

# A General Approximation for the Distribution of Count Data

Edward Fox, Bezalel Gavish, John Semple\*

Cox School of Business

Southern Methodist University

Dallas, TX - 75275

## Abstract

Under mild assumptions about the interarrival distribution, we derive a modified version of the Birnbaum-Saunders distribution, which we call the tBISA, as an approximation for the true distribution of count data. The free parameters of the tBISA are the first two moments of the underlying interarrival distribution. We show that the density for the sum of tBISA variables is available in closed form. This density is determined using the tBISA's moment generating function, which we introduce to the literature. The tBISA's moment generating function additionally reveals a new mixture interpretation that is based on the inverse Gaussian and gamma distributions. We then show that the tBISA can fit count data better than the distributions commonly used to model demand in economics and business. In numerical experiments and empirical applications, we demonstrate that modeling demand with the tBISA can lead to better economic decisions.

*Keywords:* Birnbaum-Saunders; inverse Gaussian; gamma; confluent hypergeometric functions; inventory model.

2000 Mathematics Subject Classification: Primary 62E99, Secondary 91B02.

---

\*Corresponding author: [jsemple@cox.smu.edu](mailto:jsemple@cox.smu.edu)

# 1. INTRODUCTION

It is often necessary to count the number of arrivals or events during an interval of time. In many applications, we can exploit the relationship between count data and the underlying interarrival times. For example, customer purchase data captured by point-of-sale systems can be used to estimate the distribution of demand, i.e. the count of transactions.

Previous research relating count and interarrival distributions relied on parametric assumptions about the interarrival times. By far the most common such assumption is that interarrival times are exponentially distributed, which implies Poisson counts. Despite its attractive simplicity, the Poisson distribution is not particularly robust because it accommodates only equi-dispersed count data. To improve its flexibility for count applications, several modifications to the Poisson distribution have been proposed (see [18] for a summary). Unfortunately, these modified Poisson distributions sacrifice the linkage to the interarrival distribution. More recently, Winkelmann [19] investigated the count distribution implied by gamma-distributed interarrival times (which nests exponential), and McShane et al. [13] determined the count distribution implied by Weibull-distributed interarrival times. Both papers keep the theoretical linkage between interarrivals and counts but yield different count distributions.

In this paper, we propose a more general approximation for the distribution of count data which also retains the linkage to the interarrival distribution. We assume only that interarrival times are *i.i.d.* with finite mean and variance. Under these mild assumptions, we show that a modified version of the Birnbaum-Saunders (BISA) distribution [2], which we call the tBISA, offers a robust fit for count data in a number of applications. It is flexible, permitting under-dispersed, equi-dispersed, and over-dispersed data; it is positively skewed, consistent with nearly all count data; and the logarithm of a tBISA random variable has a symmetric, unimodal distribution, which explains why the logarithm of count data usually exhibits the same property (and hence why the log transform is used in practice). In addition, the duration of the counting period (e.g., daily, weekly, yearly) can be changed without collecting new data.

The BISA was originally derived by Birnbaum and Saunders as an approximate distribution for the number of cycles until failure of a material specimen. It has subsequently been used almost exclusively as a lower-tail approximation in reliability analysis where the focus is on low cycle counts leading to material failure. In contrast, the tBISA is intended as an upper-tail approximation in systems analysis where the focus is on high counts exceeding system capacity (e.g., inventory and production problems).

The tBISA's free parameters are the first two moments of the interarrival distribution. These parameters can be estimated from count data, as is commonly done with other count distributions, or they can be estimated directly from interarrival data. The latter approach presents some special opportunities: (i) the count distribution's true shape can be determined more rapidly because a single count of size  $m$  represents  $m-1$  interarrivals; (ii) in cases where counts are censored because of an upper bound (as would be the case when capacity is exceeded or inventory is exhausted), the underlying interarrival data are not censored so valid parameter estimates for the tBISA distribution

can still be obtained from interarrival data. We demonstrate both estimation approaches.

Because many applications require that counts be summed, we investigate the additive properties of the tBISA. For example, the interarrival distribution may change (e.g., by time-of-day, day-of-the-week or season), thereby violating the assumption that arrival times are identically distributed. Another example involves dynamic inventory models. Determining the optimal policy parameters in some dynamic inventory models requires aggregating demand over the number of periods in the delivery lag.

Determining the sum of tBISA random variables requires that we derive the BISA's moment generating function (mgf), which appears to have been previously undiscovered (interestingly, the mgf of the log-BISA, also called the sinh-normal, is known, albeit in terms of modified Bessel functions of the third kind [15]). The BISA mgf reveals that the distribution can be represented as a mixture, in equal proportions, of (i) an inverse Gaussian and (ii) the same inverse Gaussian plus an independent gamma distribution with shape parameter  $k = 1/2$ . This result isolates and clarifies the difference between the BISA distribution and the inverse Gaussian. Our mixture interpretation implies that a reciprocal inverse Gaussian is the sum of an inverse Gaussian and an independent gamma with shape parameter  $k = 1/2$ . Thus, adding this independent gamma to an inverse Gaussian has the peculiar effect of taking its reciprocal. This mixture interpretation is quite different from that obtained by Desmond [4], who showed that the BISA is a mixture of an inverse Gaussian and a reciprocal inverse Gaussian. We generalize our mixture result and use it to obtain a closed-form expression for the density of the sum of  $n$  independent BISA random variables. This closed-form expression provides an analytical and computational advantage over some other distributions commonly used in counting problems, e.g., the lognormal.

## 2. A GENERAL APPROXIMATION FOR THE DISTRIBUTION OF COUNT DATA

Let  $X_i$  be the  $i^{th}$  interarrival time (the time between arrivals  $i-1$  and  $i$ ), a nonnegative continuous random variable. We assume interarrival times are *i.i.d.*, though we will relax this assumption later. Assuming the first interarrival is measured with respect to time  $t = 0$  (as in [17]), the cumulative probability of count  $C$  being  $n$  or less in the time interval  $[0, T]$  is given by

$$\Pr(C \leq n) = \Pr\left(\sum_{i=1}^{n+1} X_i > T\right) = 1 - \Pr\left(\sum_{i=1}^{n+1} X_i \leq T\right). \quad (1)$$

Although the count distribution is known to be asymptotically normal  $N(T/\mu, T\sigma^2/\mu^3)$  as  $T$  goes to infinity (e.g., see Theorem 3.3.5 of [17]), there is an approximate intermediate distribution that applies when  $T$  is finite but large.

Assume the interarrival distribution has mean  $\mu$  and standard deviation  $\sigma$ ; then by (1) and the

central limit theorem, the probability of the count  $C$  being  $n$  or less is

$$\begin{aligned} \Pr(C \leq n) &= \Pr\left(\sum_{i=1}^{n+1} X_i > T\right) = \Pr\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n+1}} > \frac{T/(n+1) - \mu}{\sigma/\sqrt{n+1}}\right) \\ &\approx 1 - \Phi\left(\frac{T/(n+1) - \mu}{\sigma/\sqrt{n+1}}\right) \end{aligned} \quad (2)$$

where  $\Phi(\cdot)$  is the cumulative distribution function for the standard normal. Approximating the discrete count  $n$  with a continuous variable  $x \geq 0$ , we obtain the density

$$\frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{[T - (x+1)\mu]}{\sigma\sqrt{x+1}}\right)^2\right) \cdot \left(\frac{T + (x+1)\mu}{(x+1)^{3/2}}\right). \quad (3)$$

By comparison, Birnbaum and Saunders [2] use  $n$  instead of  $(n+1)$  when modeling the number of cycles until failure (this is because  $n = 0$  is not a possibility in their model; it is in ours), so their density is

$$\frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{[T - x\mu]}{\sigma\sqrt{x}}\right)^2\right) \cdot \left(\frac{T + x\mu}{x^{3/2}}\right). \quad (4)$$

Birnbaum and Saunders also define new parameters  $\alpha = \sigma/\sqrt{\mu T}$ ,  $\beta = T/\mu$ , and rewrite their density in terms of these parameters. Because  $\mu$  and  $\sigma$  are observable in many count applications and because we intend to exploit the connection between interarrival and count distributions, we will retain the parameters  $\mu$ ,  $\sigma$  and  $T$ .

The probability of the count equaling  $n$  using (3) corresponds to the area between  $n$  and  $n-1$ , which is not centered at  $n$  but instead at  $n - 1/2$ . This can be corrected by shifting (3) to the right by  $1/2$  unit. We refer to the continuity-corrected density, which retains the linkage to the timing distribution, as the tBISA

$$\frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{[T - (x+.5)\mu]}{\sigma\sqrt{x+.5}}\right)^2\right) \cdot \left(\frac{T + (x+.5)\mu}{(x+.5)^{3/2}}\right). \quad (5)$$

The continuity correction is especially important in count problems like those in the following section, where precise numerical comparisons are desired.

