The Astrophysical Journal, 629:873–880, 2005 August 20

NUMERICAL BIASES ON INITIAL MASS FUNCTION DETERMINATIONS CREATED BY BINNING

J. Maíz Apellániz[1] and L. Úbeda
Space Telescope Science Institute,[2] 3700 San Martin Drive, Baltimore, MD 21218; jmaiz@stsci.edu
Received 2005 February 28; accepted 2005 April 28

ABSTRACT

We detect and quantify significant numerical biases in the determination of the slope of power laws with Salpeter (or similar) indices from uniformly binned data using χ^2 minimization. The biases are caused by the correlation between the number of stars per bin and the assigned weights and are especially important when the number of stars per bin is small. This result implies the existence of systematic errors in the values of IMFs calculated in this way. We propose as an alternative using variable-size bins and dividing the stars evenly among them. Such variable-size bins yield very small biases that are only weakly dependent on the number of stars per bin. Furthermore, we show that they allow for the calculation of reliable IMFs with only a small total number of stars. Therefore, they are a preferred alternative to the standard uniform-size binning.

Subject headings: methods: numerical — methods: statistical — stars: luminosity function, mass function

1. INTRODUCTION

The data explosion of the last decade created by the growth in telescope diameters, number of space missions, and detector surface area has produced a comparable growth in the number of objects that can be observed with a single exposure or analyzed in a single paper. In the field of stellar astronomy, this has translated into thousands, or larger numbers, of stars being included in a single color-magnitude diagram or a single luminosity or mass function. Indeed, many articles aim at measuring the initial mass function (IMF) of a cluster or of a stellar population by measuring the luminosity of each star, converting the results into masses (using isochrone fitting or directly counting the number of stars between evolutionary tracks), and fitting a power law or similar function to the masses organized in bins, using χ^2 minimization.

A known numerical effect takes place when data are binned and a function is fitted to the outcome: a bias in the derived parameters can be (and usually is) present if there is a low number of objects in some of the bins (see, e.g., Bevington & Robinson 1992; Nousek & Shue 1989). The problem originates in the strong anticorrelation between the data and the weights in χ^2 minimization for data with Poisson uncertainties (Wheaton et al. 1995); the same should be true for binomial uncertainties, which in most cases can be well approximated by Poisson ones (see, e.g., Bevington & Robinson 1992). The bias can be especially large when some bins have 0 or 1 count; in the former case some algorithms set the weights to correspond to that of 1 count, but that does not eliminate the effect. Furthermore, the bias can also be present when the number of objects per bin is not too small if some bins are more heavily populated than others (Wheaton et al. 1995). The problems associated with this bias are usually correctly dealt with in some astronomical fields, such as in high-energy astronomy, but it is surprising to find that many of the articles on the IMF appear to ignore them. Thus, it is not strange to find articles in which an IMF is calculated with some bins having just a few stars and other bins with tens or hundreds of them (see Humphreys & McElroy 1984; Massey et al. 1989; Béjar et al. 2001 for some examples). Such a binning scheme introduces biases in the calculated IMF slopes (Kroupa 2001; Elmegreen 2004).

There are different approaches to minimizing or eliminating binning biases. Kearns et al. (1995) suggest using the fitted number of counts (instead of the real number of counts) to calculate the weight of each bin, a method that is iterative by nature since one does not know the fitted number of counts a priori. An alternative recommended by D'Agostino & Stephens (1986) involves two measures:

1. Define bins of variable size in x and adjust them in such a way that each one of them has approximately the same number of objects. The reasoning behind this idea is to assign the same statistical weight to each bin and thus minimize biases.
2. If N is the total number of objects, then divide the data into $\approx 2N^{2/5}$ bins.

In this article we set out to quantify the importance of binning biases for the determination of the IMF by means of a series of simple numerical experiments. We point out that our results should be relevant not only to IMF calculations but to similar problems as well, such as the calculation of cluster mass or luminosity distributions. We start by using "standard" uniform-size bins, and then we explore the use of nonuniform ones, as suggested by D'Agostino & Stephens (1986).

2. DEFINITIONS AND EXPERIMENTS

Following the Scalo (1986) notation, we define the IMF for stars $f(m)$ such that $f(m)\,dm$ is the number of stars formed at the same time in some volume of space with mass in the interval m to $m + dm$. Throughout this article, we assume that the true IMF follows a power-law distribution:

$$\frac{dn}{dm} = f(m) = Am^{\gamma}, \tag{1}$$

where the Salpeter slope is given by $\gamma = -2.35$ (Salpeter 1955).

The number of stars created in the interval $[m_a, m_b]$ can be obtained by integrating equation (1):

$$\int_{m_a}^{m_b} \frac{dn}{dm}\,dm = A \int_{m_a}^{m_b} m^{\gamma}\,dm \tag{2}$$

[1] Affiliated with the Space Telescope Division of the European Space Agency, European Space Research and Technology Centre, Noordwijk, Netherlands.

