Modeling the Mass Function of Stellar Clusters Using the Modified Lognormal Power-Law Probability Distribution Function

Deepakshi Madaan
The University of Western Ontario

Supervisor
Dr. Shantanu Basu
The University of Western Ontario

Graduate Program in Applied Mathematics
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science
© Deepakshi Madaan 2016

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Applied Mathematics Commons, and the Stars, Interstellar Medium and the Galaxy Commons

Recommended Citation

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.
Abstract

We use the Modified Lognormal Power-law (MLP) probability distribution function to model the behaviour of the mass function (MF) of young and populous stellar populations in different environments. We begin by modeling the MF of NGC1711, a simple stellar population (SSP) in the Large Magellanic Cloud as a pilot case. We then use model selection criterion to rank different candidate models. Using the MLP we find that the stellar catalogue of NGC1711 follows a pure power-law behaviour below the completeness limit with the slope $\alpha = 2.75$ for $dN/d\ln m \propto m^{-\alpha+1}$ in the mass range $0.89 \, M_\odot$ to $7.75 \, M_\odot$. Furthermore, we explore that the MLP takes a truncated form for fixed stopping time for accretion. By using model selection criterion, we conclude that the MLP serves as the most useful candidate to model lognormal, power-law or hybrid behaviour of the MF.

Keywords: stellar clusters, star formation, luminosity function, mass function, data analysis, Magellanic Clouds, model selection, regression
Co-Authorship Statement

This thesis has been written by Deepakshi Madaan under the supervision of Dr. Shantanu Basu. The work in chapter two is also in collaboration with Sophia Lianou and is nearing submission to Monthly Notices of the Royal Astronomical Society. The work in chapter three is in preparation for submission.
Acknowledgements

First and foremost, I am extremely thankful to my supervisor Dr. Shantanu Basu for his immense support and his belief in my ability when I felt otherwise sometimes. I not only have learnt so much under his guidance, but have had as much fun learning as possible. His words and motivation, I know would help me go a long way. I am really thankful to him for also making all of us, the research group, feel like at home with the many get togethers and trips. I am also very thankful to Sophia Lianou, for having found not just a mentor/collaborator but also a friend in her. She is an inspiring woman who helps me do my work with more dedication and enthusiasm. Not to forget my research mates: Sayantan, Pranav, Mark, and Andy who have made this process so much easier and help me find a home away from home. Most importantly, I couldn’t have achieved any of this without the unconditional support of my family and the two friends: Alina and Nashi who have been through thin and thick with me. I would also like to extent my heartiest thanks to Audrey, Cinthia, Brian and Dr. Wahl for helping me in the transition and answering my never-ending inquiries. Dr. Corless, Dr. Barmby, Dr. Sigut, Dr. Houde, and Dr. Reid, all have helped me to engage in rigorous learning and critical thinking and I am really thankful for that. Lastly, I just want to say: Mom, Dad, didi and bhai, I hope I made you proud.
Contents

Abstract ii

Co-Authorship Statement ii

Acknowledgements iv

List of Figures vii

List of Tables viii

List of Appendices ix

1 Introduction 1

1.1 Mass Function .................................................. 2
    1.1.1 From Luminosity to Mass ................................ 3
1.2 IMF Models .................................................... 3
1.3 Modified Lognormal Power-Law Probability Distribution Function ........ 5
    1.3.1 Physical Motivation and Derivation ....................... 6
    1.3.2 PDF and Properties ...................................... 7
1.4 Young stellar populations ................................... 8
1.5 Model Fitting and Model Selection ................................ 9
    1.5.1 Non-Linear Regression .................................... 9
    1.5.2 Maximum Likelihood Estimation ........................... 10
1.6 Thesis Statement and Contribution ............................ 11
2 Modeling the Mass Function of NGC 1711 as a Case Study using the MLP
2.1 From Luminosity to Mass
2.1.1 Luminosity Function
2.1.2 Mass Luminosity Relation
2.2 Mass Function
2.3 MLP modeling of the Mass Function
2.3.1 Fitting Results
2.3.2 Truncated MLP
2.3.3 MLP as a hybrid
2.4 Summary

Bibliography

3 Model Selection: Which Model to choose for the IMF of Young Stellar Clusters?
3.1 Observational data
3.2 IMF Models and Parameter Estimation
3.3 Model Selection using Information Criterion
3.4 Discussion

Bibliography

4 Conclusion

Bibliography

A Derivation of the MLP

B Joint Likelihood Function

Curriculum Vitae
## List of Figures

2.1 Color Magnitude Diagram ........................................ 20
2.2 Apparent Magnitude Histogram ................................. 21
2.3 Luminosity Function Histogram ................................. 22
2.4 Isochrones Plot .................................................. 23
2.5 Mass-Magnitude Relation Plot .................................. 24
2.6 Mass Function Histogram ........................................ 24
2.7 Log Mass Function Plot ......................................... 25
2.8 MLP fit for Padova isochrone ................................... 26
2.9 MLP fit for MESA isochrone .................................... 26
2.10 Residuals Plot ................................................... 28
2.11 QQ Plot .......................................................... 28
2.12 Truncated MLP .................................................. 29
2.13 MLP Hybrid Plot ............................................... 31
3.1 ONC plot ........................................................ 41
3.2 NGC 1711 plot ................................................... 41
3.3 NGC 6611 plot ................................................... 42
3.4 NGC 2024 plot ................................................... 42
List of Tables

3.1 Chabrier IMF best fit parameters table ................................................. 40
3.2 Kroupa IMF best fit parameters table .................................................... 40
3.3 MLP best fit parameters table ................................................................. 40
3.4 ONC likelihood table ............................................................................. 44
3.5 NGC 1711 likelihood table ..................................................................... 44
3.6 NGC 6611 likelihood table ..................................................................... 44
3.7 NGC 2024 likelihood table ..................................................................... 45
List of Appendices

Appendix A Derivation of the MLP ........................................... 52
Appendix B Joint Likelihood Function ................................. 53
Chapter 1

Introduction

For any star with a given chemical composition, its mass determines its structure and evolution by the Vogt-Russell theorem [5]. Once the mass of the star is known, we can find various stellar properties such as luminosity, radius, and radiation spectrum. Also, various integrated properties of any group of stars, i.e. a star cluster or a galaxy, depend on how stellar masses are distributed into different mass intervals [54]. Hence, it is necessary to study the distribution of stellar masses at birth, known as the initial mass function (IMF), in different environments to understand stellar population evolution and further galaxy evolution.

The question of universality of the IMF in different environments, i.e. whether the shape of the IMF is universal for stellar populations formed under different cloud conditions, is one of the most fundamental questions in astrophysics today [46]. There are various schools of thought that favor universality [5, 33] while many others argue otherwise [18, 41]. Dib, in 2014, does a thorough investigation of the universality of the IMF using Bayesian statistics with a sample of eight young stellar clusters. He concludes that the shape of the IMF does depend on the environment of star formation [18]. Even though various studies have been done using Bayesian statistics to study universality of the IMF, the problem of model fitting and model selection (i.e. which model best represents the underlying population) has rarely been comprehensively addressed. Model selection is used to study and compare different cosmolog-
ical models [40, 39, 57, 60], but no such study for model selection of IMF models has yet been done.

1.1 Mass Function

The distribution of stellar masses at birth in a star forming event into different mass intervals is called the initial mass function (IMF). The process of star formation is a highly complex and difficult-to-predict transformation of molecular clouds in the interstellar medium, controlled by various physical mechanisms such as self-gravity, turbulence, and magnetic fields [32, 46, 8, 44]. Thus, the process of star formation can be considered as a stochastic process and the mass of a star a continuous random variable [54]. Therefore, we can model the fraction of stars in each mass interval i.e the mass function (MF) as a probability density function (pdf).

If mass \( m \) of a star is considered as a continuous random variable which is distributed according to a pdf \( f(m) \), assuming the pdf is independent of space and time, then \( f(m)dm \) gives the number of stars in some volume of space in the interval \([m, m + dm]\) [53, 54, 13],

\[
f(m) = \frac{d(N/V)}{dm},
\]

where \( N = \) Number of stars in the interval \([m, m + dm]\), \( V = \) Volume.

The usual practice is to divide the intervals into log masses i.e. take the pdf as \( f(\log m) \) which is called the MF i.e. the mass function,

\[
f(\log m) = \frac{d(N/V)}{d\log m},
\]

\[
f(\log m) = \frac{d(N/V)}{dm} \frac{dm}{d\log m},
\]

\[
f(\log m) = (\ln 10)m f(m),
\]

For simplicity, we take the MF as \( f(ln m) \) for our study i.e. \( f(ln m) = m f(m) \). \( f(log m) \) gives the number of stars in some volume of space in the interval \([log m, log m + d log m]\).
1.2 IMF Models

1.1.1 From Luminosity to Mass

The main source of direct and accurate measurement of the dynamical mass of any star is studying a binary system [2, 59]. The data on stellar masses are usually acquired from eclipsing binaries whereby using light curves and radial velocity measurements along with Kepler’s laws one can accurately determine the masses of individual stars [21, 61]. For stars that are not formed in binaries, masses are obtained from a luminosity to mass conversion. Mass is obtained from luminosity using theoretical stellar evolutionary models that give mass-luminosity relations (MLR) or mass-magnitude\(^1\) relations (MMR) [52, 55, 56, 27, 26, 25]. Once these models are computed, they are checked against the observed dynamical mass in binaries for authenticity. A number of such evolutionary\(^2\)/isochrone models\(^3\) have been derived over the years such as Baraffe [4], D’Antona & Mazzitelli [17], Padova [9], Geneva [36], MESA [51]. All these models differ in various physical inputs and initial conditions.