A variety of shapes are possible for (5). For example, taking  $\mu = 20$  and  $T = 500$ , we have plotted the density for  $\sigma = 10, 20, 30, 40$  in Figure 1. These plots are not meant to be exhaustive but rather to illustrate the connection between the parameters of the interarrival distribution and the shape of the count distribution. Observe that increasing the variance (or coefficient of variation) in the interarrival distribution leads to greater variance and greater positive skewness in the count distribution. Calculation of the first three moments of (5) provides a more precise description of this result. These moment calculations involve an integral approximation where  $(T - .5\mu)/\sigma\sqrt{.5}$  is replaced by  $+\infty$  in the upper limit of an integral involving the standard normal density. If  $(T - .5\mu)/\sigma\sqrt{.5}$  exceeds 3 or 4—a condition that would be met in all but the lowest count problems—

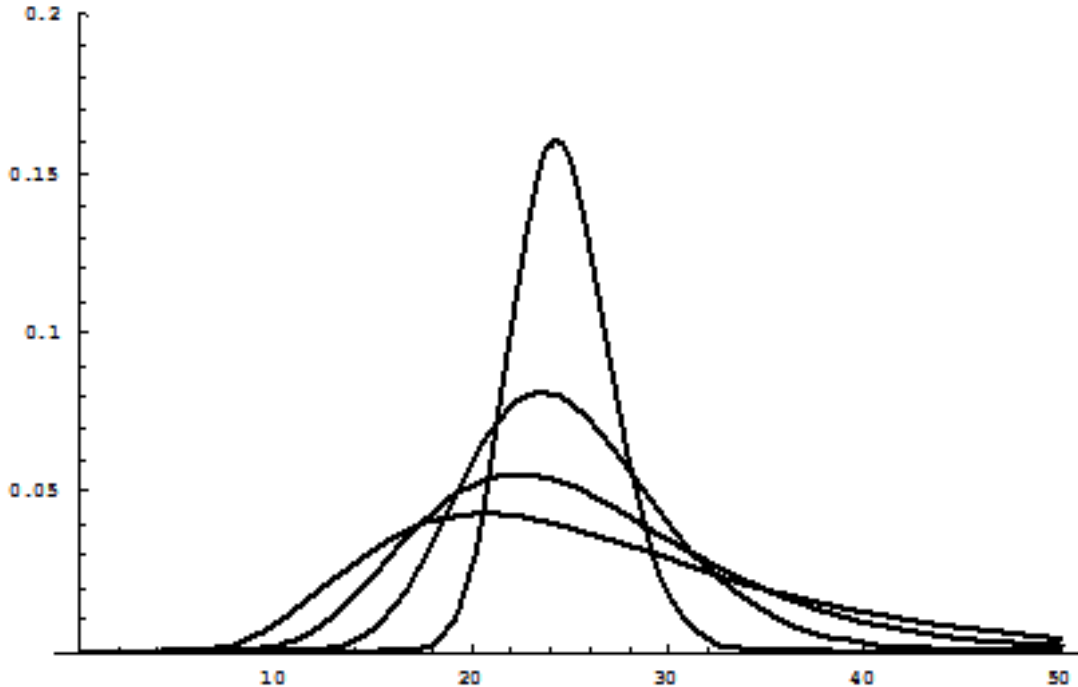


Figure 1: tBISA density for  $T = 500, \mu = 20, \sigma = 10, 20, 30, 40$

this approximation is very good. Thus, while the next proposition states approximate results, the results are nearly exact for practical purposes.

**Proposition 1.** *Let the mean and standard deviation of the (stationary) interarrival distribution be  $\mu$  and  $\sigma$ , respectively. Then the first three moments about the mean for the count distribution (5) are*

$$\begin{aligned}
 (i) \quad E(C) &\cong \frac{T}{\mu} - .5 + \frac{\sigma^2}{2\mu^2} \\
 (ii) \quad E(C - E(C))^2 &\cong \frac{5\sigma^4}{4\mu^4} + \frac{T}{\mu} \cdot \frac{\sigma^2}{\mu^2} \\
 (iii) \quad E(C - E(C))^3 &\cong \frac{11\sigma^6}{2\mu^6} + \frac{T}{\mu} \cdot \frac{\sigma^4}{\mu^4}
 \end{aligned}$$

Not surprisingly, result (i) is  $1/2$  unit less than the corresponding result in [2] while result (ii) is identical. Result (iii) can be obtained from [9] after a little algebra. We note that the moment formulas in Proposition 1 are all functions of just two fundamental quantities, the coefficient of variation of the interarrival distribution,  $\sigma/\mu$ , and the ratio  $T/\mu$ . Moreover, the moments are all increasing functions of these two terms. In particular, the third moment about the mean is always positive so the count distribution is always positively skewed.

**Proposition 2.** *The density (5) is unimodal, and its mode is less than its median which is less than its mean.*

When a tBISA random variable (5) is log-transformed, it produces a symmetric, unimodal distribution that resembles a normal distribution. This result is analogous to that obtained in [16] for the BISA distribution (4).

**Proposition 3.** *Suppose that the count  $C$  has the density (5). Then  $Y = \ln(C + .5)$  has a unimodal distribution that is symmetric about  $\ln(T/\mu)$ .*

The proof of Proposition 3 is straightforward, and the proposition provides a theoretical basis for modeling the logarithm of count data, as is customarily done in many applications in economics and business. It is worth noting, however, that the tBISA distribution retains an important advantage over logarithmic distributions—it is derived directly from the interarrival distribution whose moments define its free parameters.

### 3. SOME COMPARISONS WITH EXACT COUNT DISTRIBUTIONS

We now assess the accuracy of the tBISA approximation. Under certain assumptions, the probability that the count  $C$  equals  $n$  can be computed exactly so a comparison between the tBISA distribution (5) and a known count distribution is possible. The primary requirements for the interarrival distribution are that (i) the interarrival distribution has nonnegative support and (ii) the distribution for the sum can be determined in a convenient numerical form. We consider two such cases here. The first is a gamma interarrival process, which nests the exponential, Erlang, and chi-square as special cases. The second is a uniform interarrival process. For comparing fits, we report the mean and variance of each distribution (exact count distribution vs. tBISA) as well as the maximum absolute value of the difference,  $D_{max}$ , between the cdf of the exact count distribution and the cdf of the tBISA.

#### 3.1 Gamma Interarrivals

We follow the development of Winkelmann [19]. The time between arrivals is gamma distributed with shape parameter  $k > 0$  and scale parameter  $\theta > 0$ . The time interval is  $[0, T]$ . The mean and variance are  $k\theta$  and  $k\theta^2$ , respectively. The interarrival time has probability density

$$f(\tau; k, \theta) = \frac{1}{\theta^k \Gamma(k)} \tau^{k-1} \exp(-\tau/\theta) \quad \text{for } \tau > 0 \text{ and } k, \theta \in \mathbf{R}^+ \quad (6)$$

Define

$$G(nk, T/\theta) = \frac{1}{\Gamma(nk)} \int_0^{T/\theta} u^{nk-1} \exp(-u) du. \quad (7)$$

The count distribution on the interval  $[0, T]$  is

$$P(C = n) = G(kn, T/\theta) - G(k(n + 1), T/\theta) \quad (8)$$

for  $n = 0, 1, 2, \dots$

Figure 2 illustrates the exact count distribution for  $k = 1/2, 1, 2, \theta = 40, 20, 10$ , and  $T = 500$  as well as the tBISA approximation (5). We chose these combinations so that the mean interarrival

	Gamma Interarrivals			Uniform Interarrivals	
	$k = 1/2, \theta = 40$	$k = 1, \theta = 20$	$k = 2, \theta = 10$	$T = 5$	$T = 10$
$\mu_{count}$	25.5	25	24.75	9.667	19.660
$\mu_{tBISA}$	25.5	25	24.75	9.66667	19.6667
$\sigma_{count}$	7.053368	5	3.544361	1.886	2.602
$\sigma_{tBISA}$	7.416198	5.123475	3.579455	1.86339	2.60875
$D_{max}$	.03762	.02660	.01881	.0029	.0015

Table 1: tBISA approximation compared to exact count distributions

time is 20 in all cases; this makes the count distributions easier to compare. We observe that the tBISA distribution approximates the exact count distribution quite well overall, particularly (and importantly) in the upper tail. This is perhaps not surprising given the asymptotic nature of the central limit theorem used to derive the tBISA density. The means, standard deviations, and  $D_{max}$ 's are given in Table 1. For the exact count distribution, these values were calculated numerically; for the tBISA distribution, these values were computed using the formulas in Proposition 1. The mean of the tBISA approximation equals the true mean exactly and the standard deviations are very close. Observe that the fit of the tBISA approximation improves ( $D_{max}$  is lower) as the parameter  $k$  of the gamma interarrival distribution increases. The case  $k=1$  corresponds to the exponential distribution and so the exact count distribution is Poisson.

