Econometric Institute Report No. 9720/A

ON PURCHASE TIMING MODELS IN MARKETING

J.B.G. Frenk and S. Zhang

Econometric Institute
Erasmus University Rotterdam

May, 1997

ABSTRACT

In this paper we consider stochastic purchase timing models used in marketing for low-involvement products and show that important characteristics of those models are easy to compute. As such these calculations are based on an elementary probabilistic argument and cover not only the well-known condensed negative binomial model but also stochastic purchase timing models with other interarrival and mixing distributions.

Key words: Marketing, purchase timing model.

AMS subject classification: 90A60, 60G07.

1 Introduction

In this paper we consider purchase timing models used within the marketing literature (cf. [8]) and show by easy arguments how to compute some important characteristics of these models under various assumptions on the mixing distribution and the associated "standardized" purchase timing process. After introducing a general framework for these models we discuss a purchase timing model with an Erlang-r mixing distribution and an arbitrary point process representing this "standardized" purchase timing process. Also we consider a purchase timing model with an arbitrary mixing distribution and an Erlang-s renewal process as a "standardized" purchase timing process. For the last class of models it is relatively easy to derive analytical formulas for the important characteristics and these formulas generalize most of the results available in the literature. At the same time we show that the mathematics involved is quite elementary.

2 Purchase timing models

Let $\{X_i : i \geq 1\}$ denote a sequence of nonnegative random variables and consider the associated nonexplosive univariate point process $\{N(t) : t \geq 0\}$ given by $N(t) := \sup\{n \geq 0 : T_n \leq t\}$ with T_n, $n \in \mathcal{N}$, denoting the sum of the random variables X_i, $1 \leq i \leq n$, and $T_0 := 0$ (cf. [2]). Observe if the random variables X_i, $i = 1, ...,$ are independent and identically distributed with distribution function $F(x) := \Pr\{X_i \leq x\}$ satisfying $F(0) = 0$ the above point process represents a renewal process (cf. [11]). To model the moments of purchase timing of a customer selected at random from a population it is assumed that the interpurchase times of this random customer are given by X_i/Y, $i \in \mathcal{N}$, with Y a nonnegative random variable with distribution $G(y) := \Pr\{Y \leq y\}$. This distribution is continuous on $(0, \infty)$ and satisfies $0 \leq G(0) < 1$ and $G(\infty) = 1$. Moreover, the random variable Y representing the purchase rate parameter (cf. [8]) is independent of the sequence X_i, $i \geq 1$. Within the theory of consumer behavior the distribution G is called the mixing distribution and this distribution enables us to aggregate over the whole population of customers. Observe also that in most of the literature on consumer behavior the univariate point process $\{N(t) : t \geq 0\}$ is actually a renewal process with either an exponential or Erlang-2 interarrival distribution. Introducing now the stochastic process $\{B_t : t \geq 0\}$ given by

$$B_t := \text{ the number of purchases of a random customer up to time } t$$

it follows by the above construction that $B_t = N(Yt)$. A well-known model belonging to this class is given by the Negative Binomial model (NBD) (cf. [8, 5, 9]). In this model it is assumed that the mixing distribution is a Gamma distribution and the associated point process is a Poisson process with arrival rate 1. From a theoretical point of view important characteristics of the random variable B_t are its distribution, its first moment and generating function. To compute the distribution of B_t we observe, since the event $\{N(t) \geq k\}$, $k \in \mathcal{N}$, coincides with the event $\{T_k \leq t\}$, that

$$\Pr\{B_t \geq k\} = \Pr\{N(Yt) \geq k\} = \Pr\{T_k \leq Yt\} = \Pr\{Y \geq T_k t^{-1}\}.$$

Since G is continuous on $(0, \infty)$ and T_k is strictly positive with probability one we obtain that

$$\Pr\{B_t \geq k\} = \Pr\{Y > T_k t^{-1}\} = 1 - \mathcal{E}G(T_k t^{-1}) \tag{2.1}$$

with $\mathcal{E}$ denoting the expectation. If it happens that the considered population consists of m different classes each characterized by a different random purchase rate parameter Y_i, $i = 1, ..., m$, the mixing distribution G can be seen as a mixture of distributions. This means that there exist positive numbers $p_1, ..., p_m$ adding up to 1 with p_i representing the relative size of class i within the population and each random customer belonging to class i has a random purchase rate parameter Y_i with distribution G_i. Hence in this case the mixing distribution G is given by

$$G(y) = \sum_{i=1}^{m} p_i G_i(y)$$

or equivalently G is the distribution of the random variable Y_I where I denotes a random variable with $\Pr\{I = i\} = p_i$, $i = 1, ..., m$ and I is independent of the random variables $Y_1, ..., Y_m$. By (2.1) we now obtain that

$$\Pr\{B_t \geq k\} = \Pr\{N(Y_I t) \geq k\} = \sum_{i=1}^{m} p_i \Pr\{N(Y_i t) \geq k\} = \sum_{i=1}^{m} p_i \Pr\{B_t^{(i)} \geq k\} \tag{2.2}$$

