May 23, 2025

SA

SA 00 (2025) 1–14

Affine returns on Bernoulli trials in finance

Z.W.T. Mason

Stoch Answers Ltd, Sheffield, UK

Abstract

Offset returns of stock price movements were used to model the Paris Bourse in 1900. This first Mathematical model of Brownian motion was superseded by Geometric Brownian motion in the 1960s. i.e. the normal distribution was replaced by the log-normal and offset returns by linear returns. The market crash of 1998 caused the latter model to be questioned. This paper extends the model to affine returns, matching the average behaviour seen on the S&P 500 to calculations from averages of the daily ups and downs along with their probabilities. i.e. expected behaviour to noise. Affine returns lead to a mixture distribution consisting of two components — the log-normal distribution and one other which looks like the logit-normal distribution scaled from $[0, 1]$ to some other finite support. For the shrinking case it was shown that the distribution was not parameterisable.

1. Introduction

In his thesis Bachelier [1] derived a Mathematical model of Brownian Motion for the stock prices on the Paris Bourse, this was promptly ignored due to finance not being deemed a suitable topic for academic study. Five years later Einstein produced his model of Brownian Motion during his annus mirabilis. On the basis of a, possibly unfair, coin toss one set amount is won or a, possibly different, one is lost resulting in a binomial distribution which is approximated by a normal distribution. In the 1960s Mandelbrot [2] showed that cotton prices weren't normally distributed, similar studies for other assets followed. Geometric Brownian Motion was developed as a better model in which the stake grows or shrinks by set ratios leading to a log-normal distribution. This was used as the basis for deriving the Black-Scholes model for pricing European options. In 1997 Scholes and Merton were awarded the Nobel prize in economics, Black having died in the interim. At the time the two had left academia to work for Long Term Capital Management, the following year the US Federal Reserve intervened to save the market after the use of the model by their employer had racked up billions of dollars in losses.

This led to a re-examination of the foundations of quantitative finance. Wilmott [3] states that the returns on stocks progress from fat tailed to looking like they are normally distributed and finally to log-normal. These fat tails are what Taleb [4] calls Black Swans and not the long tails of the normal distribution which David Viniar, CFO of Goldman Sachs, apparently assumed despite their obvious impossibility:

> We were seeing things that were 25 standard deviation moves, several days in a row.

In the abstract of his book Mandelbrot [5] dismisses models based on Bernoulli trials as mere coin tossing games:

Email address: papers@stochanswers.com (Z.W.T. Mason)

Conventional financial theory assumes that variation of prices can be modeled by random processes that, in effect, follow the simplest "mild" pattern, as if each uptick or downtick were determined by the toss of a coin. What fractals show, and this book describes, is that by that standard, real prices "misbehave" very badly.

2. Affine returns on Bernoulli trials

Ignoring Mandelbrot and his propensity to see fractals everywhere in favour of Wilmott raises the question of how does the return progress from fat tailed to log-normal? The next level of complexity for a coin tossing game is to make the returns affine functions. i.e. With probability $q = 1 - p$ the payout for the next toss is a function of the stake returned at the t'th toss, S_t,

$$S_{t+1} = \alpha S_t + \gamma \tag{1}$$

and with probability p the payout is

$$S_{t+1} = \beta S_t + \delta \tag{2}$$

For $\alpha = \beta = 1$ the result is the binomial distribution for Brownian Motion which, by de Moivre-Laplace, tends to the normal distribution

$$S_t - S_0 \sim t\gamma + (\delta - \gamma)\, \mathcal{N}(tp, tpq) \tag{3}$$

hence

$$E(S_t) \sim S_0 + tq\gamma + tp\delta \sim E(S_{t-1}) + q\gamma + p\delta \tag{4}$$

For $\gamma = \delta = 0,\ 0 < \alpha < 1 < \beta$ there is the binomial distribution for Geometric Brownian Motion which tends to the log-normal distribution

$$\log\left(\frac{S_t}{S_0}\right) \sim t\log(\alpha) + \log\left(\frac{\beta}{\alpha}\right) \mathcal{N}(tp, tpq) \tag{5}$$

hence

$$E(S_t) \sim S_0 \alpha^{tq} \beta^{tp} \sim E(S_{t-1})\, \alpha^q \beta^p \tag{6}$$

The more general case of $0 < \alpha < 1,\ 0 < \beta \neq 1, 0 = \gamma < \delta$ results in an order dependent set of terms known as a non-recombinant lattice. However, the terms containing the value of S_0 can be separated out showing that the result is a mixture distribution composed of the log-normal distribution and another. To see whether this is all just idle speculation the daily opening prices from the S&P 500 index for the 5 years from the 8th of February 2013 were analysed. The daily downs were fitted to equation 1, the daily ups to equation 2. Then the expected behaviour was approximated as

$$E(S_t) \sim (\alpha^q \beta^p)^t S_0 + \frac{r^t - 1}{r - 1} p\delta \text{ where } r = q\alpha + p\beta \neq \alpha^q \beta^p \tag{7}$$

Note that, if $S_0 = 0, r = 1$, this compares to the recurrence relationship for the normal distribution – equation 4. The second term is a weighted average which doesn't compare directly to the first but, unlike the first term, for $\alpha, \beta < 1$, it doesn't shrinks away to nothing. So the model is wrong which leaves the question of whether or not the formula is useful. The resulting growth of the formula values for AIG, AIZ and ATVI are shown in figures 1 to 3 corresponding to whether the support is finite or semi-infinite (β less than or greater than one) and whether there is exponential growth or decay (r less than or greater than one). These results were taken from a handful of decent fits however the correlations may, of course, be spurious. Nine generations of histograms of the mixture distribution are shown for the selected cases in figures 4 to 6 for the reader to consider whether or not they match the behaviour seen by Wilmott. The calculations were performed in C# using arbitrarily sized rational numbers to generate each possible

