# First, second and third massive stars in Open Clusters

A. Mints

September 24, 2009

**Abstract**

## 1 Introduction

Unfortunatelly, in most cases it is not possible to get large statistics of star cluster membership, so estimation of cluster mass and/or full member number is not an easy task. Sometimes when spectoscopical data is avaliable dynamical mass estimates are applied (making use of the virial theorem). Usually, only few brightest stars are safely (???) identified as members (cite????), providing tiny amount of information about the cluster.

Recently Weidner and Kroupa ([3]) discussed the correlation between the mass of most massive star in the cluster and the total cluster mass. Maschberger and Clarke ([2]) discussed correlation between number of stars in the cluster and it's most massive member mass.

Here we will try to produce cluster mass and membership estimators using masses of three most massive members, and analyse precision of these estimators. We will concentrate on two questions:

- what data provides the most reliable information on the cluster properties;
- what is the best method to extract cluster properties from that data.

Having this to goals in mind, we will neglect two very important factors: stellar binarity and evolution. Of course, we can not weight stars directly, we can only estimate their mass, mainly by their brightness. But, again, as we are interested in the statistical side of the problem, we will postpone astrophysical difficulties for later research.

If we assume all stars in the cluster to be single, then (initial) mass of the most massive star depends only on the initial mass function (hereafter IMF). Here we will consider only one IMF — from Kroupa ([1]). It is built from several power-law parts:

$$\begin{aligned} f(m)dm &= Cm^{-\alpha}dm \\ F(m) &= \int_0^{M_{\max}} f(m)dm \end{aligned} \tag{1}$$

IMF parameters was taken from Kroupa (???):

$$\begin{aligned} \alpha_0 &= +0.30 \quad 0.01 \le m/M_\odot < 0.08 \\ \alpha_1 &= +1.30 \quad 0.08 \le m/M_\odot < 0.50 \\ \alpha_2 &= +2.35 \quad 0.50 \le m/M_\odot < M_{\max} \end{aligned} \tag{2}$$

There are still debates on the value of $M_{\max}$, so several (???) values will be considered in this paper. We will try to see, how $M_{\max}$ influences mass estimator precision.

We will write $\bar{m}_1$ for average value of $m_1$ and $\tilde{m}_1$ for median value (same for $m_2$ and $m_3$). Another possibility is to use the peak of distribution of $m_{1,2,3}$ for a given $M_{\rm cl}$ or $N$, we will designate them $\hat{m}_{1,2,3}$.

IMF described above has an average stellar mass $\bar{m} = 0.36M_\odot$ for $M_{\max} = 150M_\odot$.

# 2 Model

As in [3], three different methods for generating cluster were used:

**Random sampling** — $N$ stars are taken randomly from the IMF, with $N$ ranging from 300 to 5000.

**Constrained sampling** — in this case $M_{\rm cl}$ is fixed first, then stars are taken from the IMF until their total mass surpass $M_{\rm cl}$. Thus some spread in $N$ is expected in this sample.

**Sorted sampling** — $M_{\rm cl}$ is also fixed, then $N' = M_{\rm cl}/\bar{m}$ stars are taken from the IMF. If $M' = \sum_{N'} m_i$ smaller than $M_{\rm cl}$, then $\Delta N = (M_{\rm cl} - M')/\bar{m}$ stars are added to the cluster, repeating the procedure until cluster mass outreaches $M_{\rm cl}$. Then stellar masses are sorted. If $|M' - M_{\rm cl}|$ is larger then $|M_{\rm cl} - (M' - m_1)|$ then the heaviest star is removed from the set.

Random sampling is the least realistic model, but it is the easiest one to be modeled and described analytically.

For each set of parameters, 30000 clusters were simulated, saving for each of them five values: cluster mass $M_{\rm cl}$, number of stars in the cluster $N$, and masses of three most massive stars of the cluster — $m_1, m_2, m_3$.

The goal is to build a method to find $M_{\rm cl}$ and/or $N$, knowing $m_1, m_2$ and $m_3$. It seems natural to find functions $M_{\rm cl}(\bar{m}_{1,2,3})$, $M_{\rm cl}(\tilde{m}_{1,2,3})$ and $M_{\rm cl}(\hat{m}_{1,2,3})$ (as well as $N(\bar{m}_{1,2,3})$, $N(\tilde{m}_{1,2,3})$ and $N(\hat{m}_{1,2,3})$), we will call them hereafter *mass estimators*.