with $B_t^{(i)}$ denoting the number of purchases up to time t of a customer selected at random from class i. A special case is given by the existence of a zero and a nonzero-class within the population and by (2.2) this implies that

$$\Pr\{B_t \geq k\} = (1 - p_1)\Pr\{B_t^{(2)} \geq k\} + p_1 \delta_0(k)$$

with $\delta_0(k) = 1$ for $k = 0$ and zero otherwise and $B_t^{(2)}$ denoting the number of purchases up to time t of a customer selected at random from the non-zero class. From a theoretical point of view there seems to be no preference for a specific mixing distribution and so the selection of such a distribution is purely determined by the flexibility of the family of distributions to which this mixing distribution belongs. Since the family of Gamma distributions with scale parameter $\mu > 0$ and shape parameter $\beta > 0$ seems to be flexible enough the Gamma distribution is chosen in most of the literature (for example see [14, 3]) as a mixing distribution. If the shape parameter β is an integer r the corresponding Gamma distribution is called an Erlang-r distribution and in this case the corresponding random variable Y can be represented as the sum of r independent and exponentially distributed random variables Y_i, $i = 1, ..., r$, with the same scale parameter μ or equivalently $\mathcal{E}_r(y) := \Pr\{Y \leq y\} = \Pr\{Y_1 + \cdots + Y_r \leq y\}$. By taking finite mixtures of Erlang-r distributions with different values of r and the same scale parameter it can be shown that this class of distributions is dense in the class of all distributions on $[0, \infty)$ (cf. [1]). By this result and (2.2) it seems therefore sensible to compute for an arbitrary nonexplosive univariate point process and Erlang-r mixing distribution the probability distribution $\Pr\{B_t = k\}$, $k = 0, 1, ...$ Observe an example of such a model is given by the Condensed Negative Binomial model (cf. [9]) where the point process is a renewal process with an Erlang-2 interarrival time distribution. This is the simplest distribution with an increasing failure rate and so it incorporates the intuitive idea that the probability of a new purchase will increase with time. By relating the Erlang-r mixing distribution to the well-known Poisson process it is easy to show the following result.

Theorem 2.1 *If a purchase timing model is represented by a nonexplosive univariate point process* $\{N(t) : t \geq 0\}$ *and an Erlang-r mixing distribution with scale parameter* $\mu > 0$ *then we obtain for every* $k \geq 0$ *that* $Pr\{B_t \geq k\} = Pr\{N(Yt) \geq k\} = Pr\{M(T_k t^{-1}) \leq r - 1\}$ *with* $\{M(t) : t \geq 0\}$ *denoting a Poisson process with arrival rate* μ *and* T_k *independent of the Poisson process* $\{M(t) : t \geq 0\}$*. Moreover, it follows that*

$$Pr\{B_t \geq k\} = \sum_{j=0}^{r-1} \frac{(\mu t^{-1})^j}{j!} \mathcal{E}(\exp(-\mu t^{-1} T_k) T_k^j).$$

Proof. As already observed the random purchase parameter Y can be seen as the sum of r independent and exponentially distributed random variables Y_i with scale parameter $\mu > 0$ and so we obtain that

$$\begin{aligned} \Pr\{B_t \geq k\} &= \Pr\{N((Y_1 + \cdots + Y_r)t) \geq k\} \\ &= \Pr\{T_k \leq (Y_1 + \cdots + Y_r)t\} \\ &= \Pr\{Y_1 + \cdots + Y_r \geq T_k t^{-1}\}. \end{aligned}$$

Since the mixing Erlang-r distribution is continuous on $(0,\infty)$ and $Y_1+\cdots+Y_r$ is independent of T_k it follows that

$$\Pr\{Y_1+\cdots+Y_r \geq T_k t^{-1}\} = \Pr\{Y_1+\cdots+Y_r > T_k t^{-1}\} = \Pr\{M(T_k t^{-1}) \leq r-1\}$$

with $\{M(t): t \geq 0\}$ denoting a Poisson process with arrival rate μ and this shows the first part. Since it is well-known for a Poisson process with arrival rate $\mu > 0$ that the number of renewals in the interval $(0, T_k t^{-1})$ has a Poisson distribution with parameter $\mu T_k t^{-1}$ (cf. [11]) the second part follows.

Q.E.D.