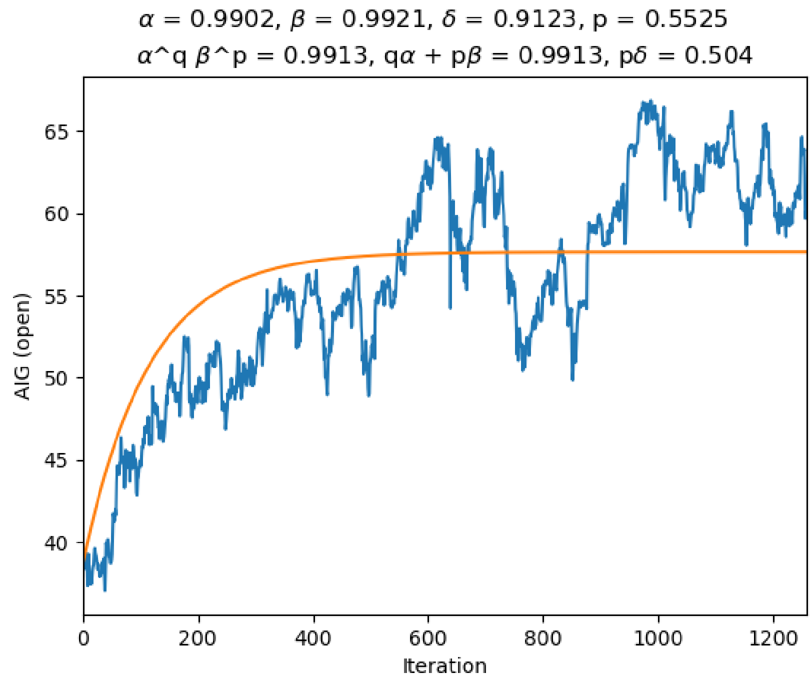

Figure 1. Fitting the noise for $\beta < 1$

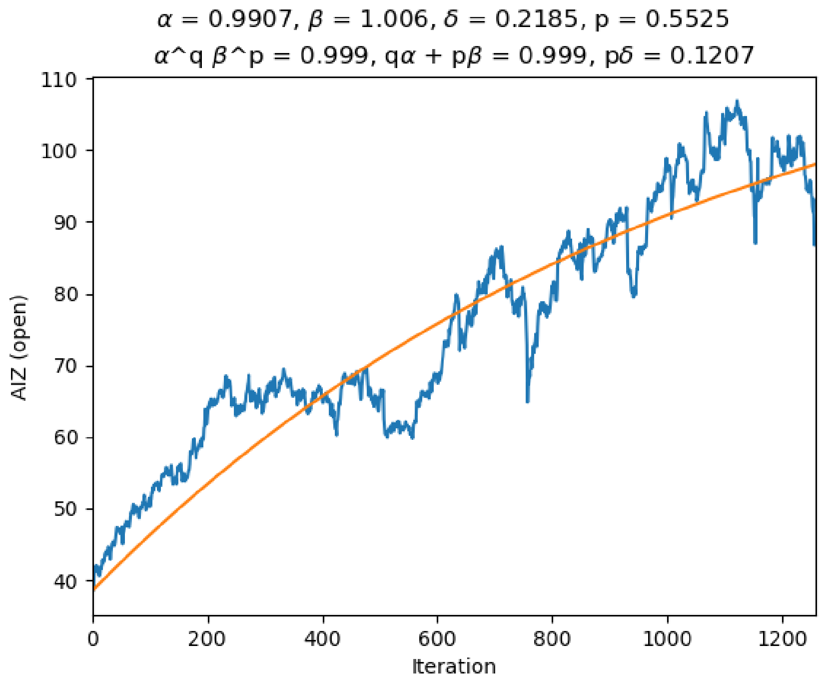

Figure 2. Fitting the noise for $\beta > 1, q\alpha + p\beta < 1$

point in a generation. Note that the reason why only 25 generations were calculated was because the algorithm is

