactive at some (unobserved) time τ in the interval $(0, T_n]$, in which case the individual-level likelihood function is
$$L(\lambda \,|\, y_n = 0, \text{inactive at } \tau \in (0, T_n]) = e^{-\lambda \tau} \,.$$

Removing the conditioning on τ yields the following expression for the individual-level likelihood function:

$$\begin{aligned} L(\lambda, \mu \,|\, y_n = 0, T_n) &= L(\lambda \,|\, y_n = 0, \tau > T_n) P(\tau > T_n \,|\, \mu) \\ &\quad + \int_0^{T_n} L(\lambda \,|\, y_n = 0, \text{inactive at } \tau \in (0, T_n]) f(\tau \,|\, \mu)\, d\tau \\ &= e^{-\lambda T_n} e^{-\mu T_n} + \int_0^{T_n} e^{-\lambda \tau} \mu e^{-\mu \tau}\, d\tau \\ &= \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu) T_n} \,. \end{aligned} \tag{6}$$

Case 2: $m = n$

Suppose we observe at least one transaction in the last discrete observation period (which implies $m = n$). The fact that $x_n > 0$ means that the customer must have been alive in the first $n-1$ periods. The likelihood of the corresponding y_{n-1} transactions in the interval $(0, T_{n-1}]$ is $\lambda^{y_{n-1}} e^{-(\lambda+\mu)T_{n-1}}$.

There are two ways by which we could observe x_n purchases in the interval $(T_{n-1}, T_n]$:

i. The customer is still active at the end of the observation period (i.e., $\tau > T_n$), in which case the nth period's contribution to the individual-level likelihood function is $\lambda^{x_n} e^{-\lambda(T_n - T_{n-1})}$.

ii. The customer became inactive at some (unobserved) time τ in the interval $(T_{n-1}, T_n]$, in which case the individual-level likelihood function is $\lambda^{x_n} e^{-\lambda(\tau - T_{n-1})}$.

Removing the conditioning on τ yields the following expression for the individual-level likelihood function:

$$\begin{aligned} L(\lambda, \mu \,|\, y_n, T_n) &= \lambda^{y_{n-1}} e^{-(\lambda+\mu)T_{n-1}} \cdot \Bigg\{ \lambda^{x_n} e^{-(\lambda+\mu)(T_n - T_{n-1})} \\ &\qquad + \int_{T_{n-1}}^{T_n} \lambda^{x_n} e^{-\lambda(\tau - T_{n-1})} \mu e^{-\mu(\tau - T_{n-1})} \, d\tau \Bigg\} \\ &= \lambda^{y_{n-1}} e^{-(\lambda+\mu)T_{n-1}} \cdot \Bigg\{ \lambda^{x_n} e^{-(\lambda+\mu)(T_n - T_{n-1})} \\ &\qquad + \frac{\lambda^{x_n} \mu}{\lambda + \mu} \Big[1 - e^{-(\lambda+\mu)(T_n - T_{n-1})} \Big] \Bigg\} \\ &= \frac{\lambda^{y_n} \mu}{\lambda + \mu} e^{-(\lambda+\mu)T_{n-1}} + \frac{\lambda^{y_n + 1}}{\lambda + \mu} e^{-(\lambda+\mu)T_n} \,. \end{aligned} \tag{7}$$

Case 3: $0 < m < n$

Finally, suppose the last observed transaction occurs before the nth period, sometime in period m. The fact that $x_m > 0$ means that the customer must have been alive in the first $m-1$ periods. The likelihood of the corresponding y_{m-1} transactions in the interval $(0, T_{m-1}]$ is $\lambda^{y_{m-1}} e^{-(\lambda+\mu)T_{m-1}}$. (By definition, $y_0 = 0$ and $T_0 = 0$.)

There are three ways by which the customer could have made x_m purchases in period m and then no purchases in the remaining $n - m$ periods (i.e., in the interval $(T_m, T_n]$):

i. The customer became inactive sometime in the mth period, in which case the incremental "contribution" to the likelihood function is

$$\begin{aligned}\mathsf{A}_1 &= \int_{T_{m-1}}^{T_m} \lambda^{x_m} e^{-\lambda(\tau - T_{m-1})} \mu e^{-\mu(\tau - T_{m-1})}\, d\tau \\ &= \frac{\lambda^{x_m}\mu}{\lambda+\mu}\Big[1 - e^{-(\lambda+\mu)(T_m - T_{m-1})}\Big]\,.\end{aligned}$$

ii. The customer was alive all through the mth period but became inactive sometime in the interval $(T_m, T_n]$, in which case the incremental "contribution" to the likelihood function is

