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 λ and μ. 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 λ. 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 ω (after which he is viewed as being inactive) is exponentially distributed with dropout rate μ:

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

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

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

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

$$g(\mu \mid s, \beta) = \frac{\beta^s \mu^{s-1} e^{-\mu\beta}}{\Gamma(s)} \,. \tag{2}$$

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

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

Suppose we know an individual's unobserved latent characteristics λ and μ. 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 during the whole interval, is $(\lambda\tau)^x e^{-\lambda\tau}/x!$. It follows that the probability of remaining alive through the interval $(t, t+\tau]$ and making x purchases is

$$\frac{(\lambda\tau)^x e^{-\lambda\tau} e^{-\mu(t+\tau)}}{x!} \,. \tag{3}$$

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

$$\begin{aligned}
&\int_t^{t+\tau} \frac{[\lambda(\omega - t)]^x e^{-\lambda(\omega-t)}}{x!} \mu e^{-\mu\omega}\, d\omega \\
&\quad = e^{-\mu t} \lambda^x \mu \int_t^{t+\tau} \frac{(\omega - t)^x e^{-(\lambda+\mu)(\omega-t)}}{x!}\, d\omega \\
&\quad = e^{-\mu t} \lambda^x \mu \int_0^{\tau} \frac{s^x e^{-(\lambda+\mu)s}}{x!}\, ds \\
&\quad = e^{-\mu t} \frac{\lambda^x \mu}{(\lambda+\mu)^{x+1}} \int_0^{\tau} \frac{(\lambda+\mu)^{x+1} s^x e^{-(\lambda+\mu)s}}{x!}\, ds
\end{aligned}$$

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

$$= e^{-\mu t} \Big(\frac{\lambda}{\lambda+\mu}\Big)^x \Big(\frac{\mu}{\lambda+\mu}\Big) \bigg[1 - e^{-(\lambda+\mu)\tau} \sum_{i=0}^{x} \frac{[(\lambda+\mu)\tau]^i}{i!}\bigg] \,. \tag{4}$$

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

$$1 - e^{-\mu t} \,. \tag{5}$$

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

$$\begin{aligned} P(X(t,t+\tau) = x \mid \lambda, \mu) = \delta_{x=0}\Big[1 - e^{-\mu t}\Big] &+ \frac{(\lambda\tau)^x e^{-\lambda\tau} e^{-\mu(t+\tau)}}{x!} \\ &+ \Big(\frac{\lambda}{\lambda+\mu}\Big)^x \Big(\frac{\mu}{\lambda+\mu}\Big) e^{-\mu t} \\ &- \Big(\frac{\lambda}{\lambda+\mu}\Big)^x \Big(\frac{\mu}{\lambda+\mu}\Big) e^{-\lambda\tau} e^{-\mu(t+\tau)} \sum_{i=0}^{x} \frac{[(\lambda+\mu)\tau]^i}{i!} \,. \end{aligned} \tag{6}$$

4 Removing the Conditioning on λ and μ

In reality, we never know an individual's latent characteristics; we therefore remove the conditioning on λ and μ by taking the expectation of (6) over the distributions of Λ and M:

$$\begin{aligned} &P(X(t,t+\tau) = x \mid r, \alpha, s, \beta) \\ &\qquad = \int_0^\infty \int_0^\infty P(X(t,t+\tau) = x \mid \lambda, \mu) g(\lambda \mid r, \alpha) g(\mu \mid s, \beta)\, d\lambda\, d\mu \,. \end{aligned} \tag{7}$$

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

$$P(X(t,t+\tau) = x \mid r, \alpha, s, \beta) = \delta_{x=0}\mathsf{A}_1 + \mathsf{A}_2 + \mathsf{A}_3 - \sum_{i=0}^{x} \frac{\tau^i}{i!}\mathsf{A}_4 \tag{8}$$

where

$$\mathsf{A}_1 = \int_0^\infty \Big[1 - e^{-\mu t}\Big] g(\mu \mid s, \beta)\, d\mu \tag{9}$$

$$\mathsf{A}_2 = \int_0^\infty \int_0^\infty \frac{(\lambda\tau)^x e^{-\lambda\tau} e^{-\mu(t+\tau)}}{x!} g(\lambda \mid r, \alpha) g(\mu \mid s, \beta)\, d\lambda\, d\mu \tag{10}$$

$$\mathsf{A}_3 = \int_0^\infty \int_0^\infty \Big(\frac{\lambda}{\lambda+\mu}\Big)^x \Big(\frac{\mu}{\lambda+\mu}\Big) e^{-\mu t} g(\lambda \mid r, \alpha) g(\mu \mid s, \beta)\, d\lambda\, d\mu \tag{11}$$

$$\mathsf{A}_4 = \int_0^\infty \int_0^\infty \Big(\frac{\lambda}{\lambda+\mu}\Big)^x \Big(\frac{\mu}{\lambda+\mu}\Big) (\lambda+\mu)^i e^{-\lambda\tau} e^{-\mu(t+\tau)} g(\lambda \mid r, \alpha) g(\mu \mid s, \beta)\, d\lambda\, d\mu \tag{12}$$