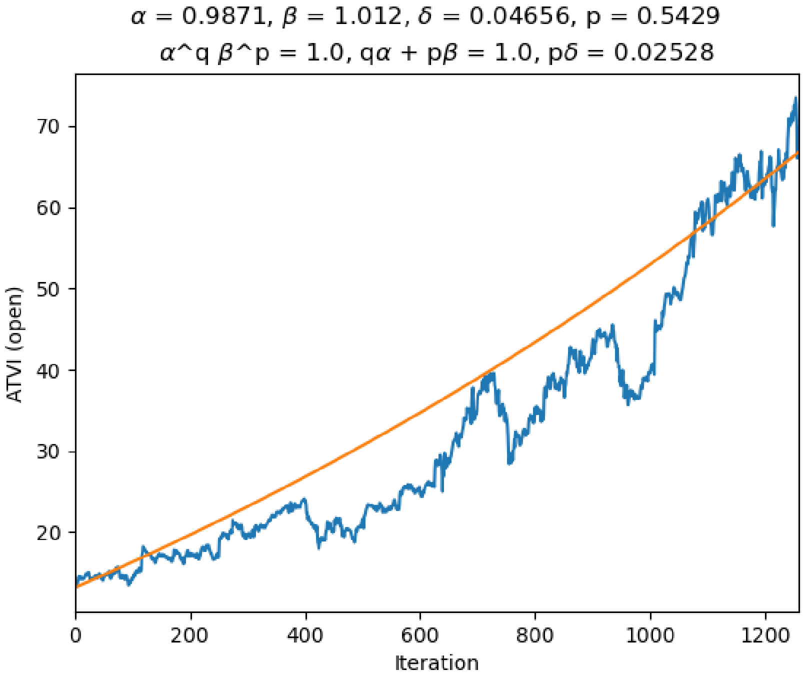

Figure 3. Fitting the noise for $\beta > 1, q\alpha + p\beta > 1$

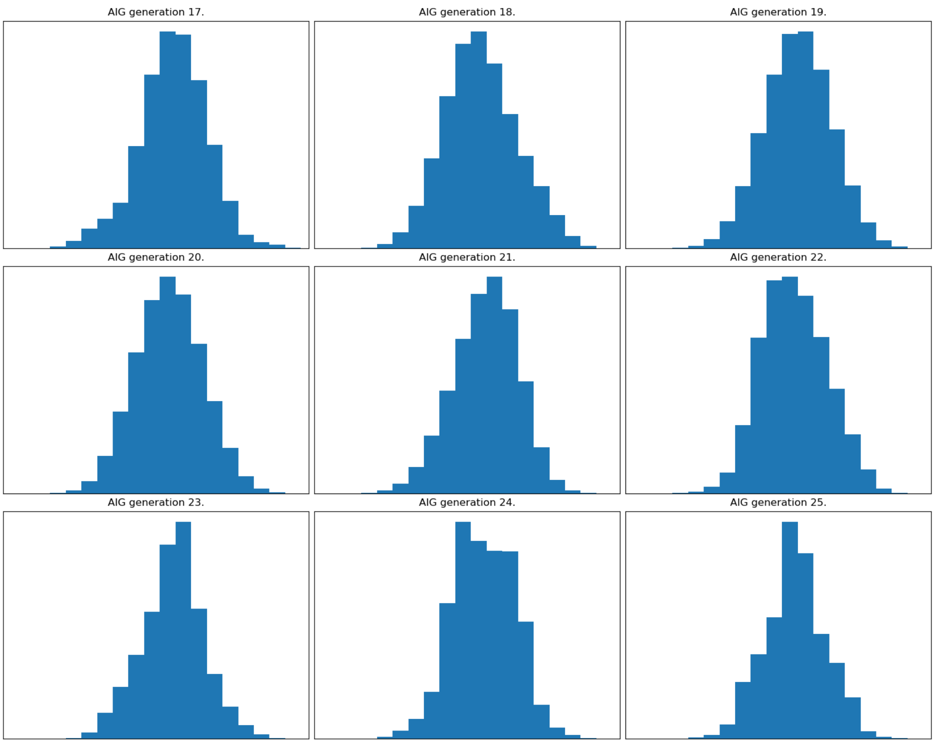

Figure 4. Mixture distributions by generation for $\beta < 1$

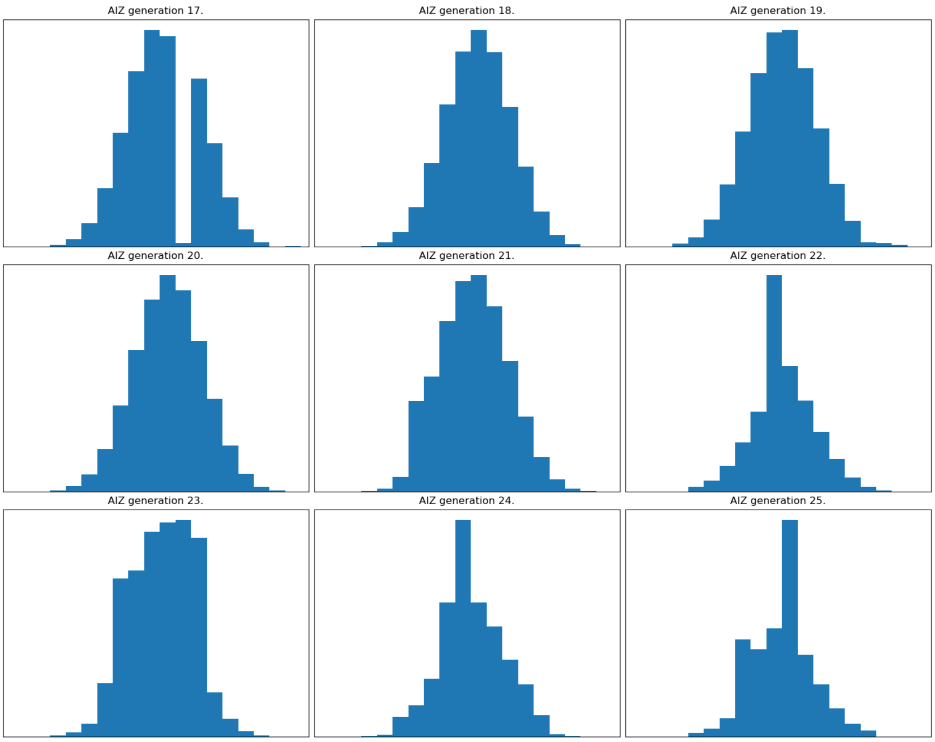

Figure 5. Mixture distributions by generation for $\beta > 1, q\alpha + p\beta < 1$

exponential in both time and storage.

The second distribution is defined on the interval

$$\left(0, \delta \frac{1-\beta^t}{1-\beta}\right) \tag{8}$$

Defining a new random variable $X_t \in [0, 1]$ as S_t divided by the support when $S_0 = 0$ results in an expected value of

$$E(X_t) = p \frac{1-\beta}{1-r} \frac{1-r^t}{1-\beta^t} \tag{9}$$

which converges to some non-zero value for finite support, $\beta < 1$, and exponential decay, $r < 1$. In general, the moments can be expressed as a weighted sum where the weights sum to one.

$$E(X_t^n) = \sum_j w_j x_j^n \tag{10}$$

Applying the transform to individual points and weights

$$\left\{\left(w_j, x_j\right)\right\} \to \left\{\left(q w_j, \alpha x_j\right)\right\} \cup \left\{\left(p w_j, \beta x_j + 1 - \beta\right)\right\} \tag{11}$$

which implies

$$E(X^n) \to qE((\alpha X)^n) + pE((\beta X + 1 - \beta)^n) \tag{12}$$

The formulae for the first four moments is compared to calculations from the actual affine terms, equations 1 and 2, extrapolated as decaying exponentially to convergence in Appendix A.

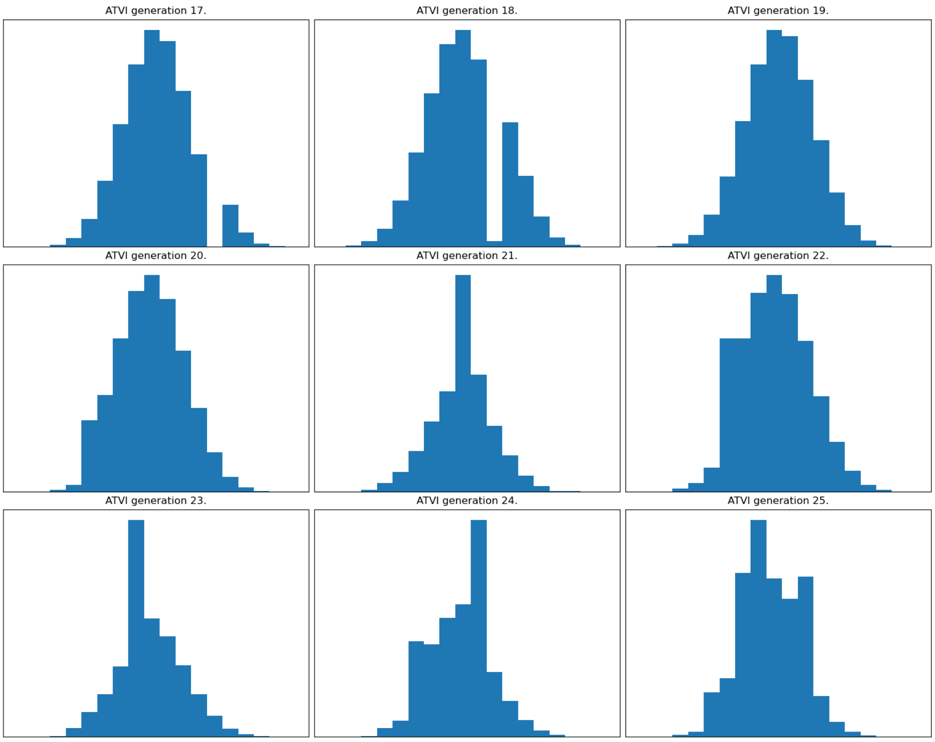

Figure 6. Mixture distributions by generation for $\beta > 1, q\alpha + p\beta > 1$