$$\begin{aligned}\mathsf{A}_2 &= \lambda^{x_m} e^{-(\lambda+\mu)(T_m - T_{m-1})} \int_{T_m}^{T_n} e^{-\lambda(\tau - T_m)} \mu e^{-\mu(\tau - T_m)}\, d\tau \\ &= \frac{\lambda^{x_m}\mu}{\lambda+\mu} e^{-(\lambda+\mu)(T_m - T_{m-1})}\Big[1 - e^{-(\lambda+\mu)(T_n - T_m)}\Big]\,.\end{aligned}$$

iii. The customer was alive all through the mth period and remained alive all through the interval $(T_m, T_n]$ making no additional purchases, in which case the incremental "contribution" to the likelihood function is

$$\mathsf{A}_3 = \lambda^{x_m} e^{-(\lambda+\mu)(T_n - T_{m-1})}\,.$$

Combining these three terms, we have

$$\begin{aligned}L(\lambda, \mu \,|\, y_m, T_n) &= \lambda^{y_{m-1}} e^{-(\lambda+\mu)T_{m-1}} \cdot \big(\mathsf{A}_1 + \mathsf{A}_2 + \mathsf{A}_3\big) \\ &= \frac{\lambda^{y_m}\mu}{\lambda+\mu} e^{-(\lambda+\mu)T_{m-1}} + \frac{\lambda^{y_m+1}}{\lambda+\mu} e^{-(\lambda+\mu)T_n}\,. \qquad (8)\end{aligned}$$

In order to create a general expression that encompasses (6)–(8), let us summarize the customer's transaction history as (x, T_{m-1}, T) where $x = \sum_{i=1}^{n} x_i$ is the number of transactions that occurred in the interval $(0, T]$, where $T = \sum_{i=1}^{n} s_i$ $(= T_n)$, and T_{m-1} is the endpoint of the interval *immediately preceding* that in which the last purchase occurred. (By definition, $T_0 = 0$; if $x = 0$, then $T_{m-1} = 0$.) We can therefore write the individual-level likelihood function as

$$L(\lambda, \mu \,|\, x, T_{m-1}, T) = \frac{\lambda^{x}\mu}{\lambda+\mu} e^{-(\lambda+\mu)T_{m-1}} + \frac{\lambda^{x+1}}{\lambda+\mu} e^{-(\lambda+\mu)T}\,. \qquad (9)$$

Note that the only difference between (9) and (1), the "full information" likelihood function, is that we use T_{m-1} in the place of t_x.

It follows that the interval-censored likelihood function for a randomly-chosen individual with transaction history (x, T_{m-1}, T) has the following form:

- if $\alpha \geq \beta$,

$$L(r,\alpha,s,\beta\,|\,x,T_{m-1},T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \times \left\{ \left(\frac{s}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,s+1;r+s+x+1;\frac{\alpha-\beta}{\alpha+T_{m-1}}\big)}{(\alpha+T_{m-1})^{r+s+x}} + \left(\frac{r+x}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,s;r+s+x+1;\frac{\alpha-\beta}{\alpha+T}\big)}{(\alpha+T)^{r+s+x}} \right\}. \quad (10)$$

- if $\alpha \leq \beta$,

$$L(r,\alpha,s,\beta\,|\,x,T_{m-1},T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)} \times \left\{ \left(\frac{s}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,r+x;r+s+x+1;\frac{\beta-\alpha}{\beta+T_{m-1}}\big)}{(\beta+T_{m-1})^{r+s+x}} + \left(\frac{r+x}{r+s+x}\right) \frac{{}_2F_1\big(r+s+x,r+x+1;r+s+x+1;\frac{\beta-\alpha}{\beta+T}\big)}{(\beta+T)^{r+s+x}} \right\}. \quad (11)$$

Given the minor change in the expression for the model likelihood function, it follows that the probability that a customer with purchase history (x, T_{m-1}, T) is alive at time T is given by

$$P(\text{alive}\,|\,r,\alpha,s,\beta,x,T_{m-1},T) = \frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s} \Big/ L(r,\alpha,s,\beta\,|\,x,T_{m-1},T)\,, \quad (12)$$

and the conditional expectation is given by

$$E[Y(t)\,|\,r,\alpha,s,\beta,x,T_{m-1},T] = \frac{\Gamma(r+x+1)}{\Gamma(r)(s-1)} \frac{\alpha^r\beta^s}{(\alpha+t)^{r+x+1}} \times \left[\frac{1}{(\beta+T)^{s-1}} - \frac{1}{(\beta+T+t)^{s-1}}\right] \Big/ L(r,\alpha,s,\beta\,|\,x,T_{m-1},T)\,. \quad (13)$$

References

Fader, Peter S. and Bruce G. S. Hardie (2005), "A Note on Deriving the Pareto/NBD Model and Related Expressions." <`http://brucehardie.com/notes/009/`>

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who Are They and What Will They Do Next?" *Management Science*, **33** (January), 1–24.