The masses of stars are derived from their luminosities using two different set of evolutionary models in Chapter 2. In Chapter 3, the data obtained for stellar masses are obtained directly from literature where various authors use different evolutionary models to obtain the MF.

1.2 IMF Models

The choice of the functional form of the pdf is significant since it is used as an important tool for various calculations in stellar population synthesis [10]. Predictions of luminosity functions of white dwarfs depend on the IMF [16]. We can also study the rate of formation of planetary nebulae using the IMF [50]. The IMF enters into the equations to study chemical evolution of galaxies [58]. Since many astrophysical studies depend on the functional form of the IMF, it is important to choose a simple analytical and integrable form of the pdf that

\(^1\) The brightness of a star measured on a logarithmic scale is the apparent magnitude of a star.

\(^2\) Evolutionary models are plotting of evolutionary tracks of stars of different masses on the Hertzsprung-Russell Diagram (H-R Diagram). H-R Diagram is a plot of effective temperature/spectral class vs. luminosity/absolute magnitude of a star.

\(^3\) Isochrones are tracks of stars on the H-R Diagram at a constant time instead for a star with a constant mass.
adequately represents the mass distribution. Not only that but more importantly the choice of
the functional form should underlie a physical motivation that bears origin from a theory of
star formation, even though the process of star formation is yet to be fully understood.

In 1955, Salpeter was the first to provide the stellar initial mass distribution with an analytic
power law pdf approximation: \( dN/dm \propto m^{-\alpha} \) with \( \alpha = 2.35 \) or \( dN/d\ln m \propto m^{-1.35} \) [53]. He
did so by studying the Luminosity function (LF) of main sequence stars of over the mass range
0.4 \( M_\odot \) < \( m \) < 10 \( M_\odot \) in the solar neighbourhood.

Subsequently, Miller and Scalo suggested a lognormal form for masses 0.1 \( M_\odot \) < \( m \) <
50 \( M_\odot \) on finding that the stellar mass distribution flattens for low mass stars [42]. Later in
1984, Zinnecker gave a theoretical explanation to Miller and Scalo’s lognormal form of the
IMF by invoking the Central Limit Theorem (CLT) [62]. According to the CLT, the sum of a
large number of independent and identically distributed random variables will follow a Gaus-
sonian distribution [1]. Since the process of star formation is a highly complex transformation
controlled by various physical processes, the formation of stellar mass can be considered as a
product of a large number of independent and identically distributed random variables deter-
mined by the processes. Thus by the CLT the log of the product of the random variables will
follow a Gaussian distribution, implying the stellar mass follows a lognormal pdf [62].

Chabrier proposed a lognormal form for the substellar and low mass stellar regime i.e. \( m \) <
1 \( M_\odot \) along with a power-law representation for intermediate and high mass stellar regime [14,
13],

\[
f(\log m) \propto \begin{cases} 
\frac{1}{\sqrt{2\pi}\sigma} \exp \left[ - \frac{(\log m - \log m_c)^2}{2\sigma^2} \right] & \text{for } m \leq 1 M_\odot \\
\frac{1}{m^{-\alpha+1}} & \text{for } m > 1 M_\odot 
\end{cases}
\]  

(1.5)

where \( \alpha > -1 \). The parameter \( m_c \) corresponds to the characteristic mass that the lognormal
takes which also is the mean of the distribution while \( \sigma \) represents the spread of the lognormal
distribution and \( \alpha \) is the slope of the power-law.

Kroupa gave a multisegment power law profile for substellar, low, and high stellar mass
1.3. Modified Lognormal Power-Law Probability Distribution Function

Basu & Jones in 2004 [7] introduced a hybrid three-parameter probability density function, the Modified Lognormal Power-Law (MLP) probability distribution function, to model the entire stellar mass regime as a single function. Except for the power law approximation introduced by Salpeter that is used to model stars above $1 \, M_\odot$, all other parameterized mass approximations need some sort of a joining condition that adds to the number of parameters involved. The MLP on the other hand doesn’t require a joining condition and is a function of only three parameters, one parameter more than the lognormal. As Chabrier suggested that the MF of any stellar population can be fitted by a lognormal distribution with a characteristic peak and turnover for sub stellar and low mass stellar regime and by a power-law for intermediate and high mass regime, the MLP can easily model the entire stellar mass regime showing both regimes [35, 34, 33].

\[
\begin{align*}
\text{f}(m) \propto m^{-\alpha} : & \quad \alpha = +0.3 \pm 0.7 \quad 0.01 \, M_\odot \leq m < 0.08 \, M_\odot \\
\text{f}(m) \propto m^{-\alpha} : & \quad \alpha = +1.3 \pm 0.5 \quad 0.08 \, M_\odot \leq m < 0.50 \, M_\odot \\
\text{f}(m) \propto m^{-\alpha} : & \quad \alpha = +2.3 \pm 0.3 \quad 0.50 \, M_\odot \leq m
\end{align*}
\]

The functional forms Salpeter, Chabrier and Kroupa are the most widely used functional forms for the MF with total number of parameters as 2, 4, and 4, respectively. With models like the Chabrier and Kroupa having to do with joining conditions or the lack of physical motivation, the need of a model to represent the MF of stellar populations as a single function with a simple analytical and integrable form bearing a physical motivation is noteworthy. Hence, the Modified Lognormal Power-law Probability (MLP) distribution function proposed by Basu & Jones in 2004 can be looked at as a good analytical approximation for the mass distribution of stellar populations [6].
lognormal-like and power law-like behaviour.

1.3.1 Physical Motivation and Derivation

Basu & Jones [7] used a statistical approach to model the subsequent growth of masses by accretion in the process of star formation, which results in a power-law distribution starting from an initial lognormal distribution.

Since the formation of a star is the result of various physical mechanisms, the mass of a star can be written as \( m = f_1 \times f_2 \times \ldots \times f_N \). Thus according to the CLT, for large \( N \), \( \ln m \) follows a Gaussian distribution [62] i.e. \( m \) follows a lognormal distribution with mean \( \mu_0 \) and standard deviation \( \sigma_0 \). Starting from an initial lognormal form, Basu and Jones explored the idea of the growth of the mass of a star due to accretion [7]:

\[
\frac{dm}{dt} = \gamma m, m(t) = m_0 \exp(\gamma t),
\]

(1.7)

\( m_0 \) is the initial mass that follows a lognormal distribution and \( \gamma \) is the growth rate. The mean of the distribution becomes \( \mu = \mu_0 + \gamma t \) while the standard deviation remains \( \sigma_0 \). Assuming an exponential distribution for accretion time i.e. \( f(t) = \delta e^{-\delta t} \) where \( \delta \) is the death rate for accretion, the pdf for stellar mass becomes\(^{4}\):

\[
\int_0^\infty \frac{1}{\sqrt{2\pi} \sigma_0 m} \exp\left(-\frac{(\ln m - \mu_0 - \gamma t)^2}{2 \sigma_0^2} \right) \delta e^{-\delta t} dt = \frac{\alpha_0}{2} \exp\left(\alpha_0 \mu_0 + \alpha_0^2 \sigma_0^2 / 2\right) m^{-(1+\alpha_0)} \times \\
\left(\text{erfc}\left(\frac{1}{\sqrt{2}} \left(\alpha_0 \sigma_0 - \frac{\ln m - \mu_0}{\sigma_0}\right)\right)\right)
\]

(1.8)

where \( \alpha_0 = \delta / \gamma \). The exponential growth of masses due to accretion gives a power-law tail to the underlying lognormal distribution of initial masses.

\(^{4}\)Refer to Appendix A for the integration.
1.3. Modified Lognormal Power-Law Probability Distribution Function

1.3.2 PDF and Properties

The MLP function is a three-parameter pdf. The three parameters of the distribution function are $\alpha_0$, $\mu_0$, and $\sigma_0$. $\alpha_0 + 1$ is the power-law index of $\frac{dN}{dm}$ for the power-law distribution: characteristic of a Pareto distribution which is used to represent pure power-law distributions. The parameters $\mu_0$ and $\sigma_0^2$ are the same as for the lognormal distribution but do not represent the mean and variance of the distribution unlike for the lognormal distribution. Parameters $\mu_0$ and $\sigma_0$ describe the shape of the lognormal-like body and $\alpha_0$ represents the power-law tail. In the limit as $\sigma_0$ tends to zero, the function behaves as a pure power-law.

