Deriving the Conditional PMF of the Pareto/NBD Model

Peter S. Fader
www.petefader.com

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

January 2014

Abstract

This note presents the derivation of an expression for the Pareto/NBD model conditional PMF, $P(X(T, T+t) = x^* \,|\, x, t_x, T)$.

1 Preliminaries

Recall the basic Pareto/NBD model results (Fader and Hardie 2005):

i) The individual-level likelihood function for someone with purchase history (x, t_x, T) is

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

$$= \frac{\lambda^x \mu}{\lambda+\mu} e^{-(\lambda+\mu)t_x} + \frac{\lambda^{x+1}}{\lambda+\mu} e^{-(\lambda+\mu)T}\,. \tag{2}$$

ii) The likelihood function for a randomly chosen individual with purchase history (x, t_x, T) is

$$L(r, \alpha, s, \beta \,|\, x, t_x, T) = \frac{\Gamma(r+x)\alpha^r \beta^s}{\Gamma(r)} \left\{ \left(\frac{s}{r+s+x} \right) \mathsf{A}_1 + \left(\frac{r+x}{r+s+x} \right) \mathsf{A}_2 \right\} \tag{3}$$

where

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

This document can be found at <http://brucehardie.com/notes/028/>.

and

$$\mathsf{A}_2 = \begin{cases} \dfrac{{}_2F_1\big(r+s+x, s; r+s+x+1; \frac{\alpha-\beta}{\alpha+T}\big)}{(\alpha+T)^{r+s+x}} & \text{if } \alpha \geq \beta \\ \dfrac{{}_2F_1\big(r+s+x, r+x+1; r+s+x+1; \frac{\beta-\alpha}{\beta+T}\big)}{(\beta+T)^{r+s+x}} & \text{if } \alpha < \beta \end{cases}$$

iii) The joint posterior distribution of Λ and M is

$$g(\lambda, \mu \,|\, r, \alpha, s, \beta; x, t_x, T) = \frac{L(\lambda, \mu \,|\, x, t_x, T) g(\lambda \,|\, r, \alpha) g(\mu \,|\, s, \beta)}{L(r, \alpha, s, \beta \,|\, x, t_x, T)}, \tag{4}$$

where

$$g(\lambda \,|\, r, \alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)} \tag{5}$$

is the gamma distribution that captures heterogeneity in transaction rates across customers, and

$$g(\mu \,|\, s, \beta) = \frac{\beta^s \mu^{s-1} e^{-\mu\beta}}{\Gamma(s)} \tag{6}$$

is the gamma distribution that captures heterogeneity in "death" rates across customers.

iv) The probability that a customer with purchase history (x, t_x, T) is "alive" at time T is the probability that the (unobserved) time at which he "dies" (ω) occurs after T. Conditional on λ and μ, this is

$$P(\Omega > T \,|\, \lambda, \mu; x, t_x, T) = \frac{\lambda^x e^{-(\lambda+\mu)T}}{L(\lambda, \mu \,|\, x, t_x, T)}. \tag{7}$$

It follows from (1) that the probability that a customer with purchase history (x, t_x, T) is "dead" at time T is

$$P(\Omega \leq T \,|\, \lambda, \mu, x, t_x, T) = \frac{\dfrac{\lambda^x \mu}{\lambda+\mu} e^{-(\lambda+\mu)t_x} - \dfrac{\lambda^x \mu}{\lambda+\mu} e^{-(\lambda+\mu)T}}{L(\lambda, \mu \,|\, x, t_x, T)}. \tag{8}$$

Removing the conditioning on λ and μ gives us

$$\begin{aligned} &P(\Omega > T \,|\, r, \alpha, s, \beta; x, t_x, T) \\ &\qquad = \frac{\Gamma(r+x)\alpha^r \beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s} \Big/ L(r, \alpha, s, \beta \,|\, x, t_x, T). \end{aligned} \tag{9}$$

As we proceed with the derivation, we will need evaluate a double integral of the following form several times:

$$\mathsf{A} = \int_0^\infty \int_0^\infty \frac{\lambda^a \mu^b}{(\lambda+\mu)^c} e^{-(\lambda+\mu)d} g(\lambda \,|\, r, \alpha) g(\mu \,|\, s, \beta) \, d\lambda \, d\mu. \tag{10}$$