If we consider the random variable $M(T_k t^{-1})$ mentioned in Theorem 2.1 and compute its probability generating function $P(z) := \mathcal{E}(z^{M(T_k t^{-1})})$, $\mid z \mid \leq 1$ then it is easy to verify that $P(z) = \mathcal{E}(\exp(-\mu t^{-1} T_k(1-z)))$ and so the distribution function of $M(T_k t^{-1})$ is a so-called Poisson mixture (cf. [10]). Moreover, if the point process $\{N(t): t \geq 0\}$ is actually a renewal process then it follows that

$$\begin{aligned} P(z) &= \mathcal{E}(\exp(-\mu t^{-1} T_k(1-z))) = \mathcal{E}(\exp(-\mu t^{-1} \sum_{i=1}^{k} X_i(1-z))) \\ &= (\mathcal{E}(\exp(-\mu t^{-1} X_1(1-z))))^k = (\mathcal{E}(z^{M(X_1 t^{-1})}))^k \end{aligned}$$

and this implies that the random variable $M(T_k t^{-1})$ can be seen as the sum of the independent and identically distributed random variables $M(X_i t^{-1})$, $i = 1, ..., k$. By this observation it follows by Theorem 2.1 that

$$\Pr\{B_t \geq k\} = \Pr\{\sum_{i=1}^{k} M(X_i t^{-1}) \leq r-1\} \tag{2.3}$$

and again the distribution of the independent and identically distributed random variables $M(X_i t^{-1})$, $i = 1, ..., k$ is a Poisson mixture.

To identify an important subclass of Poisson mixtures we introduce the following well-known discrete distribution (cf. [6]).

Definition 2.1 *A discrete random variable N defined on $\{0, 1, ...,\}$ has a geometric distribution with parameter p (Geo(p) distribution) if $Pr\{N = j\} = (1-p)p^j$, $j = 0, ...$.*

The next result shows that the random variable $M(X_i t^{-1})$ has a geometric distribution if and only if X_i has an exponential distribution.

Lemma 2.1 *Let $\{M(t) : t \geq 0\}$ be a Poisson process with arrival rate $\mu > 0$ and X a non-negative random variable independent of $\{M(t) \mid t \geq 0\}$. Then it follows that $M(Xt^{-1})$ has a geo($\frac{\lambda t}{\lambda t+\mu}$) distribution if and only if X has an exponential distribution with scale parameter λ.*

Proof. As observed the generating function of $M(Xt^{-1})$ is given by $\mathcal{E}(\exp(-\mu t^{-1}X(1-z)))$ and this yields for the random variable X exponentially distributed with parameter λ that

$$\mathcal{E}(\exp(-\mu t^{-1}X(1-z))) = \frac{\lambda}{\lambda+\mu t^{-1}(1-z)} = \frac{\frac{\lambda t}{\lambda t+\mu}}{1-\frac{\mu}{\lambda t+\mu}z}.$$

The above function is the generating function of the *geo*($\frac{\lambda t}{\lambda t+\mu}$) distribution and this shows the if-implication. To prove the reverse relation we observe for every $t > 0$ that

$$\mathcal{E}(\exp(\mu t^{-1}X(1-z))) = \frac{\lambda}{\lambda+\mu t^{-1}(1-z)}.$$

Hence for every $\alpha > 0$ the Laplace-Stieltjes transform $\mathcal{E}(\exp(-\alpha X))$ is given by $\lambda(\lambda+\alpha)^{-1}$ which denotes the Laplace-Stieltjes transform of the exponential distribution with parameter λ.

Q.E.D.

Since the negative binomial distribution (with parameters k and p) given by

$$p_j := \binom{k+j-1}{j}(1-p)^j p^k,\ j = 0, 1, ...$$

can be seen (cf. [6]) as $p_j = \Pr\{Z_1 + \cdots Z_k = j\}$ with Z_i, $i = 1, ..., k$, a sequence of independent and $Geo(p)$ distributed random variables we obtain by Lemma 2.1 and relation (2.3) the following result.

Theorem 2.2 *If a purchase timing model is represented by a renewal process with Erlang-s distributed interarrival times with scale parameter $\lambda > 0$ and an Erlang-r mixing distribution with scale parameter $\mu > 0$ then it follows for every $k \geq 1$ that*

$$Pr\{B_t \geq k\} = \sum_{j=0}^{r-1} \binom{sk+j-1}{j} (\frac{\mu}{\lambda t+\mu})^j (\frac{\lambda t}{\lambda t+\mu})^{sk}.$$

Proof. Since for every $i \geq 1$ the independent and identically distributed random variables X_i have an Erlang-s distribution it follows by relation (2.3) that

$$\begin{aligned}
\Pr\{B_t \geq k\} &= \Pr\{\sum_{i=1}^{k} M(X_i t^{-1}) \leq r-1\} \\
&= \Pr\{\sum_{i=1}^{k}\sum_{j=1}^{s} M(X_{ij} t^{-1}) \leq r-1\}
\end{aligned}$$

with X_{ij}, $i = 1,...,k$, $j = 1,...,s$, a sequence of independent and exponentially distributed random variables with scale parameter $\lambda > 0$. Moreover, the random variables $M(X_{ij}t^{-1})$, $i = 1,...,k$, $j = 1,...,s$, are independent and by Lemma 2.1 $geo(\frac{\lambda t}{\lambda t+\mu})$ distributed and this proves by the interpretation of a negative binomial distribution the desired result.