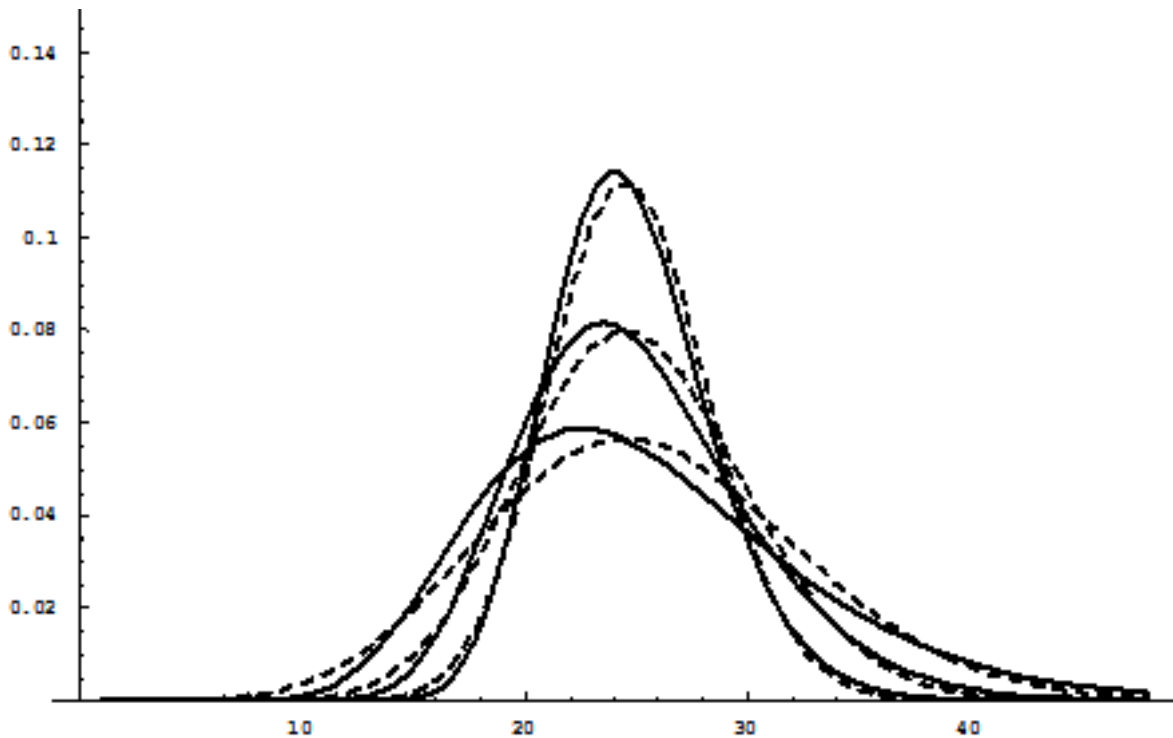


Figure 2: tBISA distribution (solid line) vs. exact count distribution (dashed line) assuming gamma interarrivals.

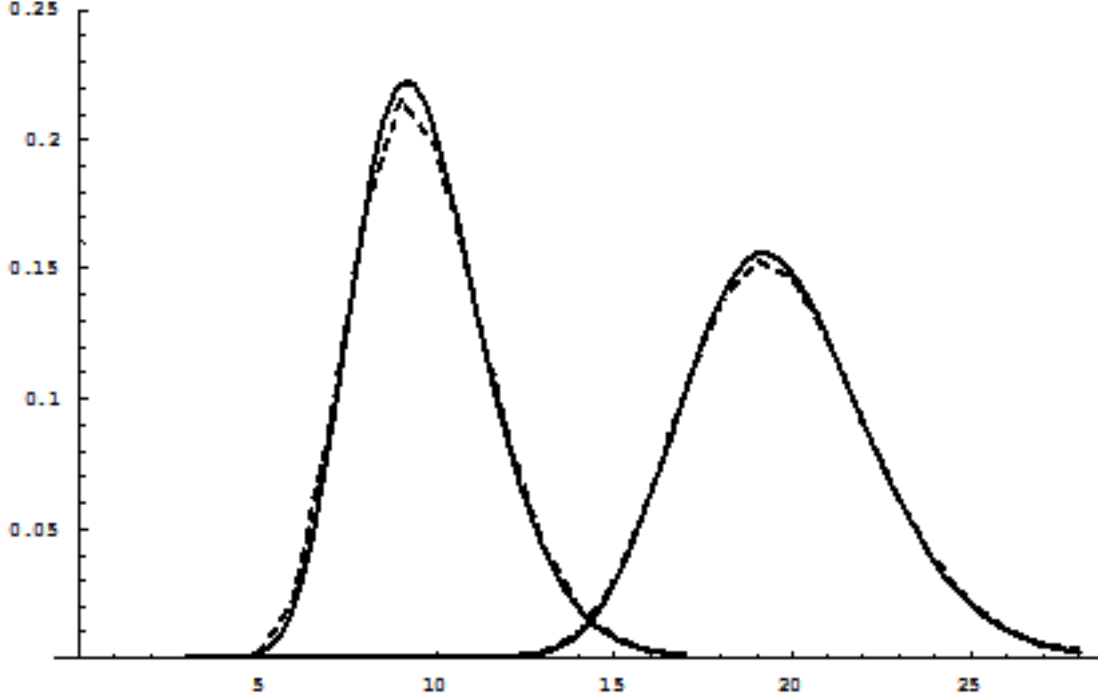


Figure 3: tBISA distribution (solid line) vs. exact count distribution (dashed line) assuming uniform interarrivals.

### 3.2 Uniform Interarrivals

Assume interarrival times are uniform  $U[0,1]$ . The mean and variance are  $1/2$  and  $1/12$ , respectively. Then the density for  $S_n = U_1 + U_2 + \dots + U_n$  is

$$f_n(x) = \frac{1}{2 \cdot (n-1)!} \sum_{k=0}^n (-1)^k \binom{n}{k} (x-k)^{n-1} \text{sgn}(x-k) \quad 0 \leq x \leq n, \quad (9)$$

which can be obtained after some algebra from Theorem 1 in [3]. From (9) one can compute the exact probability of the count equaling  $n$  for the time interval  $[0, T]$

$$P(C = n) = P(S_{n+1} \geq T) - P(S_n \geq T) = \int_T^{n+1} [f_{n+1}(x) - f_n(x)] dx. \quad (T \leq n+1) \quad (10)$$

Comparisons of the tBISA density and  $f_n(x)$  for  $T=5, 10$  are shown in Figure 3 and their fits are compared in Table 1. In both cases, the tBISA approximates the exact count distribution extremely well.



## 4. ADDITIVE PROPERTIES

In many applications, summing random counts is important. In economics and business applications, for example, the demand distribution may vary over time (e.g., by time-of-day or day-of-the-week) so demand over the specified period can be represented as the sum of demands over disjoint subintervals. Also, many inventory problems require determining the distribution of demands *summed over periods* (i.e., a lag or lead time interval).

Determining the sum of BISA random variables requires the mgf of the BISA distribution. We will use the uncorrected BISA (4) to simplify comparison with previous results. The mgf for the tBISA (5), which includes a  $1/2$  unit shift by comparison to the uncorrected BISA, would introduce an additional multiplicative factor  $e^{-t/2}$ .

The inverse Gaussian with mean  $\omega$  and (reciprocal) dispersion parameter  $\lambda$  will play a prominent role in this discussion. Its density is given by

$$f_{IG}(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left(\frac{-\lambda(x-\omega)^2}{2\omega^2 x}\right) \quad x > 0. \quad (11)$$

**Theorem 1.** *Consider a BISA random variable having density*

$$f_{BS}(x) = \frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{[T-(x)\mu]}{\sigma\sqrt{x}}\right)^2\right) \cdot \left(\frac{T+x\mu}{(x)^{3/2}}\right) \quad x > 0.$$

(a) *The mgf exists and equals*

$$M_{BS}(t) = \frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) \left[1 + \frac{1}{\sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}}\right] \quad \text{for } |t| < \frac{\mu^2}{2\sigma^2}.$$

(b) *The BISA random variable is the discrete mixture of two distributions in equal proportion. The first is an inverse Gaussian with  $\omega = T/\mu$  and  $\lambda = T^2/\sigma^2$ ; the second is the sum of an inverse Gaussian (with the same parameters) and an independent gamma distribution with shape parameter  $k = 1/2$  and scale parameter  $\theta = 2\sigma^2/\mu^2$ .*

*Proof.* Observe that the BISA density is similar in structure to the inverse Gaussian. Indeed, if we set  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$ , the associated inverse Gaussian density becomes

$$f_{IG}(x) = \frac{1}{\sqrt{2\pi}} \frac{T}{\sigma} \frac{1}{x^{3/2}} \exp\left(-\frac{1}{2}\left(\frac{T-\mu x}{\sigma\sqrt{x}}\right)^2\right). \quad (12)$$

The mgf for the inverse Gaussian,  $M_{IG}(t)$ , is known to be (see [8])

$$M_{IG}(t) = \int_0^\infty \exp(tx) \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left(\frac{-\lambda(x-\omega)^2}{2\omega^2 x}\right) dx = \exp\left(\frac{\lambda}{\omega}\left(1 - \sqrt{1 - \frac{2\omega^2 t}{\lambda}}\right)\right)$$

$$= \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) \quad \text{for } |t| < \frac{\mu^2}{2\sigma^2} \quad (13)$$

The mgf of the BISA distribution,  $M_{BS}(t)$ , can be expressed in terms of the mgf of the inverse Gaussian

$$\begin{aligned} M_{BS}(t) &= \int_0^\infty \exp(tx) \frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{T-x\mu}{\sigma\sqrt{x}}\right)^2\right) \cdot \left(\frac{T+x\mu}{x^{3/2}}\right) dx \\ &= \frac{1}{2} \int_0^\infty \exp(tx) \frac{T}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{T-x\mu}{\sigma\sqrt{x}}\right)^2\right) \frac{1}{x^{3/2}} dx \\ &\quad + \frac{\mu}{2T} \int_0^\infty \exp(tx) \frac{T}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{T-x\mu}{\sigma\sqrt{x}}\right)^2\right) \frac{1}{x^{1/2}} dx \\ &= \frac{1}{2} M_{IG}(t) + \frac{\mu}{2T} M'_{IG}(t) \end{aligned} \quad (14)$$

