Deriving an Expression for $P(X(t, t + \tau) = x)$
Under the Pareto/NBD Model

Peter S. Fader
www.petefader.com
Bruce G. S. Hardie
www.brucehardie.com
Kinshuk Jerath†

September 2006
Revised July 2020

1 Introduction

Schmittlein et al. (1987) and Fader and Hardie (2020) derive expressions for $P(X(t) = x)$, where the random variable $X(t)$ denotes the number of transactions observed in the time interval $(0, t]$, as implied by the Pareto/NBD model assumptions. In this note, we derive the corresponding expression for $P(X(t, t + \tau) = x)$, where the random variable $X(t, t + \tau)$ denotes the number of transactions observed in the time interval $(t, t + \tau]$.

In Section 2 we review the assumptions underlying the Pareto/NBD model. In Section 3, we derive an expression for $P(X(t, t + \tau) = x)$ conditional on the unobserved latent characteristics $\lambda$ and $\mu$. This conditioning is removed in Section 4.

2 Model Assumptions

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 (i.e., “die”).

ii. While alive, the number of transactions made by a customer follows a Poisson process with transaction rate $\lambda$. This implies that the probability of observing $x$ transactions in the time interval $(0, t]$ is given by

$$P(X(t) = x \mid \lambda) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \ldots.$$ 

It also implies that, assuming the customer is alive through the time interval $(t_a, t_b]$,

$$P(X(t_a, t_b) = x \mid \lambda) = \frac{[\lambda(t_b - t_a)]^x e^{-\lambda(t_b - t_a)}}{x!}, \quad x = 0, 1, 2, \ldots.$$ 

iii. A customer’s unobserved lifetime of length $\omega$ (after which he is viewed as being inactive) is exponentially distributed with dropout rate $\mu$:

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

iv. Heterogeneity in transaction rates across customers follows a gamma distribution with shape parameter $r$ and scale parameter $\alpha$:

$$g(\lambda | r, \alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda \alpha}}{\Gamma(r)}.$$  

(1)

v. Heterogeneity in dropout rates across customers follows a gamma distribution with shape parameter $s$ and scale parameter $\beta$:

$$g(\mu | s, \beta) = \frac{\beta^s \mu^{s-1} e^{-\mu \beta}}{\Gamma(s)}.$$  

(2)

vi. The transaction rate $\lambda$ and the dropout rate $\mu$ vary independently across customers.

3 $P(X(t, t + \tau) = x)$ Conditional on $\lambda$ and $\mu$

Suppose we know an individual’s unobserved latent characteristics $\lambda$ and $\mu$. For $x > 0$, there are two ways $x$ purchases could have occurred in the interval $(t, t + \tau)$:

i. The individual was alive at $t$ and remained alive through the whole interval; this occurs with probability $e^{-\mu(t+\tau)}$. The probability of the individual making $x$ purchases, given that he was alive through the whole interval, is $\left(\frac{\lambda \tau}{x!}\right)^x e^{-\lambda \tau}$. It follows that the probability of remaining alive through the interval $(t, t + \tau]$ and making $x$ purchases is

$$e^{-\mu t} \left(\frac{\lambda \mu}{x!}\right)^x e^{-\mu(x+1)}.$$  

(3)

ii. The individual was alive at $t$ but died at some point $\omega$ ($< t + \tau$), making $x$ purchases in the interval $(t, \omega]$. The probability of this occurring is

$$\int_t^{t+\tau} \frac{\lambda e^{-\lambda(t-\omega)}}{x!} \mu e^{-\mu \omega} d\omega = e^{-\mu t} \int_t^{t+\tau} \frac{(\omega - t)^x e^{-(\lambda + \mu)(\omega - t)}}{x!} d\omega.$$ 

$$= e^{-\mu t} \lambda^x \mu \int_0^\tau \frac{s^x e^{-(\lambda + \mu)s}}{x!} ds.$$ 

$$= e^{-\mu t} \lambda^x \mu \int_0^\tau \frac{(\lambda + \mu)^{x+1} s^x e^{-(\lambda + \mu)s}}{x!} ds.$$ 

which, noting that the integrand is an Erlang-$x+1$ pdf,

$$= e^{-\mu t} \left(\frac{\lambda}{\lambda + \mu}\right)^x \left(\frac{\mu}{\lambda + \mu}\right) \left[1 - e^{-(\lambda + \mu) \tau} \sum_{i=0}^x \frac{[(\lambda + \mu) \tau]^i}{i!}\right].$$  

