BETA APPROXIMATIONS FOR BRIDGE SAMPLING

Paul Glasserman
Kyoung-Kuk Kim

Graduate School of Business
3022 Broadway, Uris Hall
Columbia University
New York, NY 10027, U.S.A.

ABSTRACT

We consider the problem of simulating X conditional on the value of $X + Y$, when X and Y are independent positive random variables. We propose approximate methods for sampling $(X|X + Y)$ by approximating the fraction $(X/z|X + Y = z)$ with a beta random variable. We discuss applications to Lévy processes and infinitely divisible distributions, and we report numerical tests for Poisson processes, tempered stable processes, and the Heston stochastic volatility model.

1 INTRODUCTION

In its simplest form, the problem we consider reduces to sampling from $(X|X + Y)$, the conditional distribution of X given $X + Y$, for independent, positive random variables X and Y. This becomes practically interesting if we have an efficient simulation method for $X + Y$, but not for X or Y.

This problem arises naturally in, for example, the simulation of financial models driven by Lévy processes. (See Cont and Tankov 2004 for background on such models.) A Lévy process X_t has $X_0 = 0$ and stationary, independent increments: the increment $X_t - X_s$, $t > s$, is independent of X_s, and its distribution depends on t and s only through their difference, $t - s$. To simulate a Lévy process at a fixed set of times, $X_{t_1}, X_{t_2}, \dots, X_{t_m}$, one might simulate the increments

$$X_{t_1}, \; X_{t_2} - X_{t_1}, \; \dots, \; X_{t_m} - X_{t_{m-1}},$$

and then take their cumulative sums. Alternatively, one might start by simulating X_{t_m} and then recursively fill in intermediate values by sampling from

$$(X_{t_1}|X_{t_m}), \; (X_{t_2}|X_{t_1},X_{t_m}), \; \dots, \; (X_{t_{m-1}}|X_{t_{m-2}},X_{t_m}); \quad (1)$$

indeed, the intermediate values could be generated in any order, recursively bisecting the time interval, for example.

Conditioning a Lévy process on its endpoints produces a Lévy bridge, generalizing the notion of a Brownian bridge.

Generating a value from a Lévy bridge fits in the framework we consider. For example, writing $X = X_{t_1}$ and $Y = X_{t_m} - X_{t_1}$ puts the problem of simulating $(X_{t_1}|X_{t_m})$ in the form of sampling from $(X|X + Y)$, with X and Y independent. The same applies to simulating $(X_t|X_s,X_u)$, $s < t < u$, if we write this as

$$X_s + (X_t - X_s|(X_t - X_s) + (X_u - X_t)),$$

and use the fact that the increments are stationary and independent. We focus on the case of positive X and Y; in the Lévy context, this corresponds to a process with positive increments, called a subordinator.

There are several reasons one might use bridge sampling to simulate a Lévy process. One might start by simulating a rough approximation to the path by taking large time steps and then generate additional intermediate values where needed — for example, where the process is close to a barrier in pricing a barrier option. Bridge sampling is also useful in implementing quasi-Monte Carlo methods; this is a familiar technique in the Brownian setting and has been applied to the gamma process in Avramidis, L'Ecuyer, and Tremblay (2003). Bridge sampling is also potentially useful if, as often happens, the increment distributions are known only through their characteristic functions. In this case, one might use numerical transform inversion to simulate X_{t_m} and then use potentially faster methods to generate $X_{t_1}, \dots, X_{t_{m-1}}$ conditional on X_{t_m}.

A second application of the general problem we consider arises in simulating the Heston (1993) stochastic volatility model. Building on work of Broadie and Kaya (2006), we show in Glasserman and Kim (2008) that the Heston model can be simulated exactly by sampling from certain infinitely divisible distributions. Recall that a distribution with characteristic function ϕ is infinitely divisible if ϕ^δ is the characteristic function of a probability distribution for all rational $0 < \delta < 1$. In our application to the Heston model, this extends to all real δ, and we need to sample from the distribution associated with ϕ^δ for any value of δ — the relevant value changes as the simulation evolves.

However, these characteristic functions do not admit closed-form inverses and carrying out numerical inversion for each δ is potentially time-consuming. Instead, we can numerically invert ϕ and tabulate the distribution; sampling from ϕ^δ then reduces to bridge sampling: let X have characteristic function ϕ^δ, let Y have characteristic function $\phi^{1-\delta}$, with X and Y independent; then $X+Y$ has characteristic function ϕ, so we can generate $X+Y$ from the tabulated values of ϕ and then try to generate $(X|X+Y)$ to sample from ϕ^δ. We return to this application in Section 4.

The starting point for the approximations we consider is a standard property of gamma random variables. Let X and Y be independent with gamma distributions $X \sim \Gamma(\alpha, \theta)$ and $Y \sim \Gamma(\beta, \theta)$, meaning that they have shape parameters α and β and a common scale parameter θ. Then

$$\frac{X}{X+Y}$$

has a beta distribution with parameters α and β and is *independent* of $X+Y$. Thus, in this setting, our generic problem is solved by sampling U from the indicated beta distribution and then setting

$$(X|X+Y=z) = zU. \tag{2}$$

This idea is used for the simulation of a gamma bridge with an application in Avramidis et al. (2003).

However, a characterization result of Lukacs (1955) implies that gamma random variables provide the *only* setting in which such a procedure is exact: Lucacks shows that for two non-degenerate, independent positive random variables X and Y, the associated random variables $Z = X+Y$ and $V = Y/X$ are independent if and only if X and Y have gamma distributions with the same scale parameter, in which case $1/(1+V)$ has a beta distribution.