(Differentiation of  $M_{IG}(t)$  in equation (14) can be justified for any  $|t| < \mu^2/2\sigma^2$  by applying Lebesgue's Dominated Convergence Theorem to the difference quotients.)

$$\begin{aligned} &= \frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) + \frac{\mu}{2T} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) \left[\frac{T}{\mu} \frac{1}{\sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}}\right] \\ &= \frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) \left[1 + \frac{1}{\sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}}\right] \quad \text{for } |t| < \frac{\mu^2}{2\sigma^2} \end{aligned} \quad (15)$$

This establishes part (a). For part (b), the mgf in (a) can be written as

$$\frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) + \frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}\right)\right) \frac{1}{\sqrt{1 - 2t\frac{\sigma^2}{\mu^2}}} \quad (16)$$

The first term is  $1/2$  the mgf of an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$ . The second term is  $1/2$  the product of (i) the mgf of an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$  and (ii) the mgf of a gamma distribution with shape parameter  $k = 1/2$  and scale parameter  $\theta = 2\sigma^2/\mu^2$  (recall that the mgf of the gamma is  $(1 - \theta t)^{-k}$  for  $|t| < 1/\theta$ ). This implies the result stated in the theorem.  $\square$

By part (a) of the theorem, it can be confirmed that  $M'_{BS}(0) = T/\mu + \sigma^2/2\mu^2$  and  $M''_{BS}(0) - (M'_{BS}(0))^2 = (T/\mu)(\sigma^2/\mu^2) + 5\sigma^4/4\mu^4$ , which are the mean and variance, respectively, of the BISA distribution (4). Though the central moments have been derived before by other means (see [2], [7]) we appear to have discovered a closed-form expression for the mgf and the new mixture interpretation it implies. The mixture interpretation in part (b) of our theorem is quite different from that in [4],

which characterized the BISA distribution as a mixture, in equal proportions, of an inverse Gaussian and a reciprocal inverse Gaussian (the distribution of  $1/X$  where  $X \sim$  inverse Gaussian). Moreover, our mixture interpretation allows us to analyze sums of independent BISA random variables having different parameters  $T_i$ ,  $\mu_i$ , and  $\sigma_i$ , something Desmond's interpretation does not facilitate. Finally, our mixture result implies that the reciprocal inverse Gaussian is equivalent to the sum of an inverse Gaussian and a gamma; this will be revisited after Theorem 2.

Our discussion now turns to summing BISA random variables. The summation requires the use of *confluent hypergeometric functions*, which are general solutions of the differential equation

$$z \frac{d^2 w}{dz^2} + (b - z) \frac{dw}{dz} - aw = 0$$

introduced and analyzed by Kummer [12]. One solution is the *confluent hypergeometric function of the first kind* (also known as Kummer's function of the first kind), whose infinite series is given by

$$M(a, b, z) = 1 + \frac{a}{b}z + \frac{a(a+1)}{b(b+1)} \frac{z^2}{2!} + \frac{a(a+1)(a+2)}{b(b+1)(b+2)} \frac{z^3}{3!} + \dots \quad (17)$$

A second independent solution is the *confluent hypergeometric function of the second kind* (Kummer's function of the second kind), given by

$$U(a, b, z) = \frac{\pi}{\sin(\pi b)} \left\{ \frac{M(a, b, z)}{\Gamma(1+a-b)\Gamma(b)} - z^{1-b} \frac{M(1+a-b, 2-b, z)}{\Gamma(a)\Gamma(2-b)} \right\}. \quad (18)$$

(See Chapter 13 of [1]). Bessel functions, Hermite polynomials, Laguerre polynomials, and the error function are all special cases of confluent hypergeometric functions. These functions have recently been used to calculate closed form expressions for visiting rates [14].

In order to obtain closed-form expressions for the sum of BISA random variables, we need the following theorem, which involves the sum of two independent random variables, an inverse Gaussian and an independent gamma.

**Theorem 2.** *The sum of (i) an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$  and (ii) an independent gamma with shape parameter  $k$  ( $k = 1/2, 1, 1 1/2, 2, \dots$ ) and scale parameter  $\theta = 2\sigma^2/\mu^2$  has density*

$$f_{IG+G}(s) = \frac{T\mu^{2k}}{\sqrt{2\pi} \cdot \sigma \cdot (2\sigma^2)^k} \exp\left(-\frac{1}{2} \left(\frac{T - \mu s}{\sigma\sqrt{s}}\right)^2\right) \cdot s^{k-3/2} U(k, 3/2, T^2/2\sigma^2 s).$$

*Proof.* The density for the stated inverse Gaussian is

$$f_{IG}(x) = \frac{1}{\sqrt{2\pi}} \frac{T}{\sigma} \frac{1}{x^{3/2}} \exp\left(-\frac{1}{2} \left(\frac{T - \mu x}{\sigma\sqrt{x}}\right)^2\right) \quad (19)$$

and that for the independent gamma is

$$f_G(x) = \frac{1}{\Gamma(k) (2\sigma^2/\mu^2)^k} x^{k-1} \exp\left(-\frac{\mu^2 x}{2\sigma^2}\right). \quad (20)$$

Therefore, the sum of the random variables has a density given by the convolution  $f_{IG+G}(s) = \int_0^s f_{IG}(x) f_G(s-x) dx$  ( $s \geq 0$ )

$$= \alpha_k \int_0^s \exp\left(-\frac{1}{2} \left(\frac{T - \mu x}{\sigma \sqrt{x}}\right)^2\right) \exp\left(-\frac{\mu^2 (s-x)}{2\sigma^2}\right) (s-x)^{k-1} \frac{1}{x^{3/2}} dx \quad (21)$$

$$\alpha_k \exp\left(-\frac{1}{2} \left(\frac{\mu^2 s - 2\mu T}{\sigma^2}\right)\right) \int_0^s \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2 x}\right) (s-x)^{k-1} \frac{1}{x^{3/2}} dx \quad (22)$$

where  $\alpha_k = T / (\sqrt{2\pi} \cdot \Gamma(k) \cdot \sigma \cdot (2\sigma^2/\mu^2)^k)$ . Making the change of variable  $x = 1/(u + 1/s)$  for  $s > 0$  (when  $s = 0$  we may take  $f_{IG+G}(0) = 0$ ), (22) becomes

$$\begin{aligned} & \alpha_k \exp\left(-\frac{1}{2} \left(\frac{\mu^2 s - 2\mu T + T^2/s}{\sigma^2}\right)\right) \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2} u\right) \left(\frac{us^2}{us+1}\right)^{k-1} \left(\frac{s}{us+1}\right)^{1/2} du \\ & = \alpha_k \exp\left(-\frac{1}{2} \left(\frac{T - \mu s}{\sigma \sqrt{s}}\right)^2\right) s^{2(k-1)+1/2} \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2} u\right) u^{k-1} \left(\frac{1}{us+1}\right)^{k-1/2} du \end{aligned} \quad (23)$$

Observe that the density is the product of two functions, the first involving only  $s$  and possessing many features of an inverse Gaussian, the second involving an integral where  $s$  is a parameter. We now focus on evaluating the integral portion. Define

$$g_k(s) = \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2} u\right) u^{k-1} \left(\frac{1}{us+1}\right)^{k-1/2} du \quad (24)$$

so that

$$f_{IG+G}(s) = \alpha_k \exp\left(-\frac{1}{2} \left(\frac{T - \mu s}{\sigma \sqrt{s}}\right)^2\right) \cdot s^{2(k-1)+1/2} g_k(s). \quad (25)$$

Observe that

$$g_k(0) = \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2} u\right) u^{k-1} du = \left(\frac{2\sigma^2}{T^2}\right)^k \int_0^\infty \exp(-v) \cdot v^{k-1} dv = \left(\frac{2\sigma^2}{T^2}\right)^k \Gamma(k)$$

*for  $k = .5, 1, 1.5, \dots$ ,*

so  $g_k(s)$  is defined for  $s \geq 0$ , even though we now focus on  $s > 0$ . For each  $k$  ( $k = j/2$ ,  $j = 1, 2, 3, \dots$ ), apply the change of variable  $u = v/s$  ( $s > 0$ ) to obtain

$$g_k(s) = \int_0^\infty \exp\left(-\frac{1}{2} \frac{T^2}{\sigma^2} u\right) u^{k-1} \left(\frac{1}{us+1}\right)^{k-1/2} du = s^{-k} \int_0^\infty \exp\left(-\frac{T^2}{2\sigma^2} \frac{v}{s}\right) v^{k-1} (v+1)^{1/2-k} dv. \quad (26)$$

For  $z > 0$ ,

$$\Gamma(a)U(a, b, z) = \int_0^\infty \exp(-zt) \cdot t^{a-1} (1+t)^{b-a-1} dt, \quad (27)$$

(Formula 13.2.5 of [1], pg 505), which for  $a = k$  and  $b = 3/2$  yields

$$g_k(s) = s^{-k} \int_0^\infty \exp\left(-\frac{T^2}{2\sigma^2} \frac{v}{s}\right) \cdot v^{k-1} (v+1)^{1/2-k} dv = s^{-k} \Gamma(k)U(k, 3/2, T^2/2\sigma^2 s) \quad (28)$$

[Note: For  $k = 1/2$ , we have  $U(1/2, 3/2, T^2/2\sigma^2 s) = (\sigma\sqrt{2s})/T$ , which can be obtained by allowing “ $\Gamma(0) = \infty$ ” in the definition of  $U$  or by direct calculation of  $g_{1/2}(s)$ .] Substituting for  $\alpha_k$  and  $g_k(s)$  in 25 yields the result stated in the theorem.  $\square$

An interesting and immediate consequence of Theorem 2 is the following special result.