[2] The Space Telescope Science Institute is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

to obtain

$$n|_{m_a}^{m_b} = n(m_b) - n(m_a) = \frac{A}{\gamma+1}\left(m_b^{\gamma+1} - m_a^{\gamma+1}\right), \quad \gamma \neq -1. \tag{3}$$

Defining $\Delta m_i = m_b - m_a$ and $x_i = (m_a + m_b)/2$, it is possible to rewrite the last expression as

$$N_i \equiv n|_{m_a}^{m_b} = \frac{A}{\gamma+1}\left[\left(x_i + \frac{\Delta m_i}{2}\right)^{\gamma+1} - \left(x_i - \frac{\Delta m_i}{2}\right)^{\gamma+1}\right]. \tag{4}$$

Finally, we obtain the logarithm of both sides to obtain an expression in the form of y_i (logarithm of the number of stars in bin i) as a function of x_i (mass at the center of bin i as defined in a linear scale):

$$y_i \equiv \log_{10} N_i = \log_{10}\left\{\frac{A}{\gamma+1}\left[\left(x_i + \frac{\Delta m_i}{2}\right)^{\gamma+1} - \left(x_i - \frac{\Delta m_i}{2}\right)^{\gamma+1}\right]\right\}. \tag{5}$$

For our numerical experiments we used a random number generator to produce 1000 realizations with 1000 stars each, distributed according to equation (1) with a Salpeter slope between $m = 6.31\ M_\odot$ ($\log_{10} m/M_\odot = 0.8$) and $m = 158.49\ M_\odot$ ($\log_{10} m/M_\odot = 2.2$). The extremes were selected as typical for studies of massive stars (see, e.g., Oey & Clarke 2005), in which the IMF has been measured to be close to the Salpeter value under different circumstances, but the conclusions of this work are not expected to change if somewhat different values are used. The 1000 realizations were manipulated in two different ways to test different conditions:

1. Selecting the first 30, 100, 300, and 1000 (all) stars to simulate different total numbers of stars $N = \Sigma N_i$.
2. Binning the data into 3, 5, 10, 30, or 50 bins to test whether there is an optimal bin size. Given that the D'Agostino & Stephens (1986) recommendation yields 7.8, 12.6, 19.6, and 31.7 bins for our four values of N, the bins selected here should be able to test its validity for our experiments.

We performed experiments using three types of binning:

1. Uniform bin size in logarithmic scale (or $\log_{10}[(x_i + \Delta m_i/2)/(x_i - \Delta m_i/2)]$ constant) with the left edge of the first bin, $m_{\rm down}$, equal to $10^{0.8}\ M_\odot$ and the right edge of the last bin, $m_{\rm up}$, equal to $10^{2.2}\ M_\odot$ (i.e., the input values used for the random generation of all realizations).
2. Approximately constant number of stars in each of the bins with $m_{\rm down} = 10^{0.8}\ M_\odot$ and $m_{\rm up} = 10^{2.2}\ M_\odot$. This rule cannot be made exact, since it is not possible to divide, e.g., 100 stars in 30 bins with the same number of stars in each. In such a case, we design our bins so that they contain, e.g., 3, 4, 3, 3, 4, 3, . . . stars.
3. Same as the previous experiment but with $m_{\rm down}$ and $m_{\rm up}$ determined from the data in each realization.

The issue of the weights to be applied to the problem of IMF fitting has generated some confusion. Some authors have even decided to skip it altogether by not using any (Massey et al. 1995), but such a strategy can yield large biases. When calculating an IMF we are assuming that an underlying physical law determines the true probability distribution $f(m)$, which we could measure with arbitrary precision if an infinite number of stars were generated from it. In reality, we are limited to analyzing finite samples of size N drawn from $f(m)$. It is easy to show that, under such circumstances, the value of N_i follows a binomial distribution characterized by N and p_i, where

$$p_i = \frac{\int_{x_i-\Delta m_i/2}^{x_i+\Delta m_i/2} f(m)\, dm}{\int_{m_{\rm down}}^{m_{\rm up}} f(m)\, dm}, \tag{6}$$

which has a mean of Np_i and a variance of $Np_i(1 - p_i)$. Note that the Poisson approximation corresponds to the case in which $p_i \ll 1$; however, if a bin contains a large fraction of the objects, then such an approximation is no longer correct. The problem here, as noted by Wheaton et al. (1995), is that the true p_i needs to be determined from the true $f(m)$, which is unknown. One obvious alternative is to replace p_i by N_i/N, the value derived from the data, so that the estimate for the uncertainty associated with N_i becomes $[N_i(N - N_i)/N]^{1/2}$. That is the alternative we use, keeping in mind that this step is precisely the source of the potential biases we have previously referred to (Wheaton et al. 1995) and that we are trying to minimize by selecting bins with similar numbers of stars.

With the above considerations and given that (1) $y_i = \log_{10} N_i$ and (2) the weight w_i associated with a value y_i with uncertainty s_i (i.e., $y_i \pm s_i$)[3] when using a χ^2 minimization algorithm is $1/s_i^2$, we have

$$w_i = \frac{N_i N}{(N - N_i)(\log_{10} e)^2}, \tag{7}$$

which has the consequence of giving zero weight to bins with no stars. This has the advantage of yielding no numerical problems there, as opposed to using N_i for the independent variable, which produces infinite weights.[4]

For each of the k (k = 1–1000) realizations in each of the N + bin size combinations, the resulting data (x_i, Δm_i, y_i, w_i) were fitted using a χ^2 minimization algorithm programmed in IDL in order to obtain γ_k and its uncertainty σ_k. (The algorithm also yields A_k and its uncertainty, which is ignored here.) In order to test for a possible algorithm dependence, we calculated our results using both (1) `CURVEFIT`, which uses gradient expansion, can be found in the standard IDL distribution, and is originally based on *Numerical Recipes* (Press et al. 1986) and (2) Craig B. Markwardt's[5] `MPCURVEFIT`, which is based on MINPACK-1, available from Netlib. No significant differences were found between the two algorithms.

