Problem Set - 4
Compact Binary Evolution, Rates and Population Modelling

Mukesh Kumar Singh, Srashti Goyal, Soummyadip Basak

Problem 1:

Gaussian - Gaussian Model: The noise and observations are both normally distributed. We consider the simplest case where observed signal is dependent only on a single parameter θ. Then, the distribution of parameter:

$$p(\theta|\vec{\lambda} \equiv \{\mu, \Sigma^2\}) = \frac{1}{\sqrt{2\pi\Sigma^2}} \exp\left[-\frac{(\theta - \mu)^2}{2\Sigma^2}\right] \tag{1}$$

where $\vec{\lambda} \equiv \{\mu, \Sigma^2\}$ are the population hyper-parameters (parameters charaterising the shape of distribution of θ). The noise is also Gaussian with zero mean and variance σ^2. We can write the observed data as

$$d = \theta + n \tag{2}$$

The likleihood of data d given parameter θ is given by

$$p(d|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(d - \theta)^2}{2\sigma^2}\right] \tag{3}$$

(a) Compute the likelihood of data d given hyper-parameters $\vec{\lambda}$: $p(d|\vec{\lambda})$. *Hint: You would like to use the probability multiplicative rule here, i.e.*

$$p(d|\vec{\lambda}) = \int d\theta \; p(d|\theta)p(\theta|\vec{\lambda}) \tag{4}$$

and compute the integral analytically. We have ignored the selection effects (not all the signals are detected by the detector) here for the sake of simplicity. For N number of data points, we will have

$$p(D|\vec{\lambda}) = \prod_{i=1}^{N} \int d\theta \; p(d_i|\theta)p(\theta|\vec{\lambda}) \tag{5}$$

where $D = \{d_1, d_2, ..., d_N\}$.

(b) Assume true values of hyper-parameters as $\mu_{\text{true}} = 0.5, \Sigma_{\text{true}} = 1$. Choosing width of noise distribution as $\sigma = 0.1$, simulate the observations D for, say $N = 10^5$ samples and plot the histogram.

(c) Assuming uniform priors on μ and Σ^2, find out the posteriors on the hyper-parameters $\vec{\lambda}$. Here, given the analytical form of the likelihood, you would like to use Emcee sampler to sample the likelihood over the hyper-parameters. Plot the posteriors on hyper-parameters along with true values and 90% credible intervals to see if the true values lie inside this interval.

(d) **[Bonus]** Suppose the integration in Eq. (4) can not be performed analytically, which is the reality in most of the cases, how would you compute the likelihood/posterior over hyper-parameters? *Hint: you might like to use Monte Carlo method for integration.*

(e) **[Bonus]** Put the detection threshold $d_{\text{th}} = 0$ and see how many of the above events are detected? Plot the histogram of the detected events, i.e. $D_{\text{det}} = \{d_i\}_{i=1}^{N_{\text{det}}}$ with $d_i \geq d_{\text{th}}$.

(f) **[Bonus]** Calculate the selection function $\alpha(\vec{\lambda})$, which can be computed as

$$\alpha(\vec{\lambda}) = \int_{d_{\text{th}}}^{\infty} dd \int d\theta \; p(d|\theta)p(\theta|\vec{\lambda}) \tag{6}$$

Hint: A clever choice of the order of integration here might make your life simple to evaluate it analytically. It is worth mentioning here that most of the times this integral can not be computed analytically.

(g) [**Bonus**] Calculate the posteriors on the hyper-parameters using the detected events, i.e. after including the selection effects, using Emcee sampler. *Hint: Notice that selection effects will modify the expression for likelihood for hyper-parameters as*

$$p(D_{\text{det}}|\vec{\lambda}) = \prod_{i=1}^{N_{\text{det}}} \frac{\int d\theta \; p(d_i|\theta)p(\theta|\vec{\lambda})}{\alpha(\vec{\lambda})} \tag{7}$$

Problem 2:

GWPopulation Tutorial: Please download the Jupyter-Notebook from *LIGO-Virgo Open Data Workshop-2022* on population inference of binary black hole (BBH) mergers using GWPopulation, a python package for doing population inference.

(a) Which population model is used here? Are there other population models that could be used?

(b) Attempt the challenge question in the notebook.

(c) How does one test the different population models against the data?

(d) Make statements about the formation channels of binaries based on the above inference.

Problem 3:

Reweighting the posteriors with the population prior: By now we know that the posterior is proportional to prior times the likelihood, where the prior is choosen from an "astrophysical" expectation of the binary parameters. However, different astrophysical models predict different populations of the binary parameters(eg: power law in primary masses). As a PE run is expensive, one can simply re-weight the posteriors to go from one prior distribution to another.

In this excercise we will reweight the GW150914 posteriors on m1 and q choosing to a powerlaw population model as our prior.

$$P_2(m_1, q|d) = P_1(m_1, q|d)\frac{P_2(m_1, q)}{P_1(m_1, q)} \tag{8}$$

(a) Derive the above formula using the bayes theorem.

(b) What prior was used for the component masses? Is it uniform? *Hint*: try plotting histograms and $m_1 - m_2$ scatter. Are there any constraints on masses? What is the minimum and maximum mass?

(c) Evaluate and plot the priors in m_1 and q. Can you write an analytical expression for the prior, $P_{PE}(m_1, q)$? *Hint*: Derieve $P(q|m_1)$ and use bayes theorem.

(d) Using the population model from the last question, after fixing the hyperparameter values from the maximum likelihood values, can you write the joint probability: $P_{pop}(m_1, q)$?

$$p(m_1|\alpha, M_{\text{min}}, M_{\text{max}}) \propto \begin{cases} m_1^{-\alpha}, & M_{\text{min}} < m_1 < M_{\text{max}} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

$$p(q|\beta, M_{\text{min}}, m_1) \propto \begin{cases} q^{\beta}, & M_{\text{min}} < m_2 < m_1 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

(e) Reweight the GW150914 posteriors of m_1 and q taking the above population distribution($P_{pop}(m_1, q)$) as a prior. Plot corner of new and old posterior samples, what do you see? *Hint*: You can do rejection sampling.

(f) [**Bonus**] Evaluate $P_2(m_1, q|d)$ directly by performing KDE on the original posterior samples. Compare your results with the above re-weighting method.

(g) [**Bonus**] Use various mass population models from GWPopulation(including the above one) as prior and re-weight the posteriors.

Problem 4:

[**Bonus**] Perform a simple hierarchical inference. Using a code like GWPopulation might make this easier. A suggestion could be fitting a line to sets of normally distributed data points at different x values, where the mean of the Gaussians varies linearly with y. The population parameter would be the gradient and intercept of the line, as well as the standard deviation for the points. An extension would be to make the standard deviation a function of x too. $d(x_i) = \mu(x_i) + n(\sim \mathcal{N}(0, \sigma^2)), \mu(x_i) = mx_i + c, p(m, c|d) = ?$