ON THE NATURE OF THE STOCK MARKET: SIMULATIONS AND EXPERIMENTS

by

Hendrik J. Blok

B.Sc., University of British Columbia, 1993
M.Sc., University of British Columbia, 1995

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Department of Physics and Astronomy)

We accept this dissertation as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 2000

Appendix B

Sampling discrete processes

Frequently computer simulations generate synthetic Brownian motion via a simple random walk at discrete intervals. Sampling of such a process to get the distribution of increments $p(x)$ can be problematic because of introduced artifacts which bias the statistics.

One often-used statistic is the (excess) kurtosis, defined as

$$\mathrm{Kurt}[x] = \frac{\mu_4}{\mu_2^2} - 3 \tag{B.1}$$

where μ_k is the k'th (centered) moment of the distribution

$$\mu_k = \left\langle [x - \langle x \rangle]^k \right\rangle . \tag{B.2}$$

The kurtosis is useful because it quantifies the "weight" of the distribution tail (far from the mean). For the Gaussian the excess kurtosis is zero (because $\mu_4 = 3\mu_2^2$) compared to which a negative kurtosis indicates less weight in the tails and a positive indicates more.

Difficulties arise, however, if the Brownian motion is generated by a discrete process as will be demonstrated in two examples below. Unless great care is taken, the kurtosis may be artificially inflated by regular sampling.

B.1 Simple random walk

Consider a discrete Brownian process with normally-distributed (zero mean, unit variance) jumps at regular intervals of $\tau = 1$ (without loss of generality). If this process is sampled at regular intervals of $\Delta \neq \tau$, as demonstrated in Fig. B.1, some intervals will have more "jumps" than others so the distribution of increments will not be Gaussian.

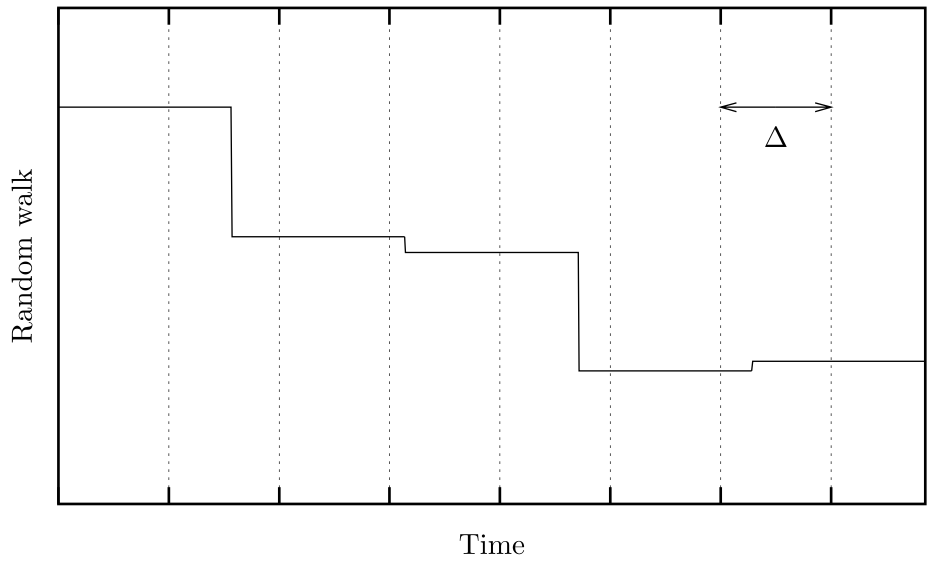

Figure B.1: When a random walk is generated at some regular interval and sampled at another, Δ, the number of jumps between samples will vary.