Nevertheless, for any positive X and Y, the ratio $X/(X+Y)$ takes values in $[0,1]$ and, as is often noted (e.g., in Springer 1979) beta distributions provide a flexible family to approximate general distributions on the unit interval. We therefore investigate the use of beta random variables to approximate $(X/z|X+Y=z)$. For gamma random variables, the distribution of this ratio is independent of z; but since this property is limited to the gamma case, we allow our choice of beta parameters to depend on z in testing approximations.

The rest of this article is organized as follows. In Section 2, we give two results that provide some insight into the effectiveness of beta approximations. Section 3 deals with the question of choosing the parameters of the beta distribution used in the approximation. Section 4 reports numerical results for several examples, including the application to the Heston model. Section 5 concludes the paper.

2 ERROR ANALYSIS

In this section, we focus on the simplest case in which U in (2) has a beta distribution that does not depend on z. Thus, we generate $X+Y$ and then generate U independently and make the approximation

$$X \approx U \cdot (X+Y) \tag{3}$$

to generate X. We stress that this involves two approximations: approximating the distribution of $X/(X+Y)$ with a beta distribution and treating this ratio as though it were independent of $X+Y$.

Our first result indicates that, as one would expect, if the positive independent random variables X and Y are close to gamma random variables with the same scale parameter, then (3) is close to exact. We make the notion of "close" precise through the ratio of densities.

Proposition 2.1 *Suppose that X and Y have densities f_X and f_Y and let*

$$h_X = f_X/g_X, \quad h_Y = f_Y/g_Y,$$

where g_X and g_Y are the densities of the gamma distributions $\Gamma(\alpha, 1)$ and $\Gamma(\beta, 1)$, for some α and β. Let U have a beta distribution with parameters α and β. Then, at any point x, the ratio of the density of $(X|X+Y=z)$ to the density of zU is, with $u_0 = x/z$, given by

$$R(u_0, z) = \frac{h_X(u_0 z)h_Y((1-u_0)z)}{\mathbb{E}[h_X(Uz)h_Y((1-U)z)]}.$$

Proof The density of $(X|X+Y=z)$ at a point x is

$$\frac{f_X(x)f_Y(z-x)}{f_{X+Y}(z)}.$$

Recall that the probability density function (PDF) of $\Gamma(k,1)$ is $x^{k-1}e^{-x}/\Gamma(k)$. So,

$$
\begin{aligned}
f_{X+Y}&(z) \\
&= \int_0^z f_X(x)f_Y(z-x)dx \\
&= \int_0^z g_X(x)g_Y(z-x)h_X(x)h_Y(z-x)dx \\
&= \frac{e^{-z}z^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du \\
&= g_{X+Y}(z)\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&\quad \times \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du
\end{aligned}
$$

and thus, with $u_0 = x/z$,

$$\frac{f_X(x)f_Y(z-x)}{f_{X+Y}(z)}$$

$$= \frac{e^{-z}}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(z-x)^{\beta-1}h_X(x)h_Y(z-x)$$

$$\div \left\{ g_{X+Y}(z)\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \right.$$

$$\left. \times \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du \right\}$$

$$= e^{-z}z^{\alpha+\beta-2}u_0^{\alpha-1}(1-u_0)^{\beta-1}h_X(u_0 z)h_Y((1-u_0)z)$$

$$\div \left\{ e^{-z}z^{\alpha+\beta-1} \right.$$

$$\left. \times \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du \right\}$$

$$= \frac{u_0^{\alpha-1}(1-u_0)^{\beta-1}h_X(u_0 z)h_Y((1-u_0)z)}{z\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du}.$$

We compare this with the density of zU, which is given by $f_U(u_0)/z$, where $U \sim Beta(\alpha,\beta)$, i.e.,

$$\frac{u_0^{\alpha-1}(1-u_0)^{\beta-1}}{z\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{u_0^{\alpha-1}(1-u_0)^{\beta-1}}{z}.$$

The ratio $R(u_0,z)$ is thus given by

$$R(u_0,z)$$

$$= \frac{u_0^{\alpha-1}(1-u_0)^{\beta-1}h_X(u_0 z)h_Y((1-u_0)z)}{z\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du}$$

$$\times \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \cdot \frac{z}{u_0^{\alpha-1}(1-u_0)^{\beta-1}}$$

$$= \frac{h_X(u_0 z)h_Y((1-u_0)z)}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}h_X(uz)h_Y((1-u)z)du}\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$= \frac{h_X(u_0 z)h_Y((1-u_0)z)}{\mathbb{E}[h_X(Uz)h_Y((1-U)z)]}. \qquad \blacksquare$$

If h_X and h_Y are close to 1, then $R(u_0,z)$ is close to 1. Also,

$$\int_0^1 f_U(u_0)R(u_0,z)du_0 = 1,$$

and this leads to the conclusion that for each z, $R(u_0,z) = 1$ for some u_0, i.e., the beta approximation is exact for some x, for every z.

Consider the application of the beta approximation to the problem of simulating a positive Lévy process (a subordinator), as discussed in the introduction. We may generate X_{t_0} (perhaps through numerical transform inversion) and then approximate X_s for some $s < t_0$ by setting $X_s = UX_{t_0}$,

with U having a beta distribution. Without loss of generality, we may take $t_0 = 1$. In the following result, we investigate how the beta approximation UX_1 to X_s works as s approaches 1. For reasons that will be illustrated in Section 3, we set $\mathbb{E}(U|X_1) = sX_1$. Denote the CDF (PDF) of X_s by F_s (f_s). Let G denote the CDF of UX_1.