Q.E.D.

Moreover, if the renewal process $\{N(t): t \geq 0\}$ has an interarrival distribution given by a finite mixture of Erlang distributions with the same scale parameter $\lambda > 0$ then it follows for every $i \geq 1$ that

$$\Pr\{X_i \leq x\} = \Pr\{\sum_{j=1}^{N_i} X_{ij} \leq x\}$$

with N_i a discrete random variable on $\{1,...,s\}$ for some finite s. Observe the sequences $\{N_i, i \geq 1\}$ and $\{X_{ij}, i \geq 1, j = 1,...,s\}$ are independent of each other and consist of independent and identically distributed random variables with N_i having an arbitrary distribution on $\{1,...,s\}$ and X_{ij} exponentially distributed with scale parameter $\lambda > 0$. By a similar argument as used in Theorem 2.2 it is easy to show that

$$\begin{aligned}
\Pr\{B_t \geq k\} &= \Pr\{\sum_{i=1}^{k}\sum_{j=1}^{N_i} M(X_{ij} t^{-1}) \leq r-1\} \\
&= \sum_{j=0}^{r-1}(\frac{\mu}{\lambda t+\mu})^j \mathcal{E}(\binom{N_1+\cdots+N_k+j-1}{j}(\frac{\lambda t}{\lambda t+\mu})^{N_1+\cdots N_k}).
\end{aligned}$$

Although the above formula can be worked out for a random variable N_i with an elementary probability generating function the resulting expression is rather complicated. Moreover, it is also not clear from a theoretical point of view why the interarrival distribution of a "standardized" purchase timing process should be a mixture of Erlang distributions and so we will only consider Erlang-s distributed interarrival times.

One might justify the use of an Erlang-s interarrival distribution by the observation that this model captures the possibility that a customer is going through s different exponential stages before buying. Another, maybe more realistic reason is given by the observation that every Erlang-s distribution, $s \geq 2$, has an increasing failure rate (cf. [4]) and is therefore a theoretical more attractive distribution to model the time between purchase moments than the exponential distribution with a constant failure rate (cf. [14]). In the remainder of this section we therefore focus our attention to purchase incidence models with a *general* mixture distribution G and a renewal process with an Erlang-s interarrival time distribution and compute for these models the probability distribution $\Pr\{B_t = k\}$, the first moment $\mathcal{E}B_t$, the generating function $\mathcal{E}(z^{B_t})$ and the conditional expectation $\mathcal{E}(B_u - B_t \mid B - t = k)$ with $u > t$. By the probabilistic interpretation of an Erlang distribution the following result follows immediately. Clearly for $s = 2$ and a Gamma mixing distribution we obtain the well-known Condensed Negative Binomial model (cf. [8, 14]).

Theorem 2.3 *If a purchase timing process is represented by a renewal process with Erlang-s distributed interarrival times with scale parameter $\lambda > 0$ and an arbitrary mixing distribution G then we obtain that*

$$Pr\{B_t \geq k\} = Pr\{N(Yt) \geq k\} = Pr\{M(Yt) \geq sk\}$$

with $\{M(t) : t \geq 0\}$ denoting a Poisson process with arrival rate $\lambda > 0$ independent of the random variable Y. Moreover, it follows that

$$Pr\{B_t \geq k\} = \sum_{j=sk}^{\infty} \frac{(\lambda t)^j}{j!} \mathcal{E}(\exp(-t\lambda Y)Y^j).$$

Proof. Since the interarrival times X_i, $i \geq 1$ are independent and Erlang-s distributed with scale parameter $\lambda > 0$ we obtain that

$$\begin{aligned}
\Pr\{B_t \geq k\} &= \Pr\{N(Yt) \geq k\} \\
&= \Pr\{X_1 + \cdots + X_k \leq Yt\} \\
&= \Pr\{Z_1 + \cdots + Z_{ks} \leq Yt\}
\end{aligned}$$

with Z_i, $i = 1, ..., sk$, a sequence of independent and exponentially distributed random variables with scale parameter $\lambda > 0$. Hence it follows that

$$\Pr\{B_t \geq k\} = \Pr\{M(Yt) \geq sk\}$$

and this implies since $\{M(t): t \geq 0\}$ is a Poisson process that

$$\Pr\{B_t \geq k\} = \sum_{j=sk}^{\infty} \frac{(\lambda t)^j}{j!} \mathcal{E}(\exp(-t\lambda Y)Y^j).$$

Q.E.D.

We will now compute the first moment $\mathcal{E}B_t$ and the generating function $\mathcal{E}(z^{B_t})$ of a purchase timing model with a renewal process consisting of Erlang-s distributed interarrival times with scale parameter λ and a general mixing distribution. As observed by Morrison and Schmittlein (cf. [9]) it is also important to consider the conditional expectation $\mathcal{E}(B_u - B_t \mid B_t = k)$ with $u > t$ and this conditional expectation will also be computed for the above purchase timing model. To compute all these important characteristics we need the following lemma which is well-known within the theory of fast Fourier transforms.