(4)

These two scenarios also hold for the case of $x = 0$ but need to be augmented by an additional reason as to why no purchases could have occurred in the interval $(t, t + \tau]$: the individual was dead at the beginning of the interval, which occurs with probability

$$1 - e^{-\mu t}.$$  

(5)

2
Combining (3)–(5) gives us the following expression for the probability of observing \( x \) purchases in the interval \((t, t + \tau]\), conditional on \( \lambda \) and \( \mu \):

\[
P(X(t, t + \tau) = x | \lambda, \mu) = \delta_{x=0} \left[ 1 - e^{-\mu t} \right] + \frac{(\lambda \tau)^x e^{-\lambda \tau} e^{-\mu (t + \tau)}}{x!} \]

\[+ \left( \frac{\lambda}{\lambda + \mu} \right)^x \left( \frac{\mu}{\lambda + \mu} \right) e^{-\mu t} \]

\[ - \left( \frac{\lambda}{\lambda + \mu} \right)^x \left( \frac{\mu}{\lambda + \mu} \right) e^{-\lambda \tau} e^{-\mu (t + \tau)} \sum_{i=0}^{x} \frac{[(\lambda + \mu) \tau]^i}{i!}.
\]

(6)

4 Removing the Conditioning on \( \lambda \) and \( \mu \)

In reality, we never know an individual’s latent characteristics; we therefore remove the conditioning on \( \lambda \) and \( \mu \) by taking the expectation of (6) over the distributions of \( \Lambda \) and \( M \):

\[
P(X(t, t + \tau) = x | r, \alpha, s, \beta) = \int_{0}^{\infty} \int_{0}^{\infty} P(X(t, t + \tau) = x | \lambda, \mu) g(\lambda | r, \alpha) g(\mu | s, \beta) \, d\lambda \, d\mu.
\]

(7)

Substituting (1), (2), and (6) in (7) gives us

\[
P(X(t, t + \tau) = x | r, \alpha, s, \beta) = \delta_{x=0} A_1 + A_2 + A_3 - \sum_{i=0}^{x} \frac{\tau^i}{i!} A_4
\]

(8)

where

\[
A_1 = \int_{0}^{\infty} \left[ 1 - e^{-\mu t} \right] g(\mu | s, \beta) \, d\mu
\]

(9)

\[
A_2 = \int_{0}^{\infty} \int_{0}^{\infty} \frac{(\lambda \tau)^x e^{-\lambda \tau} e^{-\mu (t + \tau)}}{x!} g(\lambda | r, \alpha) g(\mu | s, \beta) \, d\lambda \, d\mu
\]

(10)

\[
A_3 = \int_{0}^{\infty} \int_{0}^{\infty} \left( \frac{\lambda}{\lambda + \mu} \right)^x \left( \frac{\mu}{\lambda + \mu} \right) e^{-\mu t} g(\lambda | r, \alpha) g(\mu | s, \beta) \, d\lambda \, d\mu
\]

(11)

\[
A_4 = \int_{0}^{\infty} \int_{0}^{\infty} \left( \frac{\lambda}{\lambda + \mu} \right)^x \left( \frac{\mu}{\lambda + \mu} \right) (\lambda + \mu)^\tau e^{-\lambda \tau} e^{-\mu (t + \tau)} g(\lambda | r, \alpha) g(\mu | s, \beta) \, d\lambda \, d\mu
\]

(12)

Solving (9) and (10) is trivial:

\[
A_1 = 1 - \left( \frac{\beta}{\beta + t} \right)^x
\]

(13)

\[
A_2 = \frac{\Gamma(r + x)}{\Gamma(r)} \left( \frac{\alpha}{\alpha + \tau} \right)^x \left( \frac{\tau}{\alpha + \tau} \right)^x \left( \frac{\beta}{\beta + t + \tau} \right)^x
\]

(14)

To solve (11), consider the transformation \( Y = M/(\Lambda + M) \) and \( Z = \Lambda + M \). Using the transformation technique (Casella and Berger 2002, Section 4.3, pp. 156–162; Mood et al. 1974, Section 6.2, p. 204ff), it follows that the joint distribution of \( Y \) and \( Z \) is

\[
g(y, z | \alpha, \beta, r, s) = \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} y^{s-1} (1 - y)^{r-s-1} e^{-z (\alpha - \beta) y}.
\]

