Psych
186A

Using Sampling Techniques to do Bayesian Inference: Part I

In previous tutorial, we have looked at many examples of Bayesian inference in which we know the functional form of the prior distribution, so that we are able to compute the exact prior probability for each hypothesis.

However, in some problems, you may come across situations in which computing the exact prior probabilities is difficult. In those cases, one alternative is to use some sampling methods. Today, we are going to look at how sampling can help us do Bayesian inference, with the example we went through last week: a robot tries to locate a target object in a 1D world.

Below are the mathematical specifications:

Let t be the random variable of the true target location.
Let $\boldsymbol{x}$ be the observed locations reported from your sensor. In this example, we have four reported locations, x = [7 3 -1 8];
Assume the following:
Prior for θ is a normal distribution, with mean μ_θ and variance σ_θ^2,
i.e., $P(\theta) \sim Normal(\mu_\theta, \sigma_\theta^2)$ $\quad \mu_\theta = 0, \sigma_\theta = 8$
Likelihood of a observed locations x_i follows a normal distribution, with mean θ and variance σ_x^2,
i.e., $P(x_i|\theta) \sim Normal(\theta, \sigma_x^2)$ $\quad \sigma_x = 4$
As we have multiple observed locations, we assume that these observations are independent. The likelihood of observing the four observations in the this study would be

$$P(\boldsymbol{x}|\theta) = \prod_i P(x_i|\theta) = P(x_1|\theta)P(x_2|\theta)P(x_3|\theta)P(x_4|\theta)$$

Recall the steps for the **discretize-space method**:

1. Define the range of all possible locations t for the target to appear (i.e., all possible hypotheses)
2. Compute the prior probability for each hypothesized target location, i.e., $P(t)$
3. Compute the likelihood for the observed location x, i.e., $P(x|t)$ for each t in the range defined in step 1.
4. Obtain the posterior distribution by
   a. Multiplying the prior and likelihood probabilities, correspondingly for each location t
   b. Normalizing the products
5. Obtain an estimate from the posterior distribution. We'll use the expected value of the location using its posterior distribution.

$$\mathrm{E}[X] = \sum_{i=1}^{\infty} x_i\, p_i. \qquad\qquad \mathrm{E}[X] = \int_{\mathbb{R}} x f(x)\, dx,$$

Now, suppose we cannot carry out step 2 for some reason. Let's assume that, although we cannot compute the prior probability for each t, we are able to randomly sample some hypothesized locations t based on the prior distribution (Note: these assumptions are not valid in this problem, because we can easily compute the pdf for any normal distribution; the assumptions are just for illustration purpose).

If so, we can perform the same inference by using the **sampling method**:

1. Sample N possible locations t, i.e., obtain $t_1, t_2, \ldots, t_N$, based on the prior distribution.
2. Compute the likelihood for the observed location x, i.e., $P(x|t_i)$ for each sampled t_i.
3. Assign a weight w_i to each sampled t_i such that the weight is proportional to the likelihood, i.e., $w_i \propto P(x|t_i)$. Remember to normalize the weights: i.e., divide each of them by $\sum_{i=1}^{N} w_i$
4. Resample from the N locations you obtained step 1 based on their weights, with replacement. In other words, the probability for a certain location t_i to be resampled should be equal to its normalized weight w_i. This new sample is, in fact, a sample from the posterior distribution.
5. Obtain an estimate from the posterior distribution (to make this comparable with the analytical method above, use mean here).

In-class Exercise

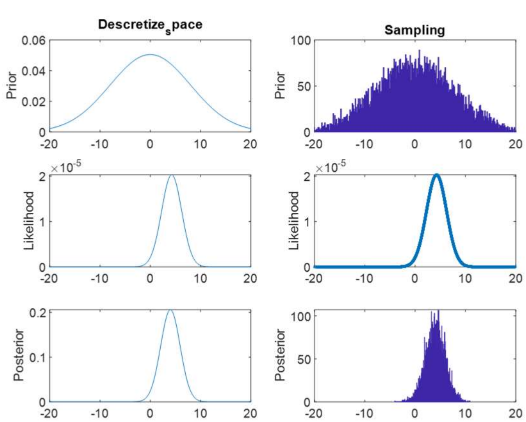
In **SamplingExercise1.m**, there are 4 sections.

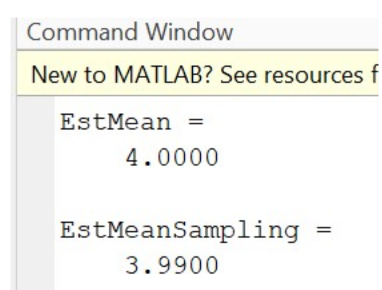
1) **Parameters**: All parameters you need are defined here. All have been defined for you with some default values. Make sure you understand the meaning of each of them.

2) **Discretize-space method**: Done for you as an example. This is the same as what we did last week. Make sure you understand what each step does and how it is done in MATLAB. The steps correspond to those described in Page 1 of this handout, under analytical method.

3) **Sampling**: *Complete this section*, based on the guidelines given in the comments. Steps correspond to those described under sampling method on Page 1. You will compute the mean value from posterior samples and display the value in the command window (note that you may get a value slightly different from the answer provided below due to the randomness of sampling).

4) **Reporting Results**: Mostly done for you. You just need to *fill in the names of the variables* you defined under the Sampling section. When you have time, try to learn those plotting techniques, which might be useful for your problem sets / project.





Optional: After completing the Sampling section, try out the following:

a) Compare values obtained from the two methods. Are they similar?

b) Make a note of the estimated value you obtain from Sampling. Rerun the program. Do you get the same value? Can you explain why?

c) Obtain fewer samples by reducing N to, say, 1000. Take the estimate from the Analytical method as a benchmark. How does the sample size N affect the accuracy and variability of your estimate?