Applied Mathematics, 2014, 5, 2077-2087
Published Online July 2014 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2014.513202

An Error Controlled Method to Determine the Stellar Density Function in a Region of the Sky

Mohammed Adel Sharaf[1], Zainab Ahmed Mominkhan[2]

[1]Department of Astronomy, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia
[2]Department of Mathematics, College of Science for Girls King Abdulaziz University, Jeddah, Saudi Arabia
Email: Sharaf_adel@hotmail.com, Zammomin @hotmail.com

Received 18 April 2014; revised 24 May 2014; accepted 9 June 2014

Abstract

In this paper, a reliable computational tool will be developed for the determination of the parameters of the stellar density function in a region of the sky with complete error controlled estimates. Of these error estimates are, the variance of the fit, the variance of the least squares solutions vector, the average square distance between the exact and the least-squares solutions, finally the variance of the density stellar function due to the variance of the least squares solutions vector. Moreover, all these estimates are given in closed analytical forms.

Keywords

Astrostatistics, Stellar Density Function, Computational Astrophysics

1. Introduction

Modern observational astronomy has been characterized by an enormous growth of data stimulated by the advent of new technologies in telescopes, detectors and computations. The new astronomical data give rise to innumerable statistical problems [1]. Moreover, empirical astrophysics researches have seen a paradigm shift in recent years in that it routinely involves data mining of large multi wavelength data sets, requiring complex automated processes that must invoke a very diverse set of statistical techniques (e.g. [2] [3]).

On the other hand, one of the most crucial pieces of information needed in astronomy is the stellar density function in a region of the sky, due to the wealth of information on galactic structure gained directly from a study of the variations in the stellar density (e.g. [4]-[6]).

Although the least-squares method is the most powerful technique that has been devised for the problems of

How to cite this paper: Sharaf, M.A. and Mominkhan, Z.A. (2014) An Error Controlled Method to Determine the Stellar Density Function in a Region of the Sky. *Applied Mathematics*, **5**, 2077-2087. http://dx.doi.org/10.4236/am.2014.513202

astrostatistics in general [7], it is at the same time exceedingly critical. This is because the least-squares method suffers from the deficiency that, its estimation procedure does not have detecting and controlling techniques for the sensitivity of the solution to the optimization criterion of the variance σ^2 is minimum. As a result, there may exist a situation in which there are many significantly different solutions that reduce the variance σ^2 to an acceptable small value.

At this stage we should point out that 1) the accuracy of the estimators and the accuracy of the fitted curve are two distinct problems; and 2) an accurate estimator will always produce small variance, but small variance does not guarantee an accurate estimator. This could be seen from Equation (2) by noting that the lower bounds for the average square distance between the exact and the least-squares values is $\sigma^2/\lambda_{\min}$ which may be large even if σ^2 is very small, depending on the magnitude of the minimum eigenvalue, $\lambda_{\min}$, of the coefficient matrix of the least-squares normal equations. Unless detecting and controlling this situation, it is not possible to make a well-defined decision about the results obtained from the applications of the least squares method.

The importance of the stellar density function as mentioned very briefly as in the above and the existing practical difficulties due to the deficiency of the error estimation and controlling had motivated our work: to develop a reliable computational tool for the determination of the parameters of the stellar density with complete error estimates. Of these error estimates are, the variance of the fit, the variance of the least squares solutions vector, the average square distance between the exact and the least-squares solutions, finally the variance of the density stellar function due to the variance of the least squares solutions vector.

By this we aim at giving an idea on what may called an "accepted solution set" for the parameters of the stellar density functions and the associated variances by the selected tolerances for the error estimates.

Before starting the analysis, it is profitable, to give brief notes on the structure of the paper as follows.

1-Using Fourier transform to obtain analytical solution of the density function;

2-Using the least squares method to find second order polynomial for each of the apparent and absolute magnitudes distributions;

3-Using steps 1 & 2, we established analytical expressions of the density function with coefficients directly obtained from observations.

2. Linear Least Squares Fit

Let y be represented by the general linear expression of the form:

$$\sum_{i=1}^{n} c_i \varphi_i(x),$$

where $\varphi's$ are linear independent functions of x. Let $\boldsymbol{c}$ be the vector of the exact values of the $c's$ coefficients and $\hat{\boldsymbol{c}}$ the least squares estimators of c obtained from the solution of the normal equations of the form $\boldsymbol{G}\hat{\boldsymbol{c}} = \boldsymbol{b}$. The coefficients matrix $\boldsymbol{G}(n\times n)$ is symmetric positive definite, that is, all its eigenvalues $\lambda_i; i = 1, 2, \cdots, n$ are positive. Let $E(z)$ denotes the expectation of z and σ^2 the variance of the fit, defined as:

$$\sigma^2 = \frac{q_n}{(N-n)}$$

where

$$q_n = \left(\mathbf{y} - \boldsymbol{\Phi}^T\hat{\boldsymbol{c}}\right)^T \left(\mathbf{y} - \boldsymbol{\Phi}^T\hat{\boldsymbol{c}}\right)$$

N is the number of observations, $\mathbf{y}$ is the vector with elements y_k and $\boldsymbol{\Phi}(n\times N)$ has elements $\Phi_{ik} = \Phi_i(x_k)$. The transpose of a vector or a matrix is indicated by the superscript "T".