Lemma 2.2 *For any real number x and integer $s = 1, 2, ...$ it follows for any integer m satisfying $1 \leq m \leq s$ that*

$$\sum_{k=1}^{\infty} \frac{x^{ks-m}}{(ks-m)!} = \frac{1}{s} \sum_{j=0}^{s-1} \theta^{jm} \exp(x\theta^j)$$

with $\theta := \exp(\frac{2\pi i}{s})$ and i the imaginary unit.

Proof.

By the Taylor expansion for $\exp(x\theta^j)$, $j = 0, ..., s-1$ we obtain that

$$\frac{1}{s} \sum_{j=0}^{s-1} \theta^{jm} \exp(x\theta^j) = \frac{1}{s} \sum_{j=0}^{s-1} \theta^{jm} \sum_{n=0}^{\infty} \frac{x^n \theta^{nj}}{n!} = \frac{1}{s} \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{j=0}^{s-1} (\theta^{m+n})^j.$$

If $m + n \geq 1$ is a multiple of s, i.e. $m + n = ks$ for some $k \in \mathcal{N}$, then clearly $\theta^{m+n} = 1$ and hence

$$\sum_{j=0}^{s-1} (\theta^{m+n})^j = s.$$

Moreover, if $m + n$ is not a multiple of s, then $\theta^{m+1} \neq 1$ and this yields by the formula for a geometric serie that

$$\sum_{j=0}^{s-1} (\theta^{m+n})^j = \frac{1 - \theta^{s(m+n)}}{1 - \theta^{m+n}} = 0.$$

Substituting this into the above double series yields the desired result.

Q.E.D.

It is now possible to compute the expectation $\mathcal{E}B_t$ of a purchase timing model with a renewal process consisting of Erlang-s interarrival times with scale parameter $\lambda > 0$ and a general mixing distribution.

Theorem 2.4 *If Y denotes the purchase rate parameter with distribution G and $\{N(t) : t \geq 0\}$ is a renewal process independent of Y with Erlang-s interarrival times with scale parameter λ then we obtain with $\theta = \exp(\frac{2\pi i}{s})$ that*

$$\mathcal{E}B_t = \frac{\lambda t}{s}\mathcal{E}Y + \frac{1-s}{2s} - \frac{1}{s}\sum_{j=1}^{s-1}\frac{\theta^j}{1-\theta^j}\mathcal{E}(\exp(-\lambda t(1-\theta^j)Y)).$$

Proof. It is well-known that

$$\mathcal{E}B_t = \sum_{k=1}^{\infty}\Pr\{B_t \geq k\} = \sum_{k=1}^{\infty}\Pr\{N(Yt) \geq t\} = \sum_{k=1}^{\infty}\Pr\{T_k \leq Yt\}.$$

Since the interarrival times X_i, $i = 1, 2, ...$, are independent and Erlang-s distributed with scale parameter $\lambda > 0$ it follows that $T_k = Z_1 + \cdots + Z_{ks}$ with Z_i, $i = 1, ..., ks$, a sequence of independent and exponentially distributed random variables with scale parameter λ and so T_k is Erlang-ks distributed. This implies by Lemma 2.2 that

$$\begin{aligned}
\mathcal{E}B_t &= \sum_{k=1}^{\infty}\mathcal{E}(\int_0^{Yt}\lambda\exp(-\lambda x)\frac{(\lambda x)^{ks-1}}{(ks-1)!}dx) \\
&= \mathcal{E}(\int_0^{Yt}\lambda\exp(-\lambda x)\sum_{k=1}^{\infty}\frac{(\lambda x)^{ks-1}}{(ks-1)!}dx) \\
&= \frac{1}{s}\sum_{j=0}^{s-1}\lambda\theta^j\mathcal{E}(\int_0^{Yt}\exp(-\lambda x(1-\theta^j))dx) \\
&= \frac{\lambda t}{s}\mathcal{E}Y + \frac{1}{s}\sum_{j=1}^{s-1}\frac{\theta^j}{1-\theta^j}(1-\mathcal{E}(\exp(-\lambda t(1-\theta^j)Y))).
\end{aligned}$$

Finally by the key renewal theorem (cf. [4]) it follows that

$$\lim_{t\uparrow\infty}\mathcal{E}B_t - \frac{\lambda t}{s}\mathcal{E}Y - \frac{1}{2}(1-s)s^{-1} = 0$$

and so the desired result is verified.

Q.E.D.

If we consider the condensed negative binomial model for which by definition the independent and identically distributed interarrival times have an Erlang-2 distribution it follows by Theorem 2.4 (take $s = 2$) that

$$\mathcal{E}B_t = \mathcal{E}(N(Yt)) = \frac{1}{2}\mathcal{E}(Y)\lambda t - \frac{1}{4}(1 - \mathcal{E}(\exp(-2\lambda tY)))$$

and this expression is still elementary. Using Lemma 2.2 one can also compute the generating function $\mathcal{E}(z^{B_t})$ of a purchase timing model with an Erlang-s interarrival time and an arbitrary mixing distribution. This will be shown in the next result.