**Corollary 1.** *The sum of (i) an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$  and (ii) an independent gamma with shape parameter  $k = 1/2$  and scale parameter  $\theta = 2\sigma^2/\mu^2$  has a reciprocal inverse Gaussian distribution.*

*Proof.* Inserting the stated parameters, the density for the sum reduces to

$$f_{IG+G}(s) = \frac{\mu}{\sqrt{2\pi} \cdot \sigma} \exp\left(-\frac{1}{2} \left(\frac{T - \mu s}{\sigma\sqrt{s}}\right)^2\right) \cdot s^{-1/2},$$

which is the density for a reciprocal inverse Gaussian as described by Desmond [4].  $\square$

Now let  $X_i$  denote a random variable following a BISA distribution (4) with parameters  $T_i$ ,  $\sigma_i$  and  $\mu_i$ . Assume the  $X_i$  are independent. While it is not true that  $\sum_{i=1}^n X_i$  follows a BISA distribution, we can derive the density for the sum under a simple, plausible property of the coefficients of variation.

**Property 1.** *There exists a positive constant  $v$  such that  $\sigma_i/\mu_i = v$  for all  $i$ .*

This property implies that, while count data may be under-dispersed, equi-dispersed or over-dispersed, that dispersion must remain constant over time. It was satisfied, for example, in our carbonated beverage demand data (discussed in section 5) when we split each 24 hour day into daytime and nighttime.

Under property 1 we have the following result.

**Theorem 3.** Let  $X_i$  be a random variable with BISA density (4) and parameters  $T_i$ ,  $\sigma_i$ , and  $\mu_i$ . Assume  $\sigma_i$  and  $\mu_i$  adhere to property 1 and the  $X_i$  are independent. Then  $\sum_{i=1}^n X_i$  has a mixture distribution whose density is given by  $f(s) = (1/2)^n f_0(s) + \sum_{j=1}^n (1/2)^n \binom{n}{j} f_j(s)$  where

$$f_j(s) = \frac{T\mu^j}{\sqrt{2\pi} \cdot 2^{j/2} \sigma^{j+1}} \exp\left(-\frac{1}{2} \left(\frac{T - \mu s}{\sigma\sqrt{s}}\right)^2\right) \cdot s^{j/2-3/2} U(j/2, /2, T^2/2\sigma^2 s)$$

and  $\mu = \sqrt{\sum_{i=1}^n \mu_i^2}$ ,  $\sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$ , and  $T = \sum_{i=1}^n \frac{T_i}{\mu_i} \mu$ .

[**Note:** For  $j=0$  we define  $U(0, /2, T^2/2\sigma^2 s) = 1$ ; for  $j=1$ , we have  $U(1/2, /2, T^2/2\sigma^2 s) = (\sigma\sqrt{2s})/T$ ]

*Proof.* The mgf for  $\sum_{i=1}^n X_i$  can be written as

$$\frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2} \left(1 - \sqrt{1 - 2t \frac{\sigma^2}{\mu^2}}\right)\right) + \frac{1}{2} \exp\left(\frac{T\mu}{\sigma^2} \left(1 - \sqrt{1 - 2t \frac{\sigma^2}{\mu^2}}\right)\right) \frac{1}{\sqrt{1 - 2t \frac{\sigma^2}{\mu^2}}} \quad (29)$$

The first term is  $1/2$  the mgf of an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$ . The second term is  $1/2$  the product of (i) the mgf of an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$  and (ii) the mgf of a gamma distribution with shape parameter  $k=1/2$  and scale parameter  $\theta = 2\sigma^2/\mu^2$  (recall that the mgf of the gamma is  $(1 - \theta t)^{-k}$  for  $|t| < 1/\theta$ ). This implies the result stated in the theorem.

$$M_{\sum X_i}(t) = (1/2)^n \prod_{i=1}^n \exp\left(\frac{T_i \mu_i}{\sigma_i^2} \left(1 - \sqrt{1 - 2t \left(\frac{\sigma_i^2}{\mu_i^2}\right)}\right)\right) \left(1 + \left(1 - 2t \frac{\sigma_i^2}{\mu_i^2}\right)^{-1/2}\right) \quad (30)$$

Since the coefficient of variation is constant, set  $v^2 = \sigma_i^2/\mu_i^2$  for all  $i$ . Then

$$M_{\sum X_i}(t) = (1/2)^n \prod_{i=1}^n \exp\left(\frac{T_i}{v^2 \mu_i} \left(1 - \sqrt{1 - 2tv^2}\right)\right) \left(1 + \left(1 - 2tv^2\right)^{-1/2}\right) \quad (31)$$

$$= \exp\left(\frac{\left(1 - \sqrt{1 - 2tv^2}\right) \sum_{i=1}^n \frac{T_i}{\mu_i}}{v^2}\right) \cdot \sum_{j=0}^n (1/2)^n \cdot \binom{n}{j} \cdot \left(1 - 2tv^2\right)^{-j/2} \quad (32)$$

$$= \sum_{j=0}^n (1/2)^n \cdot \binom{n}{j} \cdot \exp\left(\frac{\left(1 - \sqrt{1 - 2tv^2}\right) \sum_{i=1}^n \frac{T_i}{\mu_i}}{v^2}\right) \cdot \left(1 - 2tv^2\right)^{-j/2} \quad (33)$$

Define the new parameters:

$$\mu = \sqrt{\sum_{i=1}^n \mu_i^2}, \quad \sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}, \quad T = \sum_{i=1}^n \frac{T_i}{\mu_i} \mu \quad (34)$$

Observe that the new parameters satisfy  $\sigma/\mu = v$  due to property 1. Then each term in the summation of (33) (ignoring the mixture weights) takes the general form

$$\exp\left(\frac{T\mu}{\sigma^2}\left(1 - \sqrt{1 - 2tv^2}\right)\right) \cdot (1 - 2tv^2)^{-j/2}, \quad (35)$$

which is the mgf for the sum of (i) an inverse Gaussian with parameters  $\lambda = T^2/\sigma^2$  and  $\omega = T/\mu$  for  $T$ ,  $\mu$  and  $\sigma$  as defined in (34) and (ii) an independent gamma with shape parameter  $j/2$  and scale parameter  $\theta = 2\sigma^2/\mu^2 = 2v^2$ . By Theorem 2, each of these has a density  $f_j$  involving the confluent hypergeometric function of the second kind,

$$f_0(s) = \frac{T}{\sqrt{2\pi} \cdot \sigma} \exp\left(-\frac{1}{2}\left(\frac{T - \mu s}{\sigma\sqrt{s}}\right)^2\right) \cdot s^{-3/2} \quad \text{for } j = 0$$

$$f_j(s) = \frac{T\mu^j}{\sqrt{2\pi} \cdot \sigma^{j+1} \cdot 2^{j/2}} \exp\left(-\frac{1}{2}\left(\frac{T - \mu s}{\sigma\sqrt{s}}\right)^2\right) \cdot s^{j/2-3/2} U(j/2, 3/2, T^2/2\sigma^2 s) \quad \text{for } j = 1, 2, 3, \dots \quad (36)$$

The density for the sum of independent BISA random variables whose interarrival distributions have the same coefficient of variation is therefore the mixture

$$f(x) = (1/2)^n f_0(s) + \sum_{j=1}^n (1/2)^n \binom{n}{j} f_j(s). \quad (37)$$

This is a closed form representation involving confluent hypergeometric functions.  $\square$

Clearly, the shape of the final density in Theorem 3 is determined by the shape of the individual densities  $f_j(x)$ . To understand how  $T$ ,  $\mu$ , and  $\sigma$  affect the overall shape, we graphed the individual densities  $j = 0, 1, 2, 3, 4, 5$  for two numerical cases: when  $T = 500$ ,  $\mu = 20$ , and  $\sigma = 10$  (Figure 4); and when  $T = 500$ ,  $\mu = 20$ , and  $\sigma = 40$  (Figure 5). Mixing the two leftmost densities in equal proportions (.5, .5) corresponds to the BISA distribution. Mixing the three leftmost densities in proportions (.25, .50, .25) corresponds to adding two BISA distributions. Mixing the four leftmost densities in proportions (.125, .375, .375, .125) corresponds to adding three BISA distributions, etc. As one might expect, the individual densities exhibit greater spread as the coefficient of variation increases from  $v = .5$  (Figure 4) to  $v = 2$  (Figure 5). Moreover, the expected values for the  $f_j(s)$  increase with  $v$  as well. This result could be obtained directly by considering the expected value formula for a single BISA random variable (see Proposition 1).