If $m$ is the the mass of a star, the pdf of the MLP function is given in the closed form as [6]:

$$f(m) = \frac{\alpha_0}{2} \exp\left(\alpha_0\mu_0 + \frac{\alpha_0^2\sigma_0^2}{2}\right) m^{-\alpha_0-1} \times \text{erfc}\left(\frac{1}{\sqrt{2}} \left(\frac{\ln m - \mu_0}{\sigma_0}\right)\right), \ m \in [0, \infty),$$

Some properties of the MLP function are:

(i) Raw Moments:

$$E[M^k] = \frac{\alpha_0}{\alpha_0 - k} \exp\left(\frac{\sigma_0^2 k^2}{2} + \mu_0 k\right), \ \alpha_0 > k. \quad (1.9)$$

(ii) Variance:

$$\text{Var}(M) = E[M^2] - (E[M])^2 = \alpha_0 \exp(\sigma_0^2 + 2\mu_0) \left(\frac{e^{\sigma_0^2}}{\alpha_0 - 2} - \frac{\alpha_0}{(\alpha_0 - 1)^2}\right), \ \alpha_0 > 2.$$

(iii) Cumulative Distribution Function:

$$F_M(m; \alpha_0, \mu_0, \sigma_0) = \frac{1}{2} \text{erfc}\left(-\frac{\ln(m) - \mu_0}{\sqrt{2}\sigma_0}\right) - 1/2 \exp\left(\alpha_0\mu_0 + \frac{\alpha_0^2\sigma_0^2}{2}\right) m^{-\alpha_0} \text{erfc}\left(\frac{\alpha_0\sigma_0}{\sqrt{2}} - \frac{\ln(m) - \mu_0}{\sqrt{2}\sigma_0}\right).$$

(1.10)

(iv) Mode: Solving the following transcendental equation will give us the mode of the distribution

$$f'(m) = 0 \iff \text{Kerfc}(u) = e^{-u^2}, \quad (1.11)$$
where

$$K = \sigma_0 (\alpha_0 + 1) \sqrt{\frac{\pi}{2}} , u = \frac{1}{\sqrt{2}} \left( \alpha_0 \sigma_0 - \frac{\ln m - \mu_0}{\sigma_0} \right).$$  \hspace{1cm} (1.12)

### 1.4 Young stellar populations

We have initiated a study to apply the MLP function for the investigation of the IMF in young and populous star clusters. Star clusters are considered as simple stellar populations with the same chemical composition and age \[45\], making them ideal targets for IMF studies.

In chapter 2, we present a pilot study introducing our method and its application to NGC 1711, a young and populous star cluster in the Large Magellanic Cloud (LMC). LMC is a gas-rich satellite galaxy of the Milky Way located at a distance of \~50 kpc \[23\]. Overall, the star clusters in the LMC span a wide range in ages \(10^6 \text{ yr to } 10^{10} \text{ yr}\) and masses \(10 \text{ to } 10^6 \, M_\odot\) \[28\]. NGC 1711 is located in the north-west part of the LMC, below its bar. NGC 1711 is a populous young star cluster, with an age of \(10^{7.7\pm0.05} \text{ in logarithmic space, a metallicity}\footnote{Elements other than hydrogen and helium are called as metals in astrophysics. Metallicity is the ratio of metal content to hydrogen and helium content.} \text{ of } -0.57\pm0.17 \text{ dex and a reddening}\footnote{Interstellar reddening is a phenomena where stars appear more red in color due to absorption or scattering of shorter wavelengths by the gas and dust in between the light source and the observer [12].} E(B-V) \text{ of } 0.09\pm0.03 \text{ [19].}}

In chapter 3, we take masses for three young and populous clusters: Orion Nebula Cluster (ONC), NGC 2024, and NGC6611 from the literature directly. These three clusters were taken because they vary in their mass ranges and their MFs show different behaviour. ONC is the nearest cluster with plenty of massive O and B stars located at a distance of 400 pc from the sun. Various groups like Hillenbrand \[24\], Palla and Stahler \[49\], Andersen \[3\] have studied the stellar content of the ONC. We consider the census reviewed by Hillenbrand \[24\] and Da Rio et al. \[15\] to get a complete sample of mass range between \(0.029 \, M_\odot \text{ and } 45.7 \, M_\odot\). Both Hillenbrand and Da Rio et al. derived the masses using D’Antona & Mazzitelli \[17\] evolutionary tracks. NGC2024 is a young cluster rich in brown dwarfs and low mass stars. Levine et al. \[38\] conducted a near-infrared spectroscopic study of this young cluster and obtained a mass range
of $0.02\ M_\odot$ and $0.72\ M_\odot$ using the Baraffe et al. [4] evolutionary tracks. NGC6611 is also a young cluster useful to study the low mass IMF but also helps in probing higher i.e. intermediate stellar masses. Using the Baraffe et al. [4] evolutionary models, Oliveira et al. [48, 47] obtain a mass range of $0.01\ M_\odot$ to $6.04\ M_\odot$ for the age of the cluster as 2 million years.

1.5 Model Fitting and Model Selection

Given a sample of observations and a set of candidate mathematical models to describe the data set, the problem of model fitting and model selection comes into play. For model fitting one first needs to investigate whether a mathematical model can be considered as a candidate model or not which can be done using non-linear regression for a non-linear model. Regression finds best fit parameter values for the model by minimizing the sum of the squared errors. A more robust way of parameter estimation is by using maximum likelihood estimation that also aims at minimizing the residuals between the mathematical model and the underlying data set. Once the set of candidate models have been established, model selection helps in finding the best candidate model [11].

1.5.1 Non-Linear Regression

Regression is a statistical technique that analyzes the relationship between a dependent variable and several independent variables [22]. It is the process of fitting a mathematical equation of several unknown parameters to the experimental data presuming that the equation is a correct mathematical description of the underlying process [29]. The main objective of a non-linear least squares method is to estimate the unknown parameters of the mathematical equation by minimizing the sum of squares of the residual [43]. We use the Levenberg Marquardt method on the normalized MF to estimate the best-fit parameters for the MLP function.

The Levenberg Marquardt (LM) Method is an iterative process evolved from the combination of the Gauss Newton and the Steepest Descent method. It uses the advantages of the two
methods to compute the estimates for the parameters of the given equation with maximum like-
lihood whilst ignoring the limitations of both these methods. The method of Steepest Descent
is advantageous in initial iterations as it quickly moves along the direction of steepest descent
to minimize the sum of squares of the residuals, but becomes less accurate on later iterations.
Unlike the Steepest Descent method, Gauss Newton method is effective for later iterations but
may go in the wrong direction for initial iterations. Hence, the LM method jumps from the
Steepest Descent to the Gauss Newton method from initial to later iterations [37]. Likewise
the Method of Steepest Descent and Gauss Newton Method, the LM method is an iterative
process and requires an initial estimation of the parameters. It tries to find a better estimate to
the parameters by minimizing the sum of the squares of the residuals. To check whether the
algorithm gives us the best fitting parameters, an understanding of how good the fit is and the
uncertainty involved is fundamental.

1.5.2 Maximum Likelihood Estimation

Given a sample of observations $x_1, x_2, x_3, \ldots, x_n$ where $x_i's$ are independent and identically
distributed data points assumed to be taken from a pdf $f(X|\Theta)$ of $k$ unknown parameters
$\theta_1, \theta_2, \ldots, \theta_k$, the likelihood function can be defined as [30]:

$$L(\Theta | x_i) = f(x_1 | \Theta)f(x_2 | \Theta)\ldots f(x_n | \Theta) = \prod_{i=1}^{n} f(x_i | \Theta),$$

(1.13)

Note that $L(\Theta | x_i)$ is a function of the unknown parameters with data points kept fixed unlike the
pdf which is a function of observations with fixed parameter values. Maximizing the likelihood
function helps in finding the parameter values that are most likely to describe the data set.
For simplicity the log of the likelihood function is maximized i.e.:

$$\ln L(\Theta | x_i) = \ln f(x_1 | \Theta) + \ln f(x_2 | \Theta) + \ldots + \ln f(x_n | \Theta) = \sum_{i=1}^{n} \ln f(x_i | \Theta),$$

(1.14)

One can find maximum likelihood estimator for the parameters $\theta_1, \theta_2, \ldots, \theta_k$ by simulta-
neously solving for:
\[
\frac{d \ln L(\Theta | x_i)}{d\theta_j} = 0 : j = 1, \ldots, k.
\] (1.15)

For various distributions like the lognormal distribution or the Pareto distribution, functional forms can be found for the maximum likelihood estimators but for distributions that do not have an analytic form for the estimators, global optimization techniques such as simulated annealing or particle swarm are explored to find global minima for the negative-likelihood function i.e. 
\[ -\ln L(\Theta | x_i) \] which is same as finding global maxima for the likelihood function.

