
Alternatives to GBM Prices

The developments of the previous chapters have built a financial edifice upon the normal distribution, the Gaussian, primarily through the Wiener process. But there is much evidence that the world is not Gaussian, that Gaussian is only an approximation to reality. Some evidence that it is not is seen in Fig. 6.1. These depictions should be compared with Fig. 1.1 on page 2. The earlier graph portrays a stock's price through time as being continuous. But by magnifying the time scale and viewing prices over a few months we see that stock prices are occasionally discontinuous, they can suddenly change from one value to another without going through the values in-between. This often occurs between days as seen in Fig. 6.1a. By expanding the scale to the level of hours, one sees that the prices are possibly nothing but jumps, many of them small as in panel (b).

Other evidence comes from the phenomenon of *volatility smile*. According to Black-Scholes theory, for a fixed time to maturity T , the price of all options on a given stock as a function of strike price, should be calculated using the same volatility. Namely, it should be the volatility that prevails over the time horizon of the option (or at least the average of such). But this is not what is observed. Implied volatilities for puts are greater than those for calls; and the lower the strike, the greater the volatility.

And there is more. Events that should happen only rarely or, practically speaking, never, instead occur two or three times a generation. This indicates that the Gaussian is the wrong distribution, that rare events should have a higher probability of occurring. It indicates that the tails of a more accurate distribution should have more probability mass than does the Gaussian.

In this chapter we study price processes that are not Gaussian; processes that have jumps and "heavy tails." However there is a concomitant downside, namely that options can no longer be hedged, and therefore have no unique price. The term used is *incomplete market*, and it is here that we start.

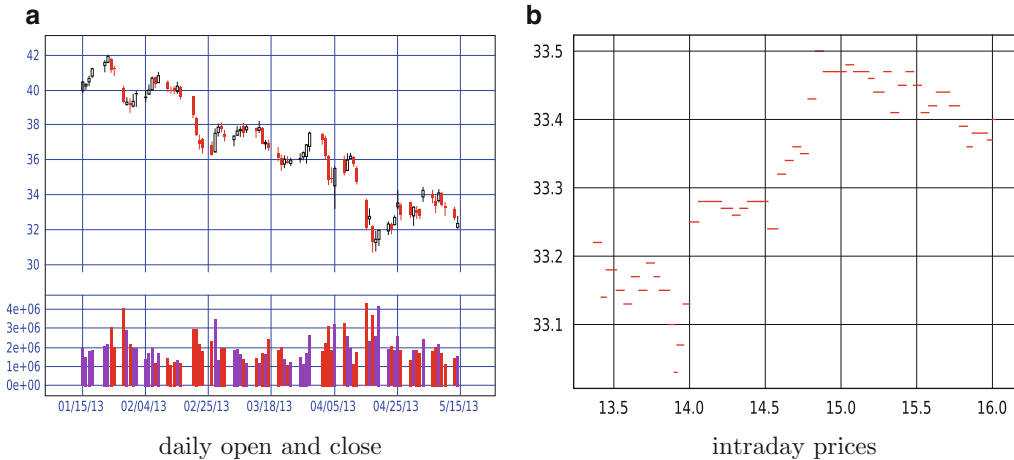


Fig. 6.1. Stock price history for SCCO over short periods of time. In (a) 4 months of prices are illustrated. Each candle's top is the day's high while the bottom is the day's low. If the price fell during the day, the candle is shown in *red*. It can be seen that frequently 1 day's candle does not overlap the next; thus the price jumped by at least the gap between the two. The *lower* part of the figure shows the volume or the number of stocks traded during the day. In (b) SCCO's intra day's prices are shown over a 3 h period during an afternoon. It is clearly seen that the price jumps almost minute by minute

6.1 Martingale Measures

Up to now we have lived in a discrete time world. Our techniques have exploited binomial lattices and GBM implemented discretely over finite increments in time. But a Wiener process is a continuous time theory. Its major accomplishment lies in showing how to define a probability or *measure* to Brownian motion paths; the object our random walks attempt to simulate. Of course, any single path has probability 0; there are, for any finite interval of time $0 \leq t \leq T$, an uncountable infinity of continuous paths X_t . But it makes sense to talk about the probability of sets of paths. For example, all paths whose Brownian particle lies between $x = 0$ and $x = 1$ when $t = T$, or in another example, all paths that were less than $x = -1$ at some time $t < T$ but finished bigger than $x = 5$ when $t = T$. And there are sets of paths for which we can assign a probability from first principles, those determined by their position at any fixed time t . We can do so because, by axiom, W_t is normally distributed with mean 0 and variance t at this time.

As we noted in Section 1.3, a Wiener process is a martingale. In this chapter we consider price processes which are not based on the Wiener process. The paths $X = \{X_t, t \geq 0\}$ of such a process must belong to some universal set Ω . And there must also be a measure or probability function defined for subsets of Ω as discussed above; the class of subsets must be closed under countable set operations (set complement, countable unions, countable intersections). And as we have seen earlier, there can be more than one probability function defined on paths. For example, for price paths there can be a historical or real-world

probability P and there can be a risk-neutral probability Q . Two probabilities are said to be *equivalent* if their sets of probability zero are identical (hence also their sets of probability one).

A stochastic process $X = \{X_t, t \geq 0\}$ in such a space Ω is a *martingale* with respect to a given measure Q if the expected future value of X at any time s is X_s ,

$$\mathbb{E}_Q(X_{s+t} | \text{the information about } X \text{ up to time } s) = X_s. \quad (6.1)$$

A measure for which the process is a martingale is called a *martingale measure*. A martingale process is something like a fair game in that a player's expected fortune at the end of the game is the same as his fortune at the start, see Chapter 7.

The importance of martingale measures is made clear in the Fundamental Theorem of Asset Pricing. In order to state it we need to define a few terms. A financial *derivative* or *contingent claim* is a security whose value depends on the value of other more basic underlying securities. Options and forwards are two examples of derivatives. A *complete market* is one for which every contingent claim has a self-financing replicating portfolio.

Theorem (Fundamental Theorem of Asset Pricing) *A discrete time pricing model has no arbitrage opportunities if and only if it has a measure for which discounted prices are a martingale. Further, the model is complete if and only if the martingale measure is unique.*

For a proof see [Rom12]. For continuous time pricing models, the theorem breaks down in that the conditions are no longer if and only if. It is still true however that the existence of a martingale measure implies there are no arbitrage opportunities and the uniqueness of the measure implies market completeness.

6.2 Incomplete Markets

In Section 3.4 we began our study of option pricing by applying the principle of no-arbitrage to a one-step price tree. Suppose now there had been three possible prices of the stock at expiry instead of just two, see Fig. 6.2.

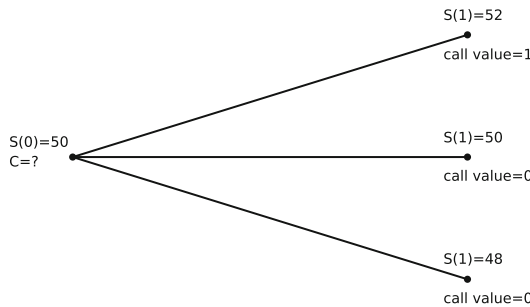


Fig. 6.2. A one step price tree with three possible prices at expiry

As before, consider a portfolio consisting of Δ shares of stock and short 1 call option struck at $K = 51$ costing C . To simplify the calculation, assume

the risk-free rate is 0; this just means money can be borrowed with no interest but still has to be paid back. From the previous analysis the number of shares to hold in the portfolio is $\Delta = 1/4$ and from (3.11) the value of the call is $12.50 - 12e^{-r_f} = 0.50$. So the initial value of the portfolio is $(\frac{1}{4})50 - 0.50 = 12$.

And in the present case this is the value of the portfolio if the stock goes up to 52 or down to 48 as before (recall $r_f = 0$). But if the final price is 50, the value of the portfolio is 12.50. Now an arbitrage is possible: borrow \$12 to set up the portfolio, pay it back if the price goes up or down, but if the price stays the same, the portfolio makes \$0.50 after retiring the loan. Thus with some positive probability, the probability of the middle branch of the tree, the portfolio makes a positive profit with no chance of losing money.