According to the least squares criterion, it could be shown that [8]

1-The estimators $\hat{\boldsymbol{c}}$ obtained by the least squares method gives the minimum of q_n.

2-The estimators $\hat{\boldsymbol{c}}$ of the coefficients $\boldsymbol{c}$, obtained by the least squares method, are unbiased; *i.e.* $E(\hat{\boldsymbol{c}}) = \boldsymbol{c}$.

3-The variance-covariance matrix $Var(\hat{\boldsymbol{c}})$ of the unbiased estimators $\hat{c}$ is given by:

$$Var(\hat{\boldsymbol{c}}) = \sigma^2 \boldsymbol{G}^{-1}, \tag{1}$$

where $\boldsymbol{G}^{-1}$ is the inverse of the matrix $\boldsymbol{G}$.

4-The average squared distance between $\hat{\boldsymbol{c}}$ and $\boldsymbol{c}$ is:

$$E\left(L^2\right)=\sigma^2\sum_{i=1}^{n}\frac{1}{\lambda_i}. \qquad (2)$$

Also it should be noted that, if the precision is measured by probable error e, then:

$$\mathrm{e}=0.6745\sigma .$$

Finally, if R is a linear function of the independent variables ξ,η,ς given by

$$R=S_1\xi+S_2\eta+S_3\varsigma , \qquad (3.1)$$

then [9]

$$\sigma_R^2=S_1^2\sigma_\xi^2+S_2^2\sigma_\eta^2+S_3^2\sigma_\varsigma^2 , \qquad (3.2)$$

where σ_R^2 is the variance of R and $\sigma_\xi^2,\sigma_\eta^2,\sigma_\varsigma^2$ are the variances of the independent variables ξ,η,ς

3. Basic Equations

3.1. The Integral Equation of the Problem

The absolute magnitude, M of a star is given in terms of the apparent magnitude m and parallax p (in second of arc) by

$$M=m+5+5\log p,$$

where M is thus defined in terms of the standard distance of 10 parsecs. We write, for convenience,

$$M_1=M-5 ,$$

so that M_1 is defined in terms of the standard distance of 1 parsec, and

$$M_1=m+5\log p .$$

In the above formulae the base of the logarithm is 10.

We shall refer to M_1 in this connection as the *modified absolute magnitude*. Also, with r measured in parsecs, we have $p=1/r$ and

$$M_1=m-5\log r .$$

Let $\Phi(M)$ be the frequency function of M and $b(m)\mathrm{d}m$ denote the total number of stars with apparent magnitude between m and $m+\mathrm{d}m$ in small region of the sky subtends a solid angle S in the distance interval r and $r+\mathrm{d}r$ where the density function is $D(r)$, then (Trumple & Weaver1953)

$$b(m)=S\int_0^\infty r^2D(r)\Phi(m-5\log r)\mathrm{d}r .$$

Let

$$\rho=-5\log r ,$$

then

$$r=10^{-\rho/5}\Rightarrow \ln r=-\left(\frac{\rho}{5}\right)\ln 10\Rightarrow r=\exp\left\{-\frac{\rho}{5}\ln 10\right\}.$$

Let

$$c=\frac{\ln 10}{5}=0.4605$$

then

$$r=\exp\{-c\rho\} .$$

Consequently

$$\mathrm{d}r=-c\exp\{-c\rho\}\mathrm{d}\rho ,$$

$$r=0 \Rightarrow \rho=\infty,$$

$$r=\infty \Rightarrow \rho=-\infty.$$

Hence

$$b(m) = cS\int_{-\infty}^{\infty}\exp\{-3c\rho\}D\left(e^{-c\rho}\right)\Phi(m+\rho)\,d\rho,$$

or

$$b(m) = \int_{-\infty}^{\infty}\Delta(\rho)\Phi(m+\rho)\,d\rho \tag{4}$$

where

$$\Delta(\rho) = cS\exp\{-3c\rho\}D\left(e^{-c\rho}\right), \tag{5}$$

$$m+\rho = M_1. \tag{6}$$

Equation (4) is the basic integral equation to be solved for the density function $D(r)$ as will be shown latter.

3.2. Maxwellian Distributions of the Magnitudes

Let the distributions of the apparent and absolute magnitudes are Maxwellian in form. We assume that

$$b(m) = a\exp\left\{-k^2(m-m_0)^2\right\}, \tag{7.1}$$

$$\Phi(M_1) = A\exp\left\{-K^2(M-M_0)^2\right\}, \tag{7.2}$$

$$K \succ k. \tag{7.3}$$

As regards Equation (7.1), this is the form found to satisfy the star counts for a given galactic latitude in the exhaustive investigation by many authors. The parameters a, k and m_0 are to be regarded as functions of galactic latitude and possibly also of galactic longitude.

Equation (7.2) must be regard as applicable only to a particular spectral type or subdivision of spectral type. In many studies of the distribution of absolute magnitudes, the separation of stars into the giant and dwarf classes is recognized, that in dealing with a given spectral type we represent the function $\Phi(M_1)$ as the sum of two Maxwellian expressions of the type (7.2). In the following analysis, we deal with a single Maxwellian function only.

The condition (7.3) implies that the dispersion about the mean is less for absolute magnitudes of a given spectral type than for the apparent magnitudes. This is in accordance with observations, for the giants or for the dwarfs.

