Brief notes on fitting micrometeoroid population models from sparse detections

John G. Baker[1,2]

[1] *Gravitational Astrophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA and*
[4] *Joint Space-Science Institute, University of Maryland, College Park, MD 20742, USA*

(Dated: July 26, 2018)

Notes on the Bayesian population inference approach.

I. A BAYESIAN APPROACH TO POPULATION MODELS

Observations of micrometeoroid impacts provide some information about the population of micrometeoroids overall. To make consistent inferences about a population from observational data including a moderate number of events, it is appropriate to take a Bayesian approach. It is useful begin by laying down the general problem of how population models can be constrained by event observations. What is an appropriate likelihood function for the analysis, and more broadly, what are the relevant features in appropriate expressions of Bayes theorem.

We express Bayes' theorem as

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \quad (1)$$

where we have used D to represent the overall set of observed data. We can go to the next level by splitting $\theta = \{\theta_P, \theta_O\}$ into "population" parameters and "observation" parameters, respectively corresponding to the parameters characterizing the micrometeoroid population, and those which we make inferences about in our observational impact analysis,

$$
\begin{aligned}
p(\theta_O, \theta_P|D) &= \frac{p(D|\theta_O, \theta_P)p(\theta_O, \theta_P)}{p(D)} \\
&= \frac{p(D|\theta_O, \theta_P)p(\theta_O|\theta_P)p(\theta_P)}{p(D)}.
\end{aligned}
$$

If we are only interested in the metaparameters θ_P, then we can marginalize over θ_O,

$$
\begin{aligned}
p(\theta_P|D) &= \int p(\theta_O, \theta_P|D)d\theta_O \\
&= \frac{p(\theta_P)}{p(D)} \int p(D|\theta_O, \theta_P)p(\theta_O|\theta_P)d\theta_O. \quad (2)
\end{aligned}
$$

The integral on the right hand side is the transdimensional integral over the full observation model space, including how many and what kinds of events were inferred to have occurred. The integral factor in (2) can be recognized as the effective likelihood function that we need for the population model inference analysis,

$$
\begin{aligned}
p(D|\theta_P) &= \int p(D|\theta_O, \theta_P)p(\theta_O|\theta_P)d\theta_O \\
&= \int p(D|\theta_O)p(\theta_O|\theta_P)d\theta_O \quad (3)
\end{aligned}
$$

The last line also encodes that the micrometeoriod impact observation likelihood depends only on the nature of the impacts θ_O and not also explicitly on other features of the population θ_P.

Consistent with our approach to micrometeoroids, we suppose that the data is broken into many small data subsets $D = \{D_\alpha\}$ which are small enough so that multiple impacts are unlikely in each bin, but large enough that the individual responses are unlikely to straddle the bin boundaries. In each data subspace, we make inferences on transdimensional observation parameters $\theta_{O,\alpha} = (n, \{\psi^{\{n\}}\} = (n, \{\psi_1, ... \psi_n\})$ where n is the number of impacts and ψ_n represents all the parameters for the nth impact. We thus perform an observational inference analysis, conducted with some concrete yet suitably generic reference assumptions about the population which we can call θ_{P0}. In practice, for the the impact analysis, we assume a reference distribution of uniform sky impact directions and log-uniform meteoriod momenta.

Bayes theorem for this inference problem is written as

$$p(n, \psi^{\{n\}}|D_\alpha, \theta_{P0}) = \frac{p(D_\alpha|n, \psi^{\{n\}})p(n, \psi^{\{n\}}|\theta_{P0})}{p(D_\alpha|\theta_{P0})}. \quad (4)$$

In the usual mechanics of Bayesian analysis then, a key result is a set S_α containing some number N_α of posterior-distributed samples. For this transdimensional problem it is useful to split these into subsets like $S_{\alpha,k}$ which includes the $N_{\alpha,k}$ samples including k events each. This set can then be applied to marginalize any function of the observational parameters over the posterior,

$$E_{D_\alpha, P0}[f] = \int d\{n, \psi^{\{n\}}\}f(n, \psi^{\{n\}})p(n, \psi^{\{n\}}|D_\alpha, \theta_{P0}) \quad (5)$$

$$= \frac{1}{N_\alpha} \sum_{s \in S_\alpha} f(n_s, \psi_s^{\{n_s\}})$$

$$= \frac{1}{N_\alpha} \sum_k \sum_{s \in S_{\alpha,k}} f(k, \psi_s^{\{k\}}). \quad (6)$$

Returning to the inference problem for the population model parameters θ_P, we can write its likelihood function

(3) in terms of contributions from each data segment

$$p(D_\alpha|\theta_P) = \prod \{\alpha\} p(D_\alpha|\theta_P)$$

$$p(D_\alpha|\theta_P) = \int d\{n, \psi^{\{n\}}\} p(D_\alpha|n, \psi^{\{n\}}) p(n, \psi^{\{n\}}|\theta_P).$$

$$= p(D_\alpha|\theta_{P0}) \int d\{n, \psi^{\{n\}}\} p(n, \psi^{\{n\}}|D_\alpha, \theta_{P0})$$

$$\times \frac{p(n, \psi^{\{n\}}|\theta_P)}{p(n, \psi^{\{n\}}|\theta_{P0})}.$$

This integral is of the same for as (5) with $f(n, \psi^{\{n\}}) = p(n, \psi^{\{n\}}|\theta_P)/p(n, \psi^{\{n\}}|\theta_{P0})$. Thus, using (6), we can write

$$p(D|\theta_P) = \frac{p(D_\alpha|\theta_{P0})}{N_\alpha} \sum_k \sum_{s \in S_{\alpha,k}} \frac{p(k, \psi_s^{\{k\}}|\theta_P)}{p(k, \psi_s^{\{k\}}|\theta_{P0})}$$