Given that the transform involves a scale and a shift the transformed probability distribution function at the tails is only due to a single instance rather than the sum of both. In particular, near $x = 0$

$$f(x) = qf(x/\alpha) \implies \log(f(x)) \sim \log(x)\log(q)/\log(\alpha) \tag{13}$$

Similarly there is another fat tail near $x = 1$

$$\log(f(x)) \sim \log(1 - x)\log(p)/\log(\beta) \tag{14}$$

The underlying histograms for the distribution of X gave the impression that the points were logit-normally distributed. i.e. The logit of the random variable is normally distributed.

$$\operatorname{logit}(X) \equiv \log\left(\frac{X}{1 - X}\right) \sim \mathcal{N}\left(\mu, \sigma^2\right) \tag{15}$$

Given how much it is overlooked the logit-normal could be considered to be the ugly swan of probability distributions. The author [6] has previously shown that, for a well-mixed epidemic, the ratio of growth to exponential growth follows this distribution resulting in the logistic growth of the standard SIR model. This reasonable worse case growth results from the spread of infection being seen to be due to sampling without replacement leading to the compounding at each generation of one hypergeometric distribution per infectee. i.e. A completely different discrete stochastic process.

The properties of the logit-normal distribution are detailed in Appendix B.

3. The logit-normal distribution for the converged case

Having calculated the distributions, the resulting points were used to calculate the first four moments and to fit values for μ and σ. As before with the first moment these values being a function of the current step they were assumed to decay exponentially to a converged value.

The general form of the moments, equation 16, is obtained by expanding and rearranging equation 12.

$$E\left(X^n\right) = p\frac{\sum_{i=0}^{n-1} {}^nC_i \left(1-\beta\right)^{n-i}\beta^i E\left(X^i\right)}{1-q\alpha^n - p\beta^n} \tag{16}$$

The first moment is

$$E\left(X\right) = p\frac{1-\beta}{1-q\alpha-p\beta} \tag{17}$$

hence

$$E\left(1-X\right) = q\frac{1-\alpha}{1-q\alpha-p\beta} \tag{18}$$

Thus, the mapping $(p,\alpha,\beta) \rightarrow (q,\beta,\alpha)$ implies $\left(\mu,\sigma^2\right) \rightarrow \left(-\mu,\sigma^2\right)$.
The degenerate cases occurs when $\alpha = \beta$ resulting in the simplified form seen in equation 19

$$E\left(X^n\right) = p\frac{\sum_{i=0}^{n-1} {}^nC_i \left(1-\beta\right)^{n-1-i}\beta^i E\left(X^i\right)}{\sum_{i=0}^{n-1}\beta^i} \tag{19}$$

When $p = q = 1/2$ the distribution is symmetric, when $\alpha = \beta = 1$ there is zero variance by equation B.12 since

$$E\left(X^n\right) = p^n \tag{20}$$

For the symmetric case the first four moments are

$$E\left(X\right) = \frac{1}{2} \tag{21}$$

$$E\left(X^2\right) = \frac{1}{2\left(1+\beta\right)} \tag{22}$$

$$E\left(X^3\right) = \frac{2-\beta}{4\left(1+\beta\right)} \tag{23}$$

$$E\left(X^4\right) = \frac{1+\beta^2-\beta^3}{2\left(1+\beta\right)^2\left(1+\beta^2\right)} \tag{24}$$

Hence by equations B.5 and B.6

$$\frac{\partial E\left(X\right)}{\partial\mu} = \frac{\partial E\left(X^2\right)}{\partial\mu} = \frac{\beta}{2\left(1+\beta\right)} \implies \frac{\partial E\left(X\right)}{\partial\sigma} = 0 \tag{25}$$

$$\frac{\partial E\left(X^3\right)}{\partial\mu} = \frac{3\beta\left(1-\beta+3\beta^2-\beta^3\right)}{4\left(1+\beta\right)^2\left(1+\beta^2\right)} \tag{26}$$

$$\frac{\partial E\left(X^2\right)}{\partial\sigma} = -\sigma\frac{\beta\left(1-5\beta+7\beta^2-5\beta^3\right)}{2\left(1+\beta\right)^2\left(1+\beta^2\right)} \tag{27}$$

The derivatives of the first moment are

$$\frac{\partial E\left(X\right)}{\partial\alpha} = -\frac{\partial E\left(X\right)}{\partial\beta} = \frac{1}{4\left(1-\beta\right)} \tag{28}$$

$$\frac{\partial E\left(X\right)}{\partial p} = 1 \tag{29}$$

The derivatives of the second moment are

$$\frac{\partial E\left(X^2\right)}{\partial \alpha} = \frac{\beta\left(3+\beta\right)}{4\left(1-\beta\right)\left(1+\beta\right)^2} \tag{30}$$

$$\frac{\partial E\left(X^2\right)}{\partial \beta} = -\frac{2+\beta+\beta^2}{4\left(1-\beta\right)\left(1+\beta\right)^2} \tag{31}$$

$$\frac{\partial E\left(X^2\right)}{\partial p} = 1 \tag{32}$$

Hence the differences in the derivatives of the second and first moments are

$$\frac{\partial E\left(X^2\right)}{\partial \alpha} - \frac{\partial E\left(X\right)}{\partial \alpha} = \frac{\partial E\left(X^2\right)}{\partial \beta} - \frac{\partial E\left(X\right)}{\partial \beta} = -\frac{1}{4\left(1+\beta\right)^2} \tag{33}$$

$$\frac{\partial E\left(X^2\right)}{\partial p} - \frac{\partial E\left(X\right)}{\partial p} = 0 \tag{34}$$

To ensure that $f = 0$ for the symmetric case it can be modelled as

$$f\left(p,\alpha,\beta\right) = \left(\alpha-\beta\right) f_1\left(p,\alpha,\beta\right) + \left(p-q\right) f_2\left(\alpha,\beta\right) \tag{35}$$

where f_2 is seen to be independent of p by observing the data. Inserting equations 25, 28, 29 and 35, into equation B.21 results in

$$\frac{\partial f}{\partial p} = 2f_2 = \frac{2\left(1-\beta\right)}{\beta} \tag{36}$$

$$\frac{\partial f}{\partial \alpha} = f_1 = \frac{1}{2\beta} \tag{37}$$

hence on $\alpha = \beta$

$$\mu = \operatorname{logit}\left(E\left(X\right)\right) + \left(\alpha-\beta\right)\frac{1}{\alpha+\beta} + \left(p-q\right)\frac{2-\alpha-\beta}{\alpha+\beta} \tag{38}$$

which doesn't agree well with the noisy data as shown in figure 7. Given equation 25 shows that $\partial E\left(X\right)/\partial\sigma = 0$ there is some confidence that the distribution is actually logit-normally distributed. The simulations produce noisy data which could explain the discrepancy. Alternatively, a different differential equation would have had to been solved if the argument of the logit was not the first moment.