(15)

Noting that the inverse of this transformation is \( \lambda = (1 - y) z \) and \( \mu = y z \), it follows that

\[
A_3 = \int_{0}^{1} \int_{0}^{\infty} y (1 - y)^x e^{-y z t} g(y, z | \alpha, \beta, r, s) \, dz \, dy
\]

\[
= \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_{0}^{1} \int_{0}^{\infty} y^x (1 - y)^{r+s-1} e^{-z (\alpha - (\beta + t)) y} \, dz \, dy
\]
\[
\frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_0^1 y^r (1 - y)^{r+s-1} \left\{ \int_0^\infty z^{r+s-1} e^{-z (\alpha - (\beta + t) y)} \, dz \right\} \, dy \\
= \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_0^1 y^r (1 - y)^{r+s-1} (\alpha - (\beta + t) y)^{-(r+s)} \, dy \\
= \frac{1}{B(r,s)} \frac{\alpha^r \beta^s}{\alpha^{r+s}} \int_0^1 y^r (1 - y)^{r+s-1} \left[ 1 - \frac{(\alpha - (\beta + t) y)}{\alpha} \right]^{-(r+s)} \, dy \\
\]

which, recalling Euler’s integral for the Gaussian hypergeometric function, \(^1\)

\[
= \frac{\alpha^r \beta^s}{\alpha^{r+s}} \frac{B(r+x, s+1)}{B(r,s)} 2F_1 \left( r+s, s+1; r+s+x+1; \frac{\alpha - (\beta + t)}{\alpha} \right). \tag{16}
\]

Looking closely at (16), we see that the argument of the Gaussian hypergeometric function, \(\frac{\alpha - (\beta + t)}{\alpha}\), is guaranteed to be bounded between 0 and 1 when \(\alpha \geq \beta + t\), thus ensuring convergence of the series representation of the function. However, when \(\alpha < \beta + t\) we can be faced with the situation where \(\frac{\alpha - (\beta + t)}{\alpha} < -1\), in which case the series is divergent.

Applying the linear transformation (Abramowitz and Stegun 1972, equation 15.3.4)

\[
2F_1(a, b; c; z) = (1 - z)^{-a} 2F_1\left( a - b; c - b; \frac{z}{1-z} \right), \tag{17}
\]

gives us

\[
A_3 = \frac{\alpha^r \beta^s}{(\beta + t)^{r+s}} \frac{B(r+x, s+1)}{B(r,s)} 2F_1 \left( r+s, r+x; r+s+x+1; \frac{\beta + t - \alpha}{\beta + t} \right). \tag{18}
\]

We note that the argument of the above Gaussian hypergeometric function is bounded between 0 and 1 when \(\alpha \leq \beta + t\). We therefore present (16) and (18) as solutions to (11), using (16) when \(\alpha \geq \beta + t\) and (18) when \(\alpha \leq \beta + t\). We can write this as

\[
A_3 = \frac{\alpha^r \beta^s B(r+x, s+1)}{B(r,s)} B_1 \tag{19}
\]

where

\[
B_1 = \begin{cases} 
2F_1 \left( r+s, s+1; r+s+x+1; \frac{\alpha - (\beta + t)}{\alpha} \right) & \text{if } \alpha \geq \beta + t \\
\alpha^{r+s} & \text{if } \alpha \leq \beta + t 
\end{cases} \tag{20}
\]

To solve (12), we also make use of the transformation \(Y = M/(\Lambda + M)\) and \(Z = \Lambda + M\). Given (15), it follows that

\[
A_4 = \frac{1}{\Gamma(r) \Gamma(s)} \int_0^1 \int_0^\infty y^r (1 - y)^{r+s-1} \left\{ \int_0^\infty z^{r+s-1} e^{-z (\alpha - (\beta + t) y)} \, dz \right\} \, dy \\
= \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_0^1 \int_0^\infty y^r (1 - y)^{r+s-1} z^{r+s-1} e^{-z (\alpha - (\beta + t) y)} \, dz \, dy \\
= \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_0^1 \int_0^\infty y^r (1 - y)^{r+s-1} \left\{ \int_0^\infty z^{r+s-1} e^{-z (\alpha - (\beta + t) y)} \, dz \right\} \, dy \\
= \frac{\alpha^r \beta^s}{\Gamma(r) \Gamma(s)} \int_0^1 \int_0^\infty y^r (1 - y)^{r+s-1} \left[ 1 - \frac{(\alpha - (\beta + t) y)}{\alpha} \right]^{-(r+s)} \, dy \\
= \frac{\Gamma(r+s+i)}{\Gamma(r) \Gamma(s) (\alpha + r)^{r+s+i}} \frac{\alpha^r \beta^s}{\alpha^{r+s+i}} \int_0^1 y^r (1 - y)^{r+s-1} \left[ 1 - \frac{(\alpha - (\beta + t) y)}{\alpha} \right]^{-(r+s+i)} \, dy \\
\]