Both simulated annealing and particle swarm can be used to solve non-linear bounded optimization problems. Simulated annealing is an adaption of the Monte Carlo method that compares the objective function evaluated at a random initial number with the objective function evaluated at a neighbouring random number until it reaches the maximum number of iterations or a tolerance condition [31]. Particle Swarm on the other hand is also an iterative method that moves around in search space according to some mathematical formulae to find the optimal solution to the problem [20].

1.6 Thesis Statement and Contribution

Our goal in Chapter 2 is to determine whether the distribution of stars in NGC 1711 can best be described by a power law, a lognormal or a hybrid function. As NGC 1711 is a young and populous star cluster, this provides us with a statistically significant sample of stars, ranging over a wide mass domain to be able to model the high mass stellar regime. In Chapter 3, we model the MF of different young and populous star clusters from the literature in various environments using different functions. We then use a model selection criterion to determine which model can best describe the underlying stellar population.
Bibliography


Chapter 2

Modeling the Mass Function of NGC 1711 as a Case Study using the MLP

In this chapter we investigate whether the MLP can be considered as a candidate model to describe the MF of a young and populous stellar cluster, NGC 1711. To do so we first need to obtain the data set for the MF to which the mathematical model is fitted. As described in Chapter 1, for stars for which we cannot make direct mass measurements theoretical MLRs are used to convert luminosity to mass. On obtaining the mass function, we also explore the behaviour of the MF using the MLP, once it is established as a candidate model.

2.1 From Luminosity to Mass

To do the conversion from luminosity to mass by using the MLR/MMR, we first obtain the luminosity function (LF). The distribution of stellar absolute magnitudes in a particular wavelength into different absolute magnitude intervals \([M_i, M_i + dM_i]\) is called the luminosity function \([5, 6]\).

\[
\frac{dN}{dm} = \frac{dN}{dM_i(m)} \left[ \frac{dm}{dM_i(m)} \right]^{-1},
\]  

(2.1)
Note $M_i$ is denoted as absolute magnitude in a particular wavelength while $m$ is representative of the star’s mass. Here $\frac{dN}{dM_i(m)}$ is the LF and $\frac{dm}{dM_i(m)}$ is the slope obtained from the mass luminosity/mass magnitude relation.

### 2.1.1 Luminosity Function

![Color Magnitude Diagram of NGC 1711](image)

Figure 2.1: Color Magnitude Diagram of NGC 1711. On the $x$-axis color i.e. $F555W - F814W$ is plotted and on the $y$ axis apparent magnitude $F814W$ is plotted. The blue selection points represent the main sequence branch that consists of 5481 points.

To obtain the LF for the LMC stellar cluster NGC 1711, we first need to correct the reduced data for field star contamination [10, 13]. The field star contamination corrected colour magnitude diagram (CMD) of the cluster is shown in Fig. 2.1. We then select the main sequence stars with a completeness factor\(^1\) of 75\%, which are highlighted in blue. For field star contamination we arbitrarily select the edges of the boxes on the spatial distribution plot of the stars.

\(^1\)Completeness factor is the ratio of the number of stars retrieved during source detection from the photometry and the number of artificial stars added. It gives a measure on how deep in magnitude can we go to successfully retrieve stars of fainter magnitude. It is a function of both magnitude and position [10].
2.1. From Luminosity to Mass

and remove them from the rest of the data. After correcting for field star contamination and incompleteness we divide the absolute magnitude domain in the I-band i.e. $M_{F814W}$ into bins of optimal size $2n^{2/5}$ with $n$ as the total number of points: 5481 to obtain the LF [9]. We select only the main sequence\(^2\) stars because only for stars that have one to one correspondence for luminosity with mass can we obtain a well known mass luminosity relation (MLR) to make the conversion from luminosity to mass.

Fig. 2.2 is the histogram obtained for the apparent magnitude in the I-band on the $x$-axis. Fig. 2.3 is the histogram for the LF plotted in logarithmic space on the $y$-axis. The brightest main sequence star in the cluster NGC 1711 is of 15.31 apparent magnitude in the I-band\(^3\) or -3.05 absolute magnitude\(^4\) in the I-band. We make the conversion from apparent to absolute

---

\(^2\)The main sequence corresponds to the locus of points on the H-R Diagram that corresponds to stars that are undergoing steady state nuclear fusion of hydrogen.

\(^3\)I-band is the infrared band on the electromagnetic spectrum which is the same as the F814 filter for the Hubble Space Telescope (HST) and V-band is the visual band corresponding to F555 filter of the HST.

\(^4\)Absolute magnitude is the logarithm of brightness of a star seen at a fixed distance of 10 parsecs. One can obtain the absolute magnitude from apparent magnitude if the distance to the star is known.
magnitude using the distance modulus as 18.25 and extinction as 0.116 in the I-band for NGC 1711 obtained from NASA/IPAC extragalactic database (NED). Counting the number of main sequence stars in each magnitude interval gives us the LF.

### 2.1.2 Mass Luminosity Relation

The conversion from luminosity to mass is done using a mass-luminosity relation (MLR). We consider two different sets of theoretical isochrones to obtain the MF of the stellar cluster. The use of two sets of isochrones is to check whether the mass function depends on the choice of the MLR used. In Fig. 2.5, Padova isochrones [3] are over plotted onto the CMD of NGC 1711 and MESA isochrone [12] corresponding to the age of \( \log (t/\text{yr}) = 7.7 \) is also over plotted onto the CMD. \( \log (t/\text{yr}) = 7.7 \pm 0.05 \) is taken to be the age of the cluster, a result found by [7] who found the age using Geneva [8] and Padua isochrones [2]. We take extinction in F555W band as 0.206 magnitude and 0.113 magnitude in the I- band giving us reddening equal to 0.093 magnitude, and the cluster’s metallicity to be -0.57 ± 0.17 dex [7]. We then
2.2 Mass Function

The MF is derived from the LF by using the two sets of MLRs. The MF is the distribution of stellar masses into solar mass bins. We use the same number of optimal bins i.e. 58 as used for the LF.
Figure 2.5: Theoretical mass-magnitude relations (MMR) for log (t/yr) = 7.70. The red line represents the MMR obtained using Padova isochrones and the blue line represents the MMR using MESA isochrones. The respective ranges of masses obtained are (i) Padova: 0.90 \( M_\odot \) to 7.63 \( M_\odot \) (ii) MESA: 0.89 \( M_\odot \) to 7.87 \( M_\odot \).

Figure 2.6: We have obtained the mass function for NGC 1711 using Padova and MESA isochrones. On the x axis we have mass in solar masses. On the y axis density is plotted.
2.3 MLP modeling of the Mass Function

2.3.1 Fitting Results

The LM algorithm for the MLP function on the MF obtained from theoretical isochrones converges to the following set of parameter values that best represent a good fit (i) fit1 (Using Padova MLR): \( \alpha_0 = 1.72, \mu_0 = -0.09 \) and \( \sigma_0 = 0.02 \) and (ii) fit2 (Using MESA MLR): \( \alpha_0 = 1.77, \mu_0 = -0.08 \) and \( \sigma_0 = 0.01 \). The set of values shows physical meaning and below we discuss how their goodness of fit test statistics qualify the parameter values to be a good fit.

The first step to understanding how good the fit is to visually see whether the graph of the fitted MLP function on the observed MF lies close to all the data points. Statistically, a model is said to be a ‘good fit’ to the data: 1) if the assumptions of the least squares method are satisfied 2) if the coefficients of the model can be obtained with minimum uncertainty 3) the model explains variability in the data 4) and the model has high certainty of predicting new observations.
Figure 2.8: Using Padova isochrones: Mass function fitted with the MLP function. The plot above represents the graph for the best fit values $\alpha_0 = 1.72$, $\mu_0 = -0.09$ and $\sigma_0 = 0.02$. We obtained the same slope i.e. $\alpha + 1 = 1.72$ value for the Pareto function as well. The best Pareto fit is over plotted in a green colour.

Figure 2.9: Using MESA isochrones: Mass Function fitted with MLP function. The plot above represents the graph for the best fit values $\alpha_0 = 1.77$, $\mu_0 = -0.08$ and $\sigma_0 = 0.01$. We obtained the same $\alpha + 1 = 1.77$ value for the Pareto function as well.
For fit 1 (Padova): $\alpha_0 = 1.72$, $\mu_0 = -0.09$ and $\sigma_0 = 0.02$ statistically, the residuals i.e. sum of squared errors (SSE) give a good indication whether the curve lies near the data points. The SSE value obtained for fit 1 is 1.25 which is low. We obtain an $R^2$ value of 0.91 which explains variability in data. An $R^2$ value closer to 1 indicates that a greater proportion of variance is accounted for by the model fit. We then look at the 95% of confidence bounds for each parameter. For $\alpha_0 = 1.72$, we have (1.58, 1.87) as the confidence interval, $\mu_0 = -0.09$ lies in (-0.15, -0.02) and $\sigma_0 = 0.02$ lies in (-0.07, 0.11). The narrowness of the 95% confidence intervals indicate that the coefficients of the model can be obtained with minimum uncertainty.