Similarly, the data in figure 8 show that σ^2 is minimised when $E\left(X\right) = 1/2$ and thus

$$\sigma^2 = g\left(\alpha,\beta\right) + \left(p\left(1-\beta\right) - q\left(1-\alpha\right)\right)^2 h\left(p,\alpha,\beta\right) \tag{39}$$

However inserting equations 27 and 33 into equation B.18 results in an equation for the derivative of the variance which seemingly has no analytical solution.

$$\left(\frac{\partial E\left(X^2\right)}{\partial \alpha} - \frac{\partial E\left(X\right)}{\partial \alpha}\right) \Big/ \frac{\partial^2 E\left(X^2\right)}{\partial \mu^2} = \frac{1+\beta^2}{2\beta\left(1-5\beta+7\beta^2-5\beta^3\right)} \tag{40}$$

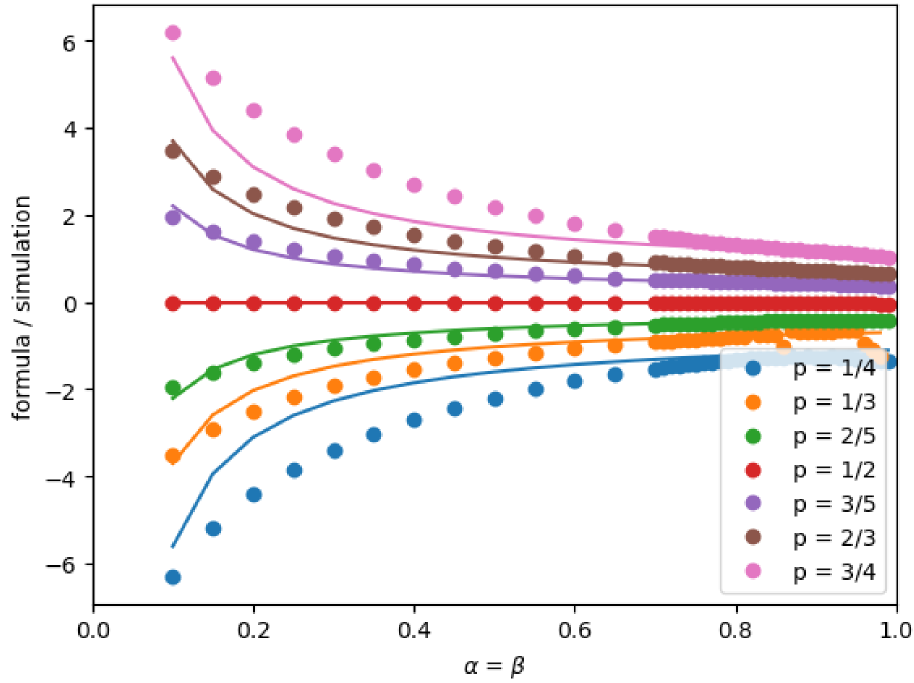

Figure 7. Comparison of mean values for $\alpha = \beta$

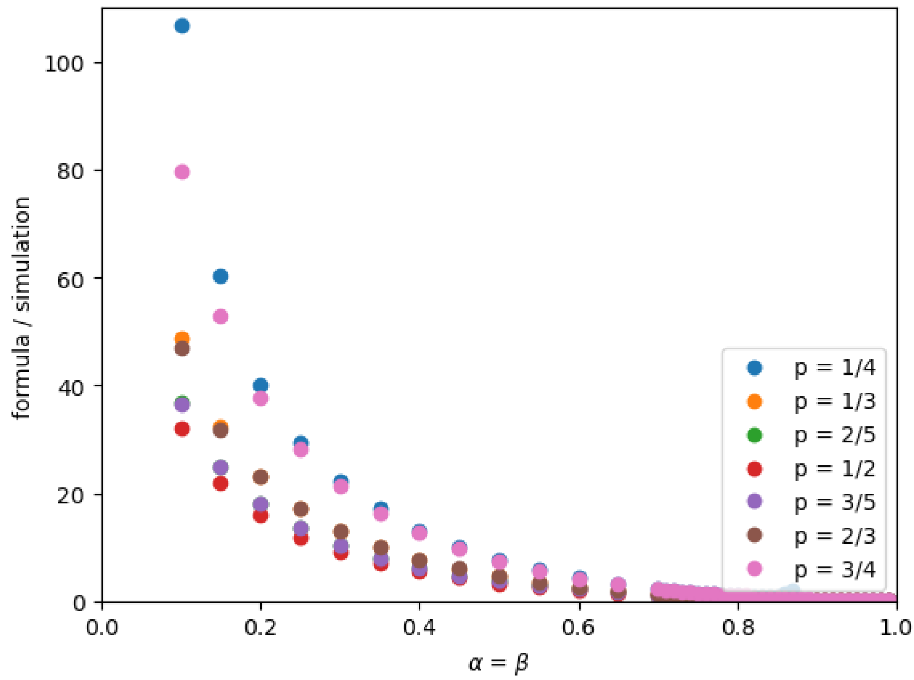

Figure 8. Comparison of variance values for $\alpha = \beta$

4. Discussion

It was shown that it is better to use a mixture distribution (due to affine returns) rather than merely the log-normal distribution (due to linear returns). This results in a model for the stock price of Geometric Brownian Motion plus an