We analyzed the results for the ensemble of 1000 realizations in each N + bin size combination. More specifically, we calculated the mean value for the power-law exponent

$$\overline{\gamma} = \frac{1}{1000}\sum_{k=1}^{1000} \gamma_k, \tag{8}$$

the mean uncertainty in the power-law exponent

$$\overline{\sigma} = \frac{1}{1000}\sum_{k=1}^{1000} \sigma_k, \tag{9}$$

[3] The expression s_i is the standard deviation derived from the parent distribution of y_i and its measured value, not from the expected one. Therefore, what we are describing here is the modified χ^2 minimization method (Wheaton et al. 1995).

[4] Note, however, that in our case w_i becomes infinite when $N_i = N$, but, of course, trying to fit a function to a histogram that has all data in a single bin is an ill-defined problem.

[5] See http://cow.physics.wisc.edu/~craigm/idl/idl.html.

TABLE 1
Results for the Uniform Bin Size Case for 3, 5, 10, 30, and 50 Bins

	$\overline{\gamma}$					$\overline{\sigma}$					b				
Stars	3	5	10	30	50	3	5	10	30	50	3	5	10	30	50
30............	−2.308	−2.207	−2.050	−1.691	−1.513	0.303	0.275	0.282	0.293	0.296	0.376	0.655	1.181	2.393	2.986
100...........	−2.343	−2.306	−2.244	−2.042	−1.895	0.155	0.144	0.148	0.154	0.156	0.176	0.376	0.772	2.058	2.988
300...........	−2.345	−2.335	−2.316	−2.231	−2.157	0.088	0.083	0.085	0.088	0.089	0.121	0.224	0.430	1.384	2.200
1000.........	−2.345	−2.344	−2.339	−2.314	−2.289	0.048	0.045	0.046	0.048	0.049	0.151	0.163	0.260	0.766	1.275

and the bias normalized with respect to the uncertainty

$$b = \frac{1}{1000}\sum_{k=1}^{1000}\frac{\gamma_k + 2.35}{\sigma_k}. \tag{10}$$

The value of b is the logical criterion to judge the existence of biases. If $|b| \ll 1$, then the fitting method will be unbiased because it will yield values that will be larger than the real one on ≈50% of the time and smaller on the other ≈50%. If, on the other hand, $|b| \sim 1$ or larger, a significant bias will exist.

Of course, we want to test whether our experiments yield biased or unbiased results in order to decide which binning technique is the optimum one. However, a fitting method can be unbiased but still yield an incorrect uncertainty estimate. A more complete test would be to analyze the distribution of the quantity $(\gamma_k + 2.35)/\sigma_k$, whose mean is given by b and whose standard deviation is given by

$$\beta = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}\left(\frac{\gamma_k + 2.35}{\sigma_k} - b\right)^2}. \tag{11}$$

A binning technique that both is unbiased and yields correct uncertainty estimates should produce a distribution for $(\gamma_k + 2.35)/\sigma_k$ that resembles a Gaussian with $b = 0$ and $\beta = 1$. Under such circumstances, we can predict that the true value of the slope of the IMF will be, e.g., within $\gamma_k - \sigma_k$ and γ_k approximately 34.1% (or, e.g., within γ_k and $\gamma_k + 2\sigma_k$ approximately 47.7%) of the times the experiment is executed. If β is significantly different from 1 but $b \approx 0$, then the technique will yield unbiased results with incorrect uncertainty estimates. If, on the other hand, $\beta \approx 1$ but b is significantly different from zero, the technique will be biased (e.g., will have a systematic error), but its uncertainty estimate (sometimes called the random error) will be correct, as previously mentioned. Of course, the worst-case scenario implies values of b and β significantly different from 0 and 1, respectively, in which case the technique produces both systematic errors and incorrect uncertainty estimates.

Finally, we expect the results from different experiments to depend on the existence of bins with few or no stars. In order to analyze that effect we define $\overline{N}_{i,\min}$ to be the mean N_i in the bin with the lowest number of counts, which for uniform-size bins and a Salpeter power law will be the rightmost one. For our variable-size bin experiments, $\overline{N}_i$ is the same for any bin so there is no need to select a specific one.

3. EXPERIMENT 1: UNIFORM-SIZE BINS

We show in Table 1 the values of $\overline{\gamma}$, $\overline{\sigma}$, and b for the first experiment, in which we use a uniform bin size. In all of the 20 cases, the value of $\overline{\gamma}$ is found out to be larger than −2.350, with the maximum at −1.513. Similarly, b is always positive. In a few cases, the bias is small, but in most of them it is quite large. The expression b is strongly anticorrelated with $\overline{N}_{i,\min}$, as we show in Figure 1. This is an expected behavior, since for bins with a small number of stars the difference between the used weight and the real weight (the one derived from the parent, not the sample, distribution) becomes larger in relative terms. Furthermore, there appears to be a critical value around $\overline{N}_{i,\min} \approx 10$: for lower values biases are quite large, while for higher ones they are small (although not always negligible). These results confirm what we mentioned in § 1: *for a binning scheme in which the number of objects per bin is highly variable, significant biases can be present even if all bins have more than one star.*

The existence of significant biases makes the use of uniform-size bins inadvisable for calculating IMFs. However, if no alternative is available, one should prefer those results based on a small number of bins, since in that case biases are smaller. Increasing the number of bins reduces the uncertainty estimate slightly but at the cost of introducing significantly larger biases (i.e., it produces smaller random uncertainties but larger systematic errors). This is another consequence of the anticorrelation between $\overline{N}_{i,\min}$ and b.