# 3 Analitycs

## Random sampling

The probability for the most massive star to have mass $m_1 \in (m, m+dm)$ writes as a probability for a given star to have mass in $(m, m+dm)$ multiplied by the probability that all other stars have masses below $m$ and by number of stars $N$ (because any of them can be the most massive one):

$$\begin{aligned} P(m_1 \in (m, m+dm)) &= N f(m)\, [F(m)]^{N-1} \\ &= N f(m) \left[1 - \int_m^{M_{\rm max}} f(m')dm'\right]^{N-1} \end{aligned} \tag{3}$$

Of course, $m_1$ should be smaller then $M_{\rm max}$.

Substituting 1 into 4 we get:

$$P(m_1 \in (m, m+dm)) = NCm^{-\alpha} \left[1 - \frac{C}{1-\alpha}\left(M_{\rm max}^{1-\alpha} - m^{1-\alpha}\right)\right]^{N-1} \tag{4}$$

If $N$ is large, than we can use exponent instead of square brackets:

$$P(m_1 \in (m, m+dm)) \simeq NCm^{-\alpha} \exp\left(-\frac{NC}{1-\alpha}\left(M_{\rm max}^{1-\alpha} - m^{1-\alpha}\right)\right) \tag{5}$$

Maximum of this distribution is located at the point

$$\hat{m}_1 = \left(\frac{NC}{\alpha}\right)^{1/(\alpha-1)} \tag{6}$$

For $\hat{m}_1 \geq M_{\rm max}$ obviously maximum is at the point $\hat{m}_1 = M_{\rm max}$. Inverting this equation we can get an estimate for $N$ and $M_{\rm cl}$ from $\hat{m}_1$:

$$\begin{aligned} N &= \frac{\hat{m}_1^{\alpha-1}\alpha}{C} \\ M_{\rm cl} &= \frac{\bar{m}\hat{m}_1^{\alpha-1}\alpha}{C} \end{aligned} \tag{7}$$

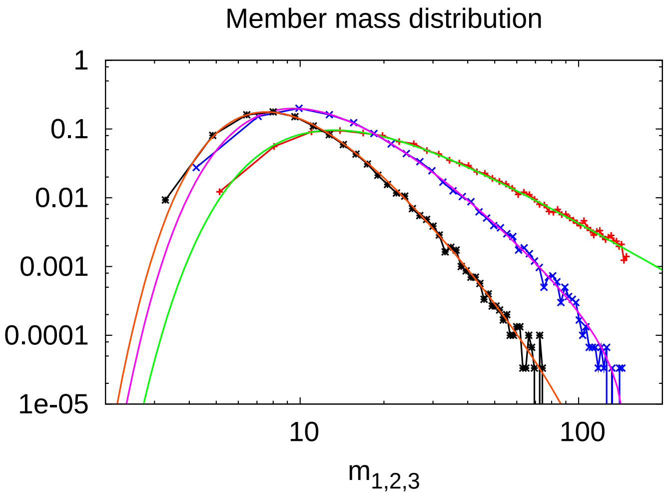


Figure 1: Distribution of $m_1$ (red), $m_2$ (blue) and $m_3$ (black) with theoretical estimates from eq. 5 and 8 (green, magenta and orange, respectively) for $N = 1000$

For $n$'th massive star, if $n \ll N$ we can use the expression:

$$P(m_n \in (m, m+dm)) \simeq (1 - F(m))^{n-1} P(m_1 \in (m, m+dm)) \tag{8}$$

Finding average and median values for the function 5 is not that easy. (???)

### Other samplings

Unfortunatelly the other samplings are too complex to be described analytically in a useful way.

# 4 Results

## 4.1 Random sampling

Random sampling model has as natural parameter number of stars in the cluster — $N$. Here range of $N$ from 300 to 5000 was studied, with 30000 clusters being simulated for each value of $N$.

For each value of $N$ distributions of $m_1, m_2$ and $m_3$ were obtained. An example of these distributions is shown on figure 4.1. Theoretical estimates from eq. 5 and 8 matches the data. Note long power-law tails of distributions, specially for $m_1$. This tail leads to significant difference between average and median value, making average much higher. Thus averages are not so good for making cluster mass estimators.

These functions are shown at figure 4.1. They can be approximated with the function of the shape:

$$\begin{aligned} M_{\rm cl}(m_{1,2,3}) &= a m_{1,2,3}^b (150 - m_{1,2,3})^c \\ N(m_{1,2,3}) &= a m_{1,2,3}^b (150 - m_{1,2,3})^c \end{aligned} \tag{9}$$

so function rises as power law for small $m$ and then "saturates" when $m$ goes to $150 M_\odot$.