4. The Normal Equations and the Associated Error Analysis

Taking the natural logarithm of Equations (7.1) and (7.2) we get,

$$y^{(i)} = C_1^{(i)} + C_2^{(i)}x + C_3^{(i)}x^2, \tag{8}$$

where

$$y^{(i)} = \begin{cases}\ln b(m), & \text{if } i=1\\ \ln\Phi(M_1), & \text{if } i=2\end{cases};\quad C_3^{(i)} = \begin{cases}-k^2, & \text{if } i=1\\ -K^2, & \text{if } i=2\end{cases}; \tag{9}$$

$$-\frac{C_2^{(i)}}{2C_3^{(i)}} = \begin{cases}m_0, & \text{if } i=1\\ M_0, & \text{if } i=2\end{cases};\ \exp\left[\frac{4C_1^{(i)}C_3^{(i)} - C_2^{(i)^2}}{4C_3^{(i)}}\right] = \begin{cases}a, & \text{if } i=1\\ A, & \text{if } i=2\end{cases}; \tag{10}$$

and x stands for m if $i=1$ and for M if $i=2$.

Since, $y^{(i)}; i=1,2$ are known from observations for all x, then according to Section2,the normal equations associated with Equations (8) are:

$$\begin{aligned}N\tilde{C}_1^{(i)} + h^{(i)}\tilde{C}_2^{(i)} + g^{(i)}\tilde{C}_3^{(i)} &= Q_1^{(i)},\\ h^{(i)}\tilde{C}_1^{(i)} + g^{(i)}\tilde{C}_2^{(i)} + f^{(i)}\tilde{C}_3^{(i)} &= Q_2^{(i)},\\ g^{(i)}\tilde{C}_1^{(i)} + f^{(i)}\tilde{C}_2^{(i)} + p^{(i)}\tilde{C}_3^{(i)} &= Q_3^{(i)},\end{aligned} \tag{11}$$

where

$$Q_1^{(i)} = \sum_{j=1}^{N} y_j^{(i)};\ \ Q_2^{(i)} = \sum_{j=1}^{N} x_j y_j^{(i)};\ \ Q_3^{(i)} = \sum_{j=1}^{N} x_j^2 y_j^{(i)} \tag{12}$$

$$h^{(i)} = \sum_{j=1}^{N} x_j;\ \ g^{(i)} = \sum_{j=1}^{N} x_j^2;\ \ f^{(i)} = \sum_{j=1}^{N} x_j^3;\ \ p^{(i)} = \sum_{j=1}^{N} x_j^4 \tag{13}$$

In the following two sections, the solutions of the normal equations for $i = 1, 2$ together with the associated error analysis will be developed in closed analytical forms.

4.1. Solutions of the Normal Equations

The solutions of the normal Equations (11) for $i = 1, 2$ are given exactly as,

$$\begin{aligned}
\tilde{C}_1^{(i)} &= \frac{1}{\Delta^{(i)}}\left(\left(f^{(i)2} - p^{(i)}g^{(i)}\right)Q_1^{(i)} + \left(p^{(i)}h^{(i)} - g^{(i)}f^{(i)}\right)Q_2^{(i)} + \left(g^{(i)2} - f^{(i)}h^{(i)}\right)Q_3^{(i)}\right),\\
\tilde{C}_2^{(i)} &= \frac{1}{\Delta^{(i)}}\left(\left(p^{(i)}h^{(i)} - g^{(i)}f^{(i)}\right)Q_1^{(i)} + \left(g^{(i)2} - p^{(i)}N\right)Q_2^{(i)} + \left(Nf^{(i)} - g^{(i)}h^{(i)}\right)Q_3^{(i)}\right),\\
\tilde{C}_3^{(i)} &= \frac{1}{\Delta^{(i)}}\left(\left(g^{(i)2} - f^{(i)}h^{(i)}\right)Q_1^{(i)} + \left(Nf^{(i)} - g^{(i)}h^{(i)}\right)Q_2^{(i)} + \left(h^{(i)2} - g^{(i)}N\right)Q_3^{(i)}\right),
\end{aligned} \tag{14}$$

where

$$\Delta^{(i)} = g^{(i)2} + h^{(i)2}p^{(i)} + f^{(i)2}N - g^{(i)}\left(2f^{(i)}h^{(i)} + NP^{(i)}\right). \tag{15}$$

4.2. Error Analysis

According to Section 2, we deduce for $i = 1, 2$, that:

1-The variance of the fit is:

$$\sigma^{(i)2} = \frac{1}{N-3}\left[\sum_{j=1}^{N} y_j^{(i)2} - N\tilde{C}_1^{(i)2} - 2\tilde{C}_1^{(i)}\tilde{C}_2^{(i)}h^{(i)} - g^{(i)}\left(\tilde{C}_2^{(i)2} + 2\tilde{C}_1^{(i)}\tilde{C}_3^{(i)}\right) - 2\tilde{C}_2^{(i)}\tilde{C}_3^{(i)}f^{(i)} - p^{(i)}\tilde{C}_3^{(i)2}\right]. \tag{16}$$

