Draft version July 6, 2021
Preprint typeset using LaTeX style emulateapj v. 12/16/11

ON REWEIGHING SINGLE-EVENT POSTERIORS WITH POPULATION PRIORS

R. Essick
Kavli Institute for Cosmological Physics, University of Chicago, 5640 S Ellis Ave, Chicago, IL 60637

M. Fishbach
Dept. of Astronomy and Astrophysics, University of Chicago, 5640 S Ellis Ave, Chicago, IL 60637

LIGO-T1900895-v2

Let us consider the posterior for individual event parameters (θ_i for $i = 1, 2, \dots, N$) and population parameters (Λ) so that

$$\theta_i \sim p(\theta|\Lambda)$$

We begin with the inhomogeneous Poisson likelihood and priors for Λ and R to obtain

$$p(\theta_1, \theta_2, \dots, \theta_N, \Lambda, R|d_1, d_2, \dots, d_N) = \frac{p(R)p(\Lambda)}{p(d_1, d_2, \dots, d_N)} e^{-R\beta(\Lambda)} \prod_{i=1}^{N} \left[p(d_i|\theta_i) R p(\theta_i|\Lambda)\right]$$

where

$$\beta(\Lambda) = \int d\vartheta \, P(\det|\vartheta) p(\vartheta|\Lambda)$$

with $P(\det|\theta)$ the probability of detecting a signal with individual event parameters θ. If we assume $p(R) \sim 1/R$, we can marginalize to obtain

$$p(\theta_1, \theta_2, \dots, \theta_N, \Lambda|d_1, d_2, \dots, d_N) = \frac{p(\Lambda)}{p(d_1, d_2, \dots, d_N)} \prod_{i=1}^{N} \left[\frac{p(d_i|\theta_i)p(\theta_i|\Lambda)}{\beta(\Lambda)}\right]$$

Now, let us continue to marginalize over all single-event parameters except for one event

$$\begin{aligned}
p(\theta_1|d_1, d_2, \dots, d_N) &= \int d\Lambda d\theta_2 \cdots d\theta_N \frac{p(\Lambda)}{p(d_1, d_2, \dots, d_N)} \prod_{i}^{N} \left[\frac{p(d_i|\theta_i)p(\theta_i|\Lambda)}{\beta(\Lambda)}\right] \\
&= \int d\Lambda \frac{p(\Lambda)}{p(d_1, d_2, \dots, d_N)} \frac{p(d_1|\theta_1)p(\theta_1|\Lambda)}{\beta(\Lambda)} \prod_{i=2}^{N} \left[\int d\vartheta \frac{p(d_i|\vartheta)p(\vartheta|\Lambda)}{\beta(\Lambda)}\right] \\
&= \int d\Lambda p(d_1|\theta_1)p(\theta_1|\Lambda) \left(\frac{p(\Lambda)}{\beta(\Lambda)p(d_1, d_2, \dots, d_N)} \prod_{i=2}^{N} \left[\int d\vartheta \frac{p(d_i|\vartheta)p(\vartheta|\Lambda)}{\beta(\Lambda)}\right]\right) \\
&= p(d_1|\theta_1) \int d\Lambda p(\theta_1|\Lambda) \left(\frac{p(\Lambda|d_2, \dots, d_N)}{\beta(\Lambda)} \frac{p(d_2, \dots, d_N)}{p(d_1, d_2, \dots, d_N)}\right)
\end{aligned}$$

We note that we do not simply compute the leave-one-out posterior predictive distribution (PPD) as the leave-one-out posterior distribution on Λ comes along with another factor the selection function $\beta(\Lambda)$. This is because we condition on having detected N events, while the leave-one-out PPD only conditions on having detected $N-1$ events.

We also note that, to avoid having to generate separate leave-one-out posteriors for Λ for each event we wish to reweigh, we can instead use

$$\frac{p(d_2, \dots, d_N)}{p(d_1, d_2, \dots, d_N)} \frac{p(\Lambda|d_2, \dots, d_N)}{\beta(\Lambda)} = \frac{p(\Lambda|d_1, d_2, \dots, d_N)}{\int d\vartheta p(d_1|\vartheta)p(\vartheta|\Lambda)}$$

to obtain

$$p(\theta_1|d_1, d_2, \dots, d_N) = p(d_1|\theta_1) \int d\Lambda \, p(\theta_1|\Lambda) \frac{p(\Lambda|d_1, d_2, \dots, d_N)}{\int d\vartheta p(d_1|\vartheta)p(\vartheta|\Lambda)}$$

Within this approach, we compute an "effective prior" and the reweigh single-event posteriors as needed. This means that we can generate a single posterior for Λ using all events and then weigh catalog samples by that posterior divided

reed.essick@gmail.com
maya.fishbach@gmail.com

by the evidence for a single event given $p(\theta|\Lambda)$. In this way, we account for the fact that we detected all N events when reweighing each individual event within the catalog.

Furthermore, we also note that one can efficiently sample from the joint posterior for Λ and any number of single event parameters for one or more events without explicitly constructing an effective prior. Specifically, if we iteratively draw Λ samples from the hyperposterior conditioned on all N events and draw a θ_i sample from the corresponding reweighed single event posterior conditioned on the drawn hyperparameters for each hyperparameter sample

$$\Lambda^{(j)} \sim p(\Lambda|d_1, d_2, \ldots, d_N)$$
$$\theta_i^{(j)} \sim p(\theta_i|d_i, \Lambda^{(j)})$$

then we effectively draw samples from the joint distribution

$$\begin{aligned}
\Lambda^{(j)}, \theta_i^{(j)} &\sim p(\Lambda|d_1, d_2, \ldots, d_N)p(\theta|d_i, \Lambda^{(j)}) \\
&\propto \left(p(\Lambda) \prod_k^N \left[\frac{\int d\vartheta p(d_k|\vartheta)p(\vartheta|\Lambda)}{\beta(\Lambda)} \right] \right) \left(\frac{p(d_i|\theta_i)p(\theta_i|\Lambda)}{\int d\vartheta p(d_i|\vartheta)p(\vartheta)|\Lambda)} \right) \\
&\propto p(d_i|\theta_i) \frac{p(\theta_i|\Lambda)p(\Lambda)}{\beta(\Lambda)} \prod_{k\neq i}^N \left[\frac{\int d\vartheta p(d_k|\vartheta)p(\vartheta|\Lambda)}{\beta(\Lambda)} \right]
\end{aligned}$$

which is as desired. That is, instead of computing a new effective prior by marginalizing over reweighed Λ samples and then reweighing single-event posteriors, this combined sampling procedure can rapdily draw joint samples for both Λ and θ_i. We also note that, in the effective prior approach, one would need to recompute different effective priors for different sets of single-event posteriors (or combinations of events), which may become computationally intractable. This joint sampling procedure can be extended to simultaneously sample hyperparameters and an arbitrary number of single-event parameters from an arbitrary number of events in a straightforward manner.