\(^1\) \(2F_1(a, b; c; z) = \frac{1}{\Gamma(a+b-c)} \int_0^1 t^{b-1}(1-t)^{c-1}(1-zt)^{-a} \, dt\), \(c > b\).
which, recalling Euler’s integral for the Gaussian hypergeometric function,

\[
\frac{\Gamma(r + s + i) \alpha^r \beta^s}{\Gamma(r + s)} \frac{B(r + x, s + 1)}{B(r, s)} \times \frac{2}{2}F_1\left(\alpha \left(\begin{array}{ll} r + s + i \vline & r + s + x + 1; \\
\alpha + \frac{\beta + t - \alpha}{\beta + t + \tau} \end{array}\right)\right).
\]  

(21)

Noting that the argument of the Gaussian hypergeometric function is only guaranteed to be bounded between 0 and 1 when \(\alpha \geq \beta + t \ (\forall \tau > 0)\), we apply the linear transformation (17), which gives us

\[
A_4 = \frac{\Gamma(r + s + i) \alpha^r \beta^s}{\Gamma(r + s)} \frac{B(r + x, s + 1)}{B(r, s)} \times 2\frac{2}{2}F_1\left(\alpha \left\{ \begin{array}{ll} r + s + i \vline & r + s + x + 1; \\
\beta + t + \tau \end{array}\right)\right),
\]  

(22)

The argument of the above Gaussian hypergeometric function is bounded between 0 and 1 when \(\alpha \leq \beta + t \ (\forall \tau > 0)\). We therefore present (21) and (22) as solutions to (12): we use (21) when \(\alpha \geq \beta + t\) and (22) when \(\alpha \leq \beta + t\). We can write this as

\[
A_4 = \frac{\alpha^r \beta^s}{\Gamma(r + s)} \frac{\Gamma(r + s + i) B(r + x, s + 1)}{B(r, s)} B_2
\]  

(23)

where

\[
B_2 = \begin{cases} 
2\frac{2}{2}F_1\left(\alpha \left\{ \begin{array}{ll} r + s + i \vline & r + s + x + 1; \\
\beta + t + \tau \end{array}\right)\right) & \text{if } \alpha \geq \beta + t \\
2\frac{2}{2}F_1\left(\alpha \left\{ \begin{array}{ll} r + s + i \vline & r + s + x + 1; \\
\alpha + \frac{\beta + t - \alpha}{\beta + t + \tau} \end{array}\right)\right) & \text{if } \alpha \leq \beta + t
\end{cases}
\]  

(24)

Substituting (13), (14), (19), and (23) in (8) yields the following expression for the probability of observing \(x\) transactions in the time interval \((t, t + \tau)\):

\[
P(X(t, t + \tau) = x \mid r, \alpha, s, \beta) = \delta_x \left[ 1 - \left(\frac{\beta}{\beta + t}\right)^x \right] + \frac{1}{\Gamma(r + x)} \left(\begin{array}{ll} \alpha \vline & r + x; \\
\alpha + \frac{\beta + t - \alpha}{\beta + t + \tau} \end{array}\right)^x \left(\begin{array}{ll} \beta \vline & \beta + t + \tau; \\
\alpha + \frac{\beta + t - \alpha}{\beta + t + \tau} \end{array}\right)^x
\]  

\[
+ \alpha^r \beta^s \frac{B(r + x, s + 1)}{B(r, s)} B_1 \sum_{i=0}^{x} \frac{\Gamma(r + s + i)}{\Gamma(r + s)} \tau^i B_2
\]  

(25)

where expressions for \(B_1\) and \(B_2\) are given in (20) and (24), respectively.

We note that for \(t = 0\), (25) reduces to the implied expression for \(P(X(\tau) = x)\) as given in Fader and Hardie (2020, equation 16).
References