Proposition 2.2 *Suppose there exist $x_0 > 0$ and $\varepsilon > 0$ such that*

$$c(x_0,\varepsilon) \equiv \sup_{s\in(1-\varepsilon,1]}\sup_{x\geq x_0} f_s(x) < \infty.$$

Then, for some $x' \geq x_0$, we have

$$\sup_{x\geq x'}|G(x) - F_s(x)| = O\left(\sqrt{1-s}\right).$$

Proof The Lévy metric $\rho(F,G)$ for two CDFs,

$$\rho(F,G)$$
$$= \inf\{\varepsilon : F(x-\varepsilon) - \varepsilon \leq G(x) \leq F(x+\varepsilon) + \varepsilon, \forall x\},$$

and the Ky Fan metric for random variables X and Y,

$$\alpha(X,Y) = \inf\{\varepsilon \geq 0 : \mathbb{P}(|X-Y| > \varepsilon) \leq \varepsilon\},$$

satisfy the following properties: if $\alpha(X,Y) = a$, then

$$\rho(F,G) \leq a, \quad \frac{a^2}{1+a} \leq \mathbb{E}\left(\frac{|X-Y|}{1+|X-Y|}\right) \leq \mathbb{E}|X-Y|.$$

See, for example, p.91 of Durrett (1996).

For $X = X_s$, $Y = UX_1$ and F, G their CDFs, we get

$$\frac{a^2}{1+a} \leq \mathbb{E}|X_s - UX_1| \leq \mathbb{E}|X_s - X_1| + \mathbb{E}|(1-U)X_1|$$

$$= \mathbb{E}X_{1-s} + \mathbb{E}(1-s)X_1 = 2(1-s)\mu.$$

This leads to

$$\rho(X_s,UX_1) \leq \alpha(X_s,UX_1)$$
$$\leq \mu(1-s) + \sqrt{\mu^2(1-s)^2 + 2\mu(1-s)}$$
$$\leq 2\mu(1-s) + \sqrt{2\mu(1-s)}.$$

If $\rho(X_s,UX_1) = \rho$, then

$$F_s(x-\rho) - \rho - F_s(x) \leq G(x) - F_s(x) \leq F_s(x+\rho) + \rho - F_s(x).$$

By the Mean Value Theorem, the right side becomes $f_s(x+h)\rho + \rho$ and the left side $-\rho f_s(x-h') - \rho$, for some positive real numbers h, $h' \in (0,\rho)$. This holds for $x \geq x' = x_0 +$

$2\mu\varepsilon + \sqrt{2\mu\varepsilon}$ with $\mu = \mathbb{E}X_1$ because, for $s \in (1-\varepsilon, 1]$,

$$
\begin{aligned}
x - \rho &\geq x_0 + 2\mu\varepsilon + \sqrt{2\mu\varepsilon} - 2\mu(1-s) - \sqrt{2\mu(1-s)} \\
&\geq x_0 > 0.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
|G(x) - F_s(x)| &\leq \max\left(f_s(x+h)+1, f_s(x-h')+1\right)\rho \\
&\leq \left(\sup_{y \geq x-\rho} f_s(y) + 1\right)\rho \leq (c(x_0, \varepsilon)+1)\rho
\end{aligned}
$$

and thus, for an interval (x', ∞) bounded away from zero, $|G(\cdot) - F_s(\cdot)|$ converges uniformly to zero at a rate of $O(\sqrt{1-s})$. ∎

If the densities f_s are uniformly bounded, one may take x' arbitrarily close to zero. The bound in this proposition is conservative in the sense that U can be any random variable in $[0,1]$ with one constraint $\mathbb{E}(U|X_1) = sX_1$.

3 CHOICE OF PARAMETERS

The main question we face in implementing a beta approximation is choosing the parameters α, β of U. In this section, we investigate this question by looking at Weibull distributions. We choose this distribution as our first example because its density resembles that of a gamma distribution, but has qualitatively different behavior.

The Weibull density with parameters k and λ is

$$
f(x; k, \lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}.
$$

For independent $X \sim Weibull(k_1, \lambda_1)$ and $Y \sim Weibull(k_2, \lambda_2)$, we want to approximate

$$
\left(\frac{X}{z} \Big| X + Y = z\right) \sim \frac{f_X(zu) f_Y(z(1-u))}{f_{X+Y}(z)} \cdot z \quad (4)
$$

by a beta random variable $U \sim Beta(\alpha, \beta)$, and then use (3). The right side of (4) can be written as

$$
C(z) u^{k_1-1}(1-u)^{k_2-1} e^{-(zu/\lambda_1)^{k_1}-(z(1-u)/\lambda_2)^{k_2}}
$$

for some function of z, $C(z)$. This suggests that we might choose $\alpha = k_1$ and $\beta = k_2$ regardless of z. Figure 1 shows the conditional densities for different z values and the PDF of $Beta(k_1, k_2)$. Clearly, the beta density gets closer to the conditional density as z decreases, but the difference becomes large as z increases. This leads to a poor approximation to the distribution of X, as shown in Figure 4.