2-The variance of the solutions are:

$$\begin{aligned}
\sigma^2_{\tilde{C}_1^{(i)}} &= \sigma^2\left(f^{(i)2} - p^{(i)}g^{(i)}\right)\big/\Delta^{(i)}\\
\sigma^2_{\tilde{C}_2^{(i)}} &= \sigma^2\left(g^{(i)2} - p^{(i)}N\right)\big/\Delta^{(i)} \quad ,\\
\sigma^2_{\tilde{C}_3^{(i)}} &= \sigma^2\left(h^{(i)2} - g^{(i)}N\right)\big/\Delta^{(i)}
\end{aligned} \tag{17}$$

3-The average squared distance between the least square solutions and the exact solutions is

$$E\left(L^2\right) = \frac{1}{\Delta^{(i)}}\left(f^{(i)2} + g^{(i)2} + h^{(i)2} - Ng^{(i)} - p^{(i)}\left(g^{(i)} + N\right)\right). \tag{18}$$

5. Analytical Expression of the Density Function *D*(*r*)

Recalling the Fourier transform $f^*(\beta)$ of the function $f(x)$ as:

$$f^*(\beta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\exp\{i\beta x\}\,\mathrm{d}x\ , \tag{19.1}$$

while its inverse is

$$f(x) = \int_{-\infty}^{\infty} f^*(\beta)\exp\{-i\,\beta\,x\}\,\mathrm{d}\beta\ . \tag{19.2}$$

Multiply Equation (4) by $\exp\{i\ \omega\ m\}$ and integrate between $-\infty$ and $+\infty$, then

$$\int_{-\infty}^{\infty} b(m)\exp\{i\ \omega\ m\}\mathrm{d}m = \int_{-\infty}^{\infty} \Delta(\rho)\exp\{-i\ \omega\ \rho\} H(\rho,\omega)\mathrm{d}\rho, \tag{20}$$

where

$$H(\rho,\omega) = \int_{-\infty}^{\infty} \Phi(m+\rho)\exp\{i\ \omega(m+\rho)\}\mathrm{d}m.$$

Let $m+\rho=\alpha$ then,

$$H(\omega) = \int_{-\infty}^{\infty} \Phi(\alpha)\exp\{i\ \omega\ \alpha\}\mathrm{d}\alpha = 2\pi\Phi^*(\omega),$$

also

$$\int_{-\infty}^{\infty} \Delta(\rho)\ \exp\{i\ \omega\ \rho\}\mathrm{d}\rho = 2\pi\Delta^*(-\omega),$$

$$\int_{-\infty}^{\infty} b(m)\ \exp\{i\ \omega\ m\}\mathrm{d}m = 2\pi b^*(\omega).$$

Then Equation (20) reduces to

$$b^*(\omega) = 2\pi\Delta^*(-\omega)\Phi^*(\omega),$$

also

$$b^*(-\omega) = 2\pi\Delta^*(\omega)\Phi^*(-\omega).$$

The inverse of Fourier transform of Δ^* is:

$$\Delta(\rho) = \int_{-\infty}^{\infty} \Delta^*(\omega)\exp\{-i\ \omega\ \rho\}\mathrm{d}\omega$$

then

$$\Delta(\rho) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{b^*(\omega)}{\Phi^*(\omega)}\exp\{i\ \omega\ \rho\}\mathrm{d}\omega, \tag{21}$$

where $\Phi^*(\omega)$ and $b^*(\omega)$ are the Fourier integrals of $\Phi(x)$ and $b(x)$ respectively where

$$\Phi^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi(\alpha)\mathrm{e}^{i\omega\alpha}\mathrm{d}\alpha\ ;\ b^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} b(\alpha)\mathrm{e}^{i\omega\alpha}\mathrm{d}\alpha$$

$b^*(\omega)$ could written as

$$b^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} b(m)\exp\{i\ \omega\ m\}\mathrm{d}m. \tag{22}$$

Using Equation (7.1) in Equation (22) the later becomes:

$$b^*(\omega) = \frac{a}{2\pi}\int_{-\infty}^{\infty} \exp\left\{-k^2(m-m_0)^2 + i\ \omega\ m\right\}\mathrm{d}m,$$

or, on setting $m-m_0=z$, we get:

$$b^*(\omega) = \frac{a}{\pi}\exp\{i\ \omega\ m_0\}\int_{0}^{\infty} \exp\left\{-k^2x^2\right\}\cos\omega z\mathrm{d}z.,$$

evaluating the integral on the right hand side we get

$$b^*(\omega) = \frac{a}{2k\sqrt{\pi}}\exp\left\{-\omega^2/4k^2 + i\ m_0\omega\right\}. \tag{23}$$

Similarly, as in deriving Equation (23) we can get for $\Phi^*(\omega)$ the expression:

$$\Phi^*(\omega) = \frac{A}{2K\sqrt{\pi}}\exp\left\{-\omega^2/4K^2 + i\,M_0\omega\right\}. \tag{24}$$

Now, substituting Equations (23) and (24) into Equation (21) we get,

$$\Delta(\rho) = \frac{aK^2}{A\sqrt{\pi}\cdot\sqrt{K^2-k^2}}\exp\left(-\frac{k^2K^2\left(m_0-M_0+\rho\right)^2}{K^2-k^2}\right).$$

Using Equation (6) and remembering that $r = \exp\{-c\rho\}$ we obtain for the density function the expression:

$$D(r) = \frac{1.22512a\ \exp\left[\dfrac{k^2K^2\left(m_0-M_0-2.17147\ln r\right)^2}{k^2-K^2}\right]K}{Ak\sqrt{\dfrac{1}{k^2}-\dfrac{1}{K^2}}\ S\ r^3}. \tag{25}$$

6. Empirical Determination of the Density Function *D*(*r*) and Its Accepted Solution Set

In what follows empirical expression of the density function $D(r)$ and its variance will be established in literal closed forms.

