

STAT 447 Assignment 5

Caden Hewlett

2024-02-07

Question 1: Sequential Updating

Consider a joint probabilistic model given by

$$\theta \sim \rho, \text{ and } (x_i | \theta) \stackrel{iid}{\sim} \nu_\theta, \text{ where } i \in \{1, 2, \dots, n\}$$

where ρ is a prior distribution for the unknown parameter θ , and $\{x_i\}_{i=1}^n$ is a sequence of observations with conditional distribution ν_θ .

Part 1

Write down the posterior distribution of θ given $\{x_i\}_{i=1}^n$.

In a more verbose sense, let Θ be the random variable for the unknown parameter, and θ be the realization. We can then write.

Prior follows $p_\Theta(\theta)$, where $p_\Theta(\theta)$ is the PMF/PDF given by ρ

$P(X = x | \theta) = p_{X|\theta}(x, \theta)$, where $p_{X|\theta}(x, \theta)$ is the PMF given by ν_θ

With this in mind, we can write the posterior of θ given $\{x_i\}_{i=1}^n$, where we describe the event that $\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n$ using Bayes' Rule.

We will begin by using the most verbose notation possible, for complete clarity.

$$P(\Theta = \theta | \{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n) = \frac{P(\Theta = \theta)P(\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n | \Theta = \theta)}{P(\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n)}$$

We start by considering the joint likelihood function.

$$P(\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n | \Theta = \theta) = P((X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1) | \Theta = \theta)$$

Then, using intersections, we can write the likelihood as:

$$P(\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n | \Theta = \theta) = P\left(\bigcap_{i=1}^n (X_i = x_i) | \Theta = \theta\right)$$

Now, we use the following property of *iid* random variables

$$\text{I.I.D} \implies \forall (i \neq j) \in [1, n], X_i \perp X_j \implies \forall (i \neq j) \in [1, n], P(X_i \cap X_j) = P(X_i)P(X_j)$$

Hence, we can write the likelihood as:

$$P(\{X_i\}_{i=1}^n = \{x_i\}_{i=1}^n | \Theta = \theta) = \prod_{i=1}^n \left(P(X_i = x_i | \Theta = \theta)\right) = \prod_{i=1}^n (\nu_\theta) = (\nu_\theta)^n$$

Which, as you can see, simplifies nicely to $(\nu_\theta)^n$.

Part 2

Suppose now we get an additional data point x_{n+1} with the same conditional distribution ν_θ . Show that using the posterior from part 1 as the *prior* and data equal to just x_{n+1} gives the same posterior distribution as redoing part 1 with the $n + 1$ data points.

Question 2: Baesian Inference in the Limit of Increasing Data