To be precise, let $\Delta \equiv n + r$ where $n \equiv \lfloor \Delta \rfloor$ is the largest integer not greater than Δ (the *floor* of Δ) and $0 \leq r < 1$ is the remainder. Then each interval will span at least n jumps, spanning $n + 1$ with the probability r. Since each jump x is normally distributed $N(x; 0, 1)$ with zero mean and unit variance, j jumps are also normally distributed with zero mean and variance j, denoted by $N(x; 0, j)$. The distribution of increments of the random walk, sampled at intervals of $\Delta = n + r$ is then given by

$$RW(x; 0, \Delta) = (1 - r)N(x; 0, n) + rN(x; 0, n + 1). \tag{B.3}$$

Calculating the first four moments of the increment distribution is very straight-forward since

$$\mu_k[RW(x; 0, \Delta)] = (1 - r)\mu_k[N(x; 0, n)] + r\mu_k[N(x; 0, n + 1)] \tag{B.4}$$

and the normal distribution has moments $\mu_1 = 0$, $\mu_2[N(x; 0, j)] = j$, $\mu_3 = 0$, and $\mu_4 = 3\mu_2^2$. Therefore, the moments of RW are

$$\mu_1 = 0 \tag{B.5}$$

$$\mu_2 = (1 - r)n + r(n + 1) = n + r = \Delta \tag{B.6}$$

$$\mu_3 = 0 \tag{B.7}$$

$$\mu_4 = 3(1 - r)n^2 + 3r(n + 1)^2 = 3(\Delta^2 + r(1 - r)). \tag{B.8}$$

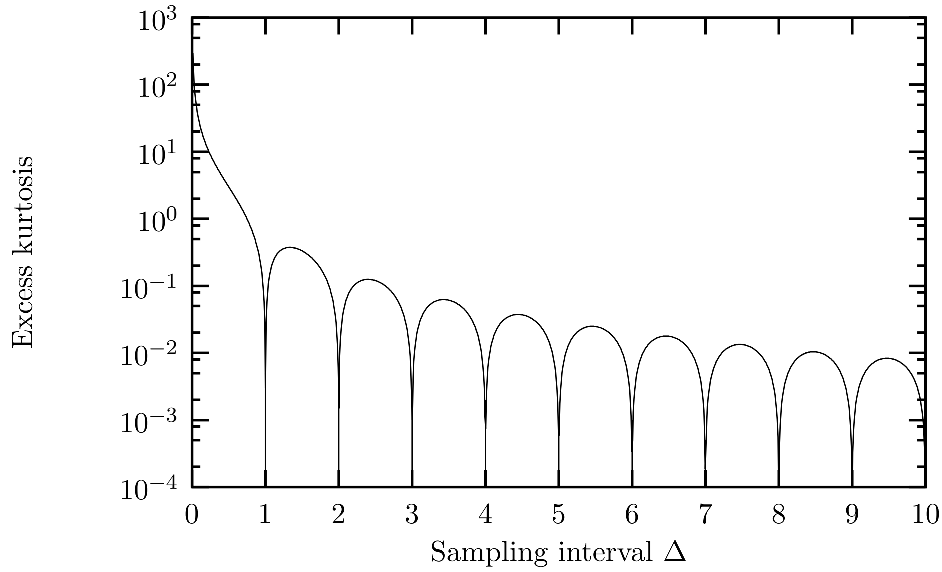

Figure B.2: The kurtosis is only zero at integer values of the sampling interval Δ and diverges as the sampling interval approaches zero.

Notice that the variance of the distribution is simply Δ, exactly the same as for *continuous* Brownian motion sampled at intervals of Δ.

In fact, all three of the lowest moments are identical to Brownian motion, lulling us into a false sense of security. However, the fourth moment differs and the excess kurtosis, which is zero for Brownian motion, is now

$$\text{Kurt}[RW(x;0,\Delta)] = 3\frac{r(1-r)}{\Delta^2} \tag{B.9}$$

which, on the surface, would seem to indicate the distribution has fat tails. The kurtosis is only zero at integer values of Δ ($r = 0$) and is a maximum for any n when $r = n/(1+2n)$ as shown in Fig. B.2.

In particular, the kurtosis diverges as the sampling rate accelerates

$$\text{Kurt}[RW] \to \frac{3}{\Delta} \text{ as } \Delta \to 0, \tag{B.10}$$

a result of the Dirac delta function $N(x;0,0)$ dominating the distribution, scaling the variance down faster than the fourth moment.

Even though all the evidence presented suggests that the distribution of increments in the random walk truly does have fat tails when sampled at non-integer intervals Δ, it is actually just an artifact of sampling.