Theorem 2.5 *If Y denotes the purchase rate parameter with distribution G and $\{N(t) : t \geq 0\}$ is a renewal process independent of Y with Erlang-s interarrival times with scale parameter $\lambda > 0$ then we obtain with $\theta = \exp(\frac{2\pi i}{s})$ that*

$$\mathcal{E}(z^{B_t}) = \frac{1}{s}(1-z)z^{-1}\sum_{j=0}^{s-1}(z^{1/s}\theta^j)(1 - z^{1/s}\theta^j)^{-1}\mathcal{E}(\exp(-\lambda t(1 - z^{1/s}\theta^j)Y))$$

for every $\mid z \mid < 1$.

Proof. Since we consider a purchase timing model with an Erlang-s renewal process and a general mixing distribution we obtain for $\mid z \mid < 1$ that

$$\begin{aligned}\sum_{k=0}^{\infty}\Pr\{B_t \geq k\}z^k &= 1 + \sum_{k=1}^{\infty}\Pr\{B_t \geq k\}z^k \\ &= 1 + \sum_{k=1}^{\infty}\Pr\{N(Yt) \geq k\}z^k \\ &= 1 + \sum_{k=1}^{\infty}\Pr\{T_k \leq Yt\}z^k\end{aligned}$$

with T_k denoting the sum of k independent and Erlang-s distributed random variables with parameter λ. Hence the distribution of T_k is an Erlang-ks distribution with scale parameter λ and this implies

$$\begin{aligned}\sum_{k=0}^{\infty}\Pr\{B_t \geq k\}z^k &= 1 + \sum_{k=1}^{\infty}\mathcal{E}(\int_0^{Yt}\lambda\exp(-\lambda x)\frac{(\lambda x)^{ks-1}}{(ks-1)!}dx)z^k \\ &= 1 + z^{1/s}\mathcal{E}(\int_0^{Yt}\lambda\exp(-\lambda x)\sum_{k=1}^{\infty}\frac{(\lambda x z^{1/s})^{ks-1}}{(ks-1)!}dx).\end{aligned}$$

Applying now Lemma 2.2 yields

$$\begin{aligned}
\sum_{k=0}^{\infty} \Pr\{B_t \geq k\} z^k &= 1 + \frac{\lambda}{s} z^{1/s} \sum_{j=0}^{s-1} \theta^j \mathcal{E}\left(\int_0^{Yt} \exp(-\lambda x) \exp(\lambda x z^{1/s} \theta^j) dx\right) \\
&= 1 + \frac{\lambda}{s} z^{1/s} \sum_{j=0}^{s-1} \theta^j \mathcal{E}\left(\int_0^{Yt} \exp(-\lambda x (1 - z^{1/s} \theta^j)) dx\right) \\
&= 1 + \frac{z^{1/s}}{s} \sum_{j=0}^{s-1} \theta^j (1 - z^{1/s} \theta^j)^{-1} (1 - \mathcal{E}(\exp(-\lambda t (1 - z^{1/s} \theta^j) Y))).
\end{aligned}$$

Since for every $\mid z \mid < 1$ it follows that

$$\lim_{t \to \infty} \mathcal{E}(\exp(-\lambda t (1 - z^{1/s} \theta^j) Y) = 0$$

for every $j = 0, ..., s-1$ we obtain by the previous equality and using $B_t \to \infty$ that

$$\begin{aligned}
\frac{1}{1-z} &= \lim_{t \to \infty} \sum_{k=0}^{\infty} \Pr\{B_t \geq k\} z^k \\
&= 1 + \frac{z^{1/s}}{s} \sum_{j=0}^{s-1} \theta^j (1 - z^{1/s} \theta^j)^{-1}.
\end{aligned}$$

This implies

$$\sum_{k=0}^{\infty} \Pr\{B_t \geq k\} z^k = \frac{1}{1-z} - \frac{z^{1/s}}{s} \sum_{j=0}^{s-1} \theta^j (1 - z^{1/s} \theta^j)^{-1} \mathcal{E}(\exp(-\lambda t (1 - z^{1/s} \theta^j) Y)).$$

Finally, since

$$\sum_{k=0}^{\infty} \Pr\{B_t \geq k\} z^k = \sum_{k=0}^{\infty} \sum_{j=k}^{\infty} \Pr\{B_t = j\} z^k = \sum_{j=0}^{\infty} \sum_{k=0}^{j} z^k \Pr\{B_t = j\} = \frac{1 - z\mathcal{E}(z^{B_t})}{1-z}$$

the desired result follows.

Q.E.D.