For fit 2 (MESA): $\alpha_0 = 1.77$, $\mu_0 = -0.08$ and $\sigma_0 = 0.01$, the SSE value obtained for is 1.34 which is low. We obtain an $R^2$ value of 0.91. Confidence bounds for each parameter are: for $\alpha_0 = 1.77$, we have (1.62, 1.92), $\mu_0 = -0.08$ lies in (-0.15, -0.01) and $\sigma_0 = 0.01$ lies in (-0.16, 0.19).

The best fit values are: Padova: $\alpha_0 = 1.72$, $\mu_0 = -0.09$ and $\sigma_0 = 0.02$ and MESA: $\alpha_0 = 1.77$, $\mu_0 = -0.08$ and $\sigma_0 = 0.01$ for $M_\odot$ vs. $\log(f(m))$. All data points lie near the best fitted curve of the MLP function and for the graph in $\log(M_\odot)$, most data points lie closer to the curve for $\log(M_\odot) < 0.8$. Deviations above that limit are discussed in next section.

As discussed in section 2, in the limit $\sigma_0$ tending to 0 the MLP function behaves as a pure power-law distribution. We obtained $\sigma_0 = 0.02$ for the Padova isochrones and $\sigma_0 = 0.01$ for the MESA isochrones implying that the MLP takes power-law behaviour. This behaviour is seen graphically in Fig. 2.8 and Fig. 2.9. We also fitted the Pareto distribution function (equation 2) to the data points and found the slope to be same for the MF obtained from the Padova and MESA.

For a non-linear least squares regression approach one should check if the uncertainties involved in the fitting are random and normally distributed i.e. the residuals obtained follow a Gaussian and do not vary systematically above or below 0. For this we look at the residual plot for the best fit parameters. From Fig. 2.10, we can see that the residuals are randomly distributed and from Fig. 2.11 i.e. Quantile-Quantile Plot of the residuals we can ascertain that
28 Chapter 2. Modeling the Mass Function of NGC 1711 as a Case Study using the MLP

Figure 2.10: Residual plot for the best fit MLP function on the mass function obtained using Padova isochrones.

Figure 2.11: Quantile-Quantile plot for MF from theoretical MLR.
the uncertainties are normally distributed [14]. A Q-Q plot, where Q stands for quantile, is a probability plot that compares if two distributions are similar or not. In the case of the error distribution, it compares with the standard normal distribution for which the points from the error distribution should lie on the straight line \( y = x \) corresponding to the standard normal.

### 2.3.2 Truncated MLP

![Figure 2.12: MF fitted with MLP function](image)

Figure 2.12: MF fitted with MLP function. The plot above represents the graph for the best fit values \( \alpha_0 = 1.55, \mu_0 = -0.08 \) and \( \sigma_0 = 0.07 \).

While deriving the pdf of the MLP distribution, Basu & Jones assumed an initial lognormal distribution with mean \( \mu_0 \) and standard deviation \( \sigma_0 \). They then assumed an exponential growth of stellar masses because of accretion assuming a linear dependence on mass for the accretion rate i.e. \( \frac{dm}{dt} = \gamma m \). This resulted the initial lognormal distribution to shift to a new mean \( \mu_0 + \gamma t \) where \( \gamma \) is the growth rate of accretion. Since there is an equally likely chance for accretion to stop any time due to various mechanisms such as ejection of the star, stellar outflows, gas swept away from the star by other stars [11], Basu & Jones assumed an exponential decay for
the distribution of lifetimes of accretion, with $\delta$ as the decay rate i.e. $f(t) = e^{-\delta t}$, and obtained the MLP function on integrating over infinite time. Accretion of mass onto the star for infinite time is not entirely physical because accretion is most likely stop at some maximum stopping time due to the dissipation of the surrounding gas. Thus, integrating equation 1.8 in the MLP derivation to a maximum stopping time $T$, we obtained the truncated MLP:

$$
\int_0^T \frac{1}{\sqrt{2\pi} \sigma_0 m} \exp\left(-\frac{(\ln m - \mu_0 - \gamma T)^2}{2\sigma_0^2}\delta e^{-\delta t} dt\right) = \frac{\alpha_0}{2} \exp\left(\frac{\alpha_0 \mu_0 + \alpha_0^2 \sigma_0^2}{2}\right)m^{-(1+\alpha_0)} \times 
\left(\text{erf}\left(\frac{1}{\sqrt{2}}\left(\alpha_0 \sigma_0 - \ln m - \mu_0 - \gamma T\right)\right) - \text{erf}\left(\frac{1}{\sqrt{2}}\left(\alpha_0 \sigma_0 - \ln m - \mu_0\right)\right)\right) \quad (2.2)
$$

Refer to Appendix A for derivation of equation above.

The truncated MLP function is a four-parameter pdf where $\mu_0$ and $\sigma_0$ describe the lognormal body, $\alpha_0 = \delta/\gamma$ gives the slope of the power-law tail and $T \propto \delta^{-1}$ i.e. the maximum stopping time describes the truncation. Using the LM Algorithm, we found that the truncated MLP distribution follows a truncation when $\gamma T = 2$. The figure below shows the truncated MLP probing the data points above $6M_\odot$ when using the truncated MLP for a fixed stopping time for accretion. The reason for it being twice will be explored in a future work.

### 2.3.3 MLP as a hybrid

Another significant purpose of using the MLP distribution function is to check whether it can work as a hybrid to model both lognormal and power law behaviour as a single function. Since our data sample was complete only to $m_{F814W} = 23$, we combined the NGC 1711 cluster data for theoretical MLR with an artificially generated data sample from the Chabrier lognormal functional form [4]. We generated an equal number (60) of data points in the range $0.06 M_\odot$ to $0.90 M_\odot$ as taken from the Chabrier function. We combined the data points by normalizing the Chabrier function from $0.06 M_\odot$ to $0.90 M_\odot$ with the obtained best fitting MLP function for the cluster data from $0.90$ to $7.64 M_\odot$ and obtained a complete data sample with low mass
Figure 2.13: We obtain the best fit parameter values $\alpha_0 = 1.57$, $\mu_0 = -2.06$ and $\sigma_0 = 0.90$ for the MLP function on the entire mass domain. The points below $0.90 \, M_\odot$ are the artificially generated data points from the Chabrier function which are taken in equal number as the data points for the resolved NGC 1711 population.
to high mass stars. Then we again used the LM method on the combined data sample and obtained the best fitting MLP function with best fitting parameter values as $\mu_0 = -2.06 \pm 0.06$, $\sigma_0 = 0.90 \pm 0.07$, and $\alpha_0 = 1.57 \pm 0.13$. Using the MLP properties [1], we found the mean of the distribution for the best fit parameter values to be $0.53 \, M_\odot$. The mean for the NGC 1711 data points alone is $2.23 \, M_\odot$. The artificially generated data points provided a lognormal body to the distribution of NGC 1711 data points thus giving a higher sigma value $\sigma_0 = 0.90$ i.e. resulting in deviation from a pure-law behaviour. It also altered the mean of the distribution where the mean of the distribution now lies in the low mass end of the stellar regime. Our aim of joining the NGC 1711 data points with Chabrier data points was also to check whether the slope of the power-law tail is affected by the lognormal body or not i.e. making the tail steeper or shallower in logarithmic space. From our fitting results $\alpha_0 = 1.57 \pm 0.13$ lies in the predicted interval of $\alpha_0 = 1.72 \pm 0.14$ for NGC 1711, hence showing not a significant effect of the lognormal body on the slope of the power-law tail.

2.4 Summary

We derived MFs for the young and populous LMC stellar cluster NGC 1711 using two different sets of theoretical isochrones. The slope and the mass range for the MF obtained seems to not depend on the underlying theoretical MLR. Using Padova we obtained the mass range of the MF to be $0.90$ to $7.63 \, M_\odot$ and using MESA we got the mass range as $0.89$ to $7.87 \, M_\odot$.

We then investigated whether the MFs showed lognormal, power law or hybrid behaviour using the MLP function along with checking whether it can be adequately used to describe the MF of the stellar cluster. Since the MFs seems to be showing pure power law behaviour, the MLP function was able to give best fit parameter value for the slope of the MFs : (i) Padova: $\alpha_0 = 1.72 \pm 0.14$ (ii) MESA: $\alpha_0 = 1.77 \pm 0.15$. In the limit where $\sigma$ tends to 0, the MFs tend to a pure power-law behaviour. We obtained sigma values as (i) Padova: $\sigma = 0.02$ (ii) MESA: $\sigma = 0.01$. The reason why the MFs of NGC 1711 seem to be having pure power law behaviour
because the data are limited to $m_{F814} \leq 23$. Since the data is only complete until 23 F814W apparent magnitude, we did not have many stellar masses less than $1M_\odot$ in our data set which is known to show lognormal behaviour in general.