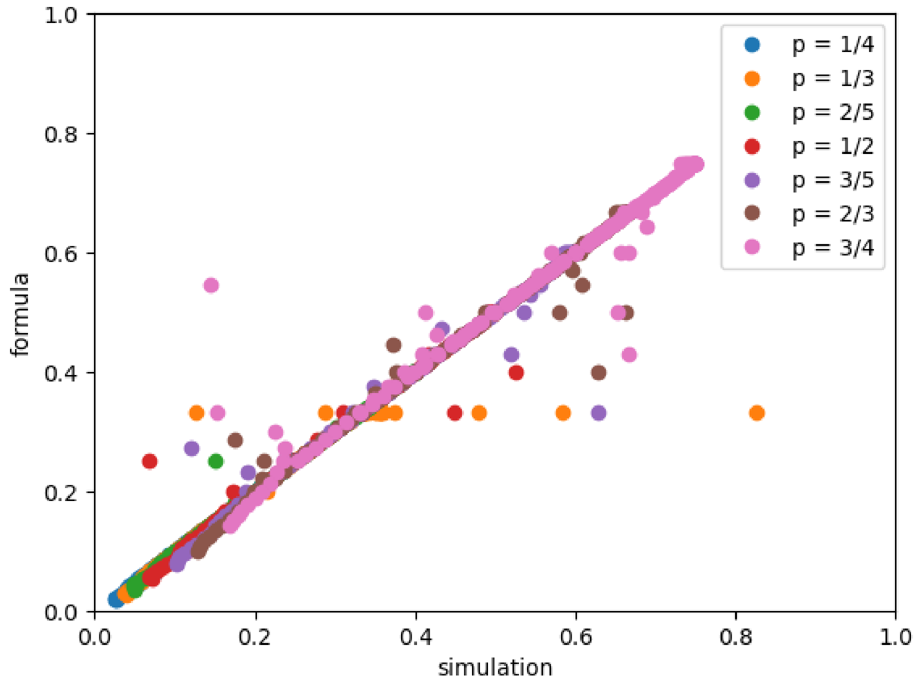

Figure A.9. Comparison of calculations to formula for $E(X)$

extra term

$$S_t = S_0\alpha^t\left(\frac{\beta}{\alpha}\right)^{tp+\sqrt{pq}W_t} + \delta\left(\frac{1-\beta^t}{1-\beta}\right)X(p,\alpha,\beta,t,W_t) \tag{41}$$

where X is taken from a probability distribution with finite support (over $[0,1]$) that is fat-tailed and is either the logit-normal or some closely related distribution. It is reasonable to assume that some function of X is normally distributed and that function is the inverse of a sigmoid function.

For the case of $\alpha,\beta < 1$, given the fit to simulated data, the moments are at least approximately known. Even so the probability density function may not even exist or, if it does, then may not be uniquely determined according to Mnatsakanov and Hakobyan [7]. It is possible to construct a set of polynomial basis functions which are all zero at the end points by using Gegenbauer polynomials with a parameter of $5/2$ and evaluate their coefficients from the functions for the moments. Any approximation to the probability density function built using these polynomials is expected to be wobbly giving rise to negative values.

Rather than guess a continuous model there is a discrete model at hand given by equations 1 and 2. This can be used to construct a non-recombinant binomial options pricing model. Such pricing models have the same issue as was seen in calculating the discrete distribution, namely that they are exponential in both time and storage. As has been mentioned for the authors 16GB machine the depth of the binary tree was limited to 25 generations. As such memory limitations may mean an implementation will not be sufficient to be useful.

Appendix A. Moments

Appendix B. The logit-normal distribution

The logit-normal distribution is defined on the interval $[0,1]$ by the probability density function

