

# Deriving an Expression for $P(X(t) = x)$ Under the Pareto/NBD Model

Peter S. Fader  
[www.petefader.com](http://www.petefader.com)

Bruce G. S. Hardie  
[www.brucehardie.com](http://www.brucehardie.com)<sup>†</sup>

September 2006

## 1 Introduction

The Pareto/NBD (Schmittlein et al. 1987, hereafter SMC) is a model for customer-base analysis in a noncontractual setting. One result presented in SMC is an expression for  $P(X(t) = x)$ , where the random variable  $X(t)$  denotes the number of transactions observed in the time interval  $(0, t]$ . This note derives an alternative expression for this quantity, one that is simpler to evaluate.

In Section 2 we review the assumptions underlying the Pareto/NBD model. In Section 3, we derive an expression for  $P(X(t) = x)$  conditional on the unobserved latent characteristics  $\lambda$  and  $\mu$ ; we remove this conditioning in Section 4. For the sake of completeness, SMC's derivation is replicated in the Appendix.

## 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.
- ii. While alive, the number of transactions made by a customer follows a Poisson process with transaction rate  $\lambda$ . Denoting the number of transactions in the time interval  $(0, t]$  by the random variable  $X(t)$ , it follows that

$$P(X(t) = x | \lambda, \text{alive at } t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \dots$$

- 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  $\alpha$ :

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

---

<sup>†</sup>© 2006 Peter S. Fader and Bruce G. S. Hardie. Document source: <<http://brucehardie.com/notes/012/>>.

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

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

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

### 3 $P(X(t) = x)$ Conditional on $\lambda$ and $\mu$

Suppose we know an individual's unobserved latent characteristics  $\lambda$  and  $\mu$ . Assuming the customer was alive at time 0, there are two ways  $x$  purchases could have occurred in the interval  $(0, t]$ :

- i. The individual remained alive through the whole interval; this occurs with probability  $e^{-\mu t}$ . The probability of the individual making  $x$  purchases, given she was alive during the whole interval, is  $(\lambda t)^x e^{-\lambda t} / x!$ . Therefore, the probability of remaining alive through the interval  $(0, t]$  and making  $x$  purchases is

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

- ii. The individual "died" at some point  $\omega$  ( $< t$ ) and made  $x$  purchases in the interval  $(0, \omega]$ . The probability of this occurring is

$$\begin{aligned} \int_0^t \frac{(\lambda\omega)^x e^{-\lambda\omega}}{x!} \mu e^{-\mu\omega} d\omega &= \lambda^x \mu \int_0^t \frac{\omega^x e^{-(\lambda+\mu)\omega}}{x!} d\omega \\ &= \frac{\lambda^x \mu}{(\lambda + \mu)^{x+1}} \int_0^t \frac{(\lambda + \mu)^{x+1} \omega^x e^{-(\lambda+\mu)\omega}}{x!} d\omega \end{aligned}$$

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

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

Combining these two scenarios gives us the following expression for the probability of observing  $x$  purchases in the interval  $(0, t]$ , conditional on  $\lambda$  and  $\mu$ :

$$P(X(t) = x | \lambda, \mu) = \frac{(\lambda t)^x e^{-(\lambda+\mu)t}}{x!} + \frac{\lambda^x \mu}{(\lambda + \mu)^{x+1}} \left[ 1 - e^{-(\lambda+\mu)t} \sum_{i=0}^x \frac{[(\lambda + \mu)t]^i}{i!} \right]. \quad (3)$$

### 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 (3) over the distributions of  $\Lambda$  and  $M$ :

$$P(X(t) = x | r, \alpha, s, \beta) = \int_0^\infty \int_0^\infty P(X(t) = x | \lambda, \mu) g(\lambda | r, \alpha) g(\mu | s, \beta) d\lambda d\mu. \quad (4)$$