We also performed a check on the values of β and found them to be close to 1 in all cases. Furthermore, a look at the sample case in Figure 2 shows that the distribution of $(\gamma_k + 2.35)/\sigma_k$ is well characterized by a Gaussian with a mean equal to b and a standard deviation of 1.0. As indicated in § 3, this means that

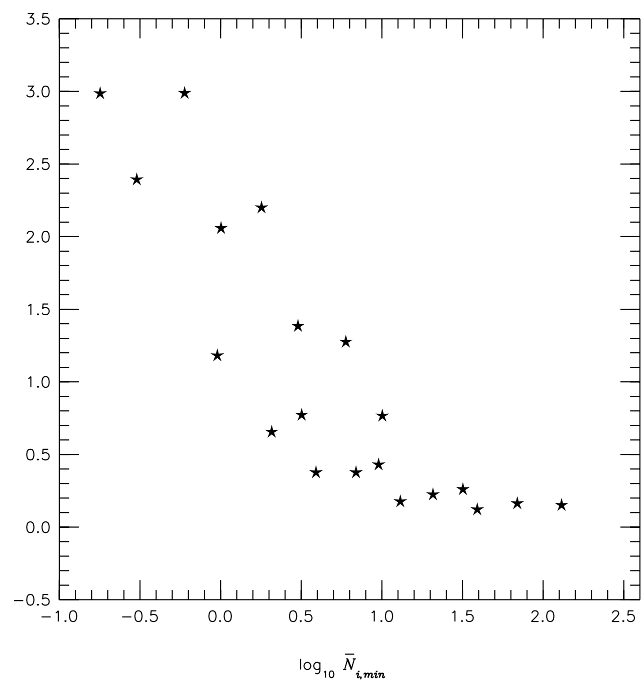

Fig. 1.—Bias as a function of $\overline{N}_{i,\min}$ for the first experiment. Note that $\overline{N}_{i,\min}$ can be smaller than 1, because it is a property derived from the parent distribution.

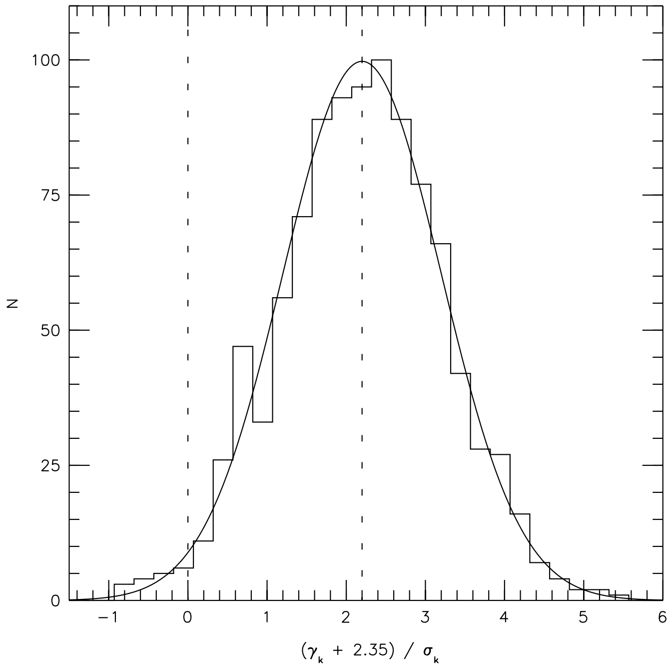

FIG. 2.—Histogram with the distribution of $(\gamma_k + 2.35)/\sigma_k$ for the 1000 realizations of the first experiment with 300 stars and 50 bins. A Gaussian distribution with mean $b = 2.200$ and dispersion of 1.0 is also plotted for comparison. The vertical lines mark the position of 0 and b.

bins with uniform size yield correct uncertainty estimates (i.e., the problem lies in the systematic, not in the random, errors).

We should point out that the existence of a bias toward a flattening of the IMF for small samples when using uniform-size bins was previously detected by Kroupa (2001), who called it a sampling bias. As we see in §§ 4 and 5, it is possible to get rid of such a bias almost completely.

4. EXPERIMENT 2: VARIABLE-SIZE BINS WITH EVENLY DIVIDED NUMBER OF STARS PER BIN

For our second experiment we adopt the first recommendation of D'Agostino & Stephens (1986) and use a variable size for our bins designed in such a way as to have a similar number of stars per bin. We do so for each of our random realizations in the following way (we use the case with 100 stars and 5 bins as an example): (1) fixing $m_{\rm down}$ and $m_{\rm up}$ to be $10^{0.8}$ and $10^{2.2}\,M_\odot$, respectively; (2) sorting the data so that $m_1 < m_2 < \ldots < m_{100}$; and (3) fixing the limits between bins i and $i+1$ to be $0.5(m_{20i} + m_{20i+1})$.

We show in Table 2 the values of $\overline{\gamma}$, $\overline{\sigma}$, and b for the second experiment. Only 19 cases were used because in one circumstance there were more bins than stars. The value of $\overline{\gamma}$ is very close to -2.35 in all cases, with a minimum of -2.403 and a maximum of -2.348. In all 19 cases $|b| \ll 1$ (maximum value of 0.062), with b being positive in some cases and negative in others. Note that the signs of $\overline{\gamma} + 2.35$ and of b can be different due to the possible existence of correlations between the values of γ_k and σ_k (see also Fig. 5).

These results indicate that using a variable bin size to include a similar number of stars in each bin is a good way of minimizing binning biases. A comparison with the previous experiment shows that this is done at no significant cost of increasing $\overline{\sigma}$. Regarding the second recommendation of D'Agostino & Stephens (1986), we only find a weak dependence of b in the number of bins. The robustness of the method is emphasized by the fact that even when 30 stars are divided into 30 bins (i.e., 1 star per bin) no significant biases are detected. This result corroborates that the existence of large biases originates in the assignment of incorrect weights to each bin and not so much by the fact that one specific bin has a low number of stars (Wheaton et al. 1995).

