Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **449,** 2413–2420 (2015) doi:10.1093/mnras/stv445

The MLP distribution: a modified lognormal power-law model for the stellar initial mass function

Shantanu Basu,[1,2]★ M. Gil[2]† and Sayantan Auddy[1]

[1]*Department of Physics and Astronomy, The University of Western Ontario, London, ON N6A 3K7, Canada*
[2]*Department of Applied Mathematics, The University of Western Ontario, London, ON N6A 5B7, Canada*

Accepted 2015 February 26. Received 2015 February 25; in original form 2014 October 22

ABSTRACT

This work explores the mathematical properties of a distribution introduced by Basu & Jones (2004), and applies it to model the stellar initial mass function (IMF). The distribution arises simply from an initial lognormal distribution, requiring that each object in it subsequently undergoes exponential growth but with an exponential distribution of growth lifetimes. This leads to a modified lognormal with a power-law (MLP) distribution, which can in fact be applied to a wide range of fields where distributions are observed to have a lognormal-like body and a power-law tail. We derive important properties of the MLP distribution, like the cumulative distribution, the mean, variance, arbitrary raw moments, and a random number generator. These analytic properties of the distribution can be used to facilitate application to modelling the IMF. We demonstrate how the MLP function provides an excellent fit to the IMF compiled by Chabrier and how this fit can be used to quickly identify quantities like the mean, median, and mode, as well as number and mass fractions in different mass intervals.

Key words: accretion, accretion discs – stars: formation – stars: luminosity function, mass function.

1 INTRODUCTION

The distribution of stellar and substellar masses at birth, the initial mass function (IMF), is a key feature of star formation. It has been studied intensely since first estimated by Salpeter (1955), who measured a power-law tail for high masses of the approximate form $\mathrm{d}N/\mathrm{d}\ln M \propto M^{-1.35}$. Subsequent work has established a shallower slope at masses less than $\mathrm{M}_{\odot}$ and a turnover at approximately $0.2\,\mathrm{M}_{\odot}$ when the masses are placed in logarithmically spaced bins. These compilations of the IMF have tended to identify a power-law profile at high masses (e.g. Scalo 1998; Kroupa 2001, 2002), although earlier work (Miller & Scalo 1979) did fit the IMF with a lognormal distribution. Chabrier (2003, see also Chabrier 2005) has compiled an IMF in the substellar and low-mass stellar regime and advocates a lognormal fit for masses up to $\mathrm{M}_{\odot}$ and a power-law fit for $M > \mathrm{M}_{\odot}$. By appealing to the Central Limit Theorem (CLT), Chabrier (2003) claims a better rationale for the lognormal fit at low masses than the approach of using broken three-component power laws (Scalo 1998; Kroupa 2001, 2002). However, Chabrier's overall fit also requires joining a lognormal with a power law at high masses, so has one joining condition instead of two. An ideal next step is to find a single function with no joining conditions that has a rationale that is at least on par with appeals to the CLT.

There are many disciplines in which a desired distribution is one that is like a lognormal at low and intermediate values, with a characteristic peak and turnover, but transitions to a power-law distribution at high values. Besides astronomy, this need has arisen in fields as diverse as biology (Limpert, Stahel & Abbt 2001), computer science (Mitzenmacher 2006), ecology (Allen, Li & Charnov 2001), and finance (Mandelbrot 1997), to name several. A review of common statistical resources reveals that very few analytic functions of this type exist, hence the attempts to fit empirical distributions by patching together different functions over different domains. Power-law distributions were first introduced by Pareto (1896) to explain the distribution of incomes seen in data from many different countries. The distribution of city sizes also shows a power-law character and regularity across many countries. It is referred to as Zipf's Law (Zipf 1949). Henceforth, we refer to pure power-law distributions as Pareto distributions, in conformity with much of the statistical literature.

In this paper, we analyse and characterize the properties of a hybrid three-parameter probability density function (pdf) introduced by Basu & Jones (2004). We feel that it can be used to fruitfully model data sets that exhibit both lognormal-like and power-law behaviour. Indeed, the modified lognormal with a power-law (MLP) distribution illustrates the fact that many generative processes that lead to a lognormal distribution, can with some modification, yield a power-law tail instead. As a result of its origin as a modified lognormal, two parameters of the MLP distribution are identified as μ_0 and σ_0^2, preserving the notation of the lognormal distribution, while the third parameter is α, the power-law index that also characterizes the

★E-mail: basu@uwo.ca
†Current affiliation: School of Information Studies, McGill University, Montreal, Canada.

Pareto distribution. However, the parameters μ_0 and σ_0^2 no longer represent the mean and variance, respectively, of the logarithm of the random variable, as they do for the lognormal distribution. This three-parameter function is a simpler and more readily usable form of a more general four-parameter pdf derived by Reed in several papers (Reed 2002; Reed 2003). The latter function arises from a stochastic growth law, Geometric Brownian Motion, rather than the pure exponential growth used by Basu & Jones (2004). An advantage of the pdf we use here is that it introduces only one additional parameter beyond that in the lognormal. As a result, it is a natural first step when fitting data that may look like a modified lognormal, and its relatively compact analytic closed-form expression makes it easy to use with common fitting techniques.

The model of Basu & Jones (2004) falls into one of the two major categories of IMF models in modern day star formation theory, that of an accretion based scenario in which temporal effects are important. The basic idea is that star formation is a killed (or stopped) process, with mass accretion terminated by events such as stellar outflows, ejection from the mass reservoir, or any other reason for emptying the mass reservoir. The distribution of accretion lifetimes then plays a key role in setting the shape of the IMF, while the Jeans mass does not. Other models in this category include those of Adams & Fatuzzo (1996), Bate & Bonnell (2005), and Myers (2009, 2014), with the latter developing a model that can fit the observed IMF by tuning several parameters. The alternate scenario is one in which the mass distribution is determined by the spatial properties of the gas distribution and by gravitational instability, so that some combination of the turbulent spectrum and the Jeans mass set the IMF, e.g. Padoan & Nordlund (2002); Hennebelle & Chabrier (2009), with the latter developing an analytic approach. Although the fitting of analytic functions to the IMF cannot alone settle which models may be most suitable, they can however greatly facilitate analysis in fields like galaxy studies, where the conclusions depend strongly on an adopted IMF model. Analytic functions also allow a simpler analysis of the effect of varying IMF parameters.