This happens because no value of Δ makes the expiry value of the portfolio equal for all three branches of the tree. So there is no one value to discount back to time 0 in order to find C .

Suppose the call is \$0.40 instead of \$0.50. Now to set up the portfolio \$12.10 will have to be borrowed. If the price goes up to 52 the stock will be worth 13, satisfying the call will cost 1 leaving only 12 to pay back the loan. Therefore the portfolio loses \$0.10. Likewise it loses the same if the price goes down to 48. If the expiry price is 50, then the stock is worth 12.50 and, after repaying the loan, the portfolio makes \$0.40. So if the call price is \$0.40 the portfolio's expectation, positive or negative, depends on the probabilities of the three outcomes.

What are those probabilities? Perhaps we can proceed by calculating the risk-neutral probability, from that find the expected payoff and discount back to get C .

Since there are three branches, the risk-neutral probability will in fact be a probability density: q_1 that the price rises, q_2 that it stays the same, and q_3 that it falls. We must have $q_1 + q_2 + q_3 = 1$ and no probability can be zero. Recall that the risk-neutral density is the one for which the expected growth of the underlying equals the risk-free rate, see page 90. Combining this expectation balance with the total density summing to 1, we have the system

$$\begin{aligned} 52q_1 + 50q_2 + 48q_3 &= 50 \\ q_1 + q_2 + q_3 &= 1. \end{aligned} \tag{6.2}$$

There is no one solution; solving in terms of q_3 we have

$$q_1 = q_3, \quad q_2 = 1 - 2q_3, \quad 0 < q_3 < \frac{1}{2}. \tag{6.3}$$

The bounds on q_3 assure that all three probabilities will be positive. To say that the expected price grows according to the risk-free rate is equivalent to saying the discounted expectation of S_t is a martingale; in this example, it will be so for any q_3 between 0 and $1/2$.

For example, choosing $q_3 = 0.4$, entails $q_1 = 0.4$ and $q_2 = 0.2$. This makes the expected call payoff equal to \$0.40, and, discounting back with $r_f = 0$, puts the price of the call at \$0.40. As analyzed above, there is no risk free profit for this value of the call.

Thus we have encountered an example of an incomplete market.

6.2.1 Pricing in an Incomplete Market

In an incomplete market there is no unique no-arbitrage price; instead there are many. In the example above, q_1 and q_2 given by (6.3) along with any choice $0 < q_3 < 1/2$ produces a martingale and with it, a no-arbitrage price for the call. The decision as to which value of q_3 to use becomes a subjective matter; a risk-averse investor would want the real-world payoff to exceed the martingale payoff.

Finally, what is the point of pricing vanilla options by a mathematical model anyway; the market already prices them. Instead, the prevailing thought is to use the price of vanilla options to determine what measure the market is using and apply those parameters in the models to calculate prices for exotic options which are only thinly traded.

The choice of measure also impacts hedging. In a complete market, continuous delta-hedging is perfect at all times and the variance of the hedge is 0. In an incomplete market, zero variance is not possible. Two possible choices are to hedge to minimize final variance or to minimize the day-to-day variance, see [Jos03] Section 15.5.

6.3 Lévy Processes

As was the case in Section 1.2, it makes sense to start with arithmetic random walks and define price processes as their geometric counterparts. In order for an arithmetic process $X = \{X_t\}$ to serve it must satisfy a very special condition, one we have used repeatedly for Brownian motion. Namely, we must be able to: (1) divide the fundamental interval, $[0, T]$ into arbitrary subintervals $\Delta t = T/n$, (2) simulate identical and independent random increments ΔX_i on each subinterval, and (3) add the increments together, $X = \sum_{i=1}^n \Delta X_i$, and get the same result statistically, that is in terms of probability density, as for any other subdivision. Such a process is said to be *infinitely divisible*.

Lévy processes are exactly those that are infinitely divisible. A Wiener process is an example of a Lévy process. Like a Wiener process, a Lévy process $L = \{L_t\}$ satisfies $L_0 = 0$ and the axioms of independent and stationary increments:

1. Every increment $L_{t+h} - L_t$ depends only on L_t and not on $L = \{L_s, 0 \leq s \leq t\}$.
2. The distribution of $L_{t+h} - L_t$ does not depend on t , it has the same distribution as L_h .¹

As we will see, a Lévy process can have jumps. By a jump we mean $\Delta L_t = \lim_{\epsilon \downarrow 0} L_{t+\epsilon} - \lim_{\epsilon \downarrow 0} L_{t-\epsilon}$. But the probability of a jump at any given value of t is 0. Note that one can always assume that a Lévy process is right continuous

¹ These conditions imply the infinite divisibility property since $L_t = L_{\frac{t}{n}} + (L_{\frac{2t}{n}} - L_{\frac{t}{n}}) + (L_{\frac{3t}{n}} - L_{\frac{2t}{n}}) + \dots + (L_t - L_{\frac{(n-1)t}{n}})$.

and has left limits at every point, $L_t = \lim_{\epsilon \downarrow 0} L_{t+\epsilon}$ and $\lim_{\epsilon \downarrow 0} L_{t-\epsilon}$ exists, the latter may be denoted as L_{t-} . Sometimes this requirement is called the *cadlag property*.² The reason for this choice is that, given a specific time in the future, say t_1 , the value of the process at t_1 cannot be predicted with complete confidence from its values at times $t < t_1$ leading up to t_1 , the process might undergo a jump at that time. If left continuity were made the choice, it would be possible to make the stated prediction.

6.3.1 The Poisson Process

Besides Wiener processes there are several known Lévy processes. The simplest is pure drift, $L_t = \mu t$. This and the Wiener process are the only two that are continuous, all others have jumps. The simplest non-continuous Lévy process is the *Poisson process* $\text{Po}(\lambda)$ (here we have put $t = 1$, because of the infinite divisibility condition the Poisson parameter for an arbitrary time t is λt). The Poisson random variable $N_t \sim \text{Po}(\lambda t)$ denotes the number of events, in our case jumps, which occur in the interval $[0, t]$. N_t is non-negative integer valued, $N_t = 0, 1, \dots$; λ is called the *intensity* parameter.

The probability density for N_t is given by

$$\Pr(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad (6.4)$$

where k is the number of jumps. The expectation, that is mean, of the Poisson random variable is λt . The variance is also λt .

The events themselves arrive at increments of time Δt according to the exponential distribution $E(\lambda)$ where λ , the same λ as above, is the *event rate*. The cumulative distribution function of $E(\lambda)$ is

$$F(t) = 1 - e^{-\lambda t}. \quad (6.5)$$

In fact N_t can be simulated by sampling $E(\lambda)$ until the time increments sum to t , a sample $N_t = k$ is returned as the greatest integer k such that

$$\sum_{i=1}^k \Delta t_i < t \quad \text{where } \Delta t_i \sim E(\lambda). \quad (6.6)$$

The Δt_i are called the *inter-arrival times*. The event arrival times themselves, t_i , are given by

$$t_i = \sum_{j=1}^i \Delta t_j, \quad i = 1, 2, \dots, k. \quad (6.7)$$

² From the French for the same phrase, ‘continue à droite et limites à gauche’. In French the spelling is càdlàg.

A sample $\Delta t_i \sim E(\lambda)$ is obtained as follows, (see (A.16))

$$\Delta t_i = \frac{-1}{\lambda} \log(1 - U) \quad \text{where } U \sim U(0, 1).^3$$

With these preliminaries in hand, the *Poisson process with drift* μ is defined by

$$L_t = \mu t + \sum_{k=1}^{N_t} J \quad (6.8)$$

where J is the fixed jump size. This is an infinitely divisible process because if $X \sim \text{Po}(\lambda_1)$ and $Y \sim \text{Po}(\lambda_2)$, then $X + Y \sim \text{Po}(\lambda_1 + \lambda_2)$, (see the Exercises). The Poisson process is always nondecreasing (if $J > 0$), that is, stays the same value or increases. In order to make the drift meaningful, one can subtract the jump size times the expected number of jumps; this gives rise to the *compensated* Poisson process with drift,

$$L_t = \mu t + \sum_{k=1}^{N_t} J - \lambda J t. \quad (6.9)$$

In Fig. 6.3a we show an instance of a compensated Poisson process. This is an *event-to-event* simulation in that time moves forward from one event to the next thus highlighting the jumps. The events are generated according to (6.7), see Algorithm 24.