5. EXPERIMENT 3: SETTING THE LOWER AND UPPER MASS LIMITS FROM THE DATA

The previous two experiments have an artificial component in them: we use for $m_{\rm down}$ and $m_{\rm up}$ the input values, i.e., the values that are attained only in a sample with an infinite number of stars. When using real data, however, those values have to be determined. An observer typically fixes the first one using incompleteness criteria (since lower mass stars usually exist but are harder to detect), and the second one is usually unknown (and, likely, a quantity one is interested in measuring, see, e.g., Oey & Clarke 2005). Therefore, we can simulate more realistic conditions by modifying the extremes of our second experiment by setting $m_{\rm down} = m_1 - 0.5(m_2 - m_1)$, $m_{\rm up} = m_N + 0.5(m_N - m_{N-1})$, which can be determined directly from the data.

We show in Table 3 the values of $\overline{\gamma}$, $\overline{\sigma}$, and b for the third experiment. Results are similar to the ones for the second experiment, with values of $\overline{\gamma}$ between -2.365 and -2.298. This is especially so for the cases with 300 and 1000 stars, as expected. Comparing results one by one, we find that $\overline{\gamma}$ is always larger here than for the previous experiment. The expression b is now always positive and the values of $|b|$ are somewhat higher than in the previous experiment. Still, all cases with more than three stars per bin have $|b| \leq 0.134$, and even the worst case has only $|b| = 0.264$. In Figure 3 we see that b is still anticorrelated with $\overline{N}_i$, the average number of stars per bin.

These results indicate that fixing the lower and upper mass limits from the data and using a variable bin size to include a similar number of stars in each bin is a practical way of minimizing binning biases. There is a price in the former of larger biases that has to be paid for the lack of knowledge of the ends of the distribution, but that price is small, as one can observe by comparing Figures 1 and 3. We should point out that the second recommendation of D'Agostino & Stephens (1986) does

TABLE 2

RESULTS FOR THE VARIABLE BIN SIZE CASE FOR 3, 5, 10, 30, AND 50 BINS

	$\overline{\gamma}$					$\overline{\sigma}$					b				
Stars	3	5	10	30	50	3	5	10	30	50	3	5	10	30	50
30............	−2.402	−2.402	−2.403	−2.394	. . .	0.285	0.285	0.288	0.292	. . .	0.012	−0.024	−0.046	−0.029	. . .
100...........	−2.365	−2.367	−2.366	−2.364	−2.366	0.152	0.152	0.154	0.157	0.156	0.015	−0.018	−0.024	−0.029	−0.030
300...........	−2.353	−2.352	−2.353	−2.355	−2.354	0.086	0.087	0.088	0.089	0.090	0.034	0.031	0.012	−0.012	−0.009
1000.........	−2.350	−2.348	−2.349	−2.349	−2.349	0.047	0.047	0.048	0.049	0.049	0.043	0.062	0.045	0.037	0.043

TABLE 3

Results for the Variable Bin Size Case with Data-determined m_{down} and m_{up} for 3, 5, 10, 30, and 50 Bins

	$\overline{\gamma}$					$\overline{\sigma}$					b				
Stars	3	5	10	30	50	3	5	10	30	50	3	5	10	30	50
30............	−2.365	−2.349	−2.329	−2.298	...	0.308	0.306	0.305	0.302	...	0.110	0.134	0.180	0.264	...
100...........	−2.359	−2.356	−2.351	−2.338	−2.335	0.155	0.156	0.157	0.157	0.158	0.047	0.046	0.071	0.143	0.161
300...........	−2.351	−2.349	−2.348	−2.348	−2.346	0.087	0.088	0.089	0.090	0.090	0.053	0.066	0.065	0.065	0.080
1000.........	−2.349	−2.348	−2.348	−2.348	−2.347	0.047	0.048	0.048	0.049	0.049	0.044	0.079	0.079	0.073	0.086

not work for this experiment: if one is interested in minimizing biases, then it is preferable to go with a low number of bins (but, as previously indicated, biases are never large anyway).

The values of β are found to be close to 1 in all cases, and the histogram for the sample case in Figure 4 shows the same behavior as that in Figure 2: the distribution of $(\gamma_k + 2.35)/\sigma_k$ is well characterized by a Gaussian with a mean equal to b (and, in this case, close to zero) and a dispersion of 1.0. Therefore, *our technique yields not only a nearly bias-free value for the slope of the IMF but also a correct estimate of its uncertainty.*

Another advantage of the method proposed here can be extracted from Table 3: a nearly bias-free measurement of the power-law exponent with an uncertainty of less than 0.2 can be obtained with only 100 stars. If the number is lowered to 30 stars, then the uncertainty in the power-law exponent is close to 0.3. Detailed results for the 30 stars + five bins case are shown in Figure 5. Note how the first histogram shows a symmetric distribution, while the second one is distinctly asymmetric. The difference is explained by the correlation ($r = -0.58667$) between γ_k and σ_k shown in the bottom plot: χ^2 fitting yields larger values of the uncertainty in the slope for lower values of the slope itself. Note that this correlation is not a problem in itself because the histogram that determines through its mean and dispersion whether the technique yields a correct estimate of the IMF slope is the first one, not the second. Therefore, we conclude that *it is possible to conduct precise studies of the mass segregation within a large cluster and to measure the IMF in a small one.*

Different binning schemes can easily yield different values for the same sample. This can be seen in Figure 6, where we show a comparison between the three experiments for a single realization. With that in mind, one wonders whether part of the variations in the IMF detected by a number of authors (see Elmegreen 2004 for a recent review) are not real but simply numerical effects introduced by the different schemes used. In order to test that, one would have to reanalyze the data in a uniform manner using an unbiased scheme, such as that presented in this article.