Substituting (1)–(3) in (4) give us

$$P(X(t) = x | r, \alpha, s, \beta) = A_1 + A_2 - \sum_{i=0}^x \frac{t^i}{i!} A_3 \quad (5)$$

where

$$A_1 = \int_0^\infty \int_0^\infty \frac{(\lambda t)^x e^{-(\lambda+\mu)t}}{x!} g(\lambda | r, \alpha) g(\mu | s, \beta) d\lambda d\mu \quad (6)$$

$$A_2 = \int_0^\infty \int_0^\infty \frac{\lambda^x \mu}{(\lambda + \mu)^{x+1}} g(\lambda | r, \alpha) g(\mu | s, \beta) d\lambda d\mu \quad (7)$$

$$A_3 = \int_0^\infty \int_0^\infty \frac{\lambda^x \mu e^{-(\lambda+\mu)t}}{(\lambda + \mu)^{x-i+1}} g(\lambda | r, \alpha) g(\mu | s, \beta) d\lambda d\mu \quad (8)$$

i. Solving (6) is trivial:

$$A_1 = \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+t}\right)^r \left(\frac{t}{\alpha+t}\right)^x \left(\frac{\beta}{\beta+t}\right)^s \quad (9)$$

ii. To solve (7), 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

$$g(y, z | \alpha, \beta, r, s) = \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} y^{s-1} (1-y)^{r-1} z^{r+s-1} e^{-z(\alpha - (\alpha-\beta)y)}. \quad (10)$$

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

$$\begin{aligned} A_2 &= \int_0^1 \int_0^\infty y(1-y)^x g(y, z | \alpha, \beta, r, s) dz dy \\ &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 \int_0^\infty y^s (1-y)^{r+x-1} z^{r+s-1} e^{-z(\alpha - (\alpha-\beta)y)} dz dy \\ &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 y^s (1-y)^{r+x-1} \left\{ \int_0^\infty z^{r+s-1} e^{-z(\alpha - (\alpha-\beta)y)} dz \right\} dy \\ &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \Gamma(r+s) \int_0^1 y^s (1-y)^{r+x-1} (\alpha - (\alpha-\beta)y)^{-(r+s)} dy \\ &= \frac{\alpha^r \beta^s}{B(r, s)} \frac{1}{\alpha^{r+s}} \int_0^1 y^s (1-y)^{r+x-1} \left[1 - \left(\frac{\alpha-\beta}{\alpha}\right)y\right]^{-(r+s)} dy \end{aligned}$$

which, recalling Euler's integral for the Gaussian hypergeometric function,<sup>1</sup>

$$= \frac{\alpha^r \beta^s}{\alpha^{r+s}} \frac{B(r+x, s+1)}{B(r, s)} {}_2F_1(r+s, s+1; r+s+x+1; \frac{\alpha-\beta}{\alpha}). \quad (11)$$

Looking closely at (11), we see that the argument of the Gaussian hypergeometric function,  $\frac{\alpha-\beta}{\alpha}$ , is guaranteed to be bounded between 0 and 1 when  $\alpha \geq \beta$  (since  $\beta > 0$ ), thus ensuring convergence of the series representation of the function. However, when  $\alpha < \beta$  we can be faced with the situation where  $\frac{\alpha-\beta}{\alpha} < -1$ , in which case the series is divergent.

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

$${}_2F_1(a, b; c; z) = (1-z)^{-a} {}_2F_1(a, c-b; c; \frac{z}{z-1}), \quad (12)$$

gives us

$$A_2 = \frac{\alpha^r \beta^s}{\beta^{r+s}} \frac{B(r+x, s+1)}{B(r, s)} {}_2F_1(r+s, r+x; r+s+x+1; \frac{\beta-\alpha}{\beta}). \quad (13)$$

We see that the argument of the above Gaussian hypergeometric function is bounded between 0 and 1 when  $\alpha \leq \beta$ . We therefore present (11) and (13) as solutions to (7): we use (11) when  $\alpha \geq \beta$  and (13) when  $\alpha \leq \beta$ .