Recall that the mgf for the tBISA introduces a factor  $e^{-t/2}$  into the expression of Theorem 1, so the mgf for the sum of  $m$  such tBISAs includes an additional factor  $e^{-mt/2}$ . This amounts to shifting all of the mixture densities in Theorem 3 to the left by  $m/2$  units. We also note that the parameters  $\mu$ ,  $\sigma$ , and  $T$  defined in Theorem 3 are not the only possible choices. These were chosen because they are easy to interpret. The proof of Theorem 3 goes through for other choices provided (i)  $(\sigma/\mu) = v$  and (ii)  $T/\mu = \sum_{i=1}^n T_i/\mu_i$ . This implies that the density in Theorem 3 is governed by two unknown parameters provided the number of terms in the sum,  $n$ , is known. Alternatively, one

could think of the parameter  $n$  as a third unknown parameter in a generalized tBISA distribution.

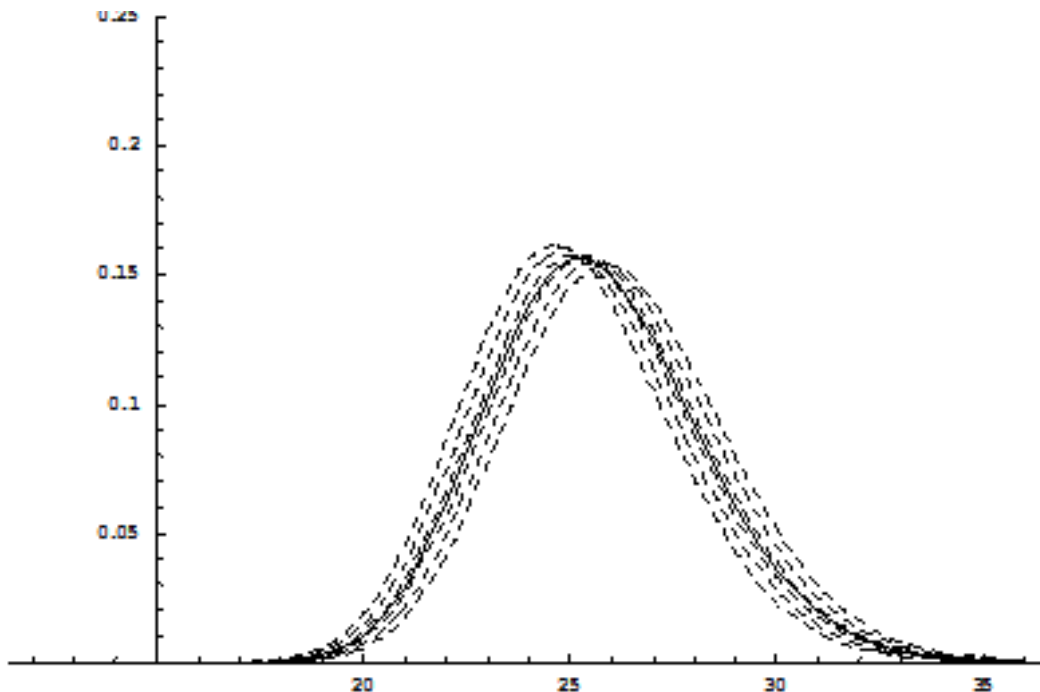


Figure 4: Mixture densities  $f_j(s)$ ,  $j = 0, 1, 2, 3, 4, 5$  (dashed lines); density of sum  $f(x)$ (solid line) for  $T = 500$ ,  $\mu = 20$ ,  $\sigma = 10$ .

## 5. APPLICATIONS

### 5.1 *An Empirical Test: Fitting the tBISA to Demand Data*

Additional tests are required to determine the suitability of the tBISA as an approximation to the distribution of count data. Our testing will focus on demand, the count of individual purchases, which is commonly analyzed in economics and business problems. Accordingly, we use the term “interpurchase” as a more descriptive synonym for “interarrival” throughout this discussion. Our first test involved fitting the tBISA to actual demand data. We obtained demand data for the best-selling carbonated beverage at a local convenience store. Three hundred and eighty-five days of data were available. We estimated the demand distribution using daily sales counts so that the input data was consistent across the candidate distributions we considered. It is interesting to note that the interpurchase distribution was not stationary over the entire day, so the assumptions under which we derived the tBISA were not, strictly speaking, met. This means the conditions for fitting the tBISA were less than ideal.

The normal and lognormal distributions are most commonly used to fit demand data in practice. We therefore fit these two distributions plus the Poisson and tBISA. All but the tBISA are easily fit using closed-form maximum likelihood estimates. The tBISA does not have closed form maximum likelihood estimates (these can be found via numerical optimization) but does have closed form



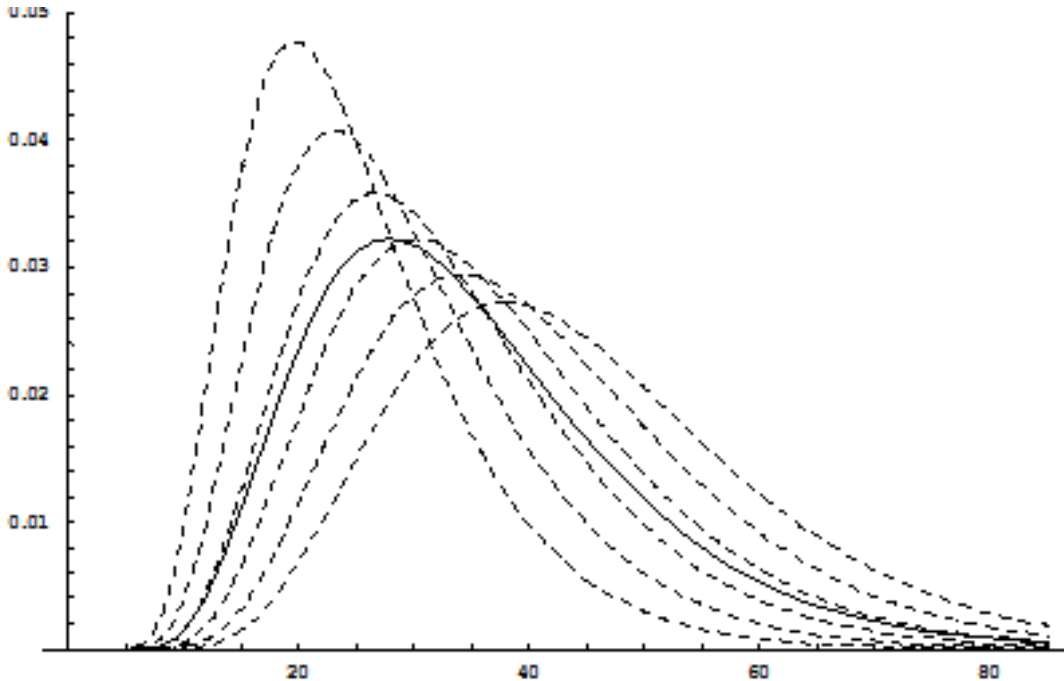


Figure 5: Mixture densities  $f_j(s)$ ,  $j = 0, 1, 2, 3, 4, 5$  (dashed lines); density of sum  $f(x)$ (solid line) for  $T = 500$ ,  $\mu = 20$ ,  $\sigma = 40$ .

method of moments estimates which we use instead (see appendix). We computed  $D_{max}$  for each distribution as compared to the empirical demand distribution. We also computed  $D_{max}$  restricted to the top decile of the empirical distribution because the upper tail of the demand distribution is typically most critical in business and economics applications. The results are summarized in Table 2, which clearly shows that the tBISA fits the carbonated beverage data better than the commonly used distributions. This is evident both for the entire distribution and for the upper tail.

	Normal	Lognormal	tBISA	Poisson
$D_{max}$	.075	.052	.042	.087
$D_{max-topdecile}$	.025	.019	.012	.068

Table 2: Goodness-of-fit for carbonated beverage demand data.

## 5.2 A Newsvendor Problem

We now consider a newsvendor application where the lognormal has been shown to fit the demand data well [5]. We first formalize how the tBISA applies to the newsvendor model.

Let the unit cost of overage be  $h$  (the per unit cost of holding excess inventory), the unit cost of shortage be  $s$  (the cost of losing a sale), and define  $\beta = s/(s + h)$ . If demand follows a tBISA distribution, the optimal newsvendor quantity  $Q$  satisfies the equation  $F_{tBISA}(Q) = \beta$  or

$$1 - \Phi([T - (Q + 1/2)\mu_I]/[\sigma_I\sqrt{Q + 1/2}]) = \beta, \quad (38)$$

where  $T$  is the time period,  $\mu_I$  is the mean interpurchase time,  $\sigma_I$  is the standard deviation of the interpurchase time, and  $\Phi$  is the cdf for the standard normal distribution. The optimal  $Q$  therefore satisfies

$$[T - (Q + 1/2)\mu_I]/[\sigma_I\sqrt{Q + 1/2}] = z_{1-\beta}, \quad (39)$$

where  $z_\beta = \Phi^{-1}(\beta)$ . Using a little algebra and the fact that  $z_{1-\beta} = -z_\beta$ , we determine that the optimal  $Q$  is

$$Q^* = T/\mu_I - 1/2 + z_\beta^2(\sigma_I/\mu_I)^2 + 1/2\sqrt{(z_\beta\sigma_I/\mu_I)^4 + 4(z_\beta\sigma_I/\mu_I)^2T/\mu_I}. \quad (40)$$

Observe that this quantity depends only on parameters of the interpurchase distribution ( $T/\mu_I$ ,  $\sigma_I/\mu_I$ ) and the same critical value one would use if the distribution of demand was assumed to be normal.