Consider the transformation $Y = M/(\Lambda + M)$ and $Z = \Lambda + M$. Noting that the Jacobian of this transformation is

$$J = \begin{vmatrix} \frac{\partial\lambda}{\partial y} & \frac{\partial\lambda}{\partial z} \\ \frac{\partial\mu}{\partial y} & \frac{\partial\mu}{\partial z} \end{vmatrix} = -z,$$

it follows from the standard transformation of random variables method (Casella and Berger 2002, Section 4.3, pp. 156–162; Mood et al. 1974, Section 6.2, p. 204ff) that the joint distribution of Y and Z is

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

We solve (10) in the following manner:

$$\begin{aligned}
\mathsf{A} &= \int_0^1 \int_0^\infty y^b (1-y)^a z^{a+b-c} e^{-zd} g(y, z \,|\, \alpha, \beta, r, s)\, dz\, dy \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 \int_0^\infty y^{s+b-1}(1-y)^{r+a-1} z^{r+s+a+b-c-1} e^{-z(\alpha+d-(\alpha-\beta)y)}\, dz\, dy \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \int_0^1 y^{s+b-1}(1-y)^{r+a-1} \left\{ \int_0^\infty z^{r+s+a+b-c-1} e^{-z(\alpha+d-(\alpha-\beta)y)}\, dz \right\} dy
\end{aligned}$$

which for $r+s+a+b>c$

$$= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \frac{\Gamma(r+s+a+b-c)}{(\alpha+d)^{r+s+a+b-c}} \int_0^1 y^{s+b-1}(1-y)^{r+a-1} \big[1 - \big(\tfrac{\alpha-\beta}{\alpha+d}\big) y\big]^{-(r+s+a+b-c)}\, dz$$

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

$$\begin{aligned}
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \frac{\Gamma(r+s+a+b-c)}{(\alpha+d)^{r+s+a+b-c}} B(r+a, s+b) \\
&\qquad \times {}_2F_1\big(r+s+a+b-c, s+b; r+s+a+b; \tfrac{\alpha-\beta}{\alpha+d}\big) \,. \qquad (11)
\end{aligned}$$

Looking closely at (11), we see that the argument z of the Gaussian hypergeometric function, $\frac{\alpha-\beta}{\alpha+d}$, 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 $\frac{\alpha-\beta}{\alpha+d} < -1$, in which case the series is divergent.

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

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

gives us

$$\begin{aligned}
\mathsf{A} &= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \frac{\Gamma(r+s+a+b-c)}{(\beta+d)^{r+s+a+b-c}} B(r+a, s+b) \\
&\qquad \times {}_2F_1\big(r+s+a+b-c, r+a; r+s+a+b; \tfrac{\beta-\alpha}{\beta+d}\big) \,. \qquad (12)
\end{aligned}$$

(We note that the argument z of the above Gaussian hypergeometric function is bounded between 0 and 1 when $\alpha < \beta$.)

We therefore present (11) and (12) as solutions to (10): we use (11) when $\alpha \geq \beta$ and (12) when $\alpha < \beta$.

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

2 Derivation

Suppose we know an individual's unobserved latent characteristics λ and μ. If the customer is alive at T (i.e., $\omega > T$), it follows from the derivation of $P(X(t) = x \,|\, \lambda, \mu)$ (Fader and Hardie 2006) that

$$\begin{aligned} &P(X(T, T+t) = x^* \,|\, \lambda, \mu, \omega > T) \\ &\quad = \frac{(\lambda t)^{x^*} e^{-(\lambda+\mu)t}}{x^*!} + \frac{\lambda^{x^*}\mu}{(\lambda+\mu)^{x^*+1}}\left[1 - e^{-(\lambda+\mu)t}\sum_{i=0}^{x^*}\frac{\big[(\lambda+\mu)t\big]^i}{i!}\right]. \end{aligned} \tag{13}$$

When $x^* > 0$,

$$\begin{aligned} &P(X(T, T+t) = x^* \,|\, \lambda, \mu; x, t_x, T) \\ &\quad = P(X(T, T+t) = x^* \,|\, \lambda, \mu, \omega > T) P(\Omega > T \,|\, \lambda, \mu; x, t_x, T)\,. \end{aligned}$$