Using the above generating function it is in principle possible to compute the first and second moment of B_t if the purchase timing model has a Erlang-r renewal process with scale parameter $\lambda > 0$ and the Laplace-Stieltjes transform $\mathcal{E}(\exp(-\alpha Y))$ of the random purchase rate parameter Y has an elementary form. Suppose now we consider a renewal process $\{N(t) : t \geq 0\}$ with an interarrival distribution having a finite second moment and an increasing failure rate and this interarrival distribution does not belong to the class of

Erlang distributions. In this case the associated renewal function $U(t) = \mathcal{E}(N(t))$ does not have an elementary expression and so it seems worthwhile at least for large values of t to approximate the renewal function $U(t)$ by its asymptotic limit $\mathcal{E}(X_1)^{-1}t+\frac{1}{2}\mathcal{E}(X_1^2)\mathcal{E}(X_1)^{-2}-1$ (cf. [4]). If this is a reasonable approximation then we may approximate $\mathcal{E}(N(Yt))$ by $t\mathcal{E}(Y)\mathcal{E}(X_1) + \frac{1}{2}\mathcal{E}(X_1^2)\mathcal{E}(X_1)^{-2} - 1$. Similarly one can show (cf. [7]) by classical renewal theoretic arguments that the asymptotic limit of the second moment $\mathcal{E}(N^2(t))$ is given by the expression

$$\frac{t^2}{(\mathcal{E}(X_1))^2} + (\frac{2\mathcal{E}(X_1^2)}{(\mathcal{E}(X_1))^2} - \frac{3}{\mathcal{E}(X_1)})t + \frac{3(\mathcal{E}(X_1^2))^2}{2(\mathcal{E}(X_1))^4} - \frac{2\mathcal{E}(X_1^3)}{3(\mathcal{E}(X_1))^3} - \frac{3\mathcal{E}(X_1^2)}{2(\mathcal{E}(X_1))^2} + 1$$

and so $\mathcal{E}N(Yt)$ can be approximated by

$$\frac{\mathcal{E}(Y^2)t^2}{\mathcal{E}(X_1))^2} + (\frac{2\mathcal{E}(X_1^2)}{(\mathcal{E}(X_1))^2} - \frac{3}{\mathcal{E}(X_1)})\mathcal{E}(Y)t + \frac{3(\mathcal{E}(X_1^2))^2}{2(\mathcal{E}(X_1))^4} - \frac{2\mathcal{E}(X_1^3)}{3(\mathcal{E}(X_1))^3} - \frac{3\mathcal{E}(X_1^2)}{2(\mathcal{E}(X_1))^2} + 1.$$

Finally we will consider for a purchase timing model with an Erlang-s renewal process and a general mixing distribution the expectation $\mathcal{E}(B_u - B_t \mid B_t = k)$. Using the well-known memoryless property of the exponential distribution it is easy to show the following result.

Theorem 2.6 *If Y denotes the purchase rate parameter with distribution G and $\{N(t) : t \geq 0\}$ is a renewal process independent of Y with Erlang-s interarrival times with scale parameter $\lambda > 0$ then we obtain that*

$$\mathcal{E}(B_u - B_t \mid B_t = k) = \frac{\lambda}{s}\mathcal{E}Y(u-t) + \frac{1}{s}\sum_{j=1}^{s-1}\frac{\theta^j}{1-\theta^j}(1 - \mathcal{E}(\exp(-\lambda(u-t)(1-\theta^j)Y)))v_j$$

with v_j, $j = 1, ..., s-1$ given by

$$v_j = \mathcal{E}(\theta^{jM(Yt)} \mid sk \leq M(Yt) \leq (k+1)s - 1\}$$

and $\{M(t) : t \geq 0\}$ is a Poisson process with arrival rate $\lambda > 0$ independent of Y.

Proof. Since the interarrival times are Erlang-s distributed with scale parameter $\lambda > 0$ the event $\{B_t = k\}$ is given by the union of the disjoint events $\{M(Yt) = sk+m\}$, $m = 0, ..., s-1$ with $M(t)$ denoting a Poisson process with arrival rate $\lambda > 0$ and so we obtain that

$$\mathcal{E}((B_u - B_t)1_{\{B_t=k\}}) = \sum_{m=0}^{s-1}\mathcal{E}((N(Yu) - N(Yt))1_{\{M(Yt)=sk+m\}}).$$