6.1. Empirical Expression

Substituting Equations (9) and (10) into Equation (25), we get for the density function $D(r)$ the empirical expression

$$D(r) = \frac{\left(1.22512\exp\left[0.25\left(4\tilde{C}_1^{(1)}-4\tilde{C}_1^{(2)}-\dfrac{\left(\tilde{C}_2^{(1)}\right)^2}{\tilde{C}_3^{(1)}}+\dfrac{\left(\tilde{C}_2^{(2)}\right)^2}{\tilde{C}_3^{(2)}}+\tilde{C}_3^{(1)}\left(4.34294\ln r+\dfrac{\tilde{C}_2^{(1)}}{\tilde{C}_3^{(1)}}-\dfrac{\tilde{C}_2^{(2)}}{\tilde{C}_3^{(2)}}\right)^2\tilde{C}_3^{(2)}\right)\sqrt{\dfrac{\tilde{C}_3^{(2)}}{\tilde{C}_3^{(1)}}}\right]\right)}{r^3S\sqrt{\dfrac{1}{\tilde{C}_3^{(2)}}-\dfrac{1}{\tilde{C}_3^{(1)}}}} \tag{26}$$

where $\tilde{C}_j^{(i)};\ j=1,2,3$ and $i=1,2$ are the solutions of the normal equations (Equations (14))

6.2. The Variance σ_D^2

Since $D(r)$ function of $\tilde{C}_j^{(i)};\ j=1,2,3$ and $i=1,2$, then what is the variance σ_D^2 due to the variances $\sigma^2_{C_j^{(i)}}$? The following analysis is devoted for the answer of this question.

Define

$$\Gamma^{(i)} = g^{(i)}p^{(i)} - f^{(i)2},\ B^{(i)} = Np^{(i)} - g^{(i)2},\ E^{(i)} = Ng^{(i)} - h^{(i)2}, \tag{27.1}$$

$$F^{(i)} = g^{(i)}h^{(i)} - Nf^{(i)},\ G^{(i)} = h^{(i)}f^{(i)} - g^{(i)2},\ H^{(i)} = g^{(i)}f^{(i)} - p^{(i)}h^{(i)}, \tag{27.2}$$

$$\alpha_k^{(i)} = \frac{1}{\Delta^{(i)}}\left(\Gamma^{(i)} + x_kH^{(i)} + x_k^2G^{(i)}\right), \tag{28.1}$$

$$\beta_k^{(i)} = \frac{1}{\Delta^{(i)}}\left(H^{(i)} + x_kB^{(i)} + x_k^2F^{(i)}\right), \tag{28.2}$$

$$\gamma_k^{(i)} = \frac{1}{\Delta^{(i)}}\left(G^{(i)} + x_kF^{(i)} + x_k^2E^{(i)}\right), \tag{28.3}$$

therefore we have

$$\Delta^{(i)} = N\ \Gamma^{(i)} + H^{(i)} h^{(i)} + G^{(i)} g^{(i)}, \tag{29.1}$$

$$0 = h^{(i)}\ \Gamma^{(i)} + H^{(i)} g^{(i)} + G^{(i)} f^{(i)}, \tag{29.2}$$

$$0 = g^{(i)}\ \Gamma^{(i)} + H^{(i)} f^{(i)} + G^{(i)} p^{(i)}, \tag{29.3}$$

$$\tilde{C}_1^{(i)} = \sum_{k=1}^{N} \alpha_k^{(i)} y_k^{(i)}, \tag{30.1}$$

$$\tilde{C}_2^{(i)} = \sum_{k=1}^{N} \beta_k^{(i)} y_k^{(i)}, \tag{30.2}$$

$$\tilde{C}_3^{(i)} = \sum_{k=1}^{N} \gamma_k^{(i)} y_k^{(i)}. \tag{30.3}$$

From Equations (29) we get

$$\sum_{k=1}^{N} \alpha_k^{(i)2} = \frac{\Gamma^{(i)}}{\Delta^{(i)}}, \tag{31.1}$$

$$\sum_{k=1}^{N} \beta_k^{(i)2} = \frac{B^{(i)}}{\Delta^{(i)}}, \tag{31.2}$$

$$\sum_{k=1}^{N} \gamma_k^{(i)2} = \frac{E^{(i)}}{\Delta^{(i)}}. \tag{31.3}$$

Multiply Equations (28.1) and (28.2) and then summing, we get

$$\sum_{k=1}^{N} \alpha_k^{(i)} \beta_k^{(i)} = \frac{H^{(i)}}{\Delta^{(i)}}, \tag{32.1}$$

similarly

$$\sum_{k=1}^{N} \alpha_k^{(i)} \gamma_k^{(i)} = \frac{G^{(i)}}{\Delta^{(i)}}, \tag{32.2}$$

$$\sum_{k=1}^{N} \gamma_k^{(i)} \beta_k^{(i)} = \frac{F^{(i)}}{\Delta^{(i)}}. \tag{32.3}$$