The paper is organized as follows. In Sections 2 and 3, we introduce the lognormal and Pareto distributions, respectively, and present some of their relevant properties. This is done for completeness of the presentation and to add context when reading the new results about the MLP distribution. In Section 4, we discuss the formulation of Basu & Jones (2004) that leads to the MLP distribution. We then examine some relevant mathematical properties of this distribution in Section 5, which includes expressions for its cumulative distribution function (cdf), mean and variance, arbitrary raw moments, and an approximation to its mode. These expressions, excluding the approximate mode, are shown to reduce to the corresponding lognormal expressions in the appropriate limit. In Section 6, we fit the MLP distribution to the IMF of Chabrier (2005) and then use the analytic function to quickly estimate some of the above described IMF properties, as well as a cumulative mass fraction. Some closing remarks are given in Section 7. The derivations of the expressions given in Section 5 are included in Section B, which also contains relevant properties of the error and complementary error functions and related integrals used in this paper.

2 THE LOGNORMAL DISTRIBUTION

According to the CLT of probability theory (Gut 2005), if $X_1, \ldots, X_n$ are identically distributed, independent random variables with mean μ and standard deviation σ, then $Z = \left(\sum X_i - n\mu\right)/(\sqrt{n}\sigma)$ converges in distribution to the standard normal variable $N(0, 1)$ with zero mean and variance of unity. In addition, the identical distribution assumption for $X_1, \ldots, X_n$ may be dropped, and the result will also follow provided certain conditions (Lindeberg's) are satisfied (Gut 2005). As pointed out in Golberg (1984) the CLT is used to partially justify why so many observable phenomena appear to be normally distributed: if the variable of interest is thought to be influenced by the sum of a large number of independent factors, then the CLT can be invoked to explain the apparent normality. However, it is frequently the case that the variables of interest are non-negative, whereas normal variables are not. One way to circumvent this complication is to have a distribution with similar properties but which is always positive. Suppose that Z above was instead assumed to be of the form $Z = X_1 \cdot X_2 \cdot \ldots \cdot X_n$. Then $W = \ln Z = \sum Y_i$, where $Y_i = \ln X_i$. Note that $Y_1, \ldots, Y_n$ satisfy the conditions of the CLT, so that W converges to a normal random variable, $W \sim N(\mu, \sigma^2)$. Moreover, if W is normal, then $Z = \mathrm{e}^W$ has a lognormal density (Aitchison & Brown 1957; Golberg 1984; Crow & Shimizu 1988),

$$f_Z(z; \mu, \sigma^2) = \frac{1}{z\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln z - \mu)^2}{2\sigma^2}\right), \quad z > 0. \tag{1}$$

Thus, under the assumptions above, Z converges to a lognormal. We list some of the relevant properties of the lognormal distribution below.

(i) Raw moments:

$$E[Z^k] = \exp\left(k\mu + \frac{1}{2}k^2\sigma^2\right); \tag{2}$$

(ii) Mode:

$$z_m = \exp(\mu - \sigma^2); \tag{3}$$

(iii) Variance:

$$\mathrm{Var}(Z) = \exp\left(2\mu + \sigma^2\right)[\exp(\sigma^2) - 1]; \tag{4}$$

(iv) Cumulative distribution function (cdf):

$$F_Z(z; \mu, \sigma^2) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln z - \mu}{\sqrt{2}\sigma}\right)\right], \tag{5}$$

where erf(z) is the error function (see equation A1 in the appendix).

3 THE PARETO DISTRIBUTION

Pareto distributions have been used extensively to model a wide variety of phenomena in the sciences and social sciences, such as the size of forest fires, the intensity of earthquakes, the citations to papers, and the population of cities. For a recent review, see Clauset, Shalizi & Newman (2009).

A random variable X has a Pareto distribution if its pdf is

$$f(x) = \frac{\alpha b^\alpha}{x^{\alpha+1}}, \quad x > b > 0 \tag{6}$$

(Pareto 1896; Mitzenmacher 2006; Seggern 1990), where $\alpha > 0$. We list some of its more relevant properties below.

(i) Raw moments:

$$E[X^n] = \frac{\alpha b^n}{\alpha - n}, \quad \alpha > n; \tag{7}$$

(ii) Mode:

$$x_m = b; \tag{8}$$

(iii) Variance:

$$\sigma^2 = \frac{\alpha b^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2; \tag{9}$$

(iv) Cumulative distribution function (cdf):

$$F_X(x; \alpha, b) = 1 - \left(\frac{b}{x}\right)^{\alpha}. \tag{10}$$

4 THE MLP DISTRIBUTION

The key element in the derivation of the MLP distribution is that even though the initial values of a random variable may have a lognormal density, later time evolution could in fact skew their distribution. In other words, the subsequent time of growth of the random variable (representing a physical quantity) is itself another random variable. Our pdf can be derived analytically on this basis using a few simplifying assumptions.

Imagine that initial values of a quantity M_0 are drawn randomly from a lognormal distribution with parameters μ_0 and σ_0. If the subsequent growth of an object with $M_0 = m_0$ is characterized by exponential growth

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \gamma\, m, \tag{11}$$

with fixed growth rate γ, then the multiplicative relation between values m of individual objects is preserved. A lognormal distribution is then maintained with the same σ_0 but the mean of the logarithmic values shifts to $\mu_0 + \gamma t$ after a fixed time t. However, we can treat the time as a random variable and draw it from an exponential pdf

$$f(t) = \delta\, \mathrm{e}^{-\delta t}, \tag{12}$$

where δ is a stopping rate. In this case, we can derive the pdf of final masses m as

$$f(m) = \int_0^{\infty} \frac{\delta \mathrm{e}^{-\delta t}}{\sqrt{2\pi}\sigma_0 m} \exp\left(-\frac{[\ln m - \mu_0 - \gamma t]^2}{2\sigma_0^2}\right) \mathrm{d}t. \tag{13}$$

Using the identity in equation (A9) we find the closed form

$$f(m) = \frac{\alpha}{2} \exp\left(\alpha\mu_0 + \alpha^2\sigma_0^2/2\right) m^{-(1+\alpha)} \times \mathrm{erfc}\left(\frac{1}{\sqrt{2}}\left(\alpha\sigma_0 - \frac{\ln m - \mu_0}{\sigma_0}\right)\right), \quad m \in [0, \infty), \tag{14}$$

where $\alpha = \delta/\gamma$.