Since we are getting an overlap of two Gaussian distributions, with variances n and $n+1$, the center of the distribution is dominated by the smaller variance

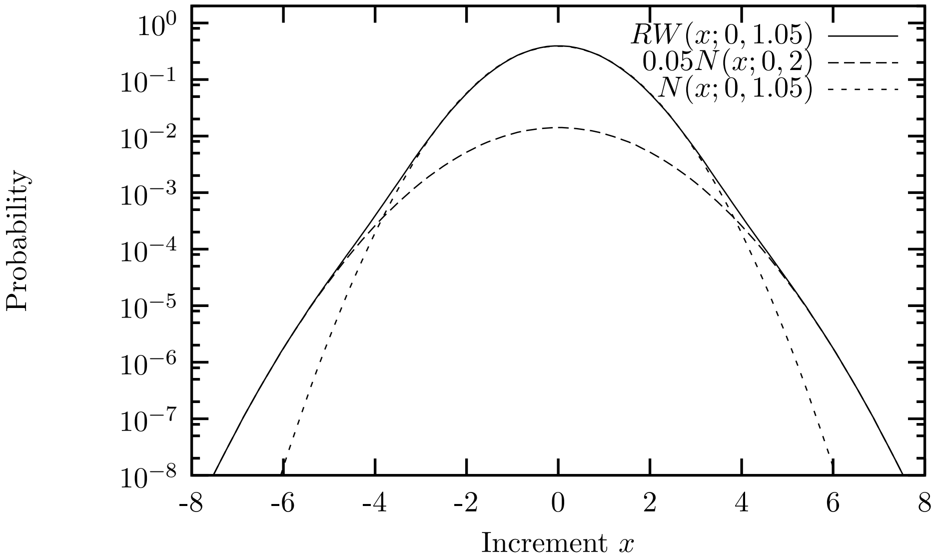

Figure B.3: The distribution of increments for the random walk appears to have fatter tails than a normal distribution with the same variance when sampled at intervals of $\Delta = 1.05$. However, the tails still drop off as e^{-x^2}.

contribution and the tails are dominated by the larger variance. Hence, the second moment of the random walk is scaled down by the smaller variance but the fourth moment is scaled up by the larger. The net effect is the illusion of fat tails in the distribution.

However, the tails of the distribution still fall off as e^{-x^2}, as demonstrated in Fig. B.3, so the term "fat tails" is misleading, usually being reserved for simple exponential or power law tails.

Notice that the center of the distribution behaves as a normal with variance Δ and the tail also behaves as a normal, with variance $n + 1$, but weighted by r. The crossover between the two regimes, after some algebra, is found to be

$$x_c = \sqrt{\frac{\left[\ln(n+1) - 2\ln(r\sqrt{\Delta})\right]\Delta(n+1)}{1-r}}. \qquad \text{(B.11)}$$

This indicates the scale of increments, $x \approx \pm x_c$, for which the distribution will appear most strongly non-Gaussian.

Next we consider a process generated at Poisson intervals rather than regular.

B.2 Poisson Brownian motion

In this section we again consider a discrete Brownian motion but, in this case, the intervals between the jumps are Poisson-distributed instead of regular. The Poisson distribution gives the probability of j events within a time interval t given an average event rate $\tau \equiv 1$ (without loss of generality),

$$P(j, t) = e^{-t} \frac{t^j}{j!}. \tag{B.12}$$

Given normally-distributed jump sizes the distribution of j jumps is $N(x; 0, j)$ so the distribution of increments of the Poisson Gaussian process at intervals of Δ is

$$PG(x; 0, \Delta) = \sum_{j=0}^{\infty} P(j, \Delta) N(x; 0, j). \tag{B.13}$$

The analytic solution for the distribution of increments is challenging but the moments of the distribution are relatively easy to compute,

$$\begin{aligned}
\mu_k[PG(x; 0, \Delta)] &= \int dx \sum_{j=0}^{\infty} P(j, \Delta) N(x; 0, j)\, x^k && \text{(B.14)} \\
&= \sum_{j=0}^{\infty} P(j, \Delta) \int dx\, N(x; 0, j)\, x^k && \text{(B.15)} \\
&= \sum_{j=0}^{\infty} P(j, \Delta)\, \mu_k[N(x; 0, j)], && \text{(B.16)}
\end{aligned}$$

depending directly on the moments of the normal distribution (which were presented in the last section).