Parameters of fits are shown at the table 4.1. For $N(\hat{m}_{1,2,3})$ equations 8 can be used. Note that $b$ is always close to $\alpha_2 - 1 = 1.35$ and $c$ is close to $-1$, although $f(x) = am^{1.35}(150 - m)^{-1}$ is a bad fit.

Now, having this approximations in hand, we can turn back to initial simulated data, and check, how good this estimations are. Namely, we will substitute $m_{1,2,3}$ for each cluster into mass estimators to get $M_{\rm cl}$, $N$ that can be compared with the real values. This will produce some distributions of estimated $M_{\rm cl}$ and $N$, together with estimate errors, as far as the "real" values are known from simulation. The sample

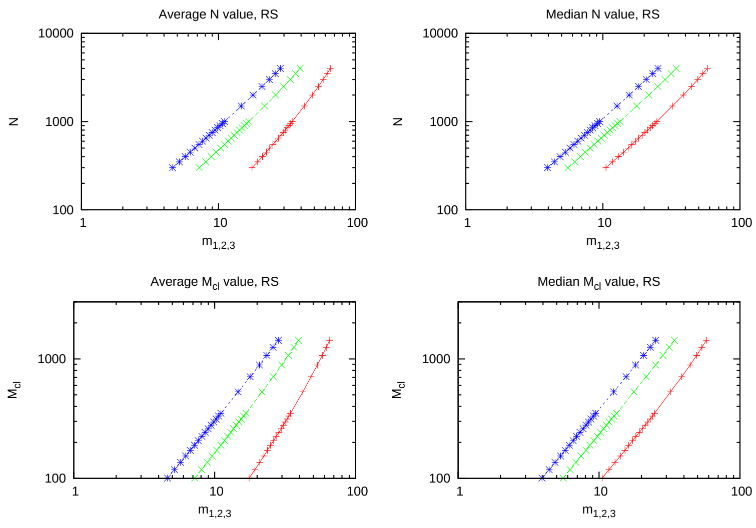


Figure 2: Median and average number of stars $N$ and cluster mass $M_{\rm cl}$ as functions of masses of three most massive cluster members $m_{1,2,3}$ (red, green and blue lines, respectively).

Table 1: Parameters of fits (see eq. 9)

| Fits based on average values | | | |
|---|---|---|---|
| Function | a | b | c |
| $M_{\rm cl}(m_1)$ | 167.90 | 1.66 | -1.08 |
| $M_{\rm cl}(m_2)$ | 1279.49 | 1.40 | -1.07 |
| $M_{\rm cl}(m_3)$ | 563.81 | 1.39 | -0.77 |
| $N(m_1)$ | 915.99 | 1.59 | -1.17 |
| $N(m_2)$ | 8159.25 | 1.34 | -1.20 |
| $N(m_3)$ | 4925.55 | 1.34 | -0.97 |
| Fits based on median values | | | |
| $M_{\rm cl}(m_1)$ | 211.92 | 1.37 | -0.80 |
| $M_{\rm cl}(m_2)$ | 228.72 | 1.37 | -0.64 |
| $M_{\rm cl}(m_3)$ | 55.20 | 1.40 | -0.27 |
| $N(m_1)$ | 913.85 | 1.32 | -0.86 |
| $N(m_2)$ | 1674.88 | 1.32 | -0.80 |
| $N(m_3)$ | 538.08 | 1.35 | -0.49 |

Table 2: Relative error (in procents) of average value of estimated values

| | Random sampling | | | Constrained sampling | | | Sorted sampling | | |
|---|---|---|---|---|---|---|---|---|---|
| Value | $f(m_1)$ | $f(m_2)$ | $f(m_3)$ | $f(m_1)$ | $f(m_2)$ | $f(m_3)$ | $f(m_1)$ | $f(m_2)$ | $f(m_3)$ |
| | | | | Median estimator | | | | | |
| N | 0.49 | 0.36 | 0.21 | 0.36 | 0.26 | 0.32 | 0.76 | 1.33 | 1.04 |
| M | 0.39 | 0.43 | 0.22 | 0.40 | 0.34 | 0.22 | 1.54 | 0.61 | 0.59 |
| | | | | Average estimator | | | | | |
| N | 45.22 | 27.38 | 18.67 | 29.49 | 20.45 | 14.35 | 24.76 | 5.65 | 1.83 |
| M | 46.11 | 27.95 | 19.05 | 29.11 | 19.89 | 13.69 | 29.52 | 11.17 | 7.34 |
| | | | | Distribution maximum estimator | | | | | |
| N | 113.43 | 56.07 | 35.90 | - | - | - | - | - | - |