Since

$$\alpha_k^{(i)} \Delta^{(i)} = \Gamma^{(i)} + x_k H^{(i)} + x_k^2 G^{(i)}, \tag{33}$$

then summing we have

$$\sum_{k=1}^{N} \alpha_k^{(i)} = 1, \tag{34.1}$$

similarly

$$\sum_{k=1}^{N} x_k \beta_k^{(i)} = 1, \tag{34.2}$$

$$\sum_{k=1}^{N} x_k^2 \gamma_k^{(i)} = 1.. \tag{34.3}$$

Multiply Equation (33) by x_k and summing we get

$$\sum_{k=1}^{N} x_k \alpha_k^{(i)} = 0, \tag{35.1}$$

similarly

$$\sum_{k=1}^{N} x_k^2 \alpha_k^{(i)} = 0, \quad (35.1)$$

$$\sum_{k=1}^{N} \beta_k^{(i)} = 0, \quad (35.2)$$

$$\sum_{k=1}^{N} x_k^2 \beta_k^{(i)} = 0, \quad (35.3)$$

$$\sum_{k=1}^{N} \gamma_k^{(i)} = 0, \quad (35.4)$$

$$\sum_{k=1}^{N} x_k \gamma_k^{(i)} = 0. \quad (35.5)$$

Since $\tilde{C}_i^{(i)}, \tilde{C}_2^{(i)}$ and $\tilde{C}_3^{(i)}$ are the least squares solutions, then the corresponding residual, $v_k^{(i)}$ is given by:

$$\tilde{C}_i^{(i)} + \tilde{C}_2^{(i)} x_k + \tilde{C}_3^{(i)} x_k^2 - y_k^{(i)} = v_k^{(i)}, \quad (36)$$

consequently,

$$\sum_{k=1}^{N} \alpha_k^{(i)} v_k^{(i)} = \sum_{k=1}^{N} \beta_k^{(i)} v_k^{(i)} = \sum_{k=1}^{N} \gamma_k^{(i)} v_k^{(i)} = 0. \quad (37)$$

According to Section 2, we have

$$C_i^{(i)} + C_2^{(i)} x_k + C_3^{(i)} x_k^2 - y_k^{(i)} = \epsilon_k^{(i)}. \quad (38)$$

where $C_j^{(i)}; j = 1,2,3$ and $i = 1,2$ are the exact values of the unknowns and $\epsilon_k^{(i)}$ is the error associated with $y_k^{(i)}$.

Multiply Equations (38) and (37) by $\alpha_k^{(i)}$, subtracting, then summing we get

$$C_1^{(i)} - \tilde{C}_1^{(i)} = \sum_{k=1}^{N} \alpha_k^{(i)} \epsilon_k^{(i)}, \quad (39.1)$$

similarly

$$C_2^{(i)} - \tilde{C}_2^{(i)} = \sum_{k=1}^{N} \beta_k^{(i)} \epsilon_k^{(i)}, \quad (39.2)$$

$$C_3^{(i)} - \tilde{C}_3^{(i)} = \sum_{k=1}^{N} \gamma_k^{(i)} \epsilon_k^{(i)}, \quad (39.3)$$

let us take the error, e, of the density function $D\left(r, \tilde{C}_1^{(i)}, \tilde{C}_2^{(i)}, \tilde{C}_3^{(i)}\right)$ in the sense

$$\begin{aligned} \mathrm{e} &= D\left(r, C_1^{(i)}, C_2^{(i)}, C_3^{(i)}\right) - D\left(r, \tilde{C}_1^{(i)}, \tilde{C}_2^{(i)}, \tilde{C}_3^{(i)}\right) \\ &= D\left(r, \tilde{C}_1^{(i)} + v_1^{(i)}, \tilde{C}_2^{(i)} + v_2^{(i)}, \tilde{C}_3^{(i)} + v_3^{(i)}\right) - D\left(r, \tilde{C}_1^{(i)}, \tilde{C}_2^{(i)}, \tilde{C}_3^{(i)}\right), \end{aligned}$$

then assuming that the errors $\epsilon_k^{(i)}$ in Equations (39) are small, then we can write e with sufficient accuracy by means of Taylor expansion as:

$$\mathrm{e} = D\left(r, \tilde{C}_1^{(i)}, \tilde{C}_2^{(i)}, \tilde{C}_3^{(i)}\right) + v_1 \left(\frac{\partial D}{\partial C_1^{(i)}}\right)_{\tilde{C}} + v_2 \left(\frac{\partial D}{\partial C_2^{(i)}}\right)_{\tilde{C}} + v_3 \left(\frac{\partial D}{\partial C_3^{(i)}}\right)_{\tilde{C}} - D\left(r, \tilde{C}_1^{(i)}, \tilde{C}_2^{(i)}, \tilde{C}_3^{(i)}\right)$$

where $v_j^{(i)} = C_j^{(i)} - \tilde{C}_j^{(i)}; \ j = 1,2,3; \ i = 1,2$,then using Equations(39) we get

$$\mathrm{e} = \sum_{k=1}^{N} \left(\alpha_k^{(i)} L^{(i)} + \beta_k^{(i)} U^{(i)} + \gamma_k^{(i)} P^{(i)} \right) \epsilon_k^{(i)}, \quad (40)$$

where

$$L^{(i)} = \left(\frac{\partial D}{\partial C_1^{(i)}}\right)_{\tilde{C}_j^{(i)}}, \quad U^{(i)} = \left(\frac{\partial D}{\partial C_2^{(i)}}\right)_{\tilde{C}_j^{(i)}}, \quad P^{(i)} = \left(\frac{\partial D}{\partial C_3^{(i)}}\right)_{\tilde{C}_j^{(i)}}.$$