From the identity $e^x \equiv \sum_j x^j / j!$, the first four moments of the Poisson Gaussian are

$$\begin{aligned}
\mu_1 &= 0 && \text{(B.17)} \\
\mu_2 &= \Delta && \text{(B.18)} \\
\mu_3 &= 0 && \text{(B.19)} \\
\mu_4 &= 3\Delta(\Delta + 1). && \text{(B.20)}
\end{aligned}$$

Again, the first three moments are unchanged from the normal distribution but the kurtosis becomes

$$\mathrm{Kurt}[PG(x; 0, \Delta)] = \frac{3}{\Delta} \tag{B.21}$$

for all Δ. (This form was also observed for the random walk in the limit $\Delta \to 0$.) So, again the kurtosis diverges as the sampling interval drops towards zero.

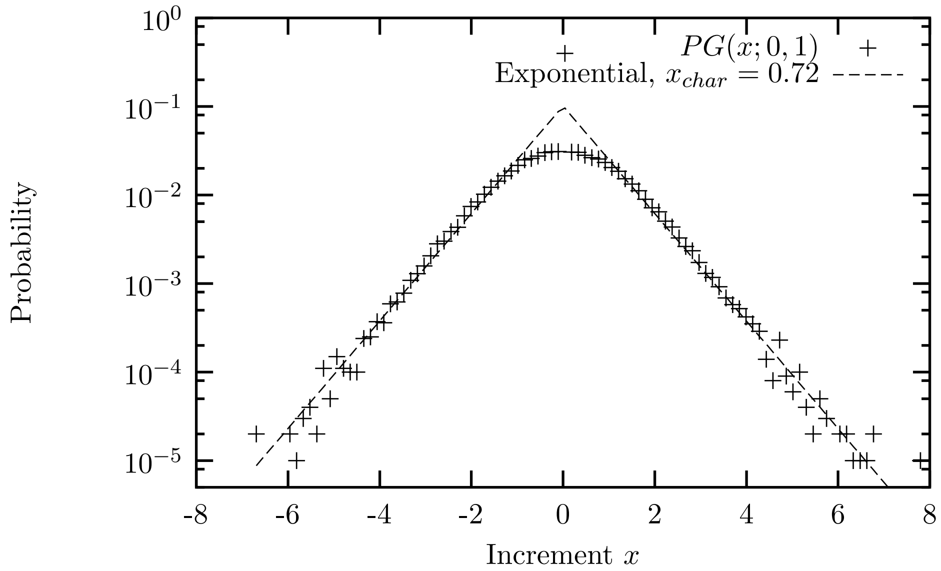

Figure B.4: Discrete Brownian motion with Poisson-distributed jump intervals has tails which fall off exponentially (with a decay constant of 0.72), instead of as e^{-x^2}, when sampled at regular intervals ($\Delta = 1$).

In the case of the random walk we found the excess kurtosis arose from the overlap of two Gaussians but the tails still fell off as e^{-x^2}. However, for Poisson Brownian motion the distribution tails are much heavier. A synthetic dataset generated from a Poisson Brownian motion sampled at intervals of $\Delta = 1$ (Fig. B.4) shows that the tails fall off only exponentially,

B.3 Sampling

Evidently, by generating a synthetic Brownian motion at non-uniform intervals, the illusion of fat tails can be achieved by simply sampling the process regularly. However, the underlying process is still generated by Gaussian-distributed jumps and, over long timescales, still looks like Brownian motion.

The easiest way to avoid these artifacts is to not sample the process in "real time" but in "event time." That is, take a single sample after each event. Then, the underlying jump process will be revealed without any complications from zero or multiple events per sample.

Unfortunately, in some cases the available data do not allow for the determination of individual events. In this case, a very high frequency sampling is recommended and all intervals with zero increment should be discarded. High frequency sampling minimizes the likelihood that multiple jumps could occur in any one interval but increases the likelihood of zero increments. By discarding these null events only the intervals with a single increment remain. (This also discards actual jump events of size zero but this should have a minimal bias on the statistics since a jump size of identically zero has a negligible probability measure.)