The turnover at $6M_\odot$ is explored using a truncated MLP where we consider a fixed stopping time for accretion. The truncated MLP was able to probe the turnover giving evidence that accretion stops at a time scale analogous to the characteristic death time.

We also investigated whether the MLP function can model hybrid i.e. both lognormal as well as power law behaviour, and also to check whether adding a lognormal body to the data has any effect on the slope of the power law tail. For that we generated artificial data points from the Chabrier lognormal function [4] and combined the data points for NGC 1711 to get a complete data set with masses ranging over low, intermediate and high mass. We took the MF using Padova isochrones. We then again fitted the MLP function using a non-linear regression approach and obtained $\alpha_0 = 1.57$ which is less steep than the one obtained only for the fit on the cluster data i.e. $\alpha_0 = 1.72$, but lies in the predicted interval for $\alpha_0 = 1.72$ i.e. $\alpha_0 = 1.72 \pm 0.14$. From this we can conclude the MLP can be used to model hybrid behaviour as a single function instead of using different functions with joining conditions. Our final conclusion is (i) NGC 1711 exhibits a power-law tail at the intermediate and high masses (ii) the MLP can be easily used to check whether the MF follows lognormal, pure power law or hybrid behaviour as a single function.
Bibliography


Chapter 3

Model Selection: Which Model to choose for the IMF of Young Stellar Clusters?

Given the plethora of observational data on stellar mass distribution (IMF) in different environments and many existing IMF models since the pioneering work of Salpeter [8] in 1955, it is important to study the statistical problem of model selection. Model selection aims to investigate: which model can be used as the best approximating model to the underlying data set? One can use the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to do model selection for the set of different candidate models, given the data set. In this chapter, we focus on the comparison of three candidate IMF models: the Modified Lognormal Power-Law (MLP) [2] probability distribution function, the Chabrier IMF [4] and the Kroupa IMF [5] using AIC and BIC, given the data set of stellar mass distribution in different environments.

To obtain the ranks for these three models on the basis of AIC/BIC, one first needs to estimate the best fit parameters of these models on the underlying distribution. Even though one can estimate parameters of the functional forms by fitting models to the observed data using a non-linear regression approach, this method involves numerical bias. Maiz Apellaniz & Ubeda showed that deriving the slope for the power-law functional form using a least-squares
minimization method with uniform binning of data has significant numerical bias [7]. The correlation between the number of stars in each bin and the weights assigned to each bin causes the bias in the determination of the slope. Hence, we use the method of maximum likelihood for the estimation of parameters for the models which like the least-squares minimization approach aims to minimize the residuals by maximizing the likelihood function but is independent of the bias due to binning.

### 3.1 Observational data

(i) Orion Nebula Cluster (ONC): The stellar population has the mass range of $0.02 \, M_\odot$ to $45.70 \, M_\odot$. This mass range contains substellar, low, intermediate as well as high mass stars hence has both a lognormal body and a power-law tail. Since the masses of the stars of the ONC population span over the entire mass regime, it is a perfect laboratory to test hybrid behaviour.

(ii) NGC 1711: We obtain the mass range of $0.89 \, M_\odot$ to $7.84 \, M_\odot$ for the stellar population. This mass range spans some low mass stars but mostly intermediate and some high mass stars. It is useful to investigate power-law behaviour of the MF.

(iii) NGC 6611: The stellar population has mass range of $0.02 \, M_\odot$ to $6.02 \, M_\odot$. The population spans over the substellar, low and intermediate mass regime but does not contain high mass stars. One can investigate the lognormal behaviour of the assumed SSP and probe some part of the power-law tail.

(iv) NGC 2024: The stellar population has the range of $0.02 \, M_\odot$ to $0.72 \, M_\odot$. The population contains only substellar and low mass stars hence helps in modeling population distribution showing lognormal behaviour alone.

### 3.2 IMF Models and Parameter Estimation

(i) Chabrier functional form (Lognormal + Power-law): this is a piecewise pdf having a log-normal form on the substellar and the low mass stellar regime i.e. below $1 \, M_\odot$ and a power-law
form on the intermediate and high mass regime i.e. above $1\,M_\odot$ [4, 3].

$$f(ln\,m) \propto \begin{cases} 
\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\left(ln\,m - \mu\right)^2}{2\sigma^2}}, & m \leq 1\,M_\odot \\
m^{-\alpha+1}, & m > 1\,M_\odot 
\end{cases}$$ \hspace{1cm} (3.1)

The functional form of the lognormal distribution is given by:

$$f(m) = \frac{1}{\sqrt{2\pi}\sigma m} e^{\frac{-\left(ln\,m - \mu\right)^2}{2\sigma^2}}, \quad m \leq 1\,M_\odot$$ \hspace{1cm} (3.2)

The lognormal function has two parameters: $\mu$ and $\sigma$. The location parameter $\mu$ is the mean of the distribution and the scale parameter $\sigma$ is the standard deviation of the distribution. The likelihood function of a lognormal distribution is given by:

$$L(\mu, \sigma|m_i) = \prod_{i=1}^{N} \frac{1}{m_i} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\left(ln\,m - \mu\right)^2}{2\sigma^2}},$$ \hspace{1cm} (3.3)

Thus the parameters that maximize the likelihood function are given by:

$$\hat{\mu} = \frac{\sum_k ln\,m_k}{n},$$ \hspace{1cm} (3.4)

$$\hat{\sigma}^2 = \frac{\sum_k \left(ln\,m_k - \hat{\mu}\right)^2}{n}. $$ \hspace{1cm} (3.5)

The functional form of the Pareto distribution is given by:

$$f(m) = (\alpha - 1)\,a^{\alpha-1}\,m^{-\alpha}, \quad m > a, \text{ where } \alpha > -1 \text{ and } a > 0$$ \hspace{1cm} (3.7)

$\alpha$ is the shape parameter for the distribution that determines the power-law tail while $a$ is the lower-limit for the distribution from which the distribution begins to show power-law be-
haviour. The likelihood function for the Pareto distribution is given by:

\[
L(\alpha|m_i) = \prod_{i=1}^{N} (\alpha - 1)^{\alpha} m_i^{-\alpha} \quad \text{for } \alpha > 1,
\]

Thus the parameter that maximizes the likelihood function is given by:

\[
\hat{\alpha} = -1 - \left[ \frac{1}{n} \sum_{i=1}^{N} \log\left( \frac{m_i}{a} \right) \right]^{-1}.
\]

(ii) Kroupa functional form (multi-segmented power-law): This is a piecewise pdf of a Pareto distribution, a Truncated-Pareto and a Power function. Essentially, they are all power-law distributions varying in either the sign of the exponent or whether they have an upper limit, lower limit or both.

\[
f(\ln m) \propto \begin{cases} 
    m^{\alpha_1+1}, & m < 0.08 M_\odot, \text{where } \alpha_1 > -1 \\
    m^{\alpha_2+1}, & 0.08 M_\odot \leq m < 0.50 M_\odot, \text{where } \alpha_2 \neq -1 \\
    m^{-\alpha_3+1}, & 0.50 M_\odot \leq m, \text{where } \alpha_3 > -1
\end{cases}
\]

The Pareto Function and Power Function represent the same distribution except \( \alpha < -1 \) for Pareto and \( \alpha > -1 \) for Power Function. The parameter that maximizes the likelihood function is given by equation 3.9. The best fit parameter value for the Truncated-Pareto distribution can only be obtained numerically by solving the following equation [10] :

\[
\ln m = -\frac{1}{\hat{\alpha} + 1} + \frac{b^{\hat{\alpha}+1} \ln b - a^{\hat{\alpha}+1} \ln a}{b^{\hat{\alpha}+1} - a^{\hat{\alpha}+1}}.
\]

where \( a \) is the lower limit of the distribution and \( b \) is the upper limit, and \( a \geq 0 \& b \geq 0 \).

(iii) MLP functional form: The functional form of the MLP is given by:

\[
f(\ln m) = \frac{\alpha_0}{2} \exp\left( \alpha_0 \mu_0 + \frac{\alpha_0^2 \sigma_0^2}{2} \right) m^{-\alpha_0} \times \text{erfc}\left( \frac{1}{\sqrt{2}} \left( \frac{\alpha_0 \sigma_0 - \ln m - \mu_0}{\sigma_0} \right) \right), \quad m \in [0, \infty). \]

\[3.12\]
Obtaining the best fit parameter that maximizes the likelihood function for the MLP is not possible analytically thus we undergo a global minimization search for $-\ln L$ i.e. the negative log-likelihood function using simulated annealing.

The following tables give the best fitting parameter values for the different stellar populations. $M_{Br}$, $M_{Br1}$ and $M_{Br2}$ represent the respective joining points.