Now, in Equation (40), e is linear function of the errors $\in_k^{(i)}$; hence, then according to Equation (3) we have

$$\sigma_D^2 = \sigma^2 \sum_{k=1}^{N} \left(\alpha_k^{(i)} L + \beta_k^{(i)} U + \gamma_k^{(i)} P\right)^2.$$

Using Equations (31), (32) and (17) we finally get

$$\sigma_D^2 = L^2 \sigma_{C_1^{(i)}}^2 + U^2 \sigma_{C_2^{(i)}}^2 + P^2 \sigma_{C_3^{(i)}}^2 + \frac{2\sigma^2}{\Delta^{(i)}} (\text{FUP+GPL+HLU}), \tag{41}$$

where $L^{(i)}$, $\text{U}^{(i)}$, $\text{P}^{(i)}, i = 1, 2$ are given as

$$L^{(i)} = D(r), \tag{42.1}$$

$$L^{(2)} = -D(r), \tag{42.1}$$

$$U^{(1)} = \left\{-0.5\tilde{C}_2^{(2)} + \left(2.17147 \ln r + 0.5 C_2^{(1)} / C_3^{(1)}\right) C_3^{(2)}\right\} D(r), \tag{42.3}$$

$$U^{(2)} = \left\{-0.5\tilde{C}_2^{(1)} + \tilde{C}_3^{(1)} \left(-2.17147 \ln r + 0.5 C_2^{(2)} / C_3^{(2)}\right)\right\} D(r), \tag{42.4}$$

$$\begin{aligned} P^{(1)} = \Bigg(& \frac{1}{\tilde{C}_3^{(1)2} \left(\tilde{C}_3^{(1)} - \tilde{C}_3^{(2)}\right) \tilde{C}_3^{(2)} \tilde{C}_3^{(1)}} \left(\tilde{C}_1^{(1)} \tilde{C}_3^{(1)2} \left(\tilde{C}_3^{(1)} - \tilde{C}_3^{(2)}\right) \tilde{C}_3^{(2)}\right. \\ & + \left(0.25\, \tilde{C}_2^{(2)2} \tilde{C}_3^{(1)3} + \tilde{C}_3^{(1)2} \left(-0.5 - 0.25 \tilde{C}_2^{(2)2} - 2.17147 \ln r \tilde{C}_2^{(2)} \tilde{C}_3^{(1)}\right) \tilde{C}_3^{(2)}\right. \\ & + \tilde{C}_3^{(1)} \left(-0.25\, \tilde{C}_2^{(1)2} + \tilde{C}_3^{(1)} \ln r \left(2.17147\, \tilde{C}_2^{(2)} + 4.71528 \tilde{C}_3^{(1)} \ln r\right)\right) \tilde{C}_3^{(2)2} \\ & \left.\left. + \left(0.25\, \tilde{C}_2^{(1)2} - 4.71528\, \tilde{C}_3^{(1)2} \ln^2 r\right) \tilde{C}_3^{(2)3}\right) \tilde{C}_3^{(1)}\right)\Bigg) D(r) \end{aligned} \tag{42.5}$$

$$\begin{aligned} P^{(2)} = \Bigg(& \frac{1}{\tilde{C}_3^{(1)} (\tilde{C}_3^{(1)} - \tilde{C}_3^{(2)}) \tilde{C}_3^{(2)3}} \left(\tilde{C}_1^{(2)} \tilde{C}_3^{(1)} \tilde{C}_3^{(2)2} \left(\tilde{C}_3^{(1)} + \tilde{C}_3^{(2)}\right) + \left(-0.25\, \tilde{C}_2^{(2)2} \tilde{C}_3^{(1)3}\right.\right. \\ & + \tilde{C}_3^{(1)2} \tilde{C}_3^{(2)} \left(1 + 0.25 \tilde{C}_2^{(2)2}\right) + \tilde{C}_3^{(1)} \left(-0.5 + 0.25\, \tilde{C}_2^{(1)2} + 2.17147 \tilde{C}_3^{(1)} \tilde{C}_2^{(1)} \ln r\right. \\ & + 4.71528\, \tilde{C}_3^{(1)2} \ln^2 r\Big) \tilde{C}_3^{(2)2} + \left(-0.25\, \tilde{C}_2^{(1)2} - 2.17147\, \tilde{C}_3^{(1)} \tilde{C}_2^{(1)} \ln r\right. \\ & \left.\left. - 4.71528\, \tilde{C}_3^{(1)2} \ln^2\right) \tilde{C}_3^{(2)3}\right) \tilde{C}_3^{(2)}\Big)\Bigg) D(r) \end{aligned} \tag{42.6}$$