6. SUMMARY AND FUTURE WORK

We conclude that the binning mechanism proposed in this paper for the fitting of power laws with Salpeter slopes yields results that (1) are nearly bias-free and (2) produce correct uncertainty estimates, as tested by our numerical simulations. On the other hand, the standard uniform-size binning introduces biases that are dependent on the number of stars per bin. The power of the technique described here extends to small samples, since we have shown that it is possible to obtain accurate values with reasonable precisions for the IMF slope even when as few as 30 stars are available for analysis.

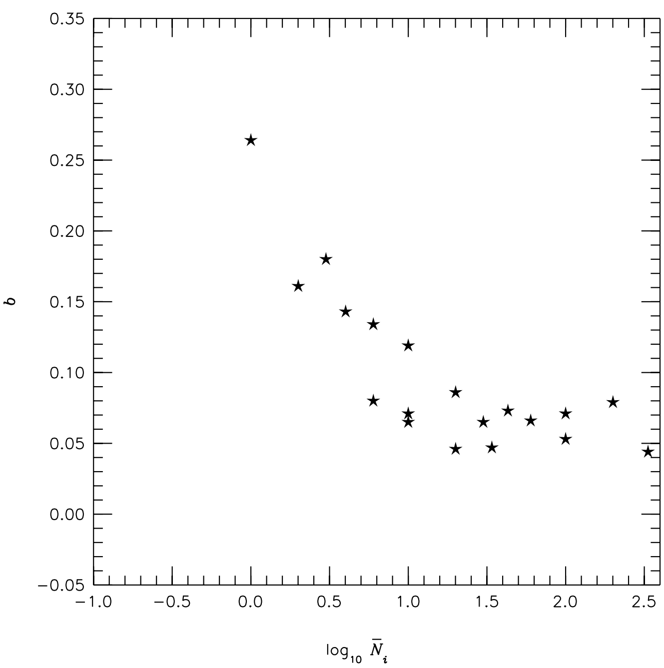

Fig. 3.—Bias as a function of $\overline{N}_i$ for the third experiment. Note that the vertical scale for the plot is $\frac{1}{10}$ that of Fig. 1.

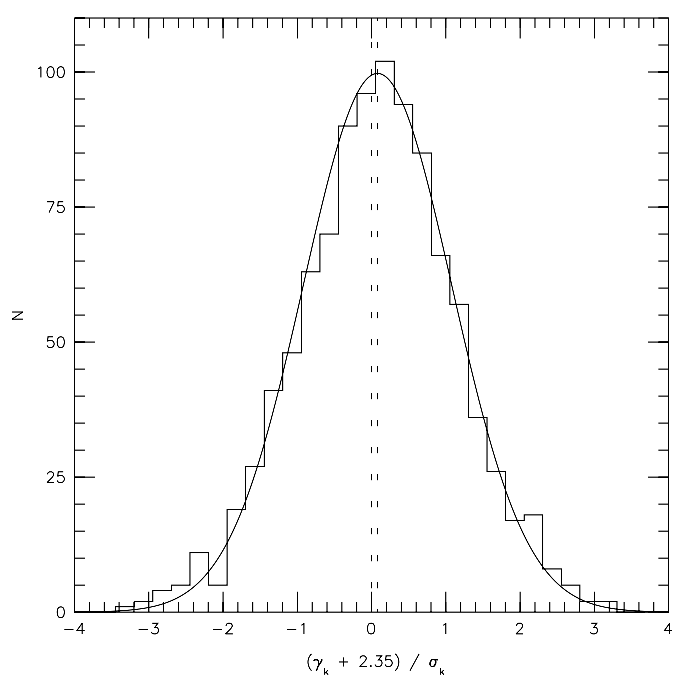

Fig. 4.—Histogram with the distribution of $(\gamma_k + 2.35)/\sigma_k$ for the 1000 realizations of the third experiment with 300 stars and 50 bins. A Gaussian distribution with mean $b = 0.090$ and dispersion of 1.0 is also plotted for comparison. The vertical lines mark the position of 0 and b.