Algorithm 24. Compensated Event-to-Event Simulation

```

inputs:  $t, \lambda, J$  (jumpsize)
X = 0;
simTime = 0;
plot(simTime, X);
arrivalArray = poissonArrivals( $\lambda t$ );  ▷use (6.7)
 $N_t$  = arrivalArray length;  ▷number of jumps
j = 0;  ▷ $N_t$  could be 0
for ...  ▷infinite loop
    j = j+1;  ▷update event index
    if  $j > N_t$  break out of loop
        ▷increment simTime
     $\Delta sT$  = arrivalArray[j] - simTime;
    simTime = simTime +  $\Delta sT$ ;  ▷move to next jump
     $\Delta X$  =  $-J\lambda\Delta sT$ ;  ▷pro-rated compensation
    X = X +  $\Delta X$ ;
    plot(simTime, X)  ▷before jump
    X = X+J;  ▷add in the jump

```

³ Note that $U < 1$ for uniform random number generators, but $U = 0$ is possible. Since $1 - U$ is uniform if U is, it is tempting to save an operation and use the latter; but this comes with the risk of computer overflow in the middle of a calculation.

```

    plot(simTime,X)    ▷vertical gap
endfor
ΔsT = t - simTime;   ▷since final jump
X = X - JλΔsT;
simTime = t;         ▷final time
plot(simTime,X);

```

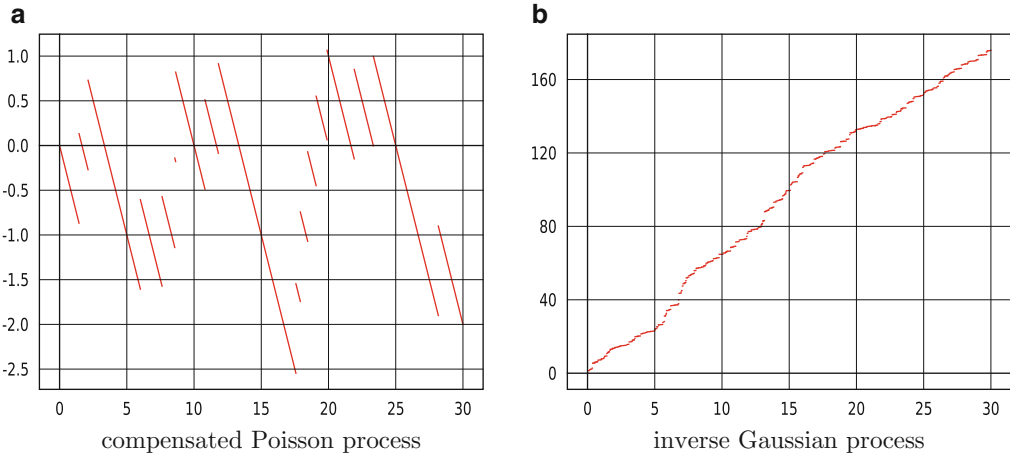


Fig. 6.3. Event to event simulations of Lévy pure jump processes

6.3.2 The Inverse Gaussian Process

The inverse Gaussian distribution has two parameters denoted by a and b . The first is a shifting parameter and has units of reciprocal time; larger a shifts the density to the right. The second is a spreading parameter, smaller b widens the density. The density itself is given by

$$f_{IG}(x; a, b) = \frac{ae^{ab}}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}\left(\frac{a^2}{x} + b^2 x\right)}, \quad x > 0. \quad (6.10)$$

Figure 6.4 shows the density function for two sets of parameters. The mean and variance of an inverse Gaussian are

$$\mu_{IG} = \frac{a}{b} \quad \text{var}_{IG} = \frac{a}{b^3}. \quad (6.11)$$

Since the density is only defined on the positive real line, the IG process is always nondecreasing; such a process is called a *subordinator*. However a process may be defined by differences of two independent inverse Gaussians to have both positive and negative jumps. We investigate this possibility in Section 6.8.

The Lévy process defined by the inverse Gaussian is a pure jump process, see Fig. 6.3b. We discuss this in the next section.

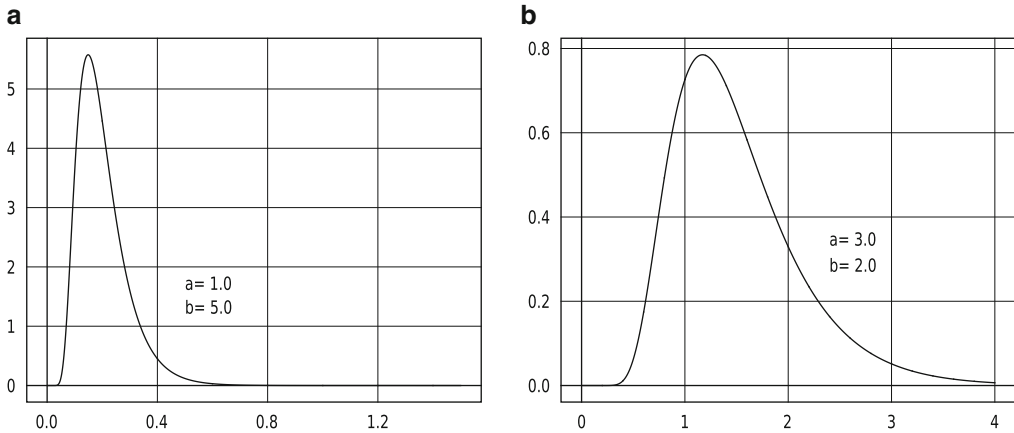


Fig. 6.4. The inverse Gaussian density for two parameter sets. Parameter a shifts the density to the right, parameter b narrows the density (for larger b). Note that the density is only defined for $x > 0$. It follows that an ARW based on this density can only move to the right

The density (6.10) gives the distribution of the end points of the process, that is at the end of the random walk, much as the normal distribution gives the end point distribution of a Brownian motion.

A random walk based on the inverse Gaussian is simulated exactly as before: the interval $[0, T]$ is subdivided into, say, n subintervals, and the update goes subinterval by subinterval, see Algorithm 25. This is a point-to-point simulation; a point-to-point path will not show jumps as they occur between the steps of the walk. Example runs of the algorithm are shown in Fig. 6.5.

Algorithm 25. Arithmetic Random Walk IG

```

inputs:  $X_0 = 0$ ,  $T$ ,  $\Delta t$ ,  $a$ ,  $b$ 
 $n = T/\Delta t$   $\triangleright$ number of iterations in time  $T$ 
for  $t = 1, \dots, n$ 
   $I \sim IG(a\Delta t, b)$   $\triangleright$ an IG sample, see A.10
   $\Delta X_t = I$ 
   $X_t = X_{t-1} + \Delta X_t$ 
endfor
 $\triangleright$ the last  $X_t$  is an outcome of  $IG(aT, b)$ 

```

6.4 Lévy Measures

Associated with each Lévy process is a unique set valued function $\nu(A)$ called the *Lévy measure*. The meaning of the measure is

$\nu(A)$ is the intensity (arrival rate) of the Poisson process for jumps of sizes in A for the path L_t , $0 \leq t \leq 1$.