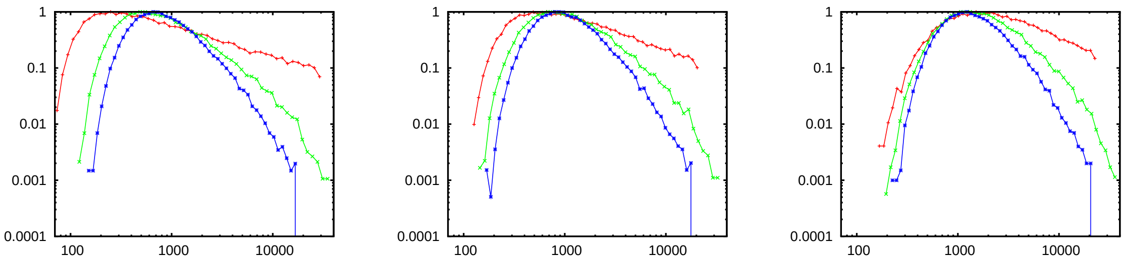


Figure 3: Estimates for a number of stars in the cluster $N(m_{1,2,3})$ (red, green and blue lines, respectively) for a cluster with pre-defined 1000 stars. Estimators are based on average values (top), median values (middle) and distribution maxima (bottom).

of the result for $N(m)$ is shown in the figure 4.1 for a cluster with pre-defined number of stars ($N = 1000$). Note, that there is a large power-law tail at the high-mass side, that is highest for $N(m_1)$ and smallest for $N(m_3)$ in all cases. Generally $N(m_3)$ shows smaller spread than other estimators. The distribution of $N(m_3)$ also peaks closer to the real value $N = 1000$ for $N(m_3)$. Table 4.1 summarise relative errors for various estimators. Median estimator seems to show the best precision. This is due to the fact, that median value is more representative for a distribution with long tails. Notice, that in all cases estimators based on $m_3$ show best results.

## 4.2 Constrained sampling

For constrained sampling almost the same algorithm was applied. the only change is that we don't have analitical formula for $\hat{m}_1, 2, 3$, so we have to use fits of the shape $f(\hat{m}) = a\hat{m}^b$.

From the figure 4.2 one can see that difference between random and constrained samplings are not very large in most cases. Distribution for constrained sampling rises and falls not slower then the one for random sampling. Faster decrease at the distribution high end for small cluster (4.2, top) is due to the fact, that as total mass comes close to desired $M_{\rm cl}$, massive stars are preferentially rejected from the sample, as far as adding them will make the cluster too massive. Obviously, this effect vanishes for higher $M_{\rm cl}$, as one can see from the bottom panel of 4.2.