Fig. 1 shows the MLP density $f(m)$ (equation 14) and a lognormal distribution for similar values of parameters. The specific values of the parameters used here are somewhat arbitrary, but correspond approximately to a best-fitting lognormal for the low-mass end of a stellar mass distribution, as modelled in Basu & Jones (2004). While μ_0 is largely a scale-dependent parameter, the other two parameters are expected to fall in the approximate range $0 \le \sigma_0 \le 1$ and $1 \le \alpha \le 2$ based on fits of lognormal and power-law distributions to a wide range of phenomena in the sciences and social sciences (Limpert et al. 2001; Clauset et al. 2009). The function derived here is related to a four-parameter distribution (Reed 2002; Reed 2003) that can be derived under the assumption of geometric Brownian motion, $\mathrm{d}m = \gamma\, m\, \mathrm{d}t + a\, m\, \mathrm{d}w$, which is a stochastic growth law where $\mathrm{d}w$ represents $R(0, 1)$, the uniform random variate in the interval [0, 1], and a is an amplitude of the fluctuations. The resulting pdf is a double-tailed Pareto distribution, with coefficients that are roots of a quadratic equation. The pdf we derive here results from a simpler model and has the advantage of a single-expression closed form. It may also more easily correspond to many data sets that seemingly warrant only one additional parameter (beyond the original two of the lognormal distribution) in order to get a reasonable fit and quantify the data.

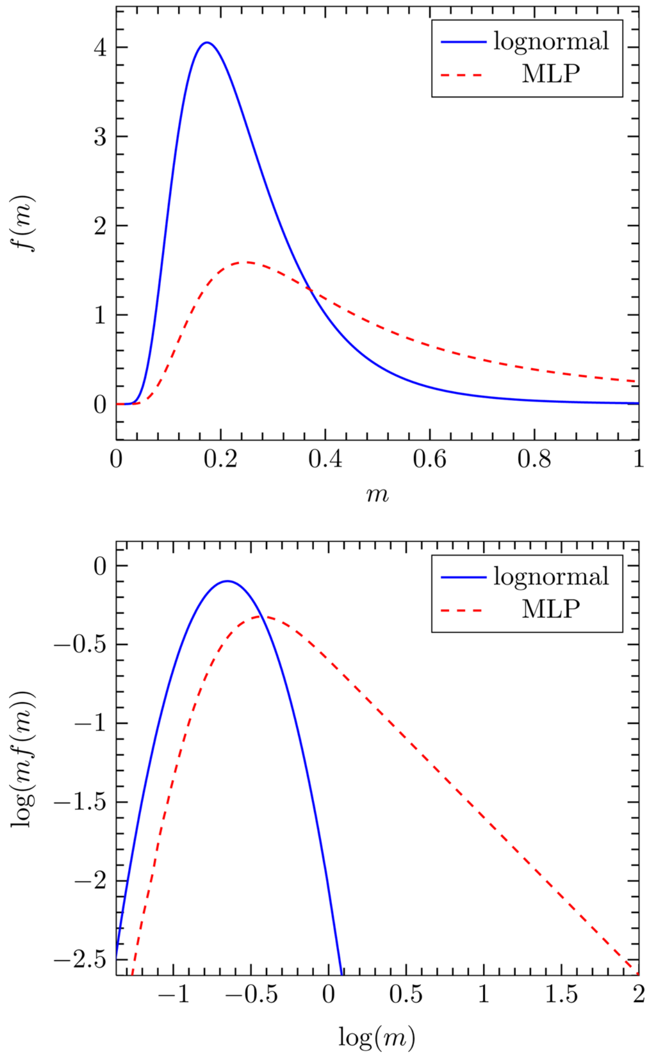

Figure 1. Comparison of a lognormal density function with $\mu = -1.5$ and $\sigma = 0.5$ with the MLP density function with $\mu_0 = -1.5$, $\sigma_0 = 0.5$, and $\alpha = 1$. The mass is normalized by a solar mass $\mathrm{M}_{\odot}$.

5 RELEVANT MATHEMATICAL PROPERTIES OF THE MLP DISTRIBUTION

A few relevant mathematical properties of the MLP distribution are examined below, leaving most of the necessary calculations in Section B. Let M denote a random variable with an MLP distribution of density $f(m)$.

5.1 Cumulative distribution

The MLP cdf,

$$F(m) = \int_{-\infty}^{m} f(t)\mathrm{d}t, \tag{15}$$

is given by

$$F(m) = \frac{1}{2}\mathrm{erfc}\left(-\frac{\ln(m)-\mu_0}{\sqrt{2}\sigma_0}\right) - \frac{1}{2}\exp\left(\alpha\mu_0+\frac{\alpha^2\sigma_0^2}{2}\right) m^{-\alpha}\mathrm{erfc}\left(\frac{\alpha\sigma_0}{\sqrt{2}}-\frac{\ln(m)-\mu_0}{\sqrt{2}\sigma_0}\right). \quad (16)$$

Note that as $m \to 0$, we can see by the result in equation (B3) that the behaviour of $F(m)$ is dominated by the first term, which is exactly the cdf for a lognormal random variable (equation 5) with parameters μ_0 and σ_0.

5.2 Mean, variance, and raw moments

The kth raw moment of M, defined as the expectation value of M^k,

$$E[M^k] = \int_0^\infty m^k f(m)\mathrm{d}m\,, \quad (17)$$

exists if and only if $\alpha > k$, in which case it is given by

$$E[M^k] = \frac{\alpha}{\alpha-k}\exp\left(\frac{\sigma_0^2 k^2}{2}+\mu_0 k\right),\ \alpha > k. \quad (18)$$

Note that this expression is exactly the formula for the raw moments of a lognormal distribution (equation 2) with parameters μ_0 and σ_0, scaled by the factor $\alpha/(\alpha - k)$, and in the limit as $\alpha \to \infty$, the expressions are identical. This is consistent with the derivation of the MLP distribution in Section 4. The limit $\alpha \to \infty$ corresponds to $\gamma \to 0$ for finite death rate δ, so that the drift term vanishes, and the distribution remains a lognormal with mean $\mu = \mu_0$. Assuming $\alpha > k$, we can obtain the following expressions, including the mean and variance of the distribution:

$$E[M] = \frac{\alpha}{\alpha-1}\exp\left(\frac{\sigma_0^2}{2}+\mu_0\right),\ \alpha > 1, \quad (19)$$

$$E[M^2] = \frac{\alpha}{\alpha-2}\exp\left(2[\sigma_0^2+\mu_0]\right),\ \alpha > 2,$$

$$\mathrm{Var}(M) = E[M^2] - (E[M])^2 \quad (20)$$

$$= \alpha\exp(\sigma_0^2+2\mu_0)\left(\frac{\mathrm{e}^{\sigma_0^2}}{\alpha-2}-\frac{\alpha}{(\alpha-1)^2}\right),\ \alpha > 2. \quad (21)$$