Alternatively, we can choose α, β such that $U \cdot (X+Y)$ matches the first two moments of X, with U independent
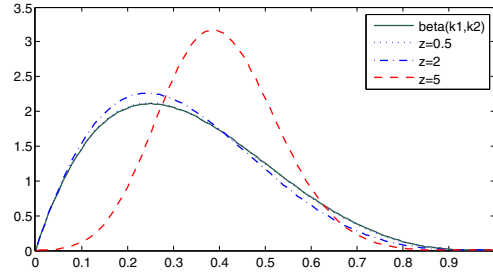

Figure 1: Beta approximation to $(X/z|X+Y = z)$ with $k_1 = 2$, $\lambda_1 = 2$, $k_2 = 4$, and $\lambda_2 = 3$

of X and Y. A simple calculation yields

$$
\begin{aligned}
\mathbb{E}U &= \frac{\mathbb{E}X}{\mathbb{E}X + \mathbb{E}Y}, \\
Var(U) &= \frac{Var(X) - Var(X+Y) \times (\mathbb{E}U)^2}{\mathbb{E}((X+Y)^2)}.
\end{aligned}
$$

For the Weibull distribution, we have $\mathbb{E}X = \lambda_1\Gamma(1+1/k_1)$, $Var(X) = \lambda_1^2\Gamma(1+2/k_1) - (\mathbb{E}X)^2$. From

$$
\mathbb{E}U = \frac{\alpha}{\alpha+\beta}, \quad Var(U) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)},
$$

we choose α, β and the results are shown in Figure 2. As we fix α, β, the conditional PDFs are quite different from the beta density for selected z values.
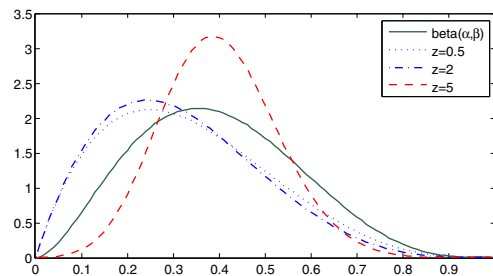

Figure 2: Beta approximation to $(X/z|X+Y = z)$ with two-moment matching

This observation leads us to consider the third and last approach, namely, matching the conditional mean and variance of $(X/z|X+Y = z)$, dropping the assumption that U is independent of $X+Y$. In our example, this requires numerical computation of

$$
\mathbb{E}\left(\frac{X}{z} \Big| X+Y = z\right), \quad Var\left(\frac{X}{z} \Big| X+Y = z\right);
$$

we choose parameters $\alpha = \alpha(z)$, $\beta = \beta(z)$ as functions of z such that $U \sim Beta(\alpha, \beta)$ matches these conditional moments. Results are shown in Figure 3. The solid lines are the true conditional densities for different z values and the dotted lines are the beta densities with corresponding moments. This choice of parameters gives a better approximation to the CDF of X, as shown in Figure 4.
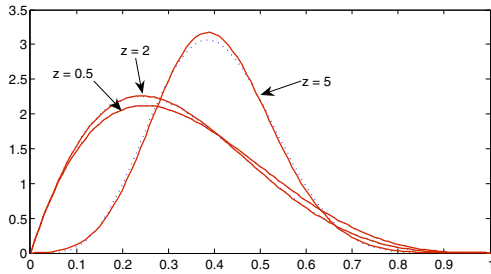

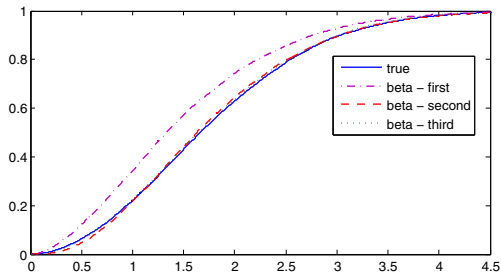Figure 3: Beta approximation to $(X/z | X + Y = z)$, conditional moment matching

Figure 4: Beta approximation $X \approx U \cdot (X + Y)$ for Weibull distributions

These experiments suggest that we choose α and β by matching the first two conditional moments of $(X/(X+Y) | X+Y)$. The availability of an explicit expression for α, β is case-dependent, but for a subordinator X_t, we can say a little more about those parameters, as in the next lemma.

Lemma 3.1   *When we apply the beta approximation to X_s for $s \in (0,t)$ given X_t by conditional two-moment matching, the parameters $\alpha(z)$ and $\beta(z)$ satisfy*

$$
\begin{aligned}
\frac{\alpha}{\alpha + \beta} &= \frac{s}{t}, \\
\frac{1}{\alpha + \beta + 1} &= \left[\frac{s}{t} \left(1 - \frac{s}{t}\right) \right]^{-1} Var\left(\frac{X_s}{z} \Big| X_t = z \right). \quad (5)
\end{aligned}
$$

These parameters result in unconditional two-moment matching: $\mathbb{E}X_s = \mathbb{E}(UX_t)$ and $Var(X_s) = Var(UX_t)$.