$$= \frac{p(D_\alpha|\theta_{P0})}{N_\alpha} \sum_k \frac{p(k|\theta_P)}{p(k|\theta_{P0})} \sum_{s \in S_{\alpha,k}} \frac{p(\psi_s^{\{k\}}|k, \theta_P)}{p(\psi_s^{\{k\}}|k, \theta_{P0})}$$

$$(7)$$

Plugging everything back into (2), and taking the log we get

$$\ln p(\theta_P|D) = \sum_\alpha \ln \left[\frac{1}{N_\alpha} \sum_k \frac{p(k|\theta_P)}{p(k|\theta_{P0})} \sum_{s \in S_{\alpha,k}} \frac{p(\psi_s^{\{k\}}|k, \theta_P)}{p(\psi_s^{\{k\}}|k, \theta_{P0})} \right]$$

$$+ \ln p(\theta_P) + \ln \frac{p(D|\theta_{P0})}{p(D|\theta_P)} \quad (8)$$

If rates are sufficiently small that we can expect no more than one impact per bin, then this becomes with some simplification of the notation,

$$\ln p(\theta_P|D) = \sum_\alpha \ln \left[\frac{N_{\alpha,0}}{N_\alpha} \frac{p(0|\theta_P)}{p(0|\theta_{P0})} \right.$$

$$+ \frac{p(1|\theta_P)}{p(1|\theta_{P0})} \frac{1}{N_\alpha} \sum_{s \in S_{\alpha,k}} \frac{p(\psi_s|1, \theta_P)}{p(\psi_s|1, \theta_{P0})} \right]$$

$$+ \ln p(\theta_P) + \text{const} \quad (9)$$

The last piece needed is an explicit form for the expected impact probabilites. For these we just assume Poisson probabilty distribution (truncated at 1 event) based on the rates we have from the micrometeoroid mod-

els:

$$p(1, \psi|\theta_P) = r_\alpha(\psi, \theta_P)$$

$$p(1|\theta_P) = \int r_\alpha(\psi, \theta_P) d\psi = \bar{r}_\alpha(\theta_P)$$

$$p(\psi|1, \theta_P) = \hat{r}_\alpha(\psi, \theta_P) = \frac{r_\alpha(\psi, \theta_P)}{r_\alpha(\theta_P)}$$

$$p(0|\theta_P) = 1 - p(1|\theta_P)$$

where $r_\alpha(\theta_P)$ is the expected rate events occuring (in LPF frame observation parameters) in segment α given the metaparameters θ_P. We can write the time-bin LPF-frame rate r_α in terms of the physical micrometeoriod fluxes $F(\bar{\theta}, \bar{\phi}, \bar{P}, \theta_P)$ by

$$r_\alpha(\psi, \theta_P) = T_\alpha A_{LPF} \frac{\partial(\bar{\theta}, \bar{\phi}, \bar{P})}{\partial \psi} F(\bar{\theta}(\psi, t_\alpha), \bar{\phi}(\psi, t_\alpha), \bar{P}(\psi), \theta_P)$$

$$(10)$$

where T_α is the duration of the observation segment in time, A_{LPF} is the spacecraft area, and the derivative factor is the Jacobian of the transformation from LPF parameters to the population model dimensions $\{\bar{\theta}, \bar{\phi}, \bar{P}\}$ at observation time t_α.

We can rewrite the log-posterior (9) as

$$\ln p(\theta_P|D) = \sum_\alpha \ln \left[(1 - \epsilon_\alpha)(1 - \bar{r}_\alpha(\theta_P)) \right.$$

$$\left. + \bar{r}_\alpha(\theta_P)\epsilon_\alpha \frac{1}{N_{\alpha,1}} \sum_{s \in S_{\alpha,k}} \frac{\hat{r}_\alpha(\psi_s|1, \theta_P)}{p(\psi_s|1, \theta_{P0})} \right]$$

$$+ \ln p(\theta_P) + \text{const} \quad (11)$$

where we have defined $\epsilon_\alpha = N_{\alpha,1}/N_\alpha$ as the impact fraction for the data segment.

We note that the population is naturally described in terms of several constituent subpopulations, each with their own fluxes. We can combine these in a general way by writing

$$F(\bar{\theta}, \bar{\phi}, \bar{P}, \theta_P) = \sum_\lambda c_\lambda F_\lambda(\bar{\theta}, \bar{\phi}, \bar{P}).$$

in this case $\theta_P \equiv c_\lambda$. If the component fluxes are already consistently weighted with the reight rates, the we would expect to find $c_\lambda \approx 1$ for each λ.

If we are only interested in the relative rates, then it is natural to work with normalized fluxes:

$$\hat{F}_\lambda(\bar{\theta}, \bar{\phi}, \bar{P}) = \frac{F_\lambda(\bar{\theta}, \bar{\phi}, \bar{P})}{\bar{F}_\lambda}$$

$$\bar{F}_\lambda(\bar{\theta}, \bar{\phi}, \bar{P}) = \int F_\lambda(\bar{\theta}, \bar{\phi}, \bar{P}) d\bar{\theta} d\bar{\phi} d\bar{P}$$

$$\hat{F}(\bar{\theta}, \bar{\phi}, \bar{P}, \theta_P) = \sum_\lambda \hat{c}_\lambda \hat{F}_\lambda(\bar{\theta}, \bar{\phi}, \bar{P}) \quad (12)$$

with $\sum \hat{c}_\lambda = 1$. In this case, we can substitute our best estimate for the detection rate in for the model event rate $\bar{r}_\alpha(\theta_P)$ in (11).