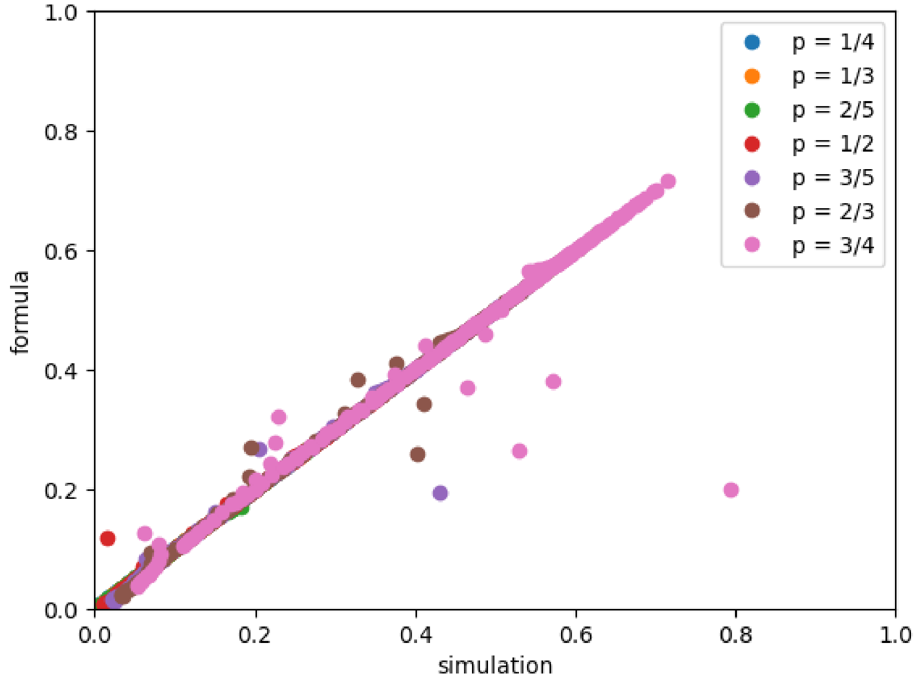

Figure A.10. Comparison of calculations to formula for $E\left(X^2\right)$

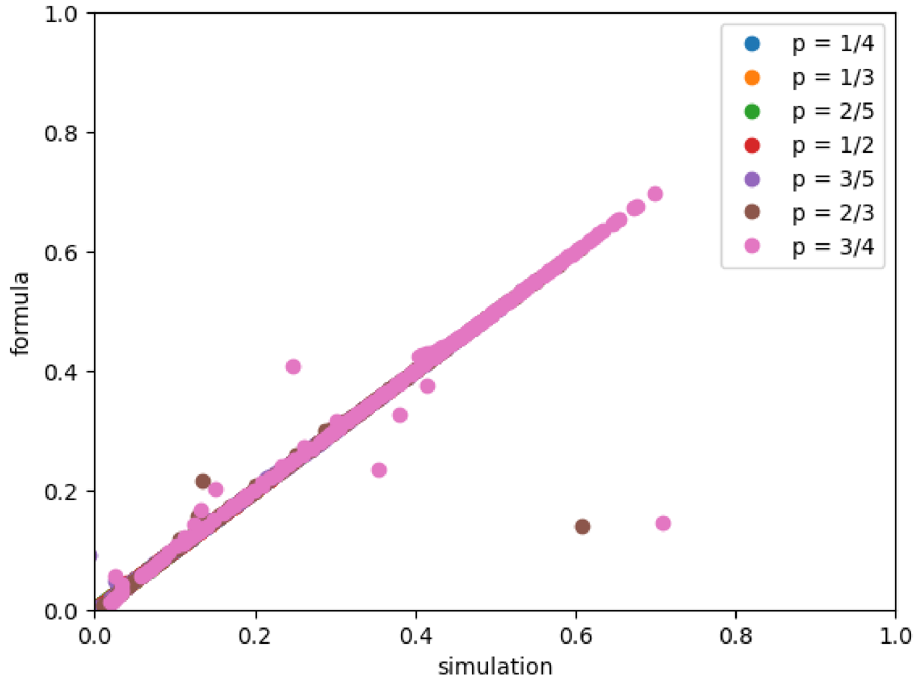

Figure A.11. Comparison of calculations to formula for $E\left(X^3\right)$

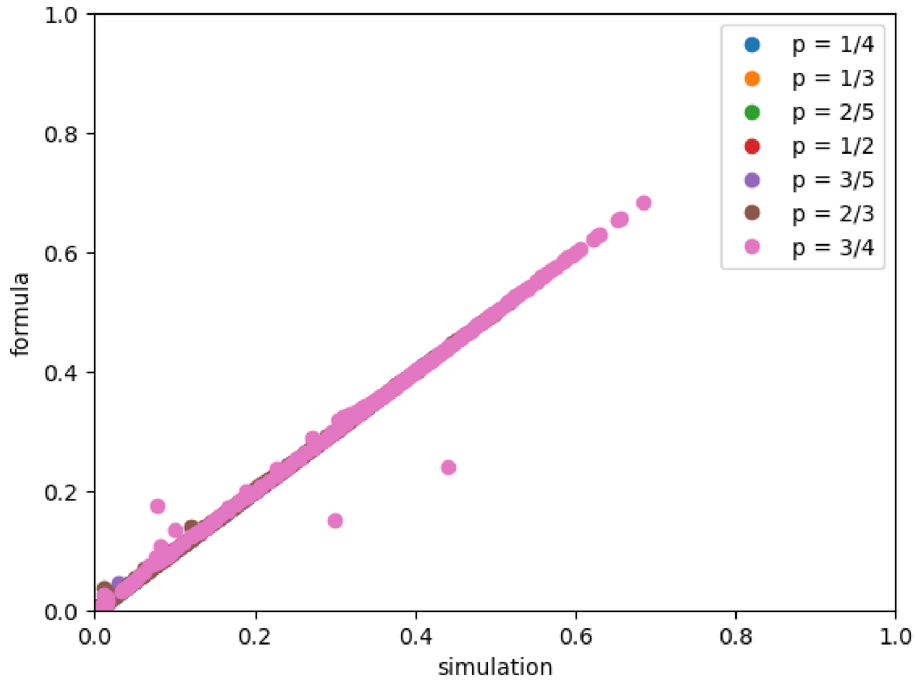

Figure A.12. Comparison of calculations to formula for $E\left(X^4\right)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{x(1-x)}\exp\left(-\frac{(\operatorname{logit}(x)-\mu)^2}{2\sigma^2}\right) \tag{B.1}$$

As noted by Frederic and Lad [8] transforming $x \to 1-x$ has the effect that $\left(\mu,\sigma^2\right) \to \left(-\mu,\sigma^2\right)$. Johnson [9] derived a formula for the moments which can be evaluated numerically by, for example, Gauss-Hermite quadrature there being no analytical solution for the general case.

$$E\left(X^n\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(1+e^{-\mu-\sigma z}\right)^{-n} e^{-z^2/2}dz \tag{B.2}$$

This formula uses the logistic function which is the inverse of the logit function. The logistic function is a simple transform of the hyperbolic tangent.

$$\left(1+e^{-x}\right)^{-1} = \frac{1}{2}\left(1+\tanh\left(\frac{x}{2}\right)\right) \tag{B.3}$$

hence the alternative form

$$E\left(X^n\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{1}{2}\tanh\left(\frac{\mu+\sigma z}{2}\right)+\frac{1}{2}\right)^n e^{-z^2/2}dz \tag{B.4}$$

Holmes-Schofield [10] differentiated the formula for the moments with respect to μ

$$\frac{\partial E\left(X^n\right)}{\partial\mu} = n\left(E\left(X^n\right)-E\left(X^{n+1}\right)\right) \tag{B.5}$$

The moments can also be differentiated with respect to σ

$$\frac{\partial E\left(X^n\right)}{\partial\sigma} = \sigma\frac{\partial^2 E\left(X^n\right)}{\partial\mu^2} \tag{B.6}$$

Taking the derivatives of two subsequent moments with respect to some parameter p results in

$$\frac{\partial E\left(X^n\right)}{\partial p}=\frac{\partial E\left(X^n\right)}{\partial \mu}\frac{\partial \mu}{\partial p}+\frac{\partial E\left(X^n\right)}{\partial \sigma}\frac{\partial \sigma}{\partial p} \tag{B.7}$$

$$\frac{\partial E\left(X^{n+1}\right)}{\partial p}=\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}\frac{\partial \mu}{\partial p}+\frac{\partial E\left(X^{n+1}\right)}{\partial \sigma}\frac{\partial \sigma}{\partial p} \tag{B.8}$$

hence equations for the derivatives of μ and σ can be derived as

$$\frac{\partial E\left(X^n\right)}{\partial p}\frac{\partial E\left(X^{n+1}\right)}{\partial \sigma}-\frac{\partial E\left(X^{n+1}\right)}{\partial p}\frac{\partial E\left(X^n\right)}{\partial \sigma}=\frac{\partial \mu}{\partial p}\left(\frac{\partial E\left(X^n\right)}{\partial \mu}\frac{\partial E\left(X^{n+1}\right)}{\partial \sigma}-\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}\frac{\partial E\left(X^n\right)}{\partial \sigma}\right) \tag{B.9}$$

and

$$\frac{\partial E\left(X^n\right)}{\partial p}\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}-\frac{\partial E\left(X^{n+1}\right)}{\partial p}\frac{\partial E\left(X^n\right)}{\partial \mu}=\frac{\partial \sigma}{\partial p}\left(\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}\frac{\partial E\left(X^n\right)}{\partial \sigma}-\frac{\partial E\left(X^n\right)}{\partial \mu}\frac{\partial E\left(X^{n+1}\right)}{\partial \sigma}\right) \tag{B.10}$$