Proof   Suppose $s = mt/n$, for some positive integers $m < n$. Then,

$$ Y^1 + Y^2 + \cdots + Y^n \sim X_t, \quad Y^i \sim X_{t/n}. $$

Conditional on X_t, the Y^i remain identically distributed and we get $\mathbb{E}(\sum_{i=1}^m Y^i | X_t) = mX_t/n$. For any real s, we can find a sequence of rational numbers $\{q_k\}_{k \geq 1}$ converging to s/t. Because Lévy processes are stochastically continuous, $|X_s - X_{q_k t}| \to 0$ in probability, so $\mathbb{E}(X_{q_k t} | X_t) = q_k X_t$ implies $\mathbb{E}(X_s | X_t) = sX_t/t$. For the conditional variance matching $Var(X_s/X_t | X_t) = Var(U | X_t)$, we note that

$$ Var(U | X_t) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{s}{t}\left(1 - \frac{s}{t}\right)\frac{1}{\alpha+\beta+1} $$

from which (5) is immediate. To see the last statement,

$$ \mathbb{E}(UX_t) = \mathbb{E}\left(\mathbb{E}(U|X_t)X_t\right) = \mathbb{E}\left(\mathbb{E}(X_s/X_t|X_t)X_t\right) = \mathbb{E}X_s. $$

Also, from the conditional variance formula, we have

$$
\begin{aligned}
Var(UX_t) \\
= Var(\mathbb{E}(U|X_t)X_t) + \mathbb{E}(Var(U|X_t)X_t^2) \\
= Var(\mathbb{E}(X_s/X_t|X_t)X_t) + \mathbb{E}(Var(X_s/X_t|X_t)X_t^2) \\
= Var(X_s). \quad \blacksquare
\end{aligned}
$$

We cannot expect to have a closed-form expression for the conditional variance of $(X_s/X_t | X_t)$ in general. When it is unknown, we approximate the right side of (5) with some $g(s,t,z)$ and choose $\alpha(z)$, $\beta(z)$ accordingly. When we want to match the unconditional variance, the conditional variance formula gives us

$$ Var(\mathbb{E}(U|X_t)X_t) + \mathbb{E}(Var(U|X_t)X_t^2) = Var(X_s). $$

Since $Var(X_s) = sVar(X_t)/t$, we get

$$
\begin{aligned}
\mathbb{E}(Var(U|X_t)X_t^2) &= \frac{s}{t}\left(1 - \frac{s}{t}\right)\mathbb{E}\left(\frac{X_t^2}{\alpha+\beta+1}\right) \\
&= \frac{s}{t}\left(1 - \frac{s}{t}\right)Var(X_t).
\end{aligned}
$$

This leads to $\mathbb{E}(g(s,t,X_t)X_t^2) = Var(X_t)$ and this equation could be viewed as the minimal requirement for any choice of a functional form for $g(s,t,z)$ as an approximation to the true function. If we take $g(s,t,z)$ to be independent of z, then the resulting beta approximation becomes that of the second approach in the Weibull case. More specifically,

$$ g(s,t) = \frac{Var(X_t)}{\mathbb{E}X_t^2} = \frac{Var(X_t)}{Var(X_t) + (\mathbb{E}X_t)^2} = \frac{\sigma^2}{\sigma^2 + t\mu^2} $$

where $\sigma^2 = Var(X_1)$ and $\mu = \mathbb{E}X_1$. This, in turn, yields

$$\alpha = s\frac{\mu^2}{\sigma^2}, \quad \beta = (t-s)\frac{\mu^2}{\sigma^2}.$$

In our numerical examples, we show how conditional moment matching performs compared to unconditional matching via brute force numerical computation. Choosing a computationally efficient approximate $g(s,t,z)$ remains a topic for future research.

4 NUMERICAL EXAMPLES

4.1 Poisson Process

Suppose that N_t is a Poisson process with rate λ. We are interested in simulating $(N_s|N_t = n)$ for $s < t$. We know this conditional distribution is *Binomial*(n, p) with $p = s/t$. Let us define

$$X = \left(\frac{N_s}{n} | N_t = n\right) \in [0,1];$$

then it is known that

$$\mathbb{P}(X \le k/n) = I_{1-p}(n-k, k+1), \quad k = 0, \ldots, n-1,$$

where $I_x(a,b)$ is the incomplete beta function, which is the same as $\mathbb{P}(Beta(a,b) \le x)$. We want to use a beta random variable that matches the first two moments and use

$$\mathbb{P}(X \le k/n) \approx \mathbb{P}(Beta(\alpha, \beta) \le k/n + 1/2n)$$

to account for the mass of X at zero. A simple calculation yields $\alpha = (n-1)p$, $\beta = (n-1)(1-p)$.

Figure 5 shows the PDFs of X and the true binomial distribution *Binomial*$(N_t = n, p)$ with $N_t = 50, 200$ for $s = 0.1, 0.5, 0.8$ with $t = 1$ fixed. Also, Figure 6 shows the CDFs of the beta approximation to N_s and the true CDF with $\lambda = 50, 200$ for $s = 0.2, 0.5, 0.8$ with $t = 1$ fixed. In those figures, the solid lines show exact results and the dotted lines are approximations. These and other numerical experiments show that the beta approximation becomes increasingly accurate as n or λ increases for any fixed s, t.

