Likelihood function for Milky Way satellite mass/velocity/luminosity function

Andrew Benson

November 17, 2014

We assume that we have a cumulative velocity function for Milky Way satellite galaxies. (The same analysis applies for luminosity and mass functions.) This can be described as a set of tuples $\theta_i = (N_i, V_i)$, where V_i is a set of ordered velocities of the satellites ($i = 1 \ldots n$) and N_i is the cumulative number of satellites with velocities equal to or greater than V_i. Importantly, we assume no errors on the measured N_i—galaxies are either there or they are not there[1]. We want to evaluate the likelihood of this data given a model, $P(\theta_i|\mathcal{M})$, which we can then use in Bayes' law to find the likelihood of the model given the data.

According to Yu's work (check this holds up in Galacticus also), the model $N_i^\prime$ are consistent with being drawn from Poisson distributions. But, of course they are not independent since we're looking at a cumulative function. Instead, we assume that the differences between pairs of $N_i^\prime$'s are drawn from independent Poisson distributions. Given the additive nature of Poisson distributions[2] we can figure out the covariance betwen points in the model mass function. Specifically, consider the covariance matrix for two points in the mass function, i and j. Writing $N_j^\prime = N_i^\prime + \Delta N_{ij}^\prime$, and taking the model expectation for the means to be $\bar{N}_i^\prime$ and $\bar{N}_j^\prime = \bar{N}_i^\prime + \Delta\bar{N}_{ij}^\prime$, we can directly write the covariance for $N_i^\prime$ and $\Delta\bar{N}_{ij}^\prime$:

$$\mathrm{Cov}(N_i^\prime, \Delta N_{ij}^\prime) = \left(\begin{array}{cc} \bar{N}_i^\prime & 0 \\ 0 & \Delta\bar{N}_{ij}^\prime \end{array} \right), \tag{1}$$

since these two variables are independent. To find the covariance for $N_i^\prime$ and $N_j^\prime$ we construct the Jacobian to transform from ($N_i^\prime$,$\Delta N_{ij}^\prime$) to ($N_i^\prime$,$N_j^\prime$), since $(N_i^\prime, N_j^\prime) = J(N_i^\prime, \Delta N_{ij}^\prime)(N_i^\prime, \Delta N_{ij}^\prime)$:

$$J(N_i^\prime, \Delta N_{ij}^\prime) = \left(\begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array} \right). \tag{2}$$

[1] In principle for the faintest galaxies we could allow for some detection probability and/or false positive probability.

[2] That is, if x and y are independent random variables drawn from Poisson distributions with means μ_x and μ_y respectively, then $x+y$ is drawn from a Poisson distribution with mean $\mu_x + \mu_y$.

The required covariance is then:

$$\mathrm{Cov}(N_i', N_j') = J \cdot \mathrm{Cov}(N_i', \Delta N_{ij}') \cdot J^{\mathrm{T}} = \begin{pmatrix} \bar{N}_i' & \bar{N}_i' \\ \bar{N}_i' & \bar{N}_i' + \Delta\bar{N}_{ij}' \end{pmatrix} = \begin{pmatrix} \bar{N}_i' & \bar{N}_i' \\ \bar{N}_i' & \bar{N}_j' \end{pmatrix}. \tag{3}$$

By extension, the full covariance matrix can be written as:

$$\mathrm{Cov}(N_i', N_j') = \begin{pmatrix} \bar{N}_1' & \bar{N}_1' & \bar{N}_1' & \dots \\ \bar{N}_1' & \bar{N}_2' & \bar{N}_2' & \dots \\ \bar{N}_1' & \bar{N}_2' & \bar{N}_3' & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{4}$$

or, more simply,

$$C_{ij} = \bar{N}'_{\min(i,j)}. \tag{5}$$

In fact, we can make use of the full Poisson probability distribution in computing the probability of the data, if we work in terms of the difference in the mass function at each point. Noting that, by definition, $\Delta N_{i+1,i} \equiv \Delta N_{i+1} = N_{i+1} - N_i = 1$, and with the convention that $N_0 = 0$ we can write the probability of each point, independent of the others as

$$P_i = \Delta\bar{N}_i' \exp(-\Delta\bar{N}_i'). \tag{6}$$

Therefore, the log-likelihood for the full dataset is simply:

$$\log\mathcal{L} = \sum_{i=1}^{n} \log(\Delta\bar{N}_i') - \Delta\bar{N}_i' \tag{7}$$

1 Errors in Velocity

Since each velocity measurement is subject to error, we should integrate over the velocity error distributions when evaluating the likelihood. Since each velocity now becomes a random variable, this could lead to changes in the rank ordering of the galaxies, which will affect how we evaluate the likelihood. In particular, we need to know the joint probability distribution of the order statistics of the velocities. In principle, this can be found using the Bapat-Beg theorem. In practice, this is computationally intractable for the number of galaxies involved. Instead, we can average over the joint probability distribution of the order statistics using a Monte Carlo approach.