Higher moments around the mean can be computed using equation (18) with the identity

$$E[(M-E[M])^n] = \sum_{j=0}^{n}\binom{n}{j}E[M^j](-1)^{n-j}(E[M])^{n-j}. \quad (22)$$

5.3 Mode

To find the mode, that is, the value m^* that maximizes the MLP pdf in equation (14), we must solve the transcendental equation

$$f'(m) = 0 \iff K\mathrm{erfc}(u) = \mathrm{e}^{-u^2}, \quad (23)$$

where

$$K = \sigma_0(\alpha+1)\sqrt{\frac{\pi}{2}},\ u = \tfrac{1}{\sqrt{2}}\left(\alpha\sigma_0 - \tfrac{\ln m - \mu_0}{\sigma_0}\right). \quad (24)$$

Although the solution to equation (23) will generally require the implementation of numerical methods, we note that if $K \approx 1$ then $u = 0$ provides an approximate solution to equation (23), which in terms of the original parameters results in

$$m^* = \exp\left(\mu_0 + \alpha\sigma_0^2\right). \quad (25)$$

The approximation in equation (25) is useful only when the assumption $K \approx 1$ is closely met, and behaves poorly otherwise. However, when a precise numerical solution is required, one can use this approximation as a starting point in the iteration procedure being implemented. It is worth noting that even if one tried to find the peak in the space $mf(m)$ versuslog (m), the resulting equation to be solved is in fact the same, owing to the fact that $mf(m)$ is just $f(m)$ with a different value of α, and $\mathrm{d}/\mathrm{d}(\log(x)) = x(\mathrm{d}/\mathrm{d}x)$.

5.4 Random number generation

For practical purposes of comparing data sets with a model distribution, it is valuable to be able to draw a random sample from the model. Using the definition of the lognormal random variable (see Section 2), a random number drawn from a lognormal pdf (equation 1) is an element of the lognormal random variate

$$L(\mu,\sigma) \sim \exp(\mu + \sigma\,N(0,1)), \quad (26)$$

where $N(0, 1)$ is the normal random variate with zero mean and variance of unity. For drawing from an exponential pdf (equation 12), the exponential random variate is

$$E(\delta) \sim -\delta^{-1}\ln(R(0,1)), \quad (27)$$

where $R(0, 1)$ is the uniform random variate in the interval [0, 1]. The above formula can be obtained from equation (12) through the general method of calculating the cdf, inverting it, and then drawing the argument as an element of the uniform random variate $R(0, 1)$. We can use equations (26) and (27) to derive a random variate for the MLP distribution. We note that the MLP distribution (equation 14) is formally obtained from an initial lognormal pdf with mean μ_0 under the transformation $\mu_0 \to \mu_0 + \gamma\,t$, in which t is chosen randomly from an exponential distribution. Therefore, we can use equations (26) and (27) to write that the MLP random variate is

$$M(\mu_0,\sigma_0,\alpha) \sim \exp(\mu_0 + \sigma_0\,N(0,1) - \alpha^{-1}\ln(R(0,1))). \quad (28)$$

Equation (28) can be used to draw a random sample from the MLP pdf just as equation (26) can be used to draw a random sample from a lognormal pdf.

6 FITTING THE MLP FUNCTION TO THE IMF

Here, we find a set of parameters for the MLP distribution that fit the updated IMF of Chabrier (2005). We use a normalized form of equation (1) in Chabrier (2005). It is divided into logarithmic bins of mass normalized by $\mathrm{M}_\odot$ with intervals of 0.05 over the range of 0.06–100 $\mathrm{M}_\odot$ and we use a Levenberg–Marquardt least-squares minimization method to fit the MLP function. The best-fitting parameters are $\mu_0 = -2.404$, $\sigma_0 = 1.044$, and $\alpha = 1.396$. Fig. 2 shows the best-fitting MLP function along with the Chabrier function, both in normalized form. Fig. 3 shows three sets of random samples drawn from this best-fitting MLP function, using equation (28) and sample sizes of 100, 1000, and 10 000, respectively. The samples are made into histograms with logarithmic bin width of 0.2, and each histogram is represented by a different symbol. We find that samples of well over 100 are needed to adequately populate the power-law tail. For MLP samples of about 100 or less, fitting a lognormal pdf

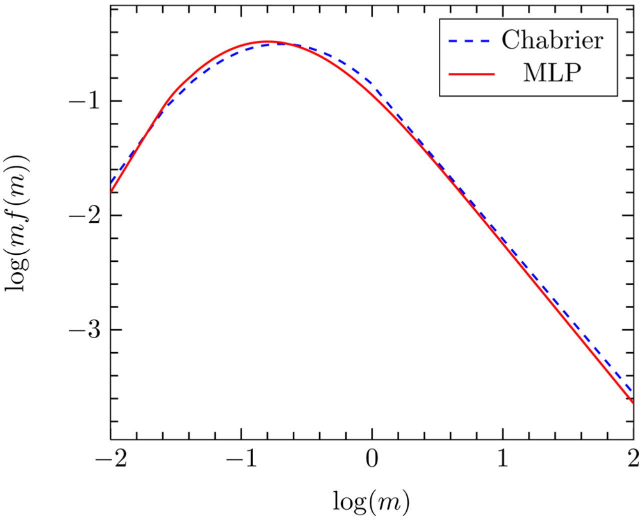

Figure 2. Comparison of the Chabrier IMF with its corresponding best-fitting MLP, where $\mu_0 = -2.404$, $\sigma_0 = 1.044$, and $\alpha = 1.396$. The mass is normalized by $\mathrm{M}_{\odot}$.