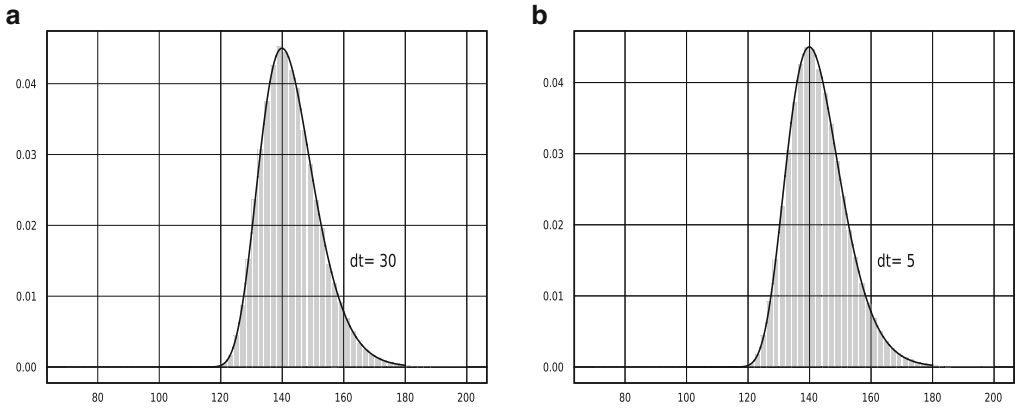


Fig. 6.5. These figures demonstrate the infinite divisibility property of the inverse Gaussian density. In (a) a $T = 30$ process is simulated in one step using Algorithm 25. In (b) the process is simulated adding six steps of size $dt = 5$. In each case the density $f_{IG}(x; 30a, b)$ is overlaid on the histogram, $a = 1, b = 0.7$

In particular if the measure is given by a density

$$\nu(dx) = h(x) dx \tag{6.12}$$

then $h(x)$ is the intensity for jumps of size x .

A Lévy measure has the same properties as a probability distribution except that it must have zero mass at the origin and its total mass may be infinite. The latter would be due to having a countable infinity of jumps of very small size. If the total mass is infinite,

$$\nu(\mathbb{R}) = \int_{-\infty}^{\infty} \nu(dx) = \infty,$$

the Lévy process has *infinite activity*. In this case there are infinitely many jumps on every interval (closed and bounded). Even if the process has infinite activity, it is always the case that it is *square summable* in the following sense

$$\int_{\mathbb{R}} \min(1, |x|^2) \nu(dx) < \infty. \tag{6.13}$$

The Lévy measure for the pure drift process and the Wiener process is null. For the Poisson process it is given by

$$\begin{aligned} \nu(A) &= \begin{cases} \lambda & \text{if } J \in A, \\ 0 & \text{otherwise} \end{cases} \\ &= \lambda \mathbb{1}_A(J). \end{aligned} \tag{6.14}$$

The Lévy measure for the inverse Gaussian process is given by a density

$$\nu_{IG}(dx) = \frac{a}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}b^2 x} dx. \tag{6.15}$$

In Fig. 6.3b we show an instance of an inverse Gaussian process. This is an event-to-event simulation and is somewhat complicated. An approximation can be made as follows. Given $\epsilon > 0$, chose positive numbers

$$\epsilon = c_0 < c_1 < \dots < c_{d+1}. \quad (6.16)$$

For each interval $[c_i, c_{i+1})$, $i = 0, \dots, d$, let $\text{Po}_i(\lambda_i)$ be an independent Poisson process with intensity given by the Lévy measure of the interval,

$$\lambda_i = \nu([c_i, c_{i+1})) = \int_{c_i}^{c_{i+1}} h(x) dx. \quad (6.17)$$

The jump size J_i should be chosen so that the variance of the Poisson process Po_i matches that part of the variance of the Lévy process for that interval,

$$J_i^2 \lambda_i = \int_{c_i}^{c_{i+1}} x^2 \nu(dx). \quad (6.18)$$

To carry out the simulation, the event times for all d processes are sampled in advance. They are then combined but with each identified to its corresponding jump size, and sorted from early to late.⁴ Then the simulation may proceed event-to-event as in Algorithm 24. When each event comes due, increment the process X_t using that event's corresponding jump size.

The above does not account for jumps of smaller size than ϵ . They may be handled, if necessary, by approximating all the small jumps by a Wiener process with drift. The parameter $\sigma(\epsilon)$ is given by

$$\sigma^2(\epsilon) = \int_0^\epsilon x^2 \nu(dx) \quad (6.19)$$

and the drift is given by

$$\mu(\epsilon) = \int_0^\epsilon x \nu(dx). \quad (6.20)$$

6.5 Jump-Diffusion Processes

By combining a Wiener process with a jump process we have what is called a jump-diffusion process. Let $F(\cdot)$ be a probability distribution (not necessarily a Lévy process) and let $J \sim F$ denote its samples. We may define a Lévy process by

$$L_t = \mu t + \sum_{k=1}^{N_t} J_k - t\lambda\mathbb{E}(J). \quad (6.21)$$

This is called a (compensated) *compound Poisson* process with drift. Just as in the compensated Poisson process of (6.9), the arrival times of the jump events

⁴ A simple technique is to maintain two arrays, one with times, t and the other with jump size, J . Now sort the times array via a rank permutation r_i so that $t_{r_1} < t_{r_2} < \dots$, see Chapter E. Then jump size J_{r_i} corresponds to time t_{r_i} .

are exponential; the only difference here is that the jumps can vary in size according to F . Since the jump sizes may vary, the compensation is determined by the average or expected jump size as shown in (6.21). The Lévy measure for a compound Poisson process is $\lambda F(dx)$.

In financial applications F is often taken to be the normal distribution. Such an example is shown in Fig. 6.6a (uncompensated in this example).

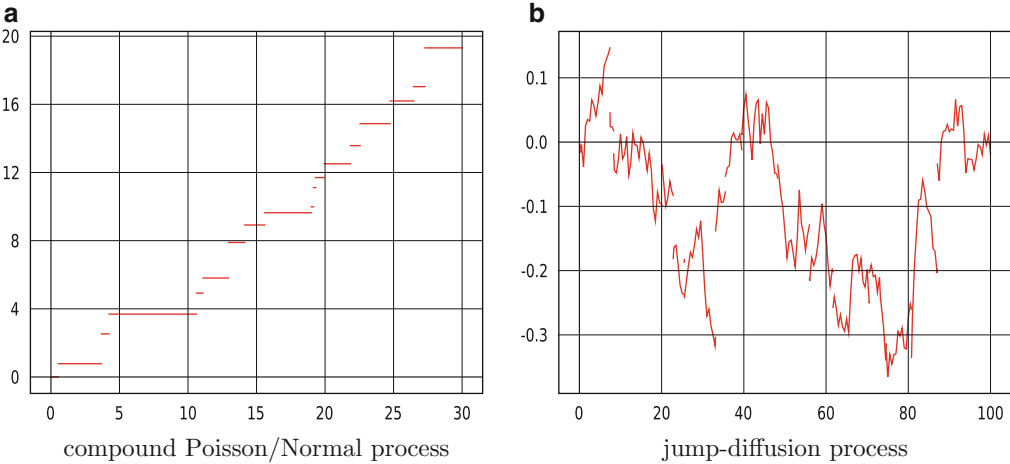


Fig. 6.6. (a) Shows an instance of an uncompensated compound Poisson process with normally distributed jump sizes. A jump-diffusion process is shown in (b), the jump sizes are random normal variates (with independent mean and variance from that of the diffusion process)

The most general Lévy process is obtained by combining all four types of processes into one: drift, diffusion (Wiener), compensated compound Poisson, and an infinite activity pure jump. The combination of the first three of these is called a *jump-diffusion* process.

$$L_t = \mu t + \sigma W_t + \left(\sum_{k=1}^{N_t} J_k - t\lambda\mathbb{E}(J_k) \right). \quad (6.22)$$

A jump-diffusion process always has finite activity. Further it is a martingale if and only if $\mu = 0$.

A jump-diffusion path is simulated from event-to-event exactly as in Algorithm 24. However, there is the additional step that the jump size be drawn from F before adding the jump to X ,

$$\begin{aligned} J &\sim F \\ X &= X + J. \end{aligned}$$

The additional difference from the cited algorithm is that, if compensation is used, the jump size to use for it is the constant expected jump size, $\mathbb{E}_F(J)$.