4.2 Tempered Stable Subordinator

A tempered stable subordinator (a special case of the CGMY process; see Carr et al. 2003) is a three-parameter Lévy process X_t with Lévy density

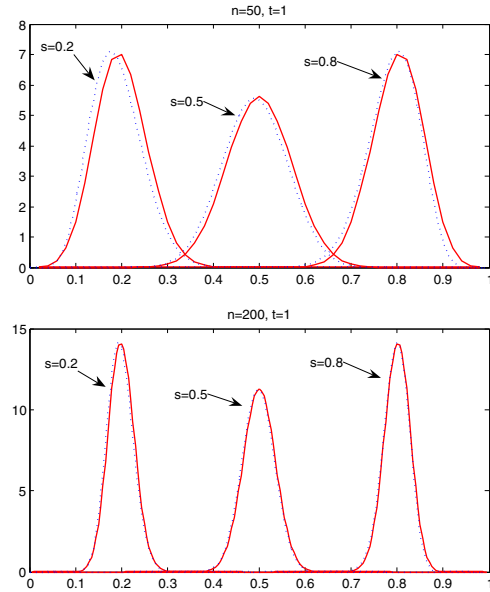$$\rho(x) = \frac{ce^{-\lambda x}}{x^{\alpha+1}}1_{x>0}, \quad c, \lambda > 0, \alpha \in [0,1).$$

Figure 5: Beta approximation to $(N_s/n|N_t = n)$ with $n = 50, 200$

This class of models has been widely studied and applied in financial engineering (see Cont and Tankov 2004 for more details). At $\alpha = 0$, the model reduces to a gamma process, and at $\alpha = 1/2$ it yields the inverse Gaussian process. A tempered stable process can be used to construct a model S_t of asset price dynamics through Brownian subordination, i.e., by setting $\log S_t = \mu t + \sigma W_{X_t}$ for appropriate μ and σ, where W is a standard Brownian motion independent of X_t. This produces the variance gamma process and the normal inverse Gaussian process when $\alpha = 0$ and 0.5, respectively. These two cases are the only instances for which a closed-form density is available. For these two cases, exact simulation schemes are known; see Chapter 6 of Cont and Tankov (2004). For general cases, we can consider using the bridge sampling method in (1).

Based on the result of Lukacs (1955), we already noted that the beta approximation is exact when $\alpha = 0$, i.e., when X_t is a gamma process, and this is the only such case. Avramidis et al. (2003) studied efficient bridge sampling algorithms for the gamma process and the variance gamma process based on this idea. For the inverse Gaussian and the normal inverse Gaussian processes, Ribeiro and Webber (2003) provide a sampling algorithm for an inverse Gaussian bridge. However, we know of no exact bridge sampling scheme for general α.

Other than $\alpha = 0$, a closed form expression for the conditional variance of $(X_s|X_t)$ for $s < t$ is not known, and the expression is quite complicated even when the density
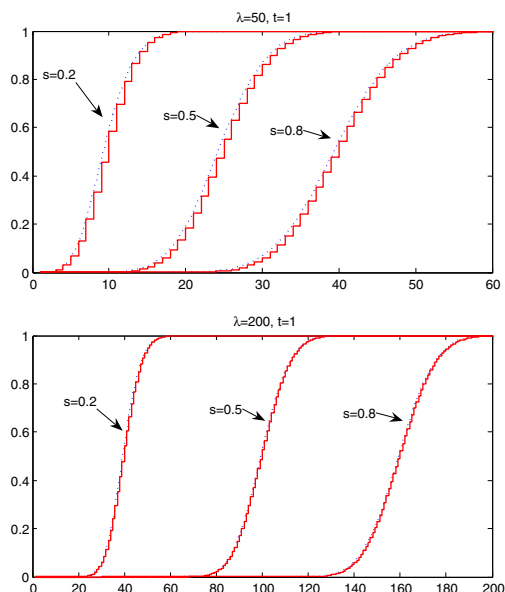
Figure 6: Beta approximation $N_s \approx UN_t$ with $\lambda = 50, 200$

function for X_t is known. When $\alpha = 0.5$,

$$f_{X_s|X_t=z}(x) \;=\; ct \cdot \frac{s}{t}\left(1 - \frac{s}{t}\right)\left(\frac{x(z-x)}{z}\right)^{-3/2}$$
$$\times \exp\left(-\pi c^2\left(\frac{s^2}{x} + \frac{(t-s)^2}{z-x} - \frac{t^2}{z}\right)\right)$$

and, after a simple change of variable, we have

$$g(s,t,z) \;=\; \int_0^1 \frac{ct}{\sqrt{z}}(u(1-u))^{-3/2}\left(u - \frac{s}{t}\right)^2$$
$$\times \exp\left(-\frac{\pi c^2 t^2}{z}\cdot\frac{(u-s/t)^2}{u(1-u)}\right)du.$$

Since a direct implementation of a numerical calculation of this $g(s,t,z)$ in the beta approximation is computationally expensive, one can adopt the second approach conducted in the Weibull example, i.e., setting $g(s,t,z)$ to be independent of z. This means that we set

$$\alpha = s\frac{\mu^2}{\sigma^2}, \quad \beta = (t-s)\frac{\mu^2}{\sigma^2}$$

where $\mu = -c\Gamma(-\alpha)\alpha\lambda^{\alpha-1}$ and $\sigma^2 = c\Gamma(-\alpha)\alpha(\alpha-1)\lambda^{\alpha-2}$. Figure 7 shows the performance of the beta approximation with $c = 2$ and $\lambda = 4$. The upper panel shows the CDFs of X_s and UX_t for different s values and fixed t and α. The lower panel shows the CDFs of X_s and UX_t for different α values with fixed s and t. The calculation of the true CDFs is conducted by implementing the algorithm

in Abate and Whitt (1992). The simulation via tabulation is done by the inverse transform method jointly with linear interpolation.
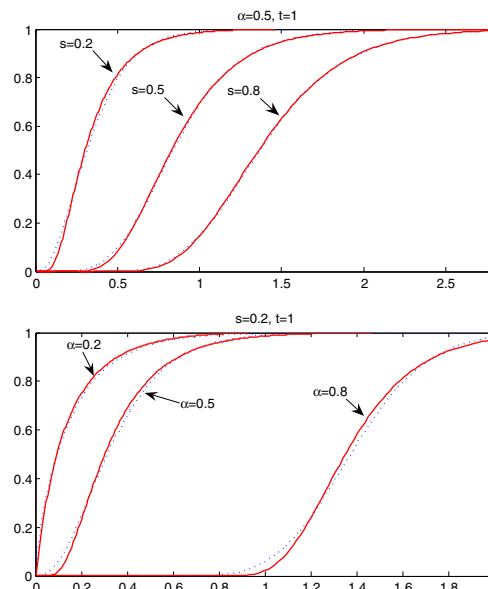