---

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

- iii. To solve (8), we also make use of the transformation  $Y = M/(\Lambda + M)$  and  $Z = \Lambda + M$ . Given (10), it follows that

$$\begin{aligned}
A_3 &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 \int_0^\infty y^s (1-y)^{r+x-1} z^{r+s+i-1} e^{-z(\alpha+t-(\alpha-\beta)y)} dz dy \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 y^s (1-y)^{r+x-1} \left\{ \int_0^\infty z^{r+s+i-1} e^{-z(\alpha+t-(\alpha-\beta)y)} dz \right\} dy \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \Gamma(r+s+i) \int_0^1 y^s (1-y)^{r+x-1} (\alpha+t-(\alpha-\beta)y)^{-(r+s+i)} dy \\
&= \frac{\alpha^r \beta^s}{B(r,s)} \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{1}{(\alpha+t)^{r+s+i}} \int_0^1 y^s (1-y)^{r+x-1} \left[1 - \left(\frac{\alpha-\beta}{\alpha+t}\right)y\right]^{-(r+s+i)} dy
\end{aligned}$$

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

$$\begin{aligned}
&= \frac{\alpha^r \beta^s}{(\alpha+t)^{r+s+i}} \frac{B(r+x, s+1)}{B(r,s)} \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \\
&\quad \times {}_2F_1\left(r+s+i, s+1; r+s+x+1; \frac{\alpha-\beta}{\alpha+t}\right). \tag{14}
\end{aligned}$$

Noting that the argument of the Gaussian hypergeometric function is only guaranteed to be bounded between 0 and 1 when  $\alpha \geq \beta$ , we apply the linear transformation (12), which gives us

$$\begin{aligned}
A_3 &= \frac{\alpha^r \beta^s}{(\beta+t)^{r+s+i}} \frac{B(r+x, s+1)}{B(r,s)} \frac{\Gamma(r+s+i)}{\Gamma(r+s)} \\
&\quad \times {}_2F_1\left(r+s+i, r+x; r+s+x+1; \frac{\beta-\alpha}{\beta+t}\right). \tag{15}
\end{aligned}$$

The argument of the above Gaussian hypergeometric function is bounded between 0 and 1 when  $\alpha \leq \beta$ . We therefore present (14) and (15) as solutions to (8), using (14) when  $\alpha \geq \beta$  and (15) when  $\alpha \leq \beta$ .

Substituting (9), (11), (13), (14), and (15) in (5), and noting that

$$\frac{\Gamma(r+s+i)}{\Gamma(r+s)} \frac{1}{i!} = \frac{1}{iB(r+s, i)},$$

yields the following expression for the distribution of the number of transactions in the interval  $(0, t]$  for a randomly-chosen individual under the Pareto/NBD model:

$$\begin{aligned}
P(X(t) = x | r, \alpha, s, \beta) &= \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+t}\right)^r \left(\frac{t}{\alpha+t}\right)^x \left(\frac{\beta}{\beta+t}\right)^s \\
&\quad + \alpha^r \beta^s \frac{B(r+x, s+1)}{B(r,s)} \left\{ B_1 - \sum_{i=0}^x \frac{t^i}{iB(r+s, i)} B_2 \right\} \tag{16}
\end{aligned}$$

where

$$B_1 = \begin{cases} \frac{{}_2F_1(r+s, s+1; r+s+x+1; \frac{\alpha-\beta}{\alpha})}{\alpha^{r+s}} & \text{if } \alpha \geq \beta \\ \frac{{}_2F_1(r+s, r+x; r+s+x+1; \frac{\beta-\alpha}{\beta})}{\beta^{r+s}} & \text{if } \alpha \leq \beta \end{cases}$$

and

$$B_2 = \begin{cases} \frac{{}_2F_1(r+s+i, s+1; r+s+x+1; \frac{\alpha-\beta}{\alpha+t})}{(\alpha+t)^{r+s+i}} & \text{if } \alpha \geq \beta \\ \frac{{}_2F_1(r+s+i, r+x; r+s+x+1; \frac{\beta-\alpha}{\beta+t})}{(\beta+t)^{r+s+i}} & \text{if } \alpha \leq \beta. \end{cases}$$

We note that this expression requires  $x+2$  evaluations of the Gaussian hypergeometric function; in contrast, SMC's expression (see the attached appendix) requires  $2(x+1)$  evaluations of the Gaussian hypergeometric function.