6.6 Application to Asset Pricing

As we learned in Chapter 1, an arithmetic random walk is an inadequate model for asset prices; a geometric walk is required. Therefore it is the log returns of the asset that must be modeled by the Lévy process

$$\frac{dS_t}{S_{t-}} = dL_t, \quad (6.23)$$

(S_{t-} is the left limit of S at t ; by the cadlag property it always exists). By conducting simulations of L_t , $0 \leq t \leq T$, as described in the previous sections, and using (6.23) we obtain a histogram approximation of the maturity price distribution and statistical information on the paths of the process leading to maturity. Of considerable importance in this regard is the *Martingale preserving Property* stating that if $(L_t)_{t \geq 0}$ is a martingale, then so is $(S_t)_{t \geq 0}$.

Generating asset prices via (6.23) is called the *stochastic exponential* method. It is the method we will use. An alternative is the *exponential-Lévy* model given by

$$S_t = S_0 e^{L_t}.$$

The two approaches are equivalent and are related by the Itô Lemma (B.11).

As was the case for diffusion increments, jump increments are taken proportional to the current asset price S . For example, $S_{\text{new}} = S_{\text{old}} J$. Then $\Delta S = S_{\text{new}} - S_{\text{old}} = S_{\text{old}}(J - 1)$. If $J > 1$ then the increment is positive. If $0 < J < 1$ then the increment is negative. And if $J < 0$ then the new price is negative; downward jumps must not exceed the current stock price. An alternative is to put $\Delta S = S_{\text{old}}(e^J - 1)$. Since $e^J > 0$ for all J , the non-negativity requirement is automatically fulfilled. By the series expansion for the exponential function, to first order, $e^J - 1 = J$. In this section we follow Merton, [Mer76], and put $\Delta S = S(J - 1)$.

Recall that, for the drift-diffusion process of Chapter 1, we were able to derive the maturity distribution analytically, see (1.18). In that case S_T is distributed lognormally. However things are not so easy for an arbitrary Lévy processes. In general the maturity distribution is the solution of the stochastic differential equation (SDE) for $\log(S_t)$ where S_t is as in (6.23). The differential of $\log(S_t)$ is given by Itô's Lemma, see (B.11), page 227. In appendix Section B.2 we solve this for the drift-diffusion process (Wiener process with drift) obtaining the lognormal as its solution. Solving it for the several processes described in the previous sections is beyond the scope of this text. Thus we will content ourselves with the simulation of the end point via small steps. In that way we generate the paths too; as we have seen, they are needed in any case for several of the exotic options.

6.6.1 Merton's Model

Besides drift-diffusion there is another process for which the end point distribution may be determined, namely for jump-diffusion processes. Let L_t be an

uncompensated jump-diffusion process and consider the product $S dL_t$ term by term. The drift and diffusion terms are $S\mu dt$ and $\sigma S dW_t$ as usual. If there is no jump at t , then the contribution from the jump term is 0. If $t = t_k$ is one of the jump event times then S jumps to SJ so the increment is $dS = SJ - S = S(J-1)$. Therefore we have

$$dS_t = S_t \mu dt + \sigma S_t dW_t + S_t \sum_{k=1}^{N_t} (J_k - 1) \delta_{t_k}(dt) \quad (6.24)$$

where the singular measure $\delta_{t_k}(dt)$ is equal to 1 if $t = t_k$ and 0 otherwise. Only one term of the sum will be non-zero for any t . Note that the value of S_t used as the multiplier for the jumps in (6.24) is the limiting value of S from the left, S_{t-} ; at an event time t_k itself, S jumps to S_{t_k} .

By an extended version of Itô's Lemma, (B.11), the differential of $\log S_t$ is given by

$$d(\log S_t) = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t + d\left(\sum_{k=1}^{N_t} \log J_k \mathbf{1}_{t_k}(t)\right). \quad (6.25)$$

The last term signifies the following: it calculates that a difference in the sum, which can not be infinitesimal, at t is $\log J_k$ if $t = t_k$ and 0 otherwise. Again, only one term is non-zero for any value of t . Integrating (6.25) we get

$$\log S_t - \log S_0 = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t + \sum_{k=1}^{N_t} \log J_k. \quad (6.26)$$

Upon exponentiation we arrive at the exponential-Lévy formulation

$$\begin{aligned} S_t &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t + \sum_{k=1}^{N_t} \log J_k} \\ &= S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \prod_{k=1}^{N_t} J_k. \end{aligned} \quad (6.27)$$

In Fig. 6.7a we show a typical path for a jump-diffusion simulation using lognormally distributed jump sizes, (b) depicts the maturity distribution. These figures were made using Algorithm 26.

Algorithm 26. Jump-Diffusion GRW, Point-to-Point Simulation

```

inputs:  $T, dt, \lambda, F(\cdot)$  (jumpsize distribution)
         $S_0, \mu, \sigma$ 
S =  $S_0$ ;
n =  $T/dt$ ;  ▷number of steps
simTime = 0;
arrivalArray = poissonArrivals( $\lambda T$ );  ▷use (6.6)
 $N_T$  = arrivalArray length;  ▷number of jumps
sDX=1;  ▷step index, point at next step
jDX = 1;  ▷jump index, point at next jump

```

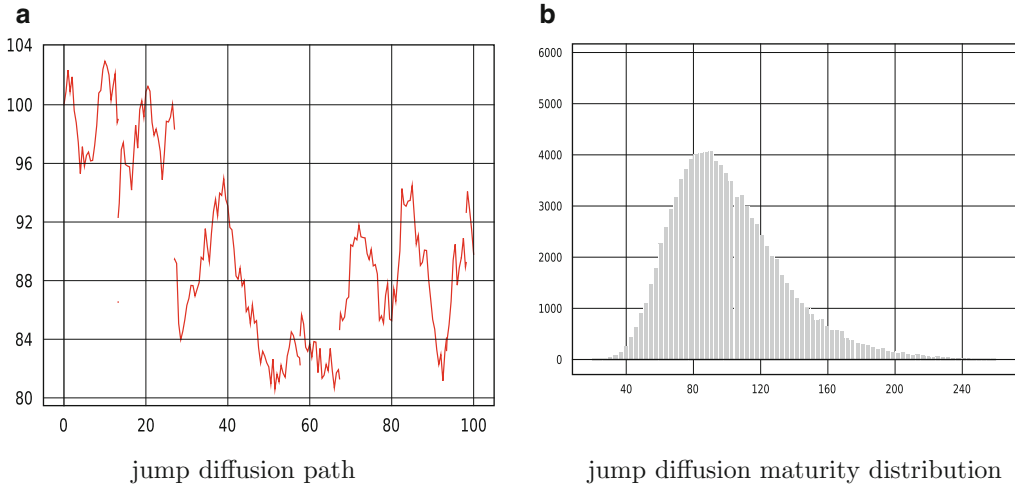


Fig. 6.7. Characteristics of a jump-diffusion geometric random walk. A typical path is shown in (a) and the end-point distribution is shown in (b). The drift-diffusion parameters are: $\mu = 3\%$ and $\sigma = 40\%$. The jumps are distributed as $LN(-0.0032, 0.08^2)$ with event rate $\lambda = 0.1$ per day

```

for ...   ▷infinite loop
  if( jDX > N_T or sDX * dt < arrivals[jDX] )
    dst = sDX*dt - simTime; //increment in simtime
    ▷do a diffusion
    dS = S(μ * dst + σ√dstZ);   ▷Z ~ N(0,1)
    S = S + dS;
    simTime = simTime+dst;   ▷update simTime
    sDX = sDX+1;   ▷point at next step
  else   ▷jump event
    ▷do a diffusion since last step
    dst = arrivals[jDX]-simTime;   ▷increment in simTime
    dS = S(μ * dst + σ√dstZ);   ▷Z ~ N(0,1)
    S = dS + S;
    simTime = arrivals[jDX];   ▷update simTime
    S = S*J;   ▷J ~ F(·), after jump price
    jDX= jDX + 1;   ▷point to next jump time
  endif
  ▷check if done
  if sDX > n break out of loop
endfor
S_T = S;

```

6.6.2 Jump-Diffusion Risk-Free Growth

In order to use a Lévy process for market predictions, the process must be a martingale. It is possible to achieve this by adjusting the drift of the process,

this is a consequence of the *Girsanov Theorem*, [Bjo04]. However, except for the Poisson pure jump and Wiener processes, the martingale measure is not unique. From the discussion in Section 6.1, this means the market is incomplete and there is no one no-arbitrage price. Notwithstanding uniqueness, next we show how to calculate a no-arbitrage drift for the jump-diffusion process.