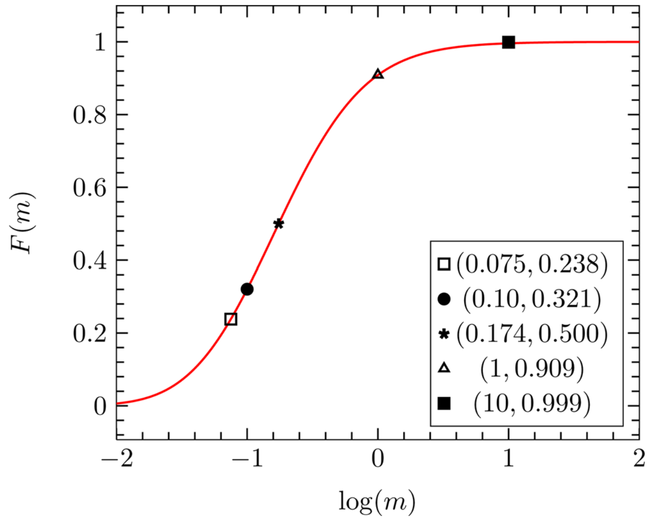

Figure 4. MLP cumulative distribution with $\mu_0 = -2.404$, $\sigma_0 = 1.044$, and $\alpha = 1.396$. The mass is normalized by $\mathrm{M}_{\odot}$. The legend illustrates interesting values of the pairs (m, $F(m)$).

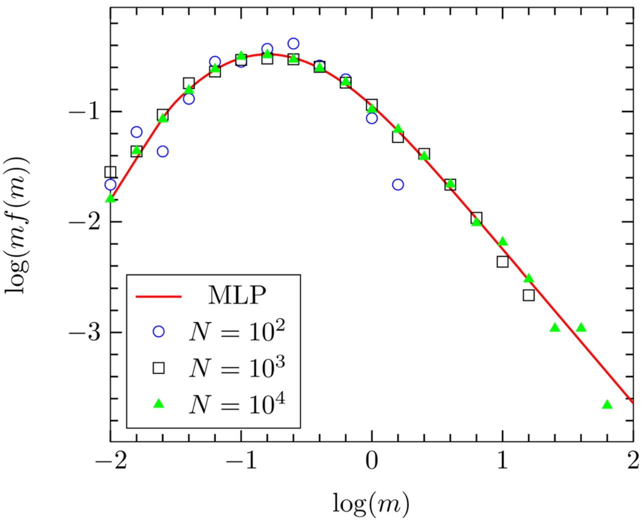

Figure 3. The MLP function with $\mu_0 = -2.404$, $\sigma_0 = 1.044$, and $\alpha = 1.396$, overlaid with histogram values for random samples drawn from the distribution with size $N = 100$ (blue circles), $N = 1000$ (black squares), and $N = 10\,000$ (green triangles). All histograms are binned in increments $\Delta \log m = 0.2$, and histogram values are the fractional number in each bin divided by $\Delta \ln m$.

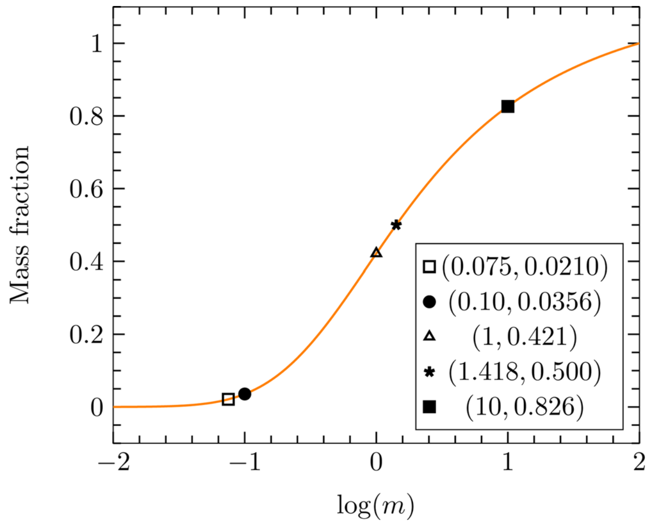

Figure 5. Mass fraction for an MLP distribution with $\mu_0 = -2.404$, $\sigma_0 = 1.044$, and $\alpha = 1.396$. The mass is normalized by $\mathrm{M}_{\odot}$. The legend illustrates interesting values of pairs of m and the mass fraction.

typically yields a better goodness-of-fit statistic than the MLP distribution, given that it also has one less fitting parameter. Some sample routines for fitting the MLP function or drawing random samples from it can be found at http://www.astro.uwo.ca/~basu/mlp.html.

With a best-fitting MLP function in hand, it is relatively easy to quantify several important properties of the IMF. Equation (19) can be used to calculate the mean mass in the distribution; it is 0.55 $\mathrm{M}_{\odot}$. If we were to truncate the distribution at 100 $\mathrm{M}_{\odot}$, the mean mass is about 10 per cent lower at 0.49 $\mathrm{M}_{\odot}$. The mode of the distribution $f(m)$ is calculated numerically to be 0.04 $\mathrm{M}_{\odot}$. However, the usual practice is to plot data as $\mathrm{d}N/\mathrm{d}\ln M \equiv m f(m)$ rather than $\mathrm{d}N/\mathrm{d}M \equiv f(m)$, so the usually quoted IMF peak is that of $m f(m)$. We find this to be 0.16 $\mathrm{M}_{\odot}$ for our best-fitting parameters. The mode of $m f(m)$ is found analogously to that for $f(m)$ described in Section 5.3, but with α lowered by one integer value.

Fig. 4 shows the cumulative MLP distribution for the best-fitting parameters. The legend panel in Fig. 4 illustrates the specific values for $F(m)$ for several relevant mass values, e.g. one solar mass, as well as the minimum stellar mass, $m = 0.075\,\mathrm{M}_{\odot}$ (Chabrier & Baraffe 2000). Although the data used to generate the IMF quoted by Chabrier (2005) does not span the entire hypothetical mass range $m \in [0, \infty]$, one can use the numbers in Fig. 4 to quickly estimate some highlights of the mass function, with interesting pairs of m, $F(m)$ marked by symbols in Fig. 4: about one quarter of objects are substellar, about one third of objects have masses less than 0.1 $\mathrm{M}_{\odot}$, the median mass is 0.17 $\mathrm{M}_{\odot}$, just over 90 per cent of objects have masses less than 1 $\mathrm{M}_{\odot}$, and only 9 per cent of objects have masses in the range 1–10 $\mathrm{M}_{\odot}$. Fig. 5 shows the mass fraction $\phi(m) = \int_0^m m' \, f(m') \, \mathrm{d}m'$ (with closed form given in equation B15) for the best-fitting parameters, normalized by the value at $m = 100\,\mathrm{M}_{\odot}$.