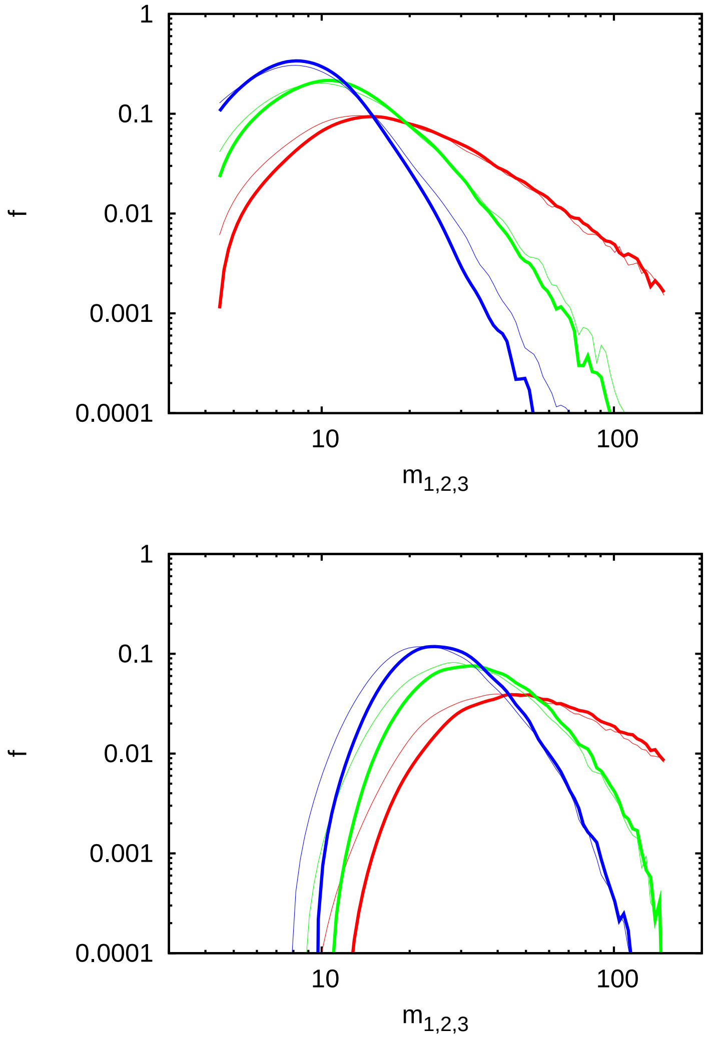


Figure 4: Normalized distribution of $m_1$ (red), $m_2$ (green) and $m_3$ (blue) for constrained sampling (thick lines) and for random sampling (thin lines, same as at figure 4.1). Top panel: $N = 1000; M_{\rm cl} \approx 300$. Bottom panel: $N = 4500; M_{\rm cl} \approx 1500$.

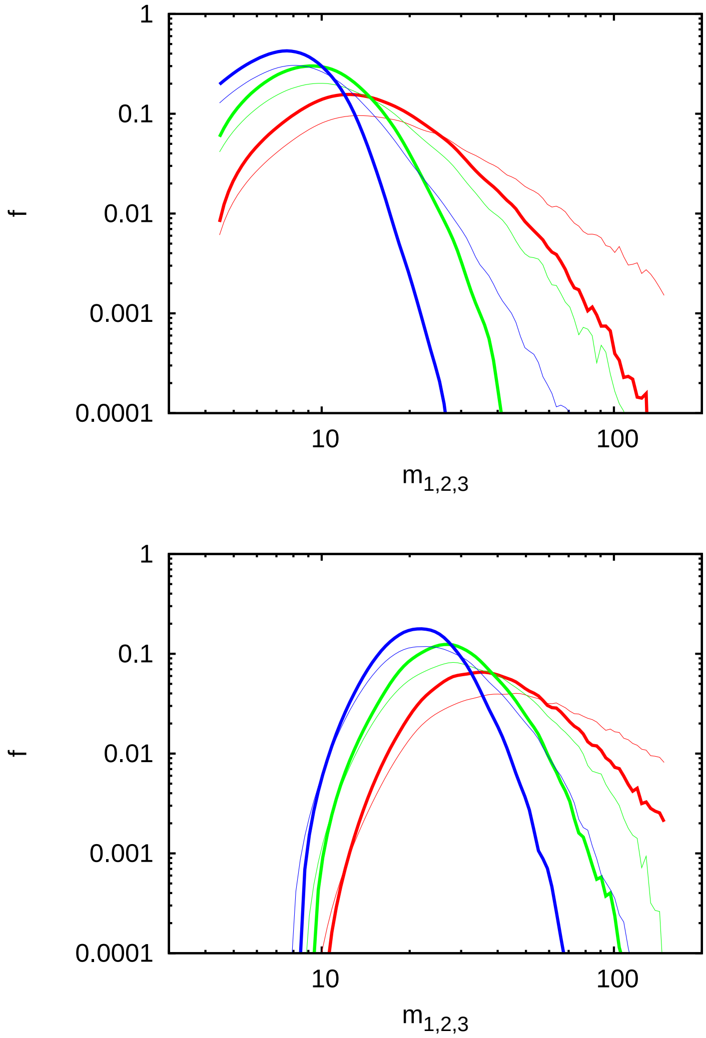


Figure 5: Normalized distribution of $m_1$ (red), $m_2$ (green) and $m_3$ (blue) for sorted sampling (thick lines) and for random sampling (thin lines, same as at figure 4.1). Top panel: $N = 1000; M_{\rm cl} \approx 300$. Bottom panel: $N = 4500; M_{\rm cl} \approx 1500$.

### 4.3 Sorted sampling

## References


[1] P. Kroupa. On the variation of the initial mass function. *mnras*, 322:231–246, April 2001.

[2] T. Maschberger and C. J. Clarke. Maximum stellar mass versus cluster membership number revisited. *mnras*, 391:711–717, December 2008.

[3] C. Weidner and P. Kroupa. The maximum stellar mass, star-cluster formation and composite stellar populations. *mnras*, 365:1333–1347, February 2006.