The infinitesimal growth rate of the jump-diffusion model may be calculated from (6.24). The drift term being constant, its expected value is itself $\mu S_t dt$, and the expected value of the diffusion term is 0 because the expected value of a Wiener process is that. With regard to the jump term, the expected value is the expected jump size times the expected arrival rate of the jumps. Since the latter arise according to a Poisson distribution with intensity λ , we may write

$$\mathbb{E}\left(S_t \sum_{k=1}^{N_t} (J_k - 1) \delta_{t_k}(dt)\right) = S_t \mathbb{E}(J - 1) (\lambda dt). \quad (6.28)$$

For example, if the jumps are distributed according to $N(\mu_J, \sigma_J^2)$, then

$$\mathbb{E}(dS_t) = \mu S dt + (\mu_J - 1) S_t \lambda dt. \quad (6.29)$$

And if they are distributed according to $LN(\alpha, \beta^2)$, then

$$\mathbb{E}(dS_t) = \mu S dt + (e^{\alpha + \frac{1}{2}\beta^2} - 1) S_t \lambda dt. \quad (6.30)$$

On the other hand, in order to be risk-neutral, the expected growth rate should be $S_t r dt$ where r is the risk-free rate. Hence in the case of normally distributed jumps

$$r = \mu + \lambda(\mu_J - 1)$$

so that

$$\mu = r - \lambda(\mu_J - 1). \quad (6.31)$$

And in the case of lognormally distributed jumps,

$$\mu = r - \lambda(e^{\alpha + \frac{1}{2}\beta^2} - 1). \quad (6.32)$$

Using these drifts in the price simulations for these jump diffusion processes is the equivalent of using the risk-free rate in the GBM simulations. One still has to discount back the option payoffs.

6.6.3 Calculating Prices for Vanilla Options

We will use the Monte Carlo method to obtain option prices by simulating the jump-diffusion model. The only change to Algorithm 26 is to add the option payoff function $G(S_T)$ (for path independent options) at the end of the loop and then discount this back to $t = 0$, see Algorithm 27.

If the jump sizes are to be normally or lognormally distributed, use (6.31) or (6.32) as appropriate for the drift.

In Fig. 6.8 we compare jump-diffusion ending price distributions for both normal and lognormal jump sizes against that of geometric Brownian motion.

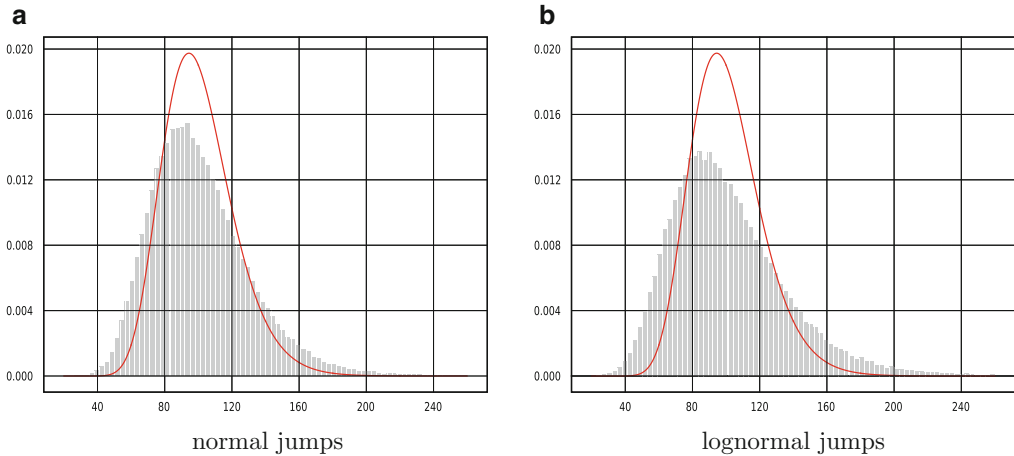


Fig. 6.8. Comparisons between GBM maturity distribution (in red) and a jump diffusion maturity histogram. For the GBM $r = 3\%$, $\text{vol} = 40\%$, $T = 100$ days. The Poisson process is $\lambda = 0.1$ per day. In (a) the jumps are $N(1, 0.06^2)$; in (b) they are $LN(-0.0032, 0.08^2)$

In each case the jump-diffusion prices show greater spread and so we can expect higher option prices as if the volatility were greater.

Algorithm 27. Monte Carlo Jump-Diffusion Pricing Algorithm

```

inputs:  $S_0, K, T, r, \sigma, nTrials$ 
 $E = 0$ 
for  $i = 1, \dots, nTrials$ 
   $S = S_0$ 
  ▷ use Algorithm 26 to generate  $S_T$ 
   $E = E + G(S_T)$ 
end for
option price =  $e^{-rT} E / nTrials$ 

```

Figure 6.9 illustrates a comparison between option prices under the Black-Scholes model and those of a jump diffusion model. As previously mentioned, the jump diffusion model is incomplete and therefore there is no unique no-arbitrage price. In the figure the risk-neutral value of (6.31) was used.

Exotic Options

Many of the exotic options go just as discussed in Chapter 4 since we are able to simulate instances of the price paths for Lévy processes. However others require some care in the use of Lévy jump process. In the case of a barrier option, a jump can carry the underlying's price across the barrier triggering the corresponding action. And again, in our shout boundary approach to shout options, a jump

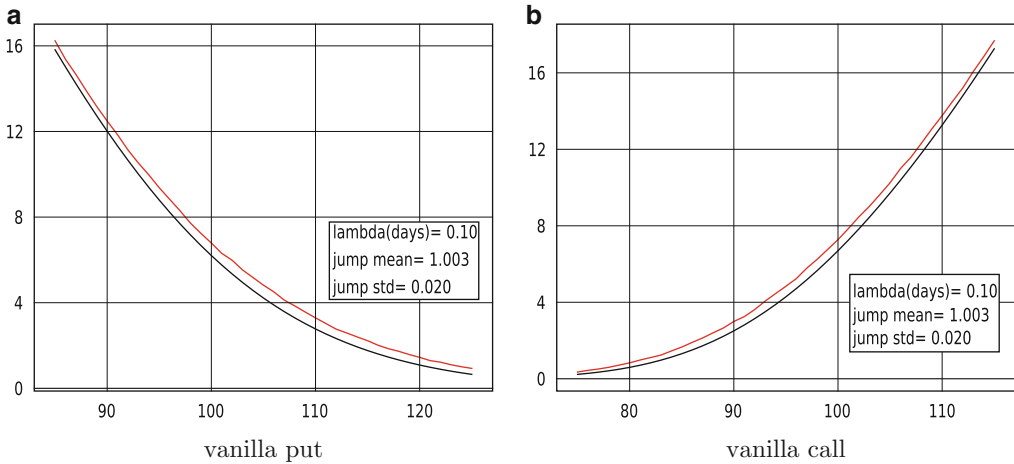


Fig. 6.9. Black-Scholes put and call values (*black*) as compared with those for the jump diffusion model (*red*) using normal distribution jumps plotted against stock price S . The option characteristics are: $K = 100$, $T = 60$ (days), $r_f = 3\%$, $\text{vol} = 40\%$. The jump parameters are as indicated. The jump diffusion ATM put costs 6.75 vs 6.20 for Black-Scholes, a 9% increase. The jump diffusion ATM call costs 7.21 vs 6.70 for Black-Scholes, an 8% increase

can carry the price across the boundary calling for a shout. These options must be simulated event-to-event and Brownian bridges must be considered between events (4.1), see [CT04].

As previously reported, simulation prices of vanilla options for a range of parameter values, for example, drifts in the jump-diffusion model, when compared to their market prices can be used to determine the exact (current) parameter values applicable. Then these values are used to calculate exotic option prices.