The equivalence of (16) and (A1), (A3), (A4) is not immediately obvious. Purely from a logical perspective, they must be equivalent. Furthermore, equivalence is observed in numerical investigations. However, we have yet to demonstrate direct equivalence of these two expressions for  $P(X(t) = x | r, \alpha, s, \beta)$ . Stay tuned.

## 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.

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.

## Appendix: SMC's Derivation of $P(X(t) = x \mid r, \alpha, s, \beta)$

SMC derive their expression for  $P(X(t) = x)$  by first integrating over  $\lambda$  and  $\mu$  and then removing the conditioning on  $\omega$ , which is the reverse of the approach used in Sections 3 and 4 above. This gives us

$$\begin{aligned}
 P(X(t) = x \mid r, \alpha, s, \beta) &= \underbrace{\frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+t}\right)^r \left(\frac{t}{\alpha+t}\right)^x}_{\text{NBD } P(X(t)=x)} \underbrace{\left(\frac{\beta}{\beta+t}\right)^s}_{P(\Omega>t)} \\
 &\quad + \int_0^t \underbrace{\frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+\omega}\right)^r \left(\frac{\omega}{\alpha+\omega}\right)^x}_{\text{NBD } P(X(\omega)=x)} \underbrace{\frac{s}{\beta} \left(\frac{\beta}{\beta+\omega}\right)^{s+1}}_{f(\omega)} d\omega \\
 &= \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+t}\right)^r \left(\frac{t}{\alpha+t}\right)^x \left(\frac{\beta}{\beta+t}\right)^s + \frac{\Gamma(r+x)}{\Gamma(r)x!} s \alpha^r \beta^s \mathbf{C} \quad (\text{A1})
 \end{aligned}$$

where

$$\mathbf{C} = \int_0^t \omega^x (\alpha + \omega)^{-(r+x)} (\beta + \omega)^{-(s+1)} d\omega. \quad (\text{A2})$$

Making the change of variable  $y = \alpha + \omega$ ,

$$\mathbf{C} = \int_{\alpha}^{\alpha+t} (y - \alpha)^x y^{-(r+x)} (y - \alpha + \beta)^{-(s+1)} dy$$

which, recalling the binomial theorem,<sup>2</sup>

$$\begin{aligned}
 &= \int_{\alpha}^{\alpha+t} \left\{ \sum_{j=0}^x \binom{x}{j} y^{x-j} (-\alpha)^j \right\} y^{-(r+x)} (y - \alpha + \beta)^{-(s+1)} dy \\
 &= \sum_{j=0}^x \binom{x}{j} (-\alpha)^j \int_{\alpha}^{\alpha+t} y^{-(r+j)} (y - \alpha + \beta)^{-(s+1)} dy \\
 &= \sum_{j=0}^x \binom{x}{j} (-\alpha)^j \left\{ \int_{\alpha}^{\infty} y^{-(r+j)} (y - \alpha + \beta)^{-(s+1)} dy - \int_{\alpha+t}^{\infty} y^{-(r+j)} (y - \alpha + \beta)^{-(s+1)} dy \right\}
 \end{aligned}$$

letting  $z = \alpha/y$  in the first integral (which implies  $dy = -dz\alpha z^{-2}$ ) and  $z = (\alpha+t)/y$  in the second integral (which implies  $dy = -dz(\alpha+t)z^{-2}$ ),

$$\begin{aligned}
 &= \sum_{j=0}^x \binom{x}{j} (-\alpha)^j \left\{ \alpha^{-(r+s+j)} \int_0^1 z^{r+s+j-1} \left(1 - \frac{\alpha-\beta}{\alpha} z\right)^{-(s+1)} dz \right. \\
 &\quad \left. - (\alpha+t)^{-(r+s+j)} \int_0^1 z^{r+s+j-1} \left(1 - \frac{\alpha-\beta}{\alpha+t} z\right)^{-(s+1)} dz \right\}
 \end{aligned}$$