This also holds when $x^* = 0$. However, in this second case we also have to account for the possibly that no purchases are observed in $(T, T+t]$ because the customer died at or before T. Therefore,

$$\begin{aligned} P(X(T, T+t) = x^* \,|\, \lambda, \mu, x, t_x, T) &= \delta_{x^*=0}\big[1 - P(\Omega > T \,|\, \lambda, \mu; x, t_x, T)\big] \\ &+ P(X(T, T+t) = x^* \,|\, \lambda, \mu, \omega > T) P(\Omega > T \,|\, \lambda, \mu; x, t_x, T)\,. \end{aligned} \tag{14}$$

Therefore,

$$\begin{aligned} P(X(T, T+t) = x^* \,|\, \lambda, \mu; x, t_x, T) &= \delta_{x^*=0}\left\{1 - \frac{\lambda^x e^{-(\lambda+\mu)T}}{L(\lambda, \mu \,|\, x, t_x, T)}\right\} \\ + \frac{1}{L(\lambda, \mu \,|\, x, t_x, T)}&\left\{\frac{t^{x^*}}{x^*!}\lambda^{x+x^*}e^{-(\lambda+\mu)(T+t)} + \frac{\lambda^{x+x^*}\mu}{(\lambda+\mu)^{x^*+1}}e^{-(\lambda+\mu)T}\right. \\ &\left. - \sum_{i=0}^{x^*}\frac{t^i}{i!}\frac{\lambda^{x+x^*}\mu}{(\lambda+\mu)^{x^*-i+1}}e^{-(\lambda+\mu)(T+t)}\right\}. \end{aligned} \tag{15}$$

We remove the conditioning on λ and μ by taking the expectation of this over the joint posterior distribution of Λ and M, (4), giving us

$$\begin{aligned} P(X(T, T+t) = x^* \,|\, r, \alpha, s, \beta; x, t_x, T) &= \delta_{x^*=0}\big\{1 - P(\Omega > T \,|\, r, \alpha, s, \beta; x, t_x, T)\big\} \\ &+ \frac{1}{L(r, \alpha, s, \beta \,|\, x, t_x, T)}\left\{\frac{t^{x^*}}{x^*!}\mathsf{B}_1 + \mathsf{B}_2 - \sum_{i=0}^{x^*}\frac{t^i}{i!}\mathsf{B}_{3i}\right\} \end{aligned} \tag{16}$$

where

$$\begin{aligned} \mathsf{B}_1 &= \int_0^\infty\int_0^\infty \lambda^{x+x^*}e^{-(\lambda+\mu)(T+t)}g(\lambda \,|\, r, \alpha)g(\mu \,|\, s, \beta)\,d\lambda\,d\mu \\ &= \left\{\int_0^\infty \lambda^{x+x^*}e^{-\lambda(T+t)}g(\lambda \,|\, r, \alpha)\,d\lambda\right\}\left\{\int_0^\infty e^{-\mu(T+t)}g(\mu \,|\, s, \beta)\,d\mu\right\} \\ &= \frac{\Gamma(r+x+x^*)}{\Gamma(r)}\frac{\alpha^r\beta^s}{(\alpha+T+t)^{r+x+x^*}(\beta+T+t)^s} \end{aligned} \tag{17}$$

$$\begin{aligned}
\mathsf{B}_2 &= \int_0^\infty \int_0^\infty \frac{\lambda^{x+x^*}\mu}{(\lambda+\mu)^{x^*+1}} e^{-(\lambda+\mu)T} g(\lambda \,|\, r,\alpha) g(\mu \,|\, s,\beta)\, d\lambda\, d\mu \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \Gamma(r+s+x) B(r+x+x^*, s+1) \\
&\quad \times \begin{cases} \dfrac{{}_2F_1\big(r+s+x, s+1; r+s+x+x*+1; \frac{\alpha-\beta}{\alpha+T}\big)}{(\alpha+T)^{r+s+x}} & \text{if } \alpha \geq \beta \\ \dfrac{{}_2F_1\big(r+s+x, r+x+x^*; r+s+x+x^*+1; \frac{\beta-\alpha}{\beta+T}\big)}{(\beta+T)^{r+s+x}} & \text{if } \alpha < \beta \end{cases}
\end{aligned} \tag{18}$$