6.7 Time Shifted Processes

In Section 6.3 we encountered an example of a subordinator, a process that is either constant or increasing. One of the main uses of such a process is to replace the smoothly moving calendar time by the subordinator process. In this way an entirely new class of Lévy processes can be generated. In finance such a process is used to simulate *business time* since businesses tend to operate from event to event. If τ_t is a subordinator and X_t an overlying Lévy process, then the *subordinated* or *time changed* process is

$$L_t = X_{\tau_t}. \quad (6.33)$$

Often a Wiener process is used as the overlying process.

Figure 6.10 shows a typical Gaussian subordinator path in (a) and the end point histogram in (b). The simulation, done via Algorithm 28, proceeds in regular time increments Δt as usual. But the Wiener process is based on an inverse Gaussian time step. The path of such a process often has large movement as

shown but they are not exactly jumps since they occur Δt time units apart. These can occur when there is a long time period commanded by the subordinator, then the Gaussian step has a chance to be large. The end point distribution is depicted in (b). It shows a narrow peak but a very wide base. A small number of large jumps in the same direction can account for this phenomenon.

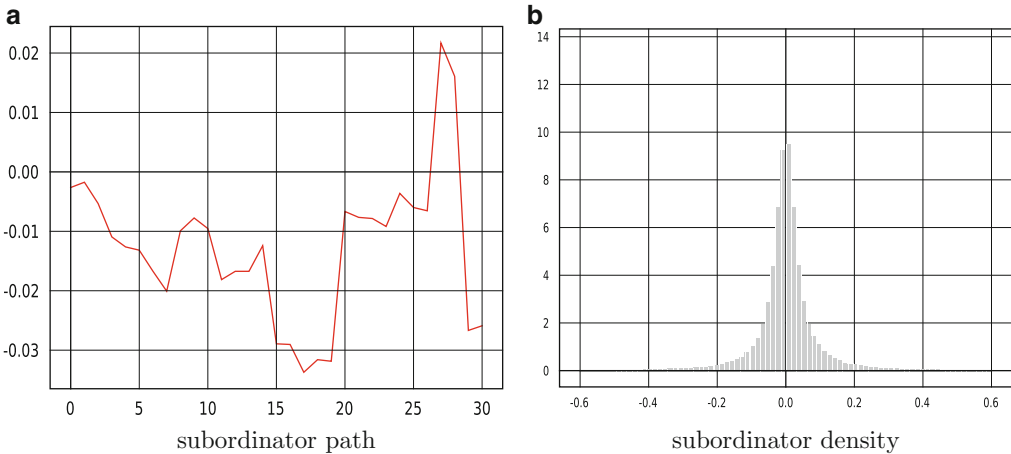


Fig. 6.10. Illustrated in (a) is a typical Gaussian subordinator path. Illustrated in (b) is an end point histogram. It shows a narrow peak but a very wide base. A small number of large jumps in the same direction can account for this phenomenon

Algorithm 28. Gaussian Subordinator Simulation

```

inputs:  $T, dt, X_0, \mu, \sigma, a, b$ 
 $X = X_0;$ 
 $n = T/dt;$   $\triangleright$ number of steps
for  $i = 1, 2, \dots, n$ 
     $\tau \sim IG(a * dt, b)$   $\triangleright$ incr. in time via subordinator
     $\Delta X = \mu\tau + \sigma\sqrt{\tau}Z;$   $\triangleright Z \sim N(0, 1)$ 
     $X = X + \Delta X;$ 
endfor
 $X_T = X;$ 

```

6.8 Heavytailed Distributions

The normal distribution is widely used in finance, but often it is only an approximation to the actual distribution of the circumstance. One piece of evidence for this is that events which should only occur once in thousands of years, instead occur 2 or 3 times in 40. Some have labeled these as “six-sigma events” since, if the normal distribution applied, their probability of occurring would be that in the upper tail six standard deviations from the mean. It can be inferred that the

actual distribution governing, for example price movements, has greater probability for extreme events than is accounted for by the normal distribution. That is to say, the tails of the distribution should be fatter.

It is for this reason that financial mathematicians study *heavytailed* distributions, densities decaying more slowly in the tails than the normal. In this section we examine two examples. The first is the widely known family of *t distributions*, sometimes known as the *Student t*. For the second example, we show that a heavytailed distribution can be constructed as the difference between two independent subordinators.

6.8.1 Student's t-Distribution

The t distribution has a single parameter, $\nu > 0$, known as the *degrees of freedom* (dof). Some members of the family are shown in Fig. 6.11.

The t probability density function is given by

$$f_{\nu}(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}. \quad (6.34)$$

In this $\Gamma(\cdot)$ is the *gamma function* defined by the integral

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt.$$

In (6.34) the gamma terms are just constants contributing to normalization.

The gamma function is an extension of the factorial function. Using integration by parts, it is easy to see that it satisfies the recursion

$$\Gamma(z + 1) = z\Gamma(z).$$

And by direct integration we get

$$\Gamma(1) = 1.$$

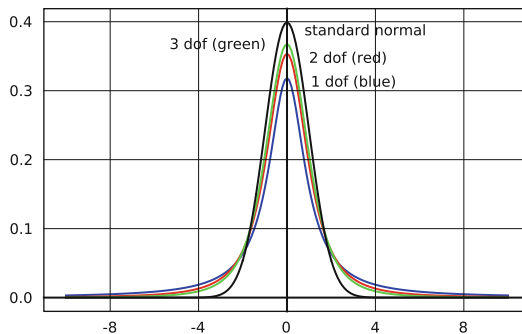


Fig. 6.11. Student t densities for degrees of freedom equal to 1, 2, 3, and infinity (the standard normal)

From these two facts it is easy to see that, for integers, gamma is the factorial function,

$$\Gamma(n) = (n - 1)! \quad n \text{ a positive integer.} \quad (6.35)$$

The only other commonly needed value of gamma is for $z = 1/2$ and that value is well-known,

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \quad (6.36)$$

With these preparations, we may write the first five members of the t_ν family (for integral ν)

$$\begin{aligned} f_1(x) &= \frac{1}{\pi(1+x^2)} \\ f_2(x) &= \frac{1}{2\sqrt{2}}(1+x^2/2)^{-3/2} \\ f_3(x) &= \frac{2}{\pi\sqrt{3}}(1+x^2/3)^{-2} \\ f_4(x) &= \frac{3}{8}(1+x^2/4)^{-5/2} \\ f_5(x) &= \frac{8}{3\pi\sqrt{5}}(1+x^2/5)^{-3}. \end{aligned}$$

The $\nu = 1$ density is also known as *Cauchy's density*. As $\nu \rightarrow \infty$ the t_ν distribution tends to the standard normal density. As seen in Fig. 6.11 the tails become less heavy as ν increases.

The t densities are symmetric about $x = 0$ and hence have mean equal to 0. The $\nu = 1$ and $\nu = 2$ densities do not have bounded square integrals and therefore their variances are infinite. But for $\nu > 2$ the variances are finite,

$$\text{var}(t_\nu) = \frac{\nu}{\nu - 2} \quad \nu = 3, 4, \dots \quad (6.37)$$

Sampling from the Student-t

The most widely used method for sampling from the t_ν distribution is due to Baily [Bai94]. It is valid for all $\nu > 0$. The Baily algorithm is a simple modification of the Marsaglia-Bray algorithm, Algorithm 2, for the standard normal.

Algorithm 29. Baily's Algorithm for t_ν Samples

```

repeat
  U ~ U(0,1); U = 2U-1;  ▷uniform on -1 to 1
  V ~ U(0,1); V = 2V-1;  ▷a point in the sqr.
until W=U2+V2<= 1
C=U2/W; R = ν(W-2/ν - 1)
T = √RC;
if( U ~ U(0,1) <.5 )
  return T
else
  return -T

```

6.8.2 Difference Subordinator Densities

Although we have only studied two Lévy densities besides the normal, namely the Poisson and the inverse Gaussian, there are many known. And we have shown how Lévy processes can be constructed as compound Poisson processes or by time change. Many of the Lévy process densities are heavytailed. Here we show another method for constructing a heavytailed density guaranteed to be infinitely divisible.