6.3. The Variances of k^2, K^2, m_0, M_0, a, A

Since each of the constants k^2, K^2, m_0, M_0, a, A is a function of the least squares solutions, the by the same arguments as for Equations (42) we get

$$\begin{aligned} \sigma_{\tilde{C}_1^{(i)}}^2 &= \sigma^2 \left(f^{(i)2} - p^{(i)} g^{(i)}\right) / \Delta^{(i)}, \\ \sigma_{\tilde{C}_2^{(i)}}^2 &= \sigma^2 \left(g^{(i)2} - p^{(i)} N\right) / \Delta^{(i)}, \\ \sigma_{\tilde{C}_3^{(i)}}^2 &= \sigma^2 \left(h^{(i)2} - g^{(i)} N\right) / \Delta^{(i)}, \end{aligned} \tag{17}$$

$$\sigma^2_{W^{(i)}} = -\sigma^2_{\tilde{C}_3^{(i)}} \tag{43.1}$$

$$\sigma^2_{T^{(i)}} = \frac{1}{4C_3^{(i)2}}\left(\sigma^2_{\tilde{C}_3^{(i)}} + \frac{C_2^{(i)2}}{C_3^{(i)2}}\sigma^2_{\tilde{C}_3^{(i)}} + \frac{2\sigma^2 C_2^{(i)2}}{\Delta^{(i)}C_3^{(3)}}\left(N f^{(i)} - g^{(i)}h^{(i)}\right)\right) \tag{43.2}$$

$$\begin{aligned}\sigma^2_{U^{(i)}} = U^{(i)2}\Big\{&\sigma^2_{\tilde{C}_1^{(i)}} + T^{(i)2}\sigma^2_{\tilde{C}_2^{(i)}} + T^{(i)4}\sigma^2_{\tilde{C}_3^{(i)}} + \frac{2\sigma^2}{\Delta^{(i)}}\Big(\left(g^{(i)}h^{(i)} - Nf^{(i)}\right)T^{(i)3} \\ &+\left(h^{(i)}f^{(i)} - g^{(i)2}\right)T^{(i)2} + \left(g^{(i)}f^{(i)} - h^{(i)}p^{(i)}\right)T^{(i)}\Big)\Big\}\end{aligned} \tag{43.3}$$

where

$$W^{(i)} = \begin{cases} k^2 \\ K^2 \end{cases}; \quad T^{(i)} = \begin{cases} m_0 \\ M_0 \end{cases}; \quad U^{(i)} = \begin{cases} a \\ A \end{cases} \tag{44}$$

6.4. An Accepted Solution Set for $D(r)$

Due to the above mentioned practical difficulties encountered in most applications of the least squares method we should at this stage reformalize the concept of an "acceptably small" variance. We may define an acceptable solution set to the determination of $D(r)$ as:

$$C = \left\{\tilde{C} : \sigma^2 \le \text{Tol}, \ \left(\sigma_C^2, \sigma_D^2, \sigma^2_{W^{(i)}}, \sigma^2_{T^{(i)}}, \sigma^2_{U^{(i)}{}^{(i)}}\right) \le \mu\right\} \tag{45}$$

where Tol and μ small numbers. In writing Equation (45) we do not mean to establish this particular definition of an acceptable solution set, as it is only intended to give the users of the least squares method for $D(r)$ some degree of concreteness to the general idea of an acceptable solution set.

5. Conclusion

In conclusion, a reliable computational tool was developed in the present paper for the determination of the parameters of the stellar density function in a region of the sky with complete error controlled estimates. Of these error estimates are, the variance of the fit, the variance of the least squares solutions vector, the average square distance between the exact and the least-squares solutions, finally the variance of the density stellar function due to the variance of the least squares solutions vector. Moreover, all these estimates are given in closed analytical forms.

References

[1] Feigelson, E.D and Babu, J.B (2012) Modern Statistical Methods for Astronomy with Applications. Cambridge University Press, Cambridge.

[2] Sharaf, M.A, Issa, I.A. and Saad, A.S. (2003) Method for the Determination of Cosmic Distances. *New Astronomy*, **8**, 15-21. http://dx.doi.org/10.1016/S1384-1076(02)00198-7

[3] Sharaf, M.A. and Sendi A.M. (2010) Computational Developments for Distance Determination of Stellar Groups. *Journal of Astrophysics and Astronomy*, **31**, 3-16. http://dx.doi.org/10.1007/s12036-010-0002-0

[4] Trumpler, R.J. and Weaver, H.F. (1953) Statistical Astronomy. Dover Publication, Inc., New York.

[5] Robinson, R.M. (1985) The Cosmological Distance Ladder. W.H. Freeman and Company, New York.

[6] Binney, J. and Merrifield, M. (1998) Galactic Astronomy. Princeton University Press, Princeton.

[7] Andreon, S. and Hurn, M. (2013) Measurement Errors and Scaling Relations in Astrophysics: A Review. *Statistical Analysis and Data Mining*: *The ASA Data Science Journal*, **6**, 15-33.

[8] Kopal, Z. and Sharaf, M.A. (1980) Linear Analysis of the Light Curves of Eclipsing Variables. *Astrophysics and Space Science*, **70**, 77-101. http://dx.doi.org/10.1007/BF00641665

[9] Bevington, P.R and Robinson, D.K. (1992) Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, New York.