$$\begin{aligned}
\mathsf{B}_{3i} &= \int_0^\infty \int_0^\infty \frac{\lambda^{x+x^*}\mu}{(\lambda+\mu)^{x^*-i+1}} e^{-(\lambda+\mu)(T+t)} g(\lambda \,|\, r,\alpha) g(\mu \,|\, s,\beta)\, d\lambda\, d\mu \\
&= \frac{\alpha^r \beta^s}{\Gamma(r)\Gamma(s)} \Gamma(r+s+x+i) B(r+x+x^*, s+1) \\
&\quad \times \begin{cases} \dfrac{{}_2F_1\big(r+s+x+i, s+1; r+s+x+x*+1; \frac{\alpha-\beta}{\alpha+T+t}\big)}{(\alpha+T+t)^{r+s+x+i}} & \text{if } \alpha \geq \beta \\ \dfrac{{}_2F_1\big(r+s+x+i, r+x+x^*; r+s+x+x^*+1; \frac{\beta-\alpha}{\beta+T+t}\big)}{(\beta+T+t)^{r+s+x+i}} & \text{if } \alpha < \beta \end{cases}
\end{aligned} \tag{19}$$

3 The Special Case of $x^* = 0$

We can derive a simpler expression for the special case of $x^* = 0$ (i.e., no purchasing in the interval $(T, T+t]$). Conditional on λ and μ, there are three ways in which a customer with purchase history (x, t_x, T) could make no purchases in the interval $(T, T+t]$:

- The customer is dead at T, which occurs with probability $P(\Omega \leq T \,|\, \lambda, \mu, x, t_x, T)$, (8), or
- The customer is alive at T, which occurs with probability $P(\Omega > T \,|\, \lambda, \mu, x, t_x, T)$, (7), and
  - he remains alive through the interval $(T, T+t]$ (with probability $e^{-\mu t}$) and makes no purchases in that interval (the Poisson probability of which is $e^{-\lambda t}$), or
  - he dies at time ω in the interval $(T, T+t]$ and makes no purchases in the interval $(T, \omega]$:

$$\begin{aligned}
\int_T^{T+t} e^{-\lambda(\omega-T)} \mu e^{-\mu(\omega-T)} d\omega &= \mu \int_0^t e^{-(\lambda+\mu)s} ds \\
&= \frac{\mu}{\lambda+\mu}\Big(1 - e^{-(\lambda+\mu)t}\Big)\,.
\end{aligned}$$

Combining terms and simplifying gives us

$$\begin{aligned}
P(X(T, T+t) = 0 \,|\, \lambda, \mu; x, t_x, T) &= \frac{1}{L(\lambda, \mu \,|\, x, t_x, T)} \\
&\quad \times \left\{ \frac{\lambda^x \mu}{\lambda+\mu} e^{-(\lambda+\mu)t_x} + \frac{\lambda^{x+1}}{\lambda+\mu} e^{-(\lambda+\mu)(T+t)} \right\}.
\end{aligned} \tag{20}$$

Taking the expectation of this over the joint posterior distribution of λ and μ, (4), it follows from (10)–(12) that

$$P(X(T,T+t)=0\,|\,r,\alpha,s,\beta;x,t_x,T)=\frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)}\times\frac{1}{L(r,\alpha,s,\beta\,|\,x,t_x,T)}\left\{\left(\frac{s}{r+s+x}\right)\mathsf{A}_1+\left(\frac{r+x}{r+s+x}\right)\mathsf{A}_3\right\},\tag{21}$$

where

$$\mathsf{A}_3=\begin{cases}\dfrac{{}_2F_1\big(r+s+x,s;r+s+x+1;\frac{\alpha-\beta}{\alpha+T+t}\big)}{(\alpha+T+t)^{r+s+x}} & \text{if } \alpha\geq\beta\\[2ex] \dfrac{{}_2F_1\big(r+s+x,r+x+1;r+s+x+1;\frac{\beta-\alpha}{\beta+T+t}\big)}{(\beta+T+t)^{r+s+x}} & \text{if } \alpha<\beta\end{cases}$$

Given (3), this can be rewritten as

$$P(X(T,T+t)=0\,|\,r,\alpha,s,\beta;x,t_x,T)=\frac{s\mathsf{A}_1+(r+x)\mathsf{A}_3}{s\mathsf{A}_1+(r+x)\mathsf{A}_2}\,.\tag{22}$$

Setting x^* to 0 in (16) does not give us (21). Are these two equations equivalent? We first note that evaluating (15) at $x^*=0$ and using (8) as the expression for $1-P(\Omega>T\,|\,\lambda,\mu;x,t_x,T)$ (and simplifying) gives us (20). Therefore (21) must be equivalent to (16) at $x^*=0$.