Let X_t and Y_t be subordinators and put Z_t equal to their difference,

$$Z_t = X_t - Y_t. \tag{6.38}$$

Then Z_t is infinitely divisible and therefore a Lévy process. For example, X and Y can be the same subordinator.

By independence, the mean of Z_t is just the difference of that of X_t and Y_t , and the variance is the sum $\text{var}(Z_t) = \text{var}(X_t) + \text{var}(Y_t)$. In Fig. 6.12 we show the difference between two inverse Gaussians for two different parameter sets. Also shown is the normal density having the same mean and variance. At about 2σ the difference density exceeds the normal showing that its tails are heavy.

Application to Asset Prices

An alternate formulation of (6.23) derives from including the risk-free rate in the stochastic exponential separately,

$$dS_t = r_f S_t dt + S_{t-} dZ_t. \tag{6.39}$$

This formulation is completely general and holds for any Lévy process Z_t . From the martingale preserving property, in this formulation the process $e^{-r_f t} S_t$ is a martingale if and only if $\mathbb{E}(Z_1) = 0$.

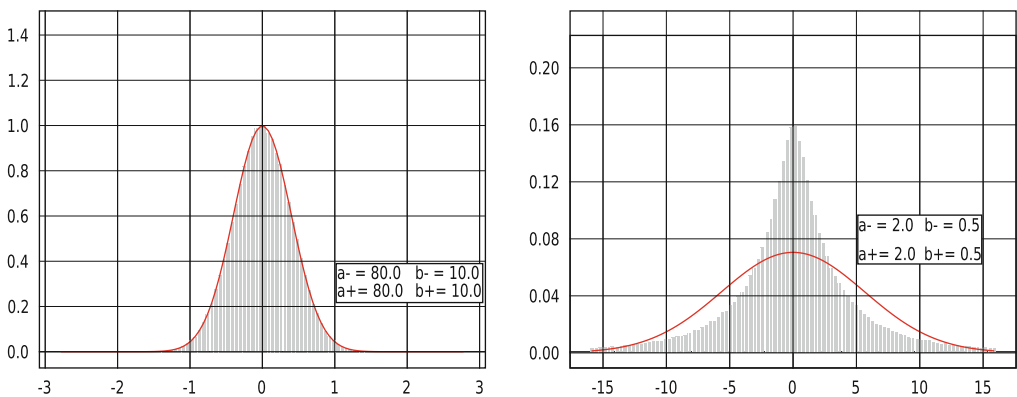


Fig. 6.12. Difference inverse Gaussian histograms for two quite different parameter sets. Overlaying each figure is the normal distribution with the same mean and variance. Both inverse Gaussians are heavytailed

Problems: Chapter 6

1. Show that if $X \sim \text{Po}(\lambda_1)$ and $Y \sim \text{Po}(\lambda_2)$, then $X + Y \sim \text{Po}(\lambda_1 + \lambda_2)$. Hint:

$$\begin{aligned} \Pr(X + Y \leq z) &= \sum_{y=0}^z \Pr(X \leq z - y | Y = y) \Pr(Y = y) \\ &= \sum_{y=0}^z \frac{\lambda_1^{z-y} e^{-\lambda_1}}{(z-y)!} \frac{\lambda_2^y e^{-\lambda_2}}{y!}. \end{aligned}$$

2. The *skew* of a random variable X is defined as

$$\text{skew} = \mathbb{E}((X - \mu_X)^3) / \text{std}_X^3. \quad (6.40)$$

Given data x_1, x_2, \dots, x_n an estimator for skew is

$$\overline{\text{skew}} = \frac{\sum_1^n (x_i - \bar{x})^3}{n\bar{s}^3}$$

where \bar{x} and \bar{s} are empirical mean and standard deviation. Being a symmetric distribution the normal has 0 skew. Calculate the empirical skew of the log returns $(\log \frac{S_{i+1}}{S_i})$, for 3 stock equities of your choice using daily prices over the last 2 years. (Use the FIMCOM database or finance.yahoo for the prices, see Section 1.7.3 page 25.)

3. The *kurtosis* of a random variable X is defined as

$$\text{kurtosis} = \mathbb{E}((X - \mu_X)^4) / \text{std}_X^4. \quad (6.41)$$

Given data x_1, x_2, \dots, x_n an estimator for kurtosis is

$$\overline{\text{kurtosis}} = \frac{\sum_1^n (x_i - \bar{x})^4}{n\bar{s}^4}$$

where \bar{x} and \bar{s} are empirical mean and standard deviation. The kurtosis of the normal distribution is 3, cf. the footnote on page 15. Calculate the empirical kurtosis of the log returns $(\log \frac{S_{i+1}}{S_i})$, for 3 stock equities of your choice using daily prices over the last 2 years. (Use the FIMCOM database or finance.yahoo for the prices, see Section 1.7.3 page 25.)

4. (a) From market price data make a graph of implied volatility σ versus strike price K for call options on the S&P-500 for expiration maturities of T on the order of 30 days (near as possible). Do the same for $T = 60$ and 90 days. You now have a *volatility surface*, implied volatility versus strike and time.
 (b) Do the same for put options.
5. The Gamma distribution, $G(\alpha, \lambda)$ has density given by

$$f_G(x : \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0. \quad (6.42)$$

Here $\Gamma(\alpha)$ is the gamma function of Section 6.8 and equals $(\alpha - 1)!$ if α is a positive integer. Show that the Gamma is infinitely divisible (empirically) by showing that

the histogram for the sum of six samples of $G(1, \lambda)$ has the same density as $G(6, \lambda)$. Note that, for α a positive integer, then

$$W = \frac{-1}{\lambda} \log\left(\prod_1^{\alpha} U_i\right) \quad U_i \sim U(0, 1) \quad (6.43)$$

is a sample from $G(\alpha, \lambda)$, [SM09].

6. (a) Make a chart similar to Fig. 6.9 showing the price of a put option using the jump diffusion model with lognormal jumps for stock prices versus the GBM model. In order to compare the results with jump diffusion using $N(\mu_J, \sigma_J^2)$ jumps, find α and β to match the mean and variance,

$$\mu_j = \mu_{LN} = e^{\alpha + \frac{1}{2}\beta^2}, \quad \sigma_J^2 = (e^{\beta^2} - 1)\mu_{LN}^2.$$

(b) Do the same for calls.

7. (a) Make a chart similar to Fig. 6.9 showing the price of a put option using a difference IG model for stock prices versus the GBM model. Use $a- = a+ = 41$ and $b- = b+ = 8$. What are the mean and variance of the difference IG process?
(b) Do the same for calls.
8. (a) Work the Bermuda option Problem 5 of Chapter 4 assuming prices follow a jump diffusion with normal sized jumps. Be sure to report your jump parameters.
(b) Repeat (a) using lognormal sized jumps.
9. Work the Bermuda option Problem 5 of Chapter 4 assuming prices follow a symmetric differential IG model, use equation (6.39). Be sure to report your model's parameters.
10. Recalculate Table 4.2 page 121 for barrier options assuming prices follow a jump diffusion model with normal sized jumps. Recall that the simulation must go event-by-event.
11. Recalculate Table 4.1 page 119 for Asian options assuming prices follow a differential IG process.
12. A portfolio consists of 100 shares each of stock A: $S_0 = 60$, $\mu = 8\%$, $\sigma = 40\%$; and B: $S_0 = 40$, $\mu = 3\%$, $\sigma = 20\%$. Their correlation is $\rho = 0.3$. After 6 months what is the probability of losing money and the expected gain of the portfolio if (a) prices follow a Gaussian GBM model? (b) a jump diffusion with normal jumps?
13. Work the VaR Problem 9 of Chapter 2 assuming prices follow jump diffusion with normal sized jumps.

In the following assume a jump diffusion model for prices with normal jumps.

14. Analyze covered calls as in Table 5.2.

15. Analyze creditspreads as in Table 5.5.

In the following assume a (symmetric) difference IG model for prices.

16. Analyze iron condors as in Table 5.9.

17. Analyze the straddle strategy as in Table 5.7.