Problem Set - 1

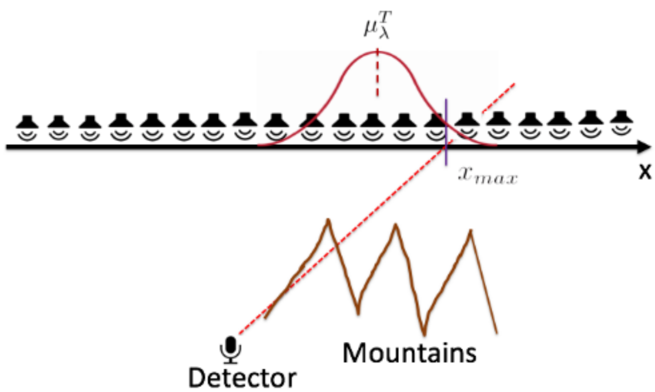
Inferring population properties of compact binary mergers

Mukesh Kumar Singh, Aditya Vijaykumar

August 11, 2021

Problem 1: A simple 1 -dimensional example

There is an infinite number of audio speakers arranged along the x-axis. An audio detector is placed at some distance as in the figure below.



When one speaker at position x emits a sound, the detector records:

$$d = x + n \tag{1}$$

where x (in meters) is the true position of the speaker, and $n \sim \mathcal{N}(0, 1)$ is a random number drawn from a normal distribution with mean 0 and standard deviation 1 meter.

Let's assume that there exists a maximum value of x such that detectors placed at $x > x_{\rm max}$ would not produce a measurable sound (this could be let's say due to mountains present between speaker and detector as shown in figure). Hence the detection criteria:

$$d \leq x_{\rm max} \leftrightarrow \rho(d) > \rho_{\rm thr} \tag{2}$$

Can you see that even a source with $x_{\rm true} = x_{\rm max}$ will produce a detectable signal 50% of the time? Using the gaussian nature of noise, we can write the likelihood for the data d given x.

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

Let us now focus on the inference of the property of the speakers, given a set of detections. We will assume that positions of the speakers that will play are scattered around some central value μ_λ^T and are drawn from a normal distribution:

$$\pi(x|\mu_\lambda, \sigma_\lambda) = \mathcal{N}(x - \mu_\lambda, \sigma_\lambda) \tag{4}$$

Someone has told us the value of σ_λ and our goal is to find out the true value of μ_λ, μ_λ^T, given the data d recorded by the detector. As you might see that there exists only one hyper parameter μ_λ (the true mean of the distribution of speakers that produce sound). The posterior on hyper parameter μ_λ (using $\vec{\lambda} = \mu_\lambda$ and $\vec{\theta} = x$ in eq. (108) of Salvatore et al):

$$p(\mu_\lambda|D) = \frac{\pi(\mu_\lambda)}{p(D|\mathcal{H}_\Lambda)} \Gamma(N^{\rm tr} - 1) \prod_{i=1}^{N^{\rm tr}} \frac{\int dx p(d_i|x)\pi(x|\mu_\lambda)}{\alpha(\mu_\lambda)} \tag{5}$$

The selection effects (detection efficiency) can be estimated by equation:

$$\alpha(\mu_\lambda) = \int dx\ \pi(x|\mu_\lambda) p(\rho_\uparrow|x) \tag{6}$$

with the probability of a source being detectable given true value of the parameter (i.e. data with $d < x_{\max}$)

$$p(\rho_\uparrow|x) = \int_{\mathcal{D}_\uparrow} \mathrm{d}d\ p(d|x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x_{\max} - x}{\sqrt{2}}\right)\right] \tag{7}$$

(a) Analytical Integration of likelihood

Using the likelihood expression (eq. [3]) for individual data d_i, the integration over $\vec{\theta} = x$ in eq. [5] can be carried out analytically rendering the hyper posterior for μ_λ upto a normalization as:

$$p(\mu_\lambda|D) \propto \pi(\mu_\lambda) \prod_{i=1}^{N^{\mathrm{tr}}} \frac{\exp\left[-\frac{(d_i-\mu_\lambda)^2}{2(1+\sigma_\lambda^2)}\right]}{\alpha(\mu_\lambda)} \tag{8}$$

Now given the analytical expression for the likelihood, a possible implementation of the algorithm to estimate the hyper parameter μ_λ would be (interested people may embark upon solving the integration and confirming if they get the above expression):

- Choose the values of $\mu_\lambda^T, \sigma_\lambda$, and $x_{\max}$ which will be used to generate the synthetic set of measurements.
- Generate N random numbers $\vec{x}_T \sim \mathcal{N}(\mu_\lambda^T, \sigma_\lambda)$, the true positions of the speakers that will be playing, and we do not have access to those directly.
- Generate N random numbers for noise $n \sim \mathcal{N}(0, 1)$, add noise to true positions $\vec{x}_T$ to get the data d.
- Select the only data points that satisfy the detection criteria:

  $$d_i \in D \quad \text{if} \quad d_i < x_{\max}$$

  where D is our catalog of N^{tr} detected sources that will be used for the inference.
- Grid the μ_λ axis with enough points, and at each point calculate selection effects $\alpha(\mu_\lambda)$. Plot it and see if this is according to our expectations.
- Pick a prior for μ_λ, flat is just fine, and calculate eq. (8). Plot the hyper posterior on μ_λ and see if this recovers the true value of μ_λ, i.e. μ_λ^T.

(b) Numerical Integration of Likelihood

In general, we will not be able to perform likelihood integration analytically as we did in the previous section, due to higher dimensionality, correlations, non-trivial priors etc. In that case, using Bayes theorem in eq. (5):

$$p(\mu_\lambda|D) \quad \propto \quad \pi(\mu_\lambda) \prod_{i=1}^{N^{\mathrm{tr}}} \frac{\int dx p(x|d_i) \frac{\pi(x|\mu_\lambda)}{\pi(x|\mathcal{H}_{\mathrm{PE}})}}{\alpha(\mu_\lambda)} \tag{9}$$

$$\propto \quad \frac{\pi(\mu_\lambda)}{\alpha(\mu_\lambda)^{N^{\mathrm{tr}}}} \prod_{i=1}^{N^{\mathrm{tr}}} \frac{1}{N_{\mathrm{samples}}} \sum_{j=1}^{N_{\mathrm{samples}}} \left.\frac{\pi(x_i^j|\mu_\lambda)}{\pi(x_i^j|\mathcal{H}_{\mathrm{PE}})}\right|_{x_i^j \sim p(x|d_i)} \tag{10}$$

where in the last equality the ratio of the population prior over sampling prior [i.e. the prior used to produce the posterior samples $p(x|d_i)$] has been evaluated at values x of size N_{samples} fairly drawn from the posterior of each trigger.