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

$$\begin{aligned}
 &= \sum_{j=0}^x \binom{x}{j} \frac{(-\alpha)^j}{r+s+j} \left\{ \frac{{}_2F_1(s+1, r+s+j; r+s+j+1; \frac{\alpha-\beta}{\alpha})}{\alpha^{r+s+j}} \right. \\
 &\quad \left. - \frac{{}_2F_1(s+1, r+s+j; r+s+j+1; \frac{\alpha-\beta}{\alpha+t})}{(\alpha+t)^{r+s+j}} \right\}. \quad (\text{A3})
 \end{aligned}$$

<sup>2</sup>  $(x+y)^r = \sum_{j=0}^x \binom{r}{k} x^k y^{r-k}$  for integer  $r \geq 0$ .

We note that the arguments of the above Gaussian hypergeometric functions are only guaranteed to be bounded between 0 and 1 when  $\alpha \geq \beta$ . We therefore revisit (A2), applying the change of variable  $y = \beta + \omega$ :

$$C = \int_{\beta}^{\beta+t} (y - \beta)^x y^{-(s+1)} (y - \beta + \alpha)^{-(r+x)} dy$$

which, recalling the binomial theorem,

$$\begin{aligned} &= \int_{\beta}^{\beta+t} \left\{ \sum_{j=0}^x \binom{x}{j} y^{x-j} (-\beta)^j \right\} y^{-(s+1)} (y - \beta + \alpha)^{-(r+x)} dy \\ &= \sum_{j=0}^x \binom{x}{j} (-\beta)^j \int_{\beta}^{\beta+t} y^{-(s+j+1-x)} (y - \beta + \alpha)^{-(r+x)} dy \\ &= \sum_{j=0}^x \binom{x}{j} (-\beta)^j \left\{ \int_{\beta}^{\infty} y^{-(s+j+1-x)} (y - \beta + \alpha)^{-(r+x)} dy \right. \\ &\quad \left. - \int_{\beta+t}^{\infty} y^{-(s+j+1-x)} (y - \beta + \alpha)^{-(r+x)} dy \right\} \end{aligned}$$

letting  $z = \beta/y$  in the first integral (which implies  $dy = -dz\beta z^{-2}$ ) and  $z = (\beta + t)/y$  in the second integral (which implies  $dy = -dz(\beta + t)z^{-2}$ ),

$$\begin{aligned} &= \sum_{j=0}^x \binom{x}{j} (-\beta)^j \left\{ \beta^{-(r+s+j)} \int_0^1 z^{r+s+j-1} \left(1 - \frac{\beta-\alpha}{\beta} z\right)^{-(r+x)} dz \right. \\ &\quad \left. - (\beta + t)^{-(r+s+j)} \int_0^1 z^{r+s+j-1} \left(1 - \frac{\beta-\alpha}{\beta+t} z\right)^{-(r+x)} dz \right\} \end{aligned}$$

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

$$\begin{aligned} &= \sum_{j=0}^x \binom{x}{j} \frac{(-\beta)^j}{r+s+j} \left\{ \frac{{}_2F_1\left(r+x, r+s+j; r+s+j+1; \frac{\beta-\alpha}{\beta}\right)}{\beta^{r+s+j}} \right. \\ &\quad \left. - \frac{{}_2F_1\left(r+x, r+s+j; r+s+j+1; \frac{\beta-\alpha}{\beta+t}\right)}{(\beta+t)^{r+s+j}} \right\}. \end{aligned} \quad (\text{A4})$$

We note that (A3) and (A4) each require  $2(x+1)$  evaluations of the Gaussian hypergeometric function.