We applied the tBISA to the semiconductor demand data used by Gallego [5]. Sample statistics for weekly demand are  $\bar{x}_D = 207$  and  $s_D^2 = 210681$ . Assuming an overage cost of  $h = \$2$  and a shortage cost of  $s = \$5$ , the optimal order quantity based on the empirical distribution of demand is approximately 100 units, which leads to an optimal profit of \$69. In contrast, the optimal order quantity based on a normal distribution leads to a loss of \$291. Gallego found the lognormal distribution was a much better alternative. Using the method of moments to fit a lognormal distribution to the demand data, he determined the optimal order quantity to be 181 with a corresponding profit of \$29—a vast improvement over the normal distribution.

Distribution	Optimal Q	Optimal Profit
Normal	467	-\$291
Lognormal	181	\$29
tBISA	137	$\geq \$50.72$
Empirical	100	\$69

Table 3: Comparison of optimal inventory levels and profits

Using the same data and cost assumptions, we found the tBISA distribution produced materially better results. As Gallego did for the lognormal distribution, we used the method of moments (see appendix) to fit the tBISA. This results in estimates of  $T/\mu_I = 2.78525$  and  $\sigma_I^2/\mu_I^2 = 409.42949$  (note that these values are calculated from the demand data, not from interpurchase times). The optimal order quantity using these estimates is  $Q^* = 137$  and the optimal profit is at least \$50.72 (this follows from concavity of the profit function; we cannot be more precise without the full dataset which is no longer available). The results are summarized in Table 3.

### 5.3 Applications to Dynamic Inventory Models

The distribution of demand also plays an essential role in more complicated models of inventory/production. In practice, the true distribution is typically unknown (see [6]) so selecting a robust approximation is important. In some inventory/production applications, one must deter-

mine aggregate demand over multiple periods and so distributions that have additive properties are preferred. To determine if the tBISA holds promise in such settings, we conduct a simulation experiment using demand generated from a gamma interpurchase distribution. This interpurchase distribution was selected because it allows for over-, under-, and equi-dispersion in the corresponding count (demand) distribution and because one can compute probabilities for the exact count distribution using the incomplete gamma (see equations 7 and 8).

The distribution of aggregate demand is a fundamental concern in dynamic inventory models. In these models, one considers the short and long term costs of inventory over a multi-period horizon. Typical inventory costs include (i) the cost of ordering/purchasing inventory, (ii) the cost of holding excess inventory, and (iii) the the cost of either backlogging an item (if excess demand is backordered) or losing a sale (if excess demand is lost). In some dynamic models, it is possible to describe in compact form the optimal order/purchase decision—otherwise termed the *optimal policy*—given the period’s starting inventory position. For example, consider an infinite horizon (with future periods discounted) with proportional order costs, full backlogging of unsatisfied demand, constant revenue per unit sold, and a known lag in delivery. In this case, the optimal policy is of the base stock type, meaning that there is a critical number  $x$  (the base stock level), and the optimal decision at the start of each period is to order enough units to bring the inventory position up to  $x$ . The optimal base stock parameter can be computed via a single equation, which highlights the impact of demand  $\xi$  and the demand density  $\phi(\xi)$  in a simple way. In the case of instantaneous delivery (zero delivery lag), the optimal base stock parameter  $x$  is the solution to

$$c \cdot (1 - \alpha) + \int_0^x h'(x - \xi)\phi(\xi)d\xi - \int_x^\infty [b'(\xi - x) + r]\phi(\xi)d\xi = 0, \quad (41)$$

where we further assumed the holding and backlogging cost functions are  $h(x - \xi) = h \cdot (x - \xi)^+$  and  $b(\xi - x) = b \cdot (\xi - x)^+$ ;  $c$  is the per-unit cost of ordering inventory;  $h$  is the cost of holding excess inventory (per-unit per-period);  $b$  is the cost of backlogging excess demand (per-unit per-period);  $r$  is the constant revenue per unit sold; and  $\alpha$  is the discount factor per period (see Karlin [11]). (Note: in contrast to Karlin, we did not obtain the revenue term  $r$  in our derivation of this equation. However, this discrepancy merely changes the backorder cost to  $b + r$  instead of  $b$ .) In the case of a two-period delivery lag, the corresponding equation becomes

$$c \cdot (1 - \alpha) + \alpha^2 \int_0^\infty \int_0^\infty L'(x - \xi - \eta)\phi(\xi)\phi(\eta)d\xi d\eta = 0 \quad (42)$$

where  $L(z) = \int_0^z h(z - \xi)\phi(\xi)d\xi + \int_z^\infty b(\xi - z)\phi(\xi)d\xi$  for  $z > 0$  and  $L(z) = \int_0^\infty b(\xi - z)\phi(\xi)d\xi$  for  $z < 0$  [10]. The number of iterated integrals increases with the delivery lag, so having a closed-form expression for the distribution of the sum of random variables offers a significant computational advantage. One can then replace the iterated integral with an equivalent expression involving only

a single integral.

We considered three possible parameter combinations for gamma distributed interurchases:  $(k, \theta) = (.5, 40)$ ,  $(1, 20)$ , and  $(2, 10)$ . Each combination implies a mean interarrival of 20; standard deviations are 28.28, 20, and 14.14, respectively. The corresponding coefficients of variation are 1.414, 1, and .707, which imply over-dispersion, equi-dispersion, and under-dispersion in their respective count distributions [19].

We calculated the optimal theoretical base stock parameters using the exact count distribution for gamma interarrivals with parameters  $(k, \theta)$  using equations (41) and (42), giving us a theoretical benchmark. To fit the alternative demand distributions, we simulated interpurchase streams from a gamma distribution with parameters  $(k, \theta)$  and tabulated the corresponding counts for each period of length  $T = 500$ . The count streams were used to fit the normal, lognormal, and tBISA distributions. Normal and lognormal distributions were fit using maximum likelihood estimates; the tBISA distribution was fit using the method of moments (see the appendix for details). The tBISA was also fit directly to interpurchase streams. tBISA distributions fit to count data were labeled tBISA-C, those fit to interpurchase data were labeled tBISA-I. All fitted distributions (normal, lognormal, tBISA-C and tBISA-I) were then used to calculate optimal base stock levels conditional on the choice of distribution. This process is identical to the *classical approach with errors in distribution* discussed in [6] and it fairly reflects the type of approximation one must make in selecting a distribution for real-world inventory applications. Because the number of periods  $n$  is an important practical issue in fitting a distribution to data, we considered five different values,  $n = 10, 25, 50, 100, 200$ . In total, three parameter combinations for  $(k, \theta)$  and five different choices for  $n$  resulted in 15 experimental cells. We simulated 50 interpurchase streams (and tabulated the corresponding count streams) for each cell.

Order, revenue, holding, and backorder (penalty) costs were taken to be  $c = 50$ ,  $r = 80$ ,  $h = 5$ , and  $b = 15$  respectively. The discount factor was  $\alpha = 1.0$ . In the case of zero delivery lag, the parameters  $(k, \theta) = (.5, 40)$ ,  $(1, 20)$ , and  $(2, 10)$  corresponded to optimal theoretical base stock levels of 38, 33, and 31 respectively. Mean absolute deviations from these theoretical base stock levels for the four fitted distributions are given in Table 4.

Observe that the tBISA-C and the tBISA-I both outperform the lognormal distribution (lower mean absolute deviation in 14 out of 15 pairwise comparisons). The tBISA-I is superior in smaller samples, especially in cases of over- or under-dispersion. This result is probably due to higher precision in the tBISA-I's estimation from interpurchase data with approximately  $T/\mu$  times as many observations as the count data stream. The tBISA-C is superior in larger samples, especially in cases of over- and equi-dispersion. Overall, our results demonstrate that the tBISA, fit to count data or interpurchase data, is a competitive alternative to the lognormal.

The normal distribution was consistently better than all competing distributions in the case of equi-dispersion. This is not so surprising given that  $(k, \theta) = (1, 20)$  corresponds to exponential interurchases and implies Poisson distributed counts; the normal distribution is known to approximate the Poisson well. Indeed, in their inventory analysis, [6] found almost no difference between

$n$	$k$	$\theta$	Normal	Lognormal	tBISA-C	tBISA-I
10	0.5	40	4.02	4.74	4.1	2.46
10	1	20	1.88	2.16	1.86	1.52
10	2	10	1.22	1.34	1.3	1.04
25	0.5	40	2.56	2.1	2.14	1.46
25	1	20	0.94	1.44	1.12	1.36
25	2	10	1.08	1.14	1.1	0.7
50	0.5	40	2.04	1.62	1.44	1.24
50	1	20	0.7	1.38	0.84	1.22
50	2	10	0.64	0.76	0.64	0.4
100	0.5	40	1.68	1.54	0.88	1.12
100	1	20	0.46	1.08	0.68	1.12
100	2	10	0.42	0.56	0.54	0.22
200	0.5	40	1.56	1.36	0.74	1.16
200	1	20	0.28	1.28	0.78	1.16
200	2	10	0.3	0.38	0.34	0.12

Table 4: Mean absolute deviations, zero period lag (averaged over 50 simulated data streams)

the inventory policy obtained from fitting a normal distribution to simulated Poisson demand data and the inventory policy obtained from fitting a Poisson distribution to the same data. The normal does not fare well against the tBISA in cases of under- and over-dispersed count data, however. This gap is particularly large compared to the tBISA-I in cases of over-dispersion or small  $n$ .