Figure 7: Beta approximation to tempered stable subordinators

The graphs show the beta approximation becoming exact as s increases to t or α decreases to 0. The first observation is already noted in Proposition 2.2 and the second one is expected because the beta approximation is exact for $\alpha = 0$. One natural question is how much more can be done beyond this $g = g(s,t)$. For example, when $\alpha = 0.5$, we can numerically compute the conditional variance of $(X_s|X_t)$ from the above formula and this can be used to determine α and β. The result is shown in Figure 8, and this conditional moment matching has a clear advantage over the unconditional moment matching. Of course, the problem is how to compute the conditional variance or at least how to approximate this conditional variance while satisfying the minimum requirement

$$\mathbb{E}(g(s,t,X_t)X_t^2) = Var(X_t).$$

Thus, the next open question is how to choose an approximate functional form instead of the exact one for general α.

4.3 Stochastic Volatility Model

Our last example is the Heston model, given by the following stochastic differential equations:

$$\frac{dS_t}{S_t} \;=\; \mu dt + \sqrt{V_t}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right),$$
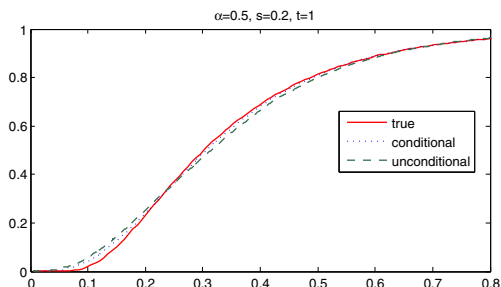$$dV_t \;=\; \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^1$$

Figure 8: Conditional moment matching vs. unconditional moment matching

Here, (W^1, W^2) is a two-dimensional standard Brownian motion and S_t stands for the asset price of interest. This model has gained wide popularity among academic researchers and practitioners. For the simulation of this model, discretization schemes such as Euler, Milstein have been proposed and used. Broadie and Kaya (2006) devised an *exact* method for simulating this process at fixed times. Their method relies on the ability to simulate

$$\left(\int_0^t V_s \, ds | V_0, V_t \right),$$

the integral of the variance process conditional on its endpoints. Broadie and Kaya (2006) derive the characteristic function for this conditional distribution and then sample through numerical transform inversion.

Their method, however, is computationally expensive mainly because it involves many computations of the modified Bessel function of the first kind. In Glasserman and Kim (2008), it is shown that

$$\left(\int_0^t V_s \, ds | V_0 = v_0, V_t = v_t \right) \sim X_1 + X_2 + \sum_{j=1}^{\eta} Z_j,$$

where the Z_j are independent copies of a random variable Z, η has a Bessel distribution with parameters $\nu = \delta/2 - 1$ and

$$z = \frac{2\kappa}{\sinh(\kappa t/2)} \sqrt{v_0 v_t},$$

and X_1, X_2, η, $Z_1, Z_2, \dots$ are independent. See Yuan and Kalbfleisch (2000) for properties of the Bessel distribution. The distribution of X_1 is determined by its Laplace transform: for $u \geq 0$,

$$\mathbb{E}e^{-uX_1} = \exp\left(\frac{(v_0 + v_t)}{\sigma^2} \left(\kappa \coth \frac{\kappa t}{2} - L \coth \frac{Lt}{2} \right) \right)$$

where $L = \sqrt{2\sigma^2 u + \kappa^2}$. Similarly, for X_2 and Z, we have

$$\mathbb{E}e^{-uX_2} = \left(\frac{L}{\kappa} \cdot \frac{\sinh \kappa t/2}{\sinh Lt/2} \right)^{\delta/2},$$

$$\mathbb{E}e^{-uZ} = \left(\frac{L}{\kappa} \cdot \frac{\sinh \kappa t/2}{\sinh Lt/2} \right)^2.$$

Note that X_2 and Z are independent of v_0, v_t which vary on every sample path in a simulation. Therefore, their distributions can be tabulated for fast simulation, as long as we fix model parameters and a time step size t. As for X_1, we observe that it is infinitely divisible with respect to $v \equiv v_0 + v_t$. Therefore, if we tabulate the distribution of X_1^* for some base number v_{base}, the problem of sampling X_1 for any $v < v_{base}$ becomes a problem of bridge sampling: X_1^* is the independent sum of X_1 with parameter v and X_1' with parameter $v_{base} - v$, and we are sampling $(X_1 | X_1 + X_1')$. We can also sample X_1 with parameter $v > v_{base}$ by summing n independent copies of X_1^*, with $n = \lfloor v/v_{base} \rfloor$, and then using bridge sampling to generate the remaining fractional piece, $\tilde{X}_1$ which has parameter $v - nv_{base}$.