To prove this, we first note that (1) leads to an alternative expression for the Pareto/NBD likelihood function (Fader and Hardie 2005):

$$L(r,\alpha,s,\beta\,|\,x,t_x,T)=\frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)}\left\{\frac{1}{(\alpha+T)^{r+x}(\beta+T)^s}+\left(\frac{s}{r+s+x}\right)\mathsf{A}_0\right\},\tag{23}$$

where for $\alpha\geq\beta$

$$\mathsf{A}_0=\frac{{}_2F_1\big(r+s+x,s+1;r+s+x+1;\frac{\alpha-\beta}{\alpha+t_x}\big)}{(\alpha+t_x)^{r+s+x}}-\frac{{}_2F_1\big(r+s+x,s+1;r+s+x+1;\frac{\alpha-\beta}{\alpha+T}\big)}{(\alpha+T)^{r+s+x}}$$

while for $\alpha<\beta$

$$\mathsf{A}_0=\frac{{}_2F_1\big(r+s+x,r+x;r+s+x+1;\frac{\beta-\alpha}{\beta+t_x}\big)}{(\beta+t_x)^{r+s+x}}-\frac{{}_2F_1\big(r+s+x,r+x;r+s+x+1;\frac{\beta-\alpha}{\beta+T}\big)}{(\beta+T)^{r+s+x}}\,.$$

Let us first consider the case of $\alpha\geq\beta$. Substituting (9) and (23) in (16) for $x^*=0$ gives us

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

Equivalence to (21) implies

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

Noting that ${}_2F_1(a,b;b;z) = (1-z)^{-a}$,

$${}_2F_1\big(r+s+x, s; r+s+x; \tfrac{\alpha-\beta}{\alpha+T+t}\big) = \left(\frac{\alpha+T+t}{\beta+T+t}\right)^s\,,$$

equivalence implies

$$(r+x)\,{}_2F_1\big(r+s+x, s; r+s+x+1; \tfrac{\alpha-\beta}{\alpha+T+t}\big)$$
$$+ s\,{}_2F_1\big(r+s+x, s+1; r+s+x+1; \tfrac{\alpha-\beta}{\alpha+T+t}\big)$$
$$- (r+s+x)\,{}_2F_1\big(r+s+x, s; r+s+x; \tfrac{\alpha-\beta}{\alpha+T+t}\big) = 0\,. \tag{24}$$

One of the so-called Gauss' relations for contiguous functions states (Abramowitz and Stegun 1972, equation 15.2.24) is

$$(c-b-1)\,{}_2F_1(a,b;c;z) + b\,{}_2F_1(a,b+1;c;z) - (c-1)\,{}_2F_1(a,b;c-1;z) = 0\,. \tag{25}$$

Therefore (24) is true, which means (16) (evaluated at $x^* = 0$) and (21) are equivalent when $\alpha \geq \beta$.

Turning to the case of $\alpha < \beta$, substituting (9) and (23) in (16) for $x^* = 0$ gives us

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

Equivalence to (21) implies

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

Using the result ${}_2F_1(a,b;b;z) = (1-z)^{-a}$, we get

$${}_2F_1\big(r+s+x, r+x; r+s+x; \tfrac{\beta-\alpha}{\beta+T+t}\big) = \left(\frac{\beta+T+t}{\alpha+T+t}\right)^{r+x}\,,$$

meaning equivalence implies

$$s\,{}_2F_1\big(r+s+x, r+x; r+s+x+1; \tfrac{\beta-\alpha}{\beta+T+t}\big)$$
$$+ (r+x)\,{}_2F_1\big(r+s+x, r+x+1; r+s+x+1; \tfrac{\beta-\alpha}{\beta+T+t}\big)$$
$$- (r+s+x)\,{}_2F_1\big(r+s+x, r+x; r+s+x; \tfrac{\beta-\alpha}{\beta+T+t}\big) = 0\,, \tag{26}$$

which we see from (25) is true. Therefore (16) (evaluated at $x^* = 0$) and (21) are also equivalent when $\alpha < \beta$.

References

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

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/>.

Fader, Peter S. and Bruce G. S. Hardie (2006), "Deriving an Expression for $P(X(t) = x)$ Under the Pareto/NBD Model," <http://brucehardie.com/notes/012/>.