<table>
<thead>
<tr>
<th>Data</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\alpha$</th>
<th>$M_{Br}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONC</td>
<td>-1.54</td>
<td>0.54</td>
<td>2.46</td>
<td>$1 M_\odot$</td>
</tr>
<tr>
<td>NGC 1711</td>
<td>-0.71</td>
<td>0.65</td>
<td>2.92</td>
<td>$1 M_\odot$</td>
</tr>
<tr>
<td>NGC 1711</td>
<td>-</td>
<td>-</td>
<td>2.90</td>
<td>0.89 $M_\odot$</td>
</tr>
<tr>
<td>NGC 6611</td>
<td>-0.32</td>
<td>0.65</td>
<td>3.89</td>
<td>$1 M_\odot$</td>
</tr>
<tr>
<td>NGC 2024</td>
<td>-2.06</td>
<td>1.0</td>
<td>-</td>
<td>$1 M_\odot$</td>
</tr>
</tbody>
</table>

Table 3.1: Best fit parameter values for the Chabrier functional form.

<table>
<thead>
<tr>
<th>Data</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha = \alpha_3$</th>
<th>$M_{Br1}$</th>
<th>$M_{Br2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONC</td>
<td>3.20</td>
<td>0.02</td>
<td>2.18</td>
<td>0.08</td>
<td>0.50</td>
</tr>
<tr>
<td>NGC 1711</td>
<td>-</td>
<td>-</td>
<td>2.90</td>
<td>-</td>
<td>0.50</td>
</tr>
<tr>
<td>NGC 1711</td>
<td>-</td>
<td>-</td>
<td>2.90</td>
<td>-</td>
<td>0.89</td>
</tr>
<tr>
<td>NGC 6611</td>
<td>0.88</td>
<td>0.37</td>
<td>3.52</td>
<td>0.08</td>
<td>0.50</td>
</tr>
<tr>
<td>NGC 2024</td>
<td>0.32</td>
<td>-1.22</td>
<td>5.03</td>
<td>0.08</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 3.2: Best fit parameter values for the Kroupa functional form.

<table>
<thead>
<tr>
<th>Data</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\alpha = \alpha_0 + 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONC</td>
<td>-2.01</td>
<td>0.35</td>
<td>2.42</td>
</tr>
<tr>
<td>NGC 1711</td>
<td>-0.11</td>
<td>0</td>
<td>2.90</td>
</tr>
<tr>
<td>NGC 6611</td>
<td>-0.97</td>
<td>0.98</td>
<td>4.59</td>
</tr>
<tr>
<td>NGC 2024</td>
<td>-2.26</td>
<td>0.98</td>
<td>max</td>
</tr>
</tbody>
</table>

Table 3.3: Best fit parameter values for the MLP functional form.

Note that: $\frac{dN}{dm} \propto m^{-\alpha}$ while $\frac{dN}{d \ln m} \propto m^{-\alpha+1}$ and $\alpha = \alpha_0 + 1$. 


3.2. IMF Models and Parameter Estimation

Figure 3.1: Logarithm of $f(ln m)$ is plotted as a function of $\log(m/M_\odot)$ for the Orion Nebula Cluster.

Figure 3.2: Logarithm of $f(ln m)$ is plotted as a function of $\log(m/M_\odot)$ for the cluster NGC 1711.
Chapter 3. Model Selection: Which Model to choose for the IMF of Young Stellar Clusters?

Figure 3.3: Logarithm of $f(ln m)$ is plotted as a function of $log(m/M_\odot)$ for the cluster NGC 6611.

Figure 3.4: Logarithm of $f(ln m)$ is plotted as a function of $log(m/M_\odot)$ for the cluster NGC 2024.
3.3 Model Selection using Information Criterion

The Chabrier IMF [4] i.e. lognormal + power-law functional form and the Kroupa IMF [5] i.e. the mutli-segmented power-law functional form are the two most commonly used IMF models up-to-date. Even though these two models can be adequately used to study the features and characteristics of the stellar mass distribution i.e. the lognormal body and the power-law tail, they require different joining conditions that add to the number of free parameters; this in return increases the complexity of these models. MLP on the other hand is a pdf of only 3 parameters that can probe both the lognormal body and the power-law tail of the MF as a single function (refer to Chapter 2).

The AIC or the BIC provide a trade off between how well the model fits the data and how complex the model is. Comparing the Chabrier, Kroupa and the MLP distribution functions on the basis of the ranks obtained using the AIC/BIC will help us find the simplest model of the competing models i.e. a model with the least number of parameters but also the one that lies very close to the data set.

We compute the AIC value by $AIC = -2 \ln L + 2k$ where $L$ is the value of the likelihood function maximized by the best fit parameters and $k$ is the number of parameters in the model [1]. The model with the minimum AIC value gives us the best model amongst the candidate set of models.

The model with the minimum BIC value also gives us the best approximating model. BIC is defined as $BIC = -2 \ln L + k \ln N$ where $N$ is the number of data points [9]. The only difference between AIC and BIC is that BIC has a more strict condition to penalize a model for overfitting. The penalty term in the AIC is $2k$ while in the BIC penalty term is $k \ln N$. For large data sets, AIC can tend to pick models with more number of parameters than the true model while BIC penalizes an overparameterized model for large data sets with a stronger penalty term [6].

For models that have similar AIC value, we find the relative likelihood and the associated Akaike weight for the model. The relative likelihood for the model is given by $exp\left(\frac{AIC_{\text{min}} - AIC_i}{2}\right)$
where $AIC_{\text{min}}$ is the lowest AIC rank for all the models, $AIC_i$ are the individual AIC ranks for the candidate model and $\Delta AIC_i$ is the difference. Akaike weight is given by $\frac{AIC_{\text{min}}-AIC_i}{\sum AIC_{\text{min}}-AIC_i}$. A higher Akaike weight tells us that the model has the most probability to be the best fit model amongst the candidate models.

The joining conditions for the Chabrier IMF and the Kroupa IMF are kept fixed as $1M_\odot$ for Chabrier and $0.08M_\odot$ and $0.50M_\odot$ respectively. The other are kept as free parameters.

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>-2 ln L</th>
<th>AIC</th>
<th>BIC</th>
<th>$\Delta AIC_i$</th>
<th>Relative likelihood</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chabrier</td>
<td>3</td>
<td>7600.3</td>
<td>7606.1</td>
<td>7623.7</td>
<td>2209.8</td>
<td>0</td>
</tr>
<tr>
<td>Kroupa</td>
<td>3</td>
<td>8696.4</td>
<td>8702.4</td>
<td>8719.9</td>
<td>330.61</td>
<td>0</td>
</tr>
<tr>
<td>MLP</td>
<td>3</td>
<td>5390.3</td>
<td>5396.3</td>
<td>5413.8</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.4: Negative likelihood function values, AIC and BIC for the ONC.

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>-2 ln L</th>
<th>AIC</th>
<th>BIC</th>
<th>$\Delta AIC_i$</th>
<th>Relative likelihood</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chabrier (1 $M_\odot$)</td>
<td>3</td>
<td>6491.8</td>
<td>6497.8</td>
<td>6517.6</td>
<td>2644.3</td>
<td>0</td>
</tr>
<tr>
<td>Chabrier (0.89 $M_\odot$)</td>
<td>3</td>
<td><strong>3847.5</strong></td>
<td><strong>3853.5</strong></td>
<td><strong>3873.3</strong></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Kroupa (0.50 $M_\odot$)</td>
<td>3</td>
<td>11957</td>
<td>11963</td>
<td>11983</td>
<td>8109.5</td>
<td>0</td>
</tr>
<tr>
<td>Kroupa (0.89 $M_\odot$)</td>
<td>3</td>
<td><strong>3847.5</strong></td>
<td><strong>3853.5</strong></td>
<td><strong>3873.3</strong></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MLP</td>
<td>3</td>
<td><strong>3847.5</strong></td>
<td><strong>3853.5</strong></td>
<td><strong>3873.3</strong></td>
<td>1</td>
<td><strong>0.97</strong></td>
</tr>
</tbody>
</table>

Table 3.5: Negative likelihood function values, AIC and BIC for the NGC 1711.

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>-2 ln L</th>
<th>AIC</th>
<th>BIC</th>
<th>$\Delta AIC_i$</th>
<th>Relative likelihood</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chabrier</td>
<td>3</td>
<td>1222.0</td>
<td>1228.0</td>
<td>1239.6</td>
<td>194.91</td>
<td>0</td>
</tr>
<tr>
<td>Kroupa</td>
<td>3</td>
<td>1141.9</td>
<td>1147.9</td>
<td>1159.6</td>
<td>114.8</td>
<td>0</td>
</tr>
<tr>
<td>MLP</td>
<td>3</td>
<td><strong>1027.1</strong></td>
<td><strong>1033.1</strong></td>
<td><strong>1044.7</strong></td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.6: Negative likelihood function values, AIC and BIC for the NGC 6611.
3.4 Discussion

The aim of the study is to provide a quantitative/statistical analysis of the comparison of different IMF models for stellar populations with different mass regimes. These MFs of the stellar populations show different behaviour depending on the range of the masses obtained.