To determine whether these patterns persist in more complex dynamic inventory formulations and to demonstrate the additive properties of the tBISA, we repeated the experiment while incorporating a two-period delivery lag. This required solving (42). The simulated data streams were identical to those used in the experiment with zero delivery lag. The optimal theoretical base stock parameters corresponding to  $(k, \theta) = (.5, 40)$ ,  $(1, 20)$ , and  $(2, 10)$  are 85, 81, and 78, respectively. The mean absolute deviations are given in Table 5.

The dominance of the tBISA-C over the lognormal is somewhat diminished, but the case for the tBISA-I over the lognormal remains fairly strong. It is superior to the lognormal in 12 out of 15 cells, and it is very close in the remaining three. Recall also that there is no closed-form expression for the sum of lognormal random variables, so we must use iterated integrals in our calculations for this distribution (see equation 42). The tBISA-I again outperforms the normal in all cases of over- and under-dispersion; it is slightly outperformed by the normal in three cases of equi-dispersion. These results suggest the tBISA is a good alternative to both the normal and lognormal in many practical situations.

## 6. CONCLUSIONS AND EXTENSIONS

We have proposed the tBISA as a candidate distribution for the modeling of count data. The tBISA distribution has several appealing properties: it can be estimated from count data or from

$N$	$k$	$\theta$	Normal	Lognormal	tBISA-C	tBISA-I
10	0.5	40	5.9	6	5.88	4.92
10	1	20	3.4	3.46	3.66	3.34
10	2	10	2.82	2.7	2.62	2.68
25	0.5	40	4.14	3.94	4.72	3.54
25	1	20	1.8	1.82	2.02	1.88
25	2	10	2.3	2.26	2.04	1.94
50	0.5	40	2.9	2.96	3.86	2.38
50	1	20	1.44	1.44	1.54	1.42
50	2	10	1.94	1.88	1.44	1.32
100	0.5	40	2.04	1.92	2.72	1.78
100	1	20	0.88	0.92	1.24	1.04
100	2	10	1.6	1.56	1.3	0.92
200	0.5	40	1.72	1.6	2.4	1.3
200	1	20	0.82	0.84	0.96	0.9
200	2	10	1.6	1.54	0.96	0.56

Table 5: Mean absolute deviations, two period lag (averaged over 50 simulated streams)

interarrival data (the latter being particularly appropriate when the count data are censored); it can be used in situations with limited count data observations; it can be adjusted to different time intervals without collecting additional data; and it has analytic properties that make it tractable in many applications, particularly those involving the cdf or sums of random variables.

One limitation of our development is that we consider only cases where the count is incremented one unit at a time. For cases in which the count increment exceeds one (group arrivals or multiple counts per arrival) a significant proportion of the time, our analysis must be modified. In the simplest case of group arrivals, suppose arrivals are either “singles” (a single arrival) or “pairs” (two simultaneous arrivals). One could then measure the interarrival time between singles (ignoring all pairs) and the interarrival time between pairs (ignoring all singles). Let  $\mu_1$  and  $\sigma_1$  represent the mean and standard deviation of the interarrival distribution for singles, and let  $X$  be the number of singles arriving during  $[0, T]$ . Let  $\mu_2$  and  $\sigma_2$  represent the mean and standard deviation of the interarrival distribution for pairs, and let  $Y$  be the number of pairs arriving during  $[0, T]$ . Then the count would be  $X+2Y$  units. If  $Y$  is a BISA random variable then so is  $2Y$  (see [9]). However,  $2Y$  has a BISA distribution (4) with parameters  $\mu = \mu_2/2$ , and  $\sigma = \sigma_2/\sqrt{2}$ . Thus, to calculate the density for  $X+2Y$  using Theorem 3, we would need property 1 to hold, which in this case means  $v = \sigma_1/\mu_1 = \sqrt{2}\sigma_2/\mu_2$ . This condition is less intuitively appealing than that for  $X + Y$ , and so Theorem 3 may be more applicable to systems in which each arrival increments the count by one unit. Using a similar approach, the model can be extended to cases in which an arrival generates multiple counts.

There are many possible extensions of this research. One would be to introduce covariates in the estimation of tBISA parameters. A second would be to address the small sample properties of the tBISA as compared to the lognormal, normal, Poisson, and other candidate distributions. A

third extension, to address nonstationarity in the interarrival distribution, would be to partition the interarrivals into distinct groups or segments. For example, interarrivals times during different parts of the day (e.g., daytime versus nighttime), different days-of-the-week (e.g., weekday versus weekend), or different seasons of the year could be partitioned and their respective count distributions fit separately. Alternatively, interarrival times could be separated based on a criterion that does not depend on time, e.g., cash customers versus credit customers (here we would measure the time between cash purchases and the time between credit purchases). In each case, the total demand would be the sum of counts for the different groups or segments. In other applications, the number of segments might not be known, in which case  $n$ , the number of segments, becomes a free parameter in Theorem 3.

## References

- [1] ABRAMOWITZ, M. AND STEGUN, I. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. US Dept. of Commerce, U.S. G.P.O., Washington, D.C.
- [2] BIRNBAUM, Z. AND SAUNDERS, S. (1969). A new family of life distributions. *Journal of Applied Probability* **6**, 319–327.
- [3] BRADLEY, D. AND GUPTA, R. (2004). On the distribution of the sum of  $n$  non-identically distributed uniform random variables. *Annals of the Institute of Statistical Mathematics* **54**, 689–700.
- [4] DESMOND, A. (1986). On the relationship between two fatigue-life models. *IEEE Transactions on Reliability* **R-35**, 167–169.
- [5] GALLEGO, G. (1995). Lecture notes in production management. Department of Industrial Engineering and Operations Research, Columbia University.
- [6] IYER, A. AND SCHRAGE, L. (1992). Analysis of the deterministic (s, s) inventory problem. *Management Science* **38**, 1299–1313.
- [7] JOHNSON, N. L. AND KOTZ, S. (1970). *Continuous Univariate Distributions-Vol. 2, First Edition*. John Wiley and Sons: New York.
- [8] JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions-Vol. 1, Second Edition*. John Wiley and Sons: New York.
- [9] JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1995). *Continuous Univariate Distributions-Vol. 2, Second Edition*. John Wiley and Sons: New York.
- [10] KARLIN, S. (1958). *Inventory Models of the Arrow-Harris-Marschak Type with Time Lag*. Stanford University Press, Stanford, CA. ch. 10, pp. 155–178.

- [11] KARLIN, S. (1960). Dynamic inventory policy with varying stochastic demands. *Management Science* **6**, 231–258.
- [12] KUMMER, E. E. (1836). Über die hypergeometrische reihe. *J. Reine Angew. Math.* **15**, 39–83, 127–172.
- [13] MCSHANE, B., ADRIAN, M., BRADLOW, E. T. AND FADER, P. S. (July 2008). Count models based on weibull interarrival times. *Journal of Business & Economic Statistics* **26**, 369–378(10).
- [14] NADARAJAH, S. AND KOTZ, S. (2008). Exact distribution of visiting rates. *Management Science* **54**, 860–862.
- [15] RIECK, J. (1999). A moment generating function with application to the birnbaum-saunders distribution. *Communications in Statistics, Theory and Methodology* **28**, 2213–2222.
- [16] RIECK, J. AND NEDELMAN, J. (1991). A log-linear model for the birnbaum-saunders distribution. *Technometrics* **33**, 51–60.
- [17] ROSS, S. (1983). *Stochastic Processes*. Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York.
- [18] WIMMER, G. AND ALTMANN, G. (1999). *Thesaurus of Univariate Discrete Probability Distributions*. STAMM Verlag: Essen, Germany.
- [19] WINKELMANN, R. (1995). Duration dependence and dispersion in count-data models. *Journal of Business and Economic Statistics* **13**, 467–474.

## A. APPENDIX

### *Fitting the BISA to Count Data Using the Method of Moments*

Let  $\bar{x}_D$  be the sample mean for period counts and let  $s_D^2 = \sum_i (x_i - \bar{x})^2/n$  be the sample variance (the denominator  $n$  is needed for the method of moments). Equating these sample moments with those of the BISA results in the equations

$$\begin{aligned}\bar{x}_D &= \frac{T}{\mu_I} + \frac{\sigma_I^2}{2\mu_I^2} - 1/2 \\ s_D^2 &= \frac{5\sigma_I^4}{4\mu_I^4} + \frac{T}{\mu_I} \cdot \frac{\sigma_I^2}{\mu_I^2}\end{aligned}\tag{43}$$

from which one obtains solutions

$$\frac{\sigma_I^2}{\mu_I^2} = \frac{2(\bar{x}_D + 1/2)}{3} \left( \sqrt{1 + \frac{3s_D^2}{(\bar{x}_D + 1/2)^2}} - 1 \right)$$



$$\frac{T}{\mu_I} = \frac{(\bar{x}_D + 1/2)}{3} \left( 4 - \sqrt{1 + 3 \frac{s_D^2}{(\bar{x}_D + 1/2)^2}} \right) \quad (44)$$

A limitation of this method is that it fails if  $s_D^2/(\bar{x}_D + 1/2)^2 \geq 5$ , thus a different estimation method (e.g., maximum likelihood) would be required. Fortunately, this violation rarely occurs in practice, and so the method of moments should be broadly applicable.