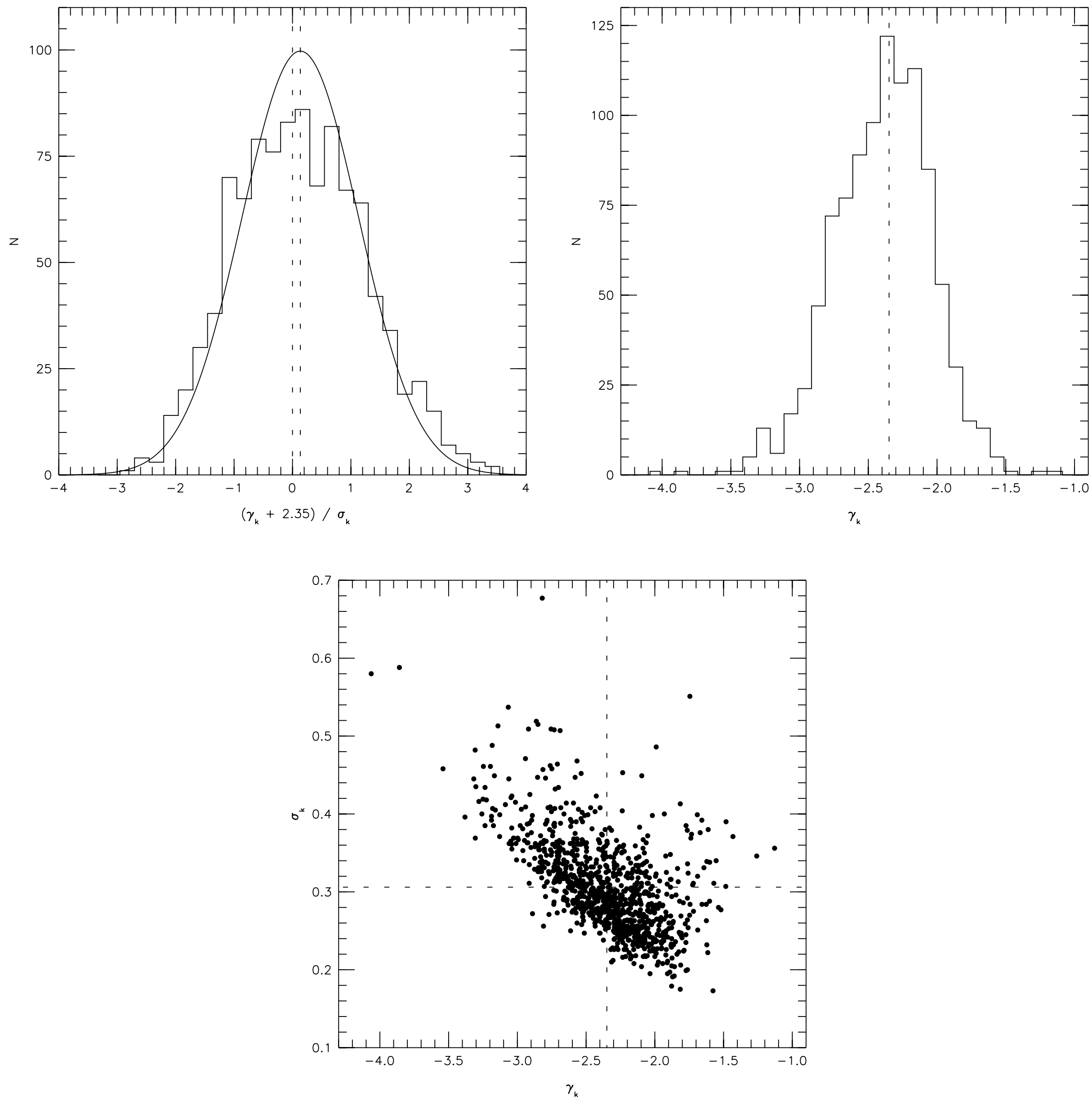

FIG. 5.—Detailed results for the 30 stars + five bins case for the third experiment. *Top left*: Histogram with the distribution of $(\gamma_k + 2.35)/\sigma_k$. A Gaussian distribution with mean $b = 0.134$ and dispersion of 1.0 is also plotted for comparison. Vertical lines mark the position of 0 and b. *Top right*: Histogram with the distribution of γ_k. A vertical line marks the position of $\overline{\gamma}$. *Bottom*: Results for γ_k and σ_k for the 1000 realizations. Lines mark the values of $\overline{\gamma}$ and $\overline{\sigma}$.