In the Lévy process context, we used bridge sampling to connect the process across different points in time. In contrast, here we have infinite divisibility with respect to a "spatial" parameter v which changes with the level of the variance process.

With $v < v_{base}$, and $\tilde{X}_1$ having parameter v, we approximate $\tilde{X}_1/X_1^*$ with a beta random variable and test the performance of the resulting Heston simulation for a European call option with the following parameters:

κ	θ	σ	ρ
6.2	0.02	0.6	-0.7

and $S_0 = 100$, strike $= 100$, maturity $= 1$ year, risk free rate $= 3\%$, $v_0 = \theta$. The model parameters are set close to the estimated values in Duffie, Pan, and Singleton (2000). Also, we set $v_{base} = \theta$ because simulated V_t values move around the long term mean θ.

In our implementation, we tabulate the distribution of X_1^* using the algorithm in Abate and Whitt (1992) and the parameters α, β are chosen to match the unconditional first two moments of $\tilde{X}_1$. We compare the results with two existing methods: the exact scheme of Broadie and Kaya (2006) and the QE method of Andersen (2005), which is an efficient discretization scheme. All computations are done using programs coded in the C programming language and compiled by Microsoft Visual C++ 6.0 in the release mode. A desktop computer with Intel Pentium 4 CPU 3.20 GHz and 1.0 GB of RAM is used. The number of simulation trials is from 10K to 10240K. The simulation biases are based on 1 billion simulation trials and the results are shown in Figure 9. The tabulation time for the beta approximation is 0.64 seconds and this is not included in the production

of the figure because the tabulation needs to be done only once in the initialization of the Monte Carlo simulation with fixed κ, σ, θ and t. As illustrated in the figure, the beta approximation performs very well in the comparison of methods based on speed and accuracy.
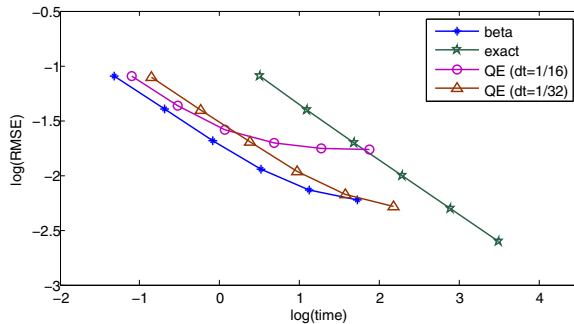

Figure 9: Beta approximation for a European call in the Heston model

5 CONCLUSION

We have proposed and tested beta approximations for bridge sampling for stochastic models such as Poisson processes, tempered stable subordinators, and the Heston stochastic volatility model. Even though the beta approximation is exact only when the random variables of interest follow gamma distributions, the approximation often works sufficiently well for practical applications. In choosing parameters for the beta distribution used in the approximation, one can use unconditional moment matching, but conditional moment matching exhibits better performance. However, conditional moment matching is computationally more demanding. Finding faster methods for selecting effective parameter choices is a topic for further investigation.

REFERENCES

Abate, J., and W. Whitt. 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems* 10:5–88.

Andersen, L. 2005. Efficient simulation of the Heston stochastic volatility model. Working Paper, Banc of America Securities.

Avramidis, A. N., P. L'Ecuyer, and P.-A. Tremblay. 2003. Efficient simulation of gamma and variance-gamma processes. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. Chick, P. J. Sanchez, D. Ferrin, and D. J. Morrice, 319–326. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Broadie, M., and O. Kaya. 2006. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54:217–231.

Carr, P., H. Geman, D. B. Madan, and M. Yor. 2003. Stochastic volatility for Lévy processes. *Mathematical Finance* 13:345–382.

Cont, R., and P. Tankov. 2004. *Financial modelling with jump processes*. New York: Chapman & Hall/CRC.

Duffie, D., J. Pan, and K. Singleton. 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68:1343–1376.

Durrett, R. 1996. *Probability: Theory and Examples*. 2nd ed. California: Duxbury Press.

Glasserman, P., and K.-K. Kim. 2008. Gamma expansion of the Heston stochastic volatility model. Working Paper, Columbia Business School.

Heston, S. L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6:327–343.

Lukacs, E. 1955. A characterization of the gamma distribution. *Annals of Mathematical Statistics* 26:319–324.

Ribeiro, C., and N. Webber. 2003. A Monte Carlo method for the normal inverse gaussian option valuation model using an inverse gaussian bridge. Working Paper, Cass Business School, City University.

Springer, M. D. 1979. *The algebra of random variables*. New York: John Wily & Sons.

Yuan, L., and J. D. Kalbfleisch. 2000. On the Bessel distribution and related problems. *Annals of the Institute of Statistical Mathematics* 52:438–447.

AUTHOR BIOGRAPHIES

PAUL GLASSERMAN is the Jack R. Anderson Professor in the Decision, Risk, and Operations Division of Columbia Business School. His research interests include variance reduction techniques and modeling and computational issues in risk management and the pricing of derivative securities. His e-mail address is `<pg20@columbia.edu>` and his web page is `<www.paulglasserman.net>`.

KYOUNG-KUK KIM works as a quantitative analyst in New York. He received a B.S. in mathematics from Seoul National University, an M.S. in mathematics from Stanford University and a Ph.D. from Columbia Business School. His research focuses on stochastic processes, simulation and risk management. His email address is `<catenoid@gmail.com>`.