Symbols denote interesting pairs of m, $\phi(m)$, and many provide an interesting contrast to their counterparts in Fig. 4. For example, only about 2 per cent of the total mass is in substellar objects, a majority of mass is tied up in objects more massive than 1 $M_{\odot}$, and half of all mass is in objects above mass 1.4 $M_{\odot}$.

The ability to model stellar populations with an IMF that has clearly identifiable number or mass fractions in different mass ranges should be a key aid to galaxy studies (e.g. Kennicutt 1998), for example in the determination of star formation rates, chemical evolution, and mass-to-light ratios.

7 CONCLUSIONS

We have derived several important properties of the MLP probability distribution function that has been recently introduced in the literature. The three-parameter MLP function has the salutary properties of a main body that resembles a lognormal, including a peak value and a decline towards low values, as well as a power-law tail at high values. This function can potentially be applied in a variety of fields where empirical distributions may have a power-law tail that coexists with a peak at lower values. The MLP distribution can also help to settle a frequent contentious question: is a data set consistent with a lognormal distribution or does it also show evidence for a power-law tail? Simultaneous fitting of the lognormal and MLP distributions to the same data sets can help to answer this question in many fields.

Comparison of real data sets with the MLP distribution is facilitated by the results presented in this paper. We have derived analytic expressions for the cdf, the mean, variance, and higher moments of the distribution. We also derived an approximation for the mode that can serve as an initial guess for non-linear solvers that can iterate to a more exact solution. The random variate of the MLP distribution has also been introduced. Together, these results can put the MLP distribution on a more equal footing with many classical distributions (e.g. lognormal, exponential, Pareto, Rayleigh) that are frequently used to fit empirical data. The use of the MLP distribution to fit an empirical data set has been demonstrated and we show how useful information about the mean, mode, median, and distributions of number or mass fractions can be easily calculated.

ACKNOWLEDGEMENTS

SB acknowledges the hospitality of the Isaac Newton Institute of Mathematical Sciences at Cambridge University during the initial stages of writing of this paper. This work was supported by an NSERC Discovery Grant to SB.

REFERENCES

Abramowitz M., Stegun I. A., 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Press, New York
Adams F. C., Fatuzzo M., 1996, ApJ, 464, 256
Aitchison J., Brown J. A. C., 1957, The Lognormal Distribution. Cambridge Univ. Press, London
Allen A. P., Li B. L., Charnov E. L., 2001, Ecology Lett., 4, 1
Basu S., Jones C. E., 2004, MNRAS, 347, L47
Bate M. R., Bonnell I. A., 2005, MNRAS, 356, 1201
Chabrier G., 2003, PASP, 115, 763
Chabrier G., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library, Vol. 327, The Initial Mass Function 50 Years Later. Springer-Verlag, Dordrecht, p. 41
Chabrier G., Baraffe I., 2000, ARA&A, 38, 337
Clauset A., Shalizi C. R., Newman M. E. J., 2009, SIAM Rev., 51, 661
Crow E. L., Shimizu K., 1988, Lognormal Distributions: Theory and Applications. CRC Press, New York, NY
Golberg M. A., 1984, An Introduction to Probability Theory with Statistical Applications. Plenum, New York
Gut A., 2005, Probability: A Graduate Course, Springer-Verlag, Berlin
Hennebelle P., Chabrier G., 2009, ApJ, 702, 1428
Kennicutt R. C., Jr 1998, in Gilmore G., Howell D., eds, ASP Conf. Ser. Vol. 142, The Stellar Initial Mass Function. Astron. Soc. Pac., San Francisco, p. 1
Kroupa P., 2001, MNRAS, 322, 231
Kroupa P., 2002, Science, 295, 82
Limpert E., Stahel W. A., Abbt M., 2001, Bioscience, 51, 341
Mandelbrot B., 1997, Fractals and Scaling in Finance. Springer-Verlag, New York, NY
Miller G. E., Scalo J. M., 1979, ApJS, 41, 513
Mitzenmacher M., 2006, Internet Math., 1, 226
Myers P. C., 2009, ApJ, 706, 1341
Myers P. C., 2014, ApJ, 781, 33
Padoan P., Nordlund Å., 2002, ApJ, 576, 870
Pareto V., 1896, Cours d'Économie Politique. Droz, Geneva
Reed W. J., 2002, J. Reg. Sci., 42, 1
Reed W. J., 2003, Physica A, 319, 469
Salpeter E. E., 1955, ApJ, 121, 161
Scalo J. M., 1998, in Gilmore G., Howell D., eds, ASP Conf. Ser. Vol. 142, The Stellar Initial Mass Function. Astron. Soc. Pac., San Francisco, p. 201
Seggern D. H. V., 1990, CRC Handbook of Mathematical Curves and Surfaces, CRC Press, Boca Raton, FL
Zipf G. K., 1949, Human Behaviour and the Principle of Least Effort, Addison-Wesley, Reading, MA

APPENDIX A: ERROR FUNCTION

A1 Definition and basic properties

The Error function, denoted by erf(x), and the complementary error function, denoted by erfc(x), are defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t\,, \tag{A1}$$

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \mathrm{e}^{-t^2} \mathrm{d}t. \tag{A2}$$

Note from the definition that erf(x) is an odd function. Also,

$$\lim_{x\to\infty} \mathrm{erf}(x) = 1 \tag{A3}$$

(Abramowitz & Stegun 1964). The derivative and integral of erf(x) are

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\mathrm{e}^{-x^2}\,, \tag{A4}$$

$$\int \mathrm{erf}(x)\mathrm{d}x = x\,\mathrm{erf}(x) + \frac{1}{\sqrt{\pi}}\mathrm{e}^{-x^2} + C\,, \tag{A5}$$

where C is an integration constant. Finally, the Taylor expansion for erf(x) is given by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)n!} = \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{3} + \frac{x^5}{10} + \cdots \right). \tag{A6}$$

A plot of erfc(x) is presented in Fig. A1.