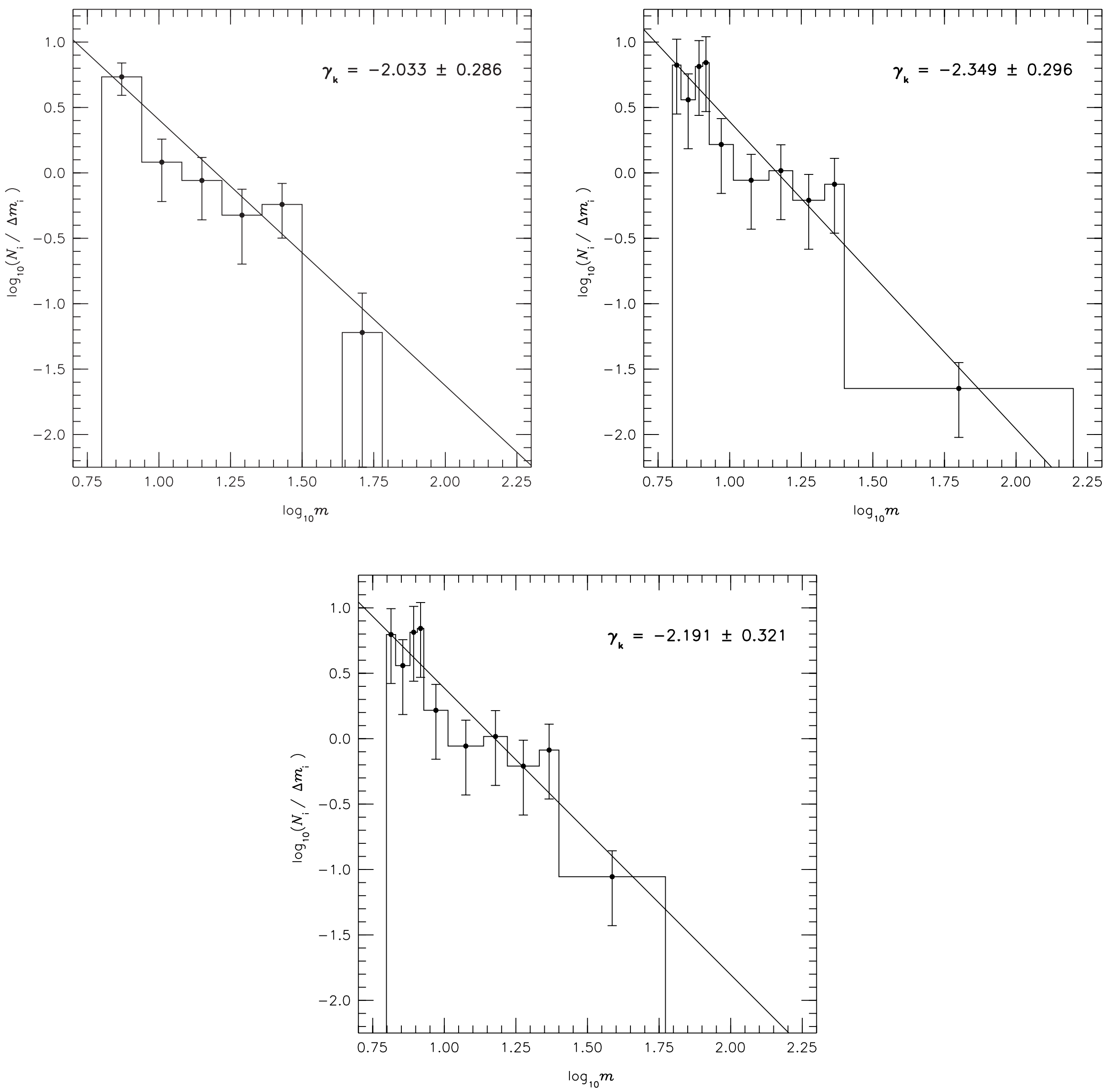

FIG. 6.—Comparison between the data and the fitted functions for one of the realizations with 30 stars and 10 bins for the three experiments in this article. The top left, top right, and bottom panels show the first, second, and third experiments, respectively. Note the differences in the size of the error bars for the histograms between the first and the last two experiments.

We are finishing an analysis of *Hubble Space Telescope* (*HST*) Wide Field Planetary Camera 2 (WFPC2) stellar photometry of the nearby dwarf starburst galaxy NGC 4214. In that article we will apply the technique described here in order to study the IMF for the massive stars in that galaxy. We will also investigate other possible sources of biases in the calculation of the IMF.

We would also like to point out that, given the purely numerical nature of our analysis, our results could be extended to other similar problems. For example, the mass function for young stellar clusters can be rather well approximated by a power law with a slope of -2.0 (see, e.g., Fall & Zhang 2001), which is quite close to -2.35, so the same type of biases should be present there as well. Note, however, that for a problem in which we expect the power law to have a radically different exponent (e.g., -10.0), the results in this paper may not apply because of the larger disparity between the values of the function at the two extremes. In general, we recommend that biases be evaluated for any function fitted to binned data through χ^2 minimization by means of specific numerical experiments similar to those in this article.

Note added in manuscript.—Throughout this paper we have assumed that, as mentioned in § 2, the distribution of stars in bin i, N_i, follows a binomial distribution. It should be noted, however, that the distribution of stars in a different bin j, N_j, is not independent of N_i. A covariance term exists because the joint distribution of $N_1, N_2, \ldots, N_{\rm nbins}$ is a multinomial (see, e.g., Lucy 2000). In our analysis we have found that the inclusion of the covariance terms in the calculation of χ^2 does not change the results significantly for our experiments. Furthermore, we would like to point out once more that the fact that the distribution of $(\gamma_k + 2.35)/\sigma_k$ is closely approximated by a Gaussian with mean 0 and standard deviation of 1 for the third experiment is a sucient proof of the validity of our method. In other words, even though in a case in which we are fitting an arbitrary function with an arbitrary binning it may be necessary to substitute χ^2 by an alternative statistic (Lucy 2000), for the special case of a power law with near-Salpeter index, $N \gtrsim 30$, and objects evenly divided among bins, such a substitution is not necessary in order to obtain a precise and unbiased estimation of γ.

We would like to thank the referee, Miguel Cerviño, for very useful suggestions and for comments that helped improved this article considerably, especially concerning the issue of weights. Support for this work was provided by NASA through grants GO-09419.01-A and AR-09553.02-A from the Space Telescope Science Institute, Inc., under NASA contract NAS5-26555.

REFERENCES

Béjar, V. J. S., et al. 2001, ApJ, 556, 830
Bevington, P. R., & Robinson, D. K. 1992, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw-Hill)
D'Agostino, R. B., & Stephens, M. A. 1986, Goodness-of-Fit Techniques (New York: Marcel Dekker)
Elmegreen, B. G. 2004, MNRAS, 354, 367
Fall, S. M., & Zhang, Q. 2001, ApJ, 561, 751
Humphreys, R. M., & McElroy, D. B. 1984, ApJ, 284, 565
Kearns, K., Primini, F., & Alexander, D. 1995, in ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 331
Kroupa, P. 2001, MNRAS, 322, 231
Lucy, L. B. 2000, MNRAS, 318, 92
Massey, P., Lang, C. C., Degioia-Eastwood, K., & Garmany, C. D. 1995, ApJ, 438, 188
Massey, P., Parker, J. W., & Garmany, C. D. 1989, AJ, 98, 1305
Nousek, J. A., & Shue, D. R. 1989, ApJ, 342, 1207
Oey, M. S., & Clarke, C. J. 2005, ApJ, 620, L43
Press, W. H., Flannery, B. P., & Teukolsky, S. A. 1986, Numerical Recipes: The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)
Salpeter, E. E. 1955, ApJ, 121, 161
Scalo, J. M. 1986, Fundam. Cosmic Phys., 11, 1
Wheaton, W. A., Dunklee, A. L., Jacobsen, A. S., Ling, J. C., Mahoney, W. A., & Radocinski, R. G. 1995, ApJ, 438, 322