The MF of ONC shows hybrid behaviour i.e. both a lognormal as well as power-law behaviour. On computing the negative log-likelihood, the AIC and the BIC value, one can conclude that the MLP is the best approximating pdf to model hybrid behaviour of the MF as compared to Chabrier IMF with a fixed breaking point at 1 $M_\odot$ or the Kroupa IMF with fixed breaking points as 0.08 $M_\odot$ and 0.50 $M_\odot$. The MLP had the lowest value for the negative log-likelihood thus showing that the model lies closer to the data than the other two models. The best fit parameters obtained for the ONC using the MLP are $\alpha_0 = 1.42$, $\mu_0 = -2.014$ and $\sigma_0 = 0.35$. Thus the best fit value for the $\alpha = \alpha_0 + 1$ for $dN/d\ln m \propto m^{-\alpha+1}$ obtained is 2.42.

The MF of NGC 1711 shows pure power-law behaviour, even though it has some low mass stars upto 0.89 $M_\odot$. MLP, Chabrier IMF with breaking point 0.89$M_\odot$ instead of 1$M_\odot$ and Kroupa IMF with breaking point 0.50$M_\odot$ gave the same negative log-likelihood value. The best fit parameter for the $\alpha$ for $dN/d\ln m \propto m^{-\alpha+1}$ is 2.90 for masses above 0.89$M_\odot$. The Chabrier IMF, the Kroupa IMF and the MLP had similar Akaike weights hence all three models are ranked equally.

The MF of NGC 6611 has a lognormal body with some power-law behaviour on the intermediate mass regime above 1$M_\odot$. The Chabrier functional form in the Fig. 3.3 is fitted keeping

<table>
<thead>
<tr>
<th>Model</th>
<th>parameters</th>
<th>-2 ln L</th>
<th>AIC</th>
<th>BIC</th>
<th>$\Delta AIC_i$</th>
<th>Relative likelihood</th>
<th>Akaike weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chabrier</td>
<td>3</td>
<td>196.69</td>
<td>202.69</td>
<td>209.39</td>
<td>0</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Kroupa</td>
<td>3</td>
<td>197.19</td>
<td>203.19</td>
<td>209.89</td>
<td>0.50</td>
<td>0.77</td>
<td>0.28</td>
</tr>
<tr>
<td>MLP</td>
<td>3</td>
<td>196.76</td>
<td>202.76</td>
<td>209.46</td>
<td>0.07</td>
<td>0.96</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Table 3.7: Negative likelihood function values, AIC and BIC for the NGC 2024.
the joining condition fixed as $1 \, M_\odot$ which is why the model departs away from the underlying data points. For the cluster NGC6611, we get the lowest AIC value for the MLP function. Best fit parameters obtained are $\alpha_0 = 4.6, \mu_0 = -0.97$ and $\sigma_0 = 0.98$.

The MF of NGC 2024 shows pure lognormal-like behaviour. We obtained the lowest negative log-likelihood value for both the Chabrier and the MLP function. On the plot for the MF of NGC 2024, the Chabrier and the MLP overlap each other. The best fit parameters obtained are $\mu_0 = -2.26$ and $\sigma_0 = 0.98$. $\alpha_0$ took the value of the upper limit of the bound constraint while optimizing the negative log-likelihood function for the MLP. This showed that the data set followed pure-lognormal behaviour in the limit of $\alpha$ becoming too large, the MLP becomes a lognormal distribution.

In general on the basis of AIC and BIC, one can statistically infer that the MLP is the best approximating model to the mass distribution of underlying stellar populations having different behaviours. The MLP having only 3 free parameters serves as a simple model which fits the observed data set very closely.
Bibliography


Chapter 4

Conclusion

Since IMF is an essential parameter in many stellar population studies, investigating the mathematical model that describes the IMF for various stellar populations in different environments is necessary [5]. Our goal was to study the MLP as a candidate model to represent the MF of young and populous stellar clusters using NGC1711 of the Large Magellanic Cloud as a pilot case and to compare it with other most commonly used IMF models: the Chabrier IMF and the Kroupa IMF, using model selection.

We first derived the MF for the NGC1711 using two different sets of theoretical isochrones, Padova isochrones [2] and MESA [4] isochrones, to understand the behaviour of the MF. Using these two isochrones we found that the mass range of the cluster lies between 0.89 $M_\odot$ to 7.75 $M_\odot$. We found that the MF obtained was independent of the underlying MLR. Then, using a non-linear regression approach we established the MLP as a candidate model to represent the MF of NGC1711 and found that the stellar cluster follows a pure-power-law behaviour. We obtained an average slope of $\alpha = 2.75$ for $dN/d\ln m \propto m^{-\alpha+1}$ in the mass range. On further investigation of the turnover at 6 $M_\odot$ of the MF of NGC1711 we found that the MLP takes a truncated form if the model is derived over a fixed stopping time for accretion. The truncated MLP was able to successfully probe the turnover at 6 $M_\odot$ giving us an indication that accretion only occurs for a finite time and thus the growth for higher masses halts after a certain time due
to various physical mechanisms. We also studied whether the presence of a lognormal body has any effect on the slope of the power-law form of MF of NGC1711 and found that the slope lies inside the predicted interval of the slope of the MF with pure-power law form. We did so by combining the data points for the stellar cluster with artificially generated data points from the lognormal distribution of the Chabrier IMF [3]. Modeling the MF of NGC1711 combined with the artificially generated data points also helped in investigating that the MLP can model hybrid behaviour i.e. both a lognormal body and a power-law tail as a single function.

Once the MLP was established as a candidate model, we used the AIC and the BIC to perform model selection [6]. We used four young and populous stellar populations: ONC, NGC1711, NGC6611 and NGC2024 to do the comparison study. The four stellar clusters showed different IMF behaviours depending on their mass range. The MF of ONC followed a hybrid behaviour i.e. a lognormal body with a power-law tail as it ranged over substellar, low, intermediate as well as high masses. NGC6611 spanned over substellar, low and intermediate masses thus having a dominant lognormal body and some power-law tail behaviour. The MF of NGC2024 had a pure lognormal form. Doing a model selection study for the MFs of stellar populations having different behaviours, we were able to investigate which model serves as the best approximating model for the MF of young and populous stellar populations showing either a pure-lognormal, a pure-power-law or hybrid behaviour. On the basis of the ranks obtained using AIC and BIC, we concluded that the MLP is the best candidate amongst the three models and can be easily used as a single function to model different IMF behaviours. Our final conclusion is that the MLP not only has a simple and analytic form that models the underlying population with less discrepancy than other models but also has an underlying physical motivation based on the exponential growth of stellar masses via accretion to explain the power-law tail behaviour [1], which the other models lack.
Bibliography


Appendix A

Derivation of the MLP

For the integration of the MLP we use:

\[
\int \exp[-(ax^2 + bx + c)]dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2 - 4ac}{4a}\right) \text{erf}\left(\sqrt{a} \left[ x + \frac{b}{2a} \right] \right) + C, \quad a > 0 \quad (A.1)
\]

where \( C \) is a constant and \( \text{erf} \) is the error function defined as:

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \quad (A.2)
\]

We also define the Complementary error function as:

\[
\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \quad (A.3)
\]

Using the error function/complementary error function we get:

\[
\int_0^\infty \exp[-(ax^2 + bx + c)]dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2 - 4ac}{4a}\right) \text{erfc}\left(\frac{b}{2\sqrt{a}}\right). \quad (A.4)
\]
Appendix B

Joint Likelihood Function

For piecewise pdf with different functional form over different domain we can define the likelihood function using an indicator variable.

\[
f(x) = \begin{cases} 
  f_1(x|\theta_1), & x \leq a \\
  f_2(x|\theta_2), & x > a 
\end{cases} \tag{B.1}
\]

Let D be an indicator variable such that \( d_i = 1 \) if \( x_i \) has \( f_1 \) pdf and \( d_i = 0 \) if \( x_i \) follows \( f_2 \) pdf. Then the joint likelihood function can be defined as:

\[
L(\theta_1, \theta_2, x) = \prod_{i=1}^{N} \left[ f(x_i|\theta_1) \right]^{d_i} \left[ f(x_i|\theta_2) \right]^{1-d_i} \tag{B.2}
\]

To find the normalization constant for the piecewise function with a joining condition we use:

\[
A \left[ \int_{0}^{a} f_1 dx + \int_{a}^{\infty} f_2 dx \right] = 1 \tag{B.3}
\]

since by the property of pdf, total probability of a pdf is 1.
Curriculum Vitae

Name: Deepakshi Madaan

Post-Secondary Education and Degrees:
- Masters of Science in Applied Mathematics
  University of Western Ontario
  London, ON.
  2014 - present
- Bachelors of Science Mathematics(Honors)
  University of Delhi
  New Delhi, India.
  2009-2013

Related Work:
- Teaching Assistant
  University of Western Ontario
  2014-2016
- Research student
  University of Delhi
  2012 - 2013

Publications:
- Madaan D., Lianou S., & Basu S.
  Modelling the mass function of stellar populations: NGC1711 as a pilot case,
  MNRAS (to be submitted).