Solving (9) and (10) is trivial:

$$\mathsf{A}_1 = 1 - \Big(\frac{\beta}{\beta+t}\Big)^s \tag{13}$$

$$\mathsf{A}_2 = \frac{\Gamma(r+x)}{\Gamma(r)x!} \Big(\frac{\alpha}{\alpha+\tau}\Big)^r \Big(\frac{\tau}{\alpha+\tau}\Big)^x \Big(\frac{\beta}{\beta+t+\tau}\Big)^s \tag{14}$$

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

$$g(y, z \mid \alpha, \beta, r, s) = \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} y^{s-1} (1-y)^{r-1} z^{r+s-1} e^{-z(\alpha-(\alpha-\beta)y)} \,. \tag{15}$$

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

$$\begin{aligned} \mathsf{A}_3 &= \int_0^1 \int_0^\infty y(1-y)^x e^{-yzt} g(y, z \mid \alpha, \beta, r, s)\, dz\, dy \\ &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 \int_0^\infty y^s (1-y)^{r+x-1} z^{r+s-1} e^{-z(\alpha-(\alpha-(\beta+t))y)}\, dz\, dy \end{aligned}$$

$$
\begin{aligned}
&= \frac{\alpha^r\beta^s}{\Gamma(r)\Gamma(s)}\int_0^1 y^s(1-y)^{r+x-1}\left\{\int_0^\infty z^{r+s-1}e^{-z(\alpha-(\alpha-(\beta+t))y)}\,dz\right\}dy\\
&= \alpha^r\beta^s\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\int_0^1 y^s(1-y)^{r+x-1}(\alpha-(\alpha-(\beta+t))y)^{-(r+s)}\,dy\\
&= \frac{1}{B(r,s)}\frac{\alpha^r\beta^s}{\alpha^{r+s}}\int_0^1 y^s(1-y)^{r+x-1}\big[1-\big(\tfrac{\alpha-(\beta+t)}{\alpha}\big)y\big]^{-(r+s)}\,dy
\end{aligned}
$$

which, recalling Euler's integral for the Gaussian hypergeometric function,[1]

$$
= \frac{\alpha^r\beta^s}{\alpha^{r+s}}\frac{B(r+x,s+1)}{B(r,s)}\,{}_2F_1\big(r+s,s+1;r+s+x+1;\tfrac{\alpha-(\beta+t)}{\alpha}\big)\,. \tag{16}
$$

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)

$$
{}_2F_1(a,b;c;z) = (1-z)^{-a}\,{}_2F_1\big(a,c-b;c;\tfrac{z}{z-1}\big)\,, \tag{17}
$$

gives us

$$
\mathsf{A}_3 = \frac{\alpha^r\beta^s}{(\beta+t)^{r+s}}\frac{B(r+x,s+1)}{B(r,s)}\,{}_2F_1\big(r+s,r+x;r+s+x+1;\tfrac{\beta+t-\alpha}{\beta+t}\big)\,. \tag{18}
$$

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

$$
\mathsf{A}_3 = \alpha^r\beta^s\frac{B(r+x,s+1)}{B(r,s)}\,\mathsf{B}_1 \tag{19}
$$

where

$$
\mathsf{B}_1 = \begin{cases} \dfrac{{}_2F_1\big(r+s,s+1;r+s+x+1;\frac{\alpha-(\beta+t)}{\alpha}\big)}{\alpha^{r+s}} & \text{if } \alpha \geq \beta+t \\[2ex] \dfrac{{}_2F_1\big(r+s,r+x;r+s+x+1;\frac{\beta+t-\alpha}{\beta+t}\big)}{(\beta+t)^{r+s}} & \text{if } \alpha \leq \beta+t \end{cases} \tag{20}
$$