The gradients of the higher moments must be a linear combination of the lower leading to a recurrence relationship for $E\left(X^{n+4}\right)$ calculated from the four previous moments

$$\begin{aligned}&\left(\frac{\partial E\left(X^n\right)}{\partial p}\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}-\frac{\partial E\left(X^n\right)}{\partial \mu}\frac{\partial E\left(X^{n+1}\right)}{\partial p}\right)\frac{\partial E\left(X^{n+2}\right)}{\partial \sigma}=\\&\left(\frac{\partial E\left(X^{n+1}\right)}{\partial \mu}\frac{\partial E\left(X^{n+2}\right)}{\partial p}-\frac{\partial E\left(X^{n+1}\right)}{\partial p}\frac{\partial E\left(X^{n+2}\right)}{\partial \mu}\right)\frac{\partial E\left(X^n\right)}{\partial \sigma}+\\&\left(\frac{\partial E\left(X^n\right)}{\partial p}\frac{\partial E\left(X^{n+2}\right)}{\partial \mu}-\frac{\partial E\left(X^n\right)}{\partial \mu}\frac{\partial E\left(X^{n+2}\right)}{\partial p}\right)\frac{\partial E\left(X^{n+1}\right)}{\partial \sigma}\end{aligned} \tag{B.11}$$

For $\sigma = 0$ the integral in equation B.2 can be evaluated analytically to produce the moments as powers of the logistic function of μ.

$$E\left(X^n\right)=\left(1+e^{-\mu}\right)^{-n}=E\left(X\right)^n \tag{B.12}$$

Calculating the moments around the mean when $\mu = 0$ shows that all the odd moments are zero. i.e. it is symmetric.

$$E\left(\left(X-\frac{1}{2}\right)^n\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{1}{2}\tanh\left(\frac{\sigma z}{2}\right)\right)^n e^{-z^2/2}dz \tag{B.13}$$

In particular

$$0=E\left(\left(X-\frac{1}{2}\right)^3\right)=E\left(X^3\right)-\frac{3}{2}E\left(X^2\right)+\frac{3}{4}E\left(X\right)-\frac{1}{8}=E\left(X^3\right)-\frac{3}{2}E\left(X^2\right)+\frac{1}{4} \tag{B.14}$$

Evaluating equation B.6 for the first moment results in

$$\frac{\partial E\left(X\right)}{\partial \sigma}=\sigma\frac{\partial^2 E\left(X\right)}{\partial \mu^2}=\sigma\left(E\left(X\right)-3E\left(X^2\right)+2E\left(X^3\right)\right) \tag{B.15}$$

hence this derivative is zero as expected and also implies

$$\frac{\partial E\left(X\right)}{\partial \mu}=\frac{\partial E\left(X^2\right)}{\partial \mu} \tag{B.16}$$

Taking the derivatives of the first two moments with respect to some parameter p results in

$$\frac{\partial E(X)}{\partial p} = \frac{\partial E(X)}{\partial \mu}\frac{\partial \mu}{\partial p} \tag{B.17}$$

and

$$\frac{\partial E\left(X^2\right)}{\partial p} = \frac{\partial E(X)}{\partial p} + \frac{\partial E\left(X^2\right)}{\partial \sigma}\frac{\partial \sigma}{\partial p} \tag{B.18}$$

Equation B.12 can be inverted to obtain

$$\mu = \operatorname{logit}(E(X)) + f \tag{B.19}$$

It is assumed that the argument of the logit function is the first moment, if another function were used then it would have to coincide with the value of the first moment when the distribution is symmetric and also when $\sigma = 0$. Similarly, f is zero under those two conditions. Note that $E(X) = \frac{1}{2}$ does not necessarily imply that the distribution is symmetric. Substituting into the derivative of the first moment results in

$$\left(\frac{1}{E(X)(1-E(X))}\frac{\partial E(X)}{\partial \mu} - 1\right)\frac{\partial E(X)}{\partial p} + \frac{\partial E(X)}{\partial \mu}\frac{\partial f}{\partial p} - \sigma\frac{\partial^2 E(X)}{\partial \mu^2}\frac{\partial \sigma}{\partial p} = 0 \tag{B.20}$$

The degenerate cases are

$$\mu = 0 \implies \left(4\frac{\partial E(X)}{\partial \mu} - 1\right)\frac{\partial E(X)}{\partial p} + \frac{\partial E(X)}{\partial \mu}\frac{\partial f}{\partial p} = 0 \tag{B.21}$$

and

$$\sigma = 0 \implies \frac{\partial f}{\partial p} = 0 \tag{B.22}$$

References

[1] L. Bachelier, Théorie de la spéculation, Annales Scientifiques de L'Ecole Normale Supérieure 17 (1900) 21–88, reprinted in P. H. Cootner (ed), 1964, The Random Character of Stock Market Prices, Cambridge, Mass. MIT Press.

[2] B. Mandelbrot, The variation of certain speculative prices, The Journal of Business 36 (4) (1963) 394–419. URL `http://www.jstor.org/stable/2350970?origin=JSTOR-pdf`

[3] P. Wilmott, Frequently Asked Questions in Quantitative Finance (2nd ed.), Wiley, Chichester, 2009.

[4] N. Taleb, The Black Swan : The Impact of the Highly Improbable, Random House, New York, 2007.

[5] B. Mandelbrot, R. Hudson, The (Mis)Behaviour of Markets: A Fractal View of Risk, Ruin and Reward, Basic Books, New York, 2004.

[6] Z. W. T. Mason, The minimal sir model, validated then extended to social networks and stochastic differential equations, TBD-doi:10.31219/osf.io/y6ckv.

[7] R. M. Mnatsakanov, A. S. Hakobyan, Recovery of distributions via moments, Lecture Notes-Monograph Series 57 (2009) 252–265. URL `http://www.jstor.org/stable/30250044`

[8] P. Frederic, F. Lad, Two moments of the logitnormal distribution, Communications in Statistics - Simulation and Computation 37 (2008) 1263–1269. doi:10.1080/03610910801983178.

[9] N. L. Johnson, Systems of frequency curves generated by methods of translation, Biometrika 36 (1-2) (1949) 149–176. arXiv:https://academic.oup.com/biomet/article-pdf/36/1-2/149/513258/36-1-2-149.pdf, doi:10.1093/biomet/36.1-2.149. URL `https://doi.org/10.1093/biomet/36.1-2.149`

[10] J. B. Holmes, M. R. Schofield, Moments of the logit-normal distribution, Communications in Statistics - Theory and Methods 51:3 (2022) 610–623. doi:10.1080/03610926.2020.1752723.