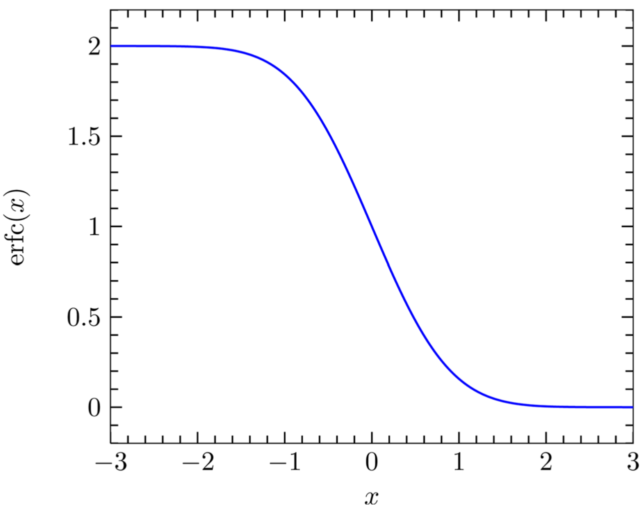

Figure A1. Complementary error function, erfc(x).

A2 Related Integrals used in our analysis

Consider the integral

$$I_1 = \int \exp[-(ax^2 + bx + c)]\mathrm{d}x\ ,\ a > 0. \tag{A7}$$

By completing the square and letting $u = \sqrt{a}\left(x + \frac{b}{2a}\right)$ it can be brought to

$$I_1 = \frac{1}{2}\sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2 - 4ac}{4a}\right)\mathrm{erf}\left(\sqrt{a}\left[x + \frac{b}{2a}\right]\right) + \tilde{C}\,, \tag{A8}$$

where $\tilde{C}$ is an integration constant. In particular, using the properties of erf(x) and its relation to erfc(x), we have

$$\int_0^\infty \exp[-(ax^2 + bx + c)]\mathrm{d}x$$
$$= \frac{1}{2}\sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2 - 4ac}{4a}\right)\mathrm{erfc}\left(\frac{b}{2\sqrt{a}}\right). \tag{A9}$$

The result in equation (A8) can also be used to find integrals of the form

$$I_2 = \int x^{-(\alpha+1)}\exp\left[-(B + C\ln x)^2\right]\mathrm{d}x\,, \tag{A10}$$

where $C \neq 0$ and $x > 0$. By using the substitution $v = B + C\ln x$, we can bring it to the form

$$I_2 = \frac{1}{C}\int \exp\left[-\left(v^2 + v\frac{\alpha}{C} - \frac{\alpha B}{C}\right)\right]\mathrm{d}v\,, \tag{A11}$$

which we can evaluate to

$$I_2 = \frac{\sqrt{\pi}}{2C}\exp\left(\frac{\alpha^2 + 4\alpha BC}{4C^2}\right)\mathrm{erf}\left(B + C\ln x + \frac{\alpha}{2C}\right) + \tilde{C}. \tag{A12}$$

APPENDIX B: DERIVATION OF THE RESULTS IN SECTION 5

B1 The MLP as a pdf

We can explicitly verify that the MLP density $f(m)$ (equation 14) is indeed a valid pdf for all $\alpha > 0$, i.e. it is positive and integrates to 1. Positivity is clear from the definition. To verify the normalization condition, let $A = (\alpha/2)\exp\left(\alpha\mu_0 + \alpha^2\sigma_0^2/2\right)$, $B = (1/\sqrt{2})\,[\alpha\sigma_0 + (\mu_0/\sigma_0)]$, and $C = -1/(\sqrt{2}\sigma_0)$. We then have

$$\int_0^\infty f(m)\mathrm{d}m = A\int_0^\infty m^{-(1+\alpha)}\mathrm{erfc}(B + C\ln m)\mathrm{d}m. \tag{B1}$$

Letting $\mathrm{d}v = m^{-(\alpha+1)}\mathrm{d}m$ and $u = \mathrm{erfc}(B + C\ln m)$, we can write this integral as

$$\int_0^\infty f(m)\mathrm{d}m = -\frac{A}{\alpha}m^{-\alpha}\mathrm{erfc}(B + C\ln m)\Big|_0^\infty$$
$$-\frac{2AC}{\alpha\sqrt{\pi}}\int_0^\infty m^{-(\alpha+1)}\exp\left[-(B + C\ln m)^2\right]\mathrm{d}m. \tag{B2}$$

Consider the first term on the right-hand side in equation (B2). As $m \to \infty$, $\mathrm{erfc}(B + C\ln m) \to 2$ (since $C < 0$ and erf(x) is odd), hence for $\alpha > 0$ the term vanishes in this limit. As $m \to 0$, the limit takes on an indeterminate form. Applying L'Hospital's rule we can see that the limit is also zero in this case:

$$\lim_{m\to 0} m^{-\alpha}[\mathrm{erfc}(B + C\ln m)]$$
$$= -\frac{2C}{\sqrt{\pi}\alpha}\lim_{m\to 0}\frac{\mathrm{e}^{-(B+C\ln m)^2}}{m^\alpha}$$
$$= -\frac{2C}{\sqrt{\pi}\alpha}\lim_{x\to\infty}\exp\left[-x^2 - \frac{\alpha}{C}(x - B)\right] = 0, \tag{B3}$$

where $x = B + C\ln m$. We can evaluate the second (integral) term with equation (A12):

$$-\frac{2CA}{\alpha\sqrt{\pi}}\int_0^\infty m^{-(\alpha+1)}\exp\left[-(B + C\ln m)^2\right]\mathrm{d}m$$
$$= -\frac{A}{\alpha}\exp\left(\frac{\alpha^2 + 4\alpha BC}{4C^2}\right)\times\mathrm{erf}\left(B + C\ln x + \frac{\alpha}{2C}\right)\Big|_0^\infty$$
$$= \frac{2A}{\alpha}\exp\left(\frac{\alpha^2 + 4\alpha BC}{4C^2}\right) = 1. \tag{B4}$$

B2 Cumulative distribution

The cdf,

$$F(m) = \int_{-\infty}^m f(t)\mathrm{d}t\,, \tag{B5}$$

and for our case $f(m)$ is defined on $[0, \infty)$. To find the closed form, we first we apply integration by parts, using also the same definitions for A, B, and C as in Section B1:

$$F(m) = -\frac{A}{\alpha}t^{-\alpha}\mathrm{erfc}(B + C\ln t)\Big|_0^m$$
$$-\frac{2AC}{\alpha\sqrt{\pi}}\int_0^m t^{-(\alpha+1)}\exp\left[-(B + C\ln t)^2\right]\mathrm{d}t. \tag{B6}$$