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

$$
\begin{aligned}
\mathsf{A}_4 &= \int_0^1\int_0^\infty y(1-y)^x z^i e^{-(1-y)z\tau}e^{-yz(t+\tau)}g(y,z\,|\,\alpha,\beta,r,s)\,dz\,dy\\
&= \frac{\alpha^r\beta^s}{\Gamma(r)\Gamma(s)}\int_0^1\int_0^\infty y^s(1-y)^{r+x-1}z^{r+s+i-1}e^{-z(\alpha+\tau-(\alpha-(\beta+t))y)}\,dz\,dy\\
&= \frac{\alpha^r\beta^s}{\Gamma(r)\Gamma(s)}\int_0^1 y^s(1-y)^{r+x-1}\left\{\int_0^\infty z^{r+s+i-1}e^{-z(\alpha+\tau-(\alpha-(\beta+t))y)}\,dz\right\}dy\\
&= \alpha^r\beta^s\frac{\Gamma(r+s+i)}{\Gamma(r)\Gamma(s)}\int_0^1 y^s(1-y)^{r+x-1}(\alpha+\tau-(\alpha-(\beta+t))y)^{-(r+s+i)}\,dy\\
&= \frac{\Gamma(r+s+i)}{\Gamma(r)\Gamma(s)}\frac{\alpha^r\beta^s}{(\alpha+\tau)^{r+s+i}}\int_0^1 y^s(1-y)^{r+x-1}\big[1-\big(\tfrac{\alpha-(\beta+t)}{\alpha+\tau}\big)y\big]^{-(r+s+i)}\,dy
\end{aligned}
$$

[1] ${}_2F_1(a,b;c;z) = \frac{1}{B(b,c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}dt\,,\ c>b\,.$

which, recalling Euler's integral for the Gaussian hypergeometric function,

$$
\begin{aligned}
&= \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{\alpha^r \beta^s}{(\alpha+\tau)^{r+s+i}} \frac{B(r+x,s+1)}{B(r,s)} \\
&\qquad \times {}_2F_1\big(r+s+i, s+1; r+s+x+1; \tfrac{\alpha-(\beta+t)}{\alpha+\tau}\big)\,. \qquad (21)
\end{aligned}
$$

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

$$
\begin{aligned}
\mathsf{A}_4 &= \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{\alpha^r \beta^s}{(\beta+t+\tau)^{r+s+i}} \frac{B(r+x,s+1)}{B(r,s)} \\
&\qquad \times {}_2F_1\big(r+s+i, r+x; r+s+x+1; \tfrac{\beta+t-\alpha}{\beta+t+\tau}\big)\,, \qquad (22)
\end{aligned}
$$

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

$$
\mathsf{A}_4 = \alpha^r \beta^s \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{B(r+x,s+1)}{B(r,s)} \mathsf{B}_2 \qquad (23)
$$

where

$$
\mathsf{B}_2 = \begin{cases}
\dfrac{{}_2F_1\big(r+s+i, s+1; r+s+x+1; \frac{\alpha-(\beta+t)}{\alpha+\tau}\big)}{(\alpha+\tau)^{r+s+i}} & \text{if } \alpha \geq \beta+t \\[2ex]
\dfrac{{}_2F_1\big(r+s+i, r+x; r+s+x+1; \frac{\beta+t-\alpha}{\beta+t+\tau}\big)}{(\beta+t+\tau)^{r+s+i}} & \text{if } \alpha \leq \beta+t
\end{cases} \qquad (24)
$$

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

$$
\begin{aligned}
&P(X(t,t+\tau) = x \mid r, \alpha, s, \beta) \\
&\qquad = \delta_{x=0}\left[1 - \Big(\frac{\beta}{\beta+t}\Big)^s\right] + \frac{\Gamma(r+x)}{\Gamma(r)x!}\Big(\frac{\alpha}{\alpha+\tau}\Big)^r \Big(\frac{\tau}{\alpha+\tau}\Big)^x \Big(\frac{\beta}{\beta+t+\tau}\Big)^s \\
&\qquad\quad + \alpha^r \beta^s \frac{B(r+x,s+1)}{B(r,s)} \left\{ \mathsf{B}_1 - \sum_{i=0}^{x} \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{\tau^i}{i!} \mathsf{B}_2 \right\} \qquad (25)
\end{aligned}
$$

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

Abramowitz, Milton and Irene A. Stegun (eds.) (1972), *Handbook of Mathematical Functions*, New York: Dover Publications.

Casella, George, and Roger L. Berger (2002), *Statistical Inference*, 2nd edition, Pacific Grove, CA: Duxbury.

Fader, Peter S. and Bruce G. S. Hardie (2020), "Deriving an Expression for $P(X(t) = x)$ Under the Pareto/NBD Model." <http://brucehardie.com/notes/012/>

Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes (1974), *Introduction to the Theory of Statistics*, 3rd edition, New York: McGraw-Hill Publishing Company.

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.