By the probabilistic interpretation of an Erlang-s distribution and the memoryless property of the exponential distribution it follows that

$$\sum_{m=0}^{s-1} \mathcal{E}((N(Yu) - N(Yt))1_{\{M(Yt)=sk+m\}}) = \sum_{m=0}^{s-1} \mathcal{E}(N^{(m)}(Y(u-t)))\Pr\{M(Yt) = sk+m\}$$

with $N^{(m)}(t)$ a delayed Erlang-s renewal process (cf. [11])) with delay distribution given by an Erlang-$(s-m)$ distribution with parameter $\lambda > 0$. Applying now Lemma 2.2 it follows for each $m = 0, ..., s-1$ that

$$\begin{aligned}
\mathcal{E}N^{(m)}(Y(u-t)) &= \sum_{k=1}^{\infty} \mathcal{E}(\int_0^{Y(u-t)} \lambda \exp(-\lambda x) \frac{(\lambda x)^{ks-m-1}}{(ks-m-1)!} dx) \\
&= \mathcal{E}(\int_0^{Y(u-t)} \lambda \exp(-\lambda x) \sum_{k=1}^{\infty} \frac{(\lambda x)^{ks-m-1}}{(ks-m-1)!} dx) \\
&= \frac{\lambda}{s}\mathcal{E}Y(u-t) + \frac{1}{s}\sum_{j=1}^{s-1} \frac{\theta^{j(m+1)}}{1-\theta^j}(1 - \mathcal{E}(\exp(-\lambda(u-t)(1-\theta^j)Y))).
\end{aligned}$$

Combining the above expressions we finally obtain that

$$\mathcal{E}(B_u - B_t \mid B_t = k) = \frac{\lambda}{s}\mathcal{E}Y(u-t) + \frac{1}{s}\sum_{j=1}^{s-1} \frac{\theta^j}{1-\theta^j}(1 - \mathcal{E}(\exp(-\lambda(u-t)(1-\theta^j)Y)))v_j$$

with v_j, $j = 1, ..., s-1$ given by

$$\begin{aligned}
v_j &= \sum_{m=0}^{s-1} \theta^{jm} \Pr\{M(Yt) = sk+m \mid ks \leq M(Yt) \leq (k+1)s - 1\} \\
&= \mathcal{E}(\theta^{j(M(Yt)-sk)} \mid sk \leq M(Yt) \leq (k+1)s - 1\} \\
&= \mathcal{E}(\theta^{jM(Yt)} \mid sk \leq M(Yt) \leq (k+1)s - 1\}
\end{aligned}$$

and so we have verified the desired result.

Q.E.D.

Although the above expression seems complicated it is not difficult to compute its value for any arbitrary s and so we can use the above formula as a before-and-after tool (cf. [9, 12]). This concludes our discussion of random purchase timing models. Observe we did not consider the relation of purchase timing models with the more general purchase timing/brand choice models (cf. [8]). However, it will be shown in a subsequent paper that there exists a unifying framework covering almost all stationary purchase timing/brand choice models considered in the marketing literature and that the main characteristics of these more general models reduce to the main characteristics of the above "single product" models. Finally we like to observe that by a simple probabilistic argument and an easy to prove equality between an infinite series and a finite serie we can compute without an extensive amount of calculations important characteristics of purchase timing models which are much more general than the models considered so far in the literature. This should be seen in contrast with the direct computation type of approach for some special subcases used in the marketing literature on this topic.

References

[1] Asmussen, S., *Applied Probability and Queues*, Wiley, New York, 1987.

[2] Brémand, P., *Point Processes and Queues (Martingale Dynamics)*, Springer-Verlag, New York, 1981.

[3] Chatfield, C., Goodhardt, G.J., *A consumer purchasing model with Erlang inter-purchase times*, Journal of the American Statistical Association 86 (344), 828-835, 1973.

[4] Chaudry, M.L., *On computations of the mean and Variance of number of the number of renewals: A unified approach*, Journal of the Operational Research Society 46, 1352-1364, 1995.

[5] Ehrenberg, A.S.C., *Repeat buying: Facts, Theory and Data*, 2nd edition, New York, Oxford University Press, 1988.

[6] Feller,W., *An introduction to Probability Theory and Its Applications*, vol 1 (third edition), Wiley, New York, 1970.

[7] De Kok, A.G., *Basics of Inventory Management: Part I. Renewal Theoretic Background*, Research Paper Department of Economics FEW 510, Tilburg University, 1991.

[8] Lilien, G.L., Kotler, P., Moorthy, K.S., *Marketing Models*, Prentice Hall, 1992.

[9] Morrison, D.G., Schmittlein D.C., *Generalizing the NBD model for customer purchases: What are the implications and is it worth the effort*, Journal of Business and Economic Statistics 6 (2), 145-159, 1988.

[10] Puri, P.S., Goldie, C.M., *Poisson mixtures and quasi-infinite divisibility of distributions*, J. Appl. Prob. 16, 138-153, 1979.

[11] Ross, S.M., *Applied Probability models with Optimization Applications*, Holden Day, San Francisco, 1970.

[12] Schmittlein, D.C., Morrison, D.G., *Prediction of future events with the condensed negative binomial distribution*, Journal of the American Statistical Association 78 (382), 449-456, 1983.

[13] Tijms, H.C., *Stochastic models (an algorithmic approach)*, Wiley, New York, 1994.

[14] Zufryden, F.S., *A composite heterogeneous model of brand choice and purchase timing behaviour*, Management Science 24 (2), 121-136, 1977.