We again use the identity in equation (A12) to evaluate the integral term and obtain

$$F(m) = -\frac{A}{\alpha}m^{-\alpha}\mathrm{erfc}(B + C\ln m)$$
$$+\frac{A}{\alpha}\exp\left(\frac{\alpha^2 + 4\alpha BC}{4C^2}\right)\mathrm{erfc}\left(B + C\ln m + \frac{\alpha}{2C}\right), \tag{B7}$$

which, upon returning to the original parameters, becomes

$$F(m) = \frac{1}{2}\mathrm{erfc}\left(-\frac{\ln(m) - \mu_0}{\sqrt{2}\sigma_0}\right) - \frac{1}{2}\exp\left(\alpha\mu_0 + \frac{\alpha^2\sigma_0^2}{2}\right) m^{-\alpha}\mathrm{erfc}\left(\frac{\alpha\sigma_0}{\sqrt{2}} - \frac{\ln(m) - \mu_0}{\sqrt{2}\sigma_0}\right). \tag{B8}$$

B3 Raw moments

Next we derive a closed form for arbitrary raw moments of the distribution, as well as an expression for its variance. Let M be an MLP random variable with pdf $f(m)$. The kth raw moment of M, defined as the expectation value of M^k, is given by

$$E[M^k] = \int_0^\infty m^k f(m)\mathrm{d}m = A\int_0^\infty m^{k-(\alpha+1)}\mathrm{erfc}(B + C\ln m)\mathrm{d}m\,, \tag{B9}$$

with A, B, and C as defined in Section B1. Before we arrive at a closed form expression for the moments, we consider the convergence of the integral in equation (B9). This integral diverges for $k \geq \alpha$. To see this, note that if $k \geq \alpha$ then $p = \alpha + 1 - k \leq 1$. Now write equation (B9) as

$$\int_0^\infty \frac{\mathrm{erfc}(B + C\ln m)}{m^p}\mathrm{d}m = \int_0^a \frac{\mathrm{erfc}(B + C\ln m)}{m^p}\mathrm{d}m + \int_a^\infty \frac{\mathrm{erfc}(B + C\ln m)}{m^p}\mathrm{d}m, \tag{B10}$$

where $a \in (0, \infty)$ such that $\mathrm{erfc}(B + C\ln m) \geq 1, \forall m \geq a$. The existence of such a is ensured by the fact that for $C < 0$, $\mathrm{erfc}(B + C\ln m)$ is a continuous strictly increasing positive function having an upper limit of 2. Then

$$\int_a^\infty \frac{1}{m^p}\mathrm{erfc}(B + C\ln m)\mathrm{d}m \geq \int_a^\infty \frac{1}{m^p}\mathrm{d}m\,, \tag{B11}$$

where the integral on the right-hand side diverges for $p \leq 1$. Thus, the integral in equation (B9) diverges for $k \geq \alpha$. Conversely, the moments are finite for $k < \alpha$. For any $a \geq 0$,

$$\int_a^\infty \frac{1}{m^p}\mathrm{erfc}(B + C\ln m)\mathrm{d}m \leq 2\int_a^\infty \frac{1}{m^p}\mathrm{d}m\,, \tag{B12}$$

which converges for $p = \alpha - k + 1 > 1$. Together with equations (B10) and (B3) this proves the existence of the moments. Suppose that $\alpha > k$. Then by simply letting $\alpha \rightarrow \alpha - k > 0$ in the integrand of equation (B1), we can see from equations (B3) and (B4) that the kth moment is given by

$$E[M^k] = \frac{2A}{\alpha - k}\exp\left(\frac{(\alpha - k)^2 + 4(\alpha - k)BC}{4C^2}\right) = \frac{\alpha}{\alpha - k}\exp\left(\frac{\sigma_0^2 k^2}{2} + \mu_0 k\right). \tag{B13}$$

B4 Mass fraction

We can obtain the expression for the mass fraction (up to a mass m),

$$\phi(m) = \int_0^m m' f(m')\mathrm{d}m' = A\int_0^m (m')^{-\alpha}\mathrm{erfc}(B + C\ln m')\mathrm{d}m',$$

using the same approaches employed in the computation of the cumulative distribution and the raw moments in the previous sections. Since the mean exists only if $\alpha > 1$, it is also natural to consider the closed-form expression for the mass fraction for $\alpha > 1$. This can be achieved by replacing all explicit appearances of α with $\alpha - 1$ in equation (B7), while maintaining the dependences in the parameters $A(\alpha)$, $B(\alpha)$, and $C(\alpha)$ in the original form, as we did for the calculation of the raw moments. Thus,

$$\phi(m) = -\frac{A}{\alpha - 1}m^{-(\alpha-1)}\mathrm{erfc}(B + C\ln m) + \left\{\frac{A}{\alpha - 1}\exp\left(\frac{(\alpha - 1)^2 + 4(\alpha - 1)BC}{4C^2}\right) \times \mathrm{erfc}\left(B + C\ln m + \frac{\alpha - 1}{2C}\right)\right\}, \alpha > 1, \tag{B14}$$

which, upon returning to the original parameters, becomes

$$\phi(m) = \frac{1}{2}\frac{\alpha}{\alpha - 1}\exp\left(\frac{\sigma_0^2}{2} + \mu_0\right)\mathrm{erfc}\left(\frac{\sigma_0}{\sqrt{2}} - \frac{\ln m - \mu_0}{\sqrt{2}\sigma_0}\right) - \left\{\frac{1}{2}\frac{\alpha}{\alpha - 1}\exp\left(\frac{\alpha^2\sigma_0^2}{2} + \alpha\mu_0\right) m^{-(\alpha-1)} \times \mathrm{erfc}\left(\frac{\alpha\sigma_0}{\sqrt{2}} - \frac{\ln m - \mu_0}{\sqrt{2}\sigma_0}\right)\right\}, \alpha > 1. \tag{B15}$$

Note that, as expected, $\phi(m)$ becomes the expression for the expectation value as $m \rightarrow \infty$ (for $\alpha > 1$).

This paper has been typeset from a TEX/LATEX file prepared by the author.