Lecture 5

Bayes From the Ground Up

Definition: MAP
Let,

$$\hat{\theta}_{MAP} = argmax_{\theta \in \Theta_0}\{P(\theta|x)\}$$

MAP stands for *maximum a posteriori* where the x in $P(\theta|x)$ is the value we get after the fact. Back when we were studying MLE (maximum likelihood function) we were attempting to find the value of theta given the data. We are looking for the same thing here, we are looking for the best value of theta, the parameter given the data. This is one of three Bayesian point estimate we will study.

$$\mathcal{F} = iid\ Bern(\theta), n = 3, x =< 0, 1, 1 >, \Theta = \{0.5, 0.75\}$$

Note that this data set, x is part of a larger data space which may be defined as follows,

$$x \in \mathcal{X} = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$$

This value $\mathcal{X}$ defines everything that can happen given this is a Bernoulli random variable and $n = 3$.

$$P(x =< 1, 1, 1 > |\theta = 0.75) = 0.75^3 = 0.422$$

$$P(x =< 1, 1, 0 > |\theta = 0.75) = 0.75^2 \times 0.25 = 0.141$$

$$...$$

$$P(x =< 1, 0, 0 > |\theta = 0.75) = 0.75 \times 0.25^2 = 0.047$$

$$...$$

$$P(x =< 0, 0, 0 > |\theta = 0.75) = 0.25^3 = 0.016$$

and for $\theta = 0.5$, we get the following

$$P(x =< 1, 1, 0 > |\theta = 0.5) = 0.5^2 \times 0.25 = 0.125$$

The probability that x is a realization of the random variable is the sum of the probabilities that x is that realization with all parameters in the parameter space. For instance,

$$P(x =< 1,0,0 >= P(x =< 1,0,0 > |\theta = 0.5) + P(x =< 1,0,0 > |\theta = 0.75) = 0.57 + 0.125$$

Furthermore, the probability that theta is a certain constant value given x is a given realization is the probability that theta is that value at the realization over the probability of the realization at all values of theta in the parameter space.

$$P(\theta|x =< 1,0,0 >) = \frac{P(\theta = 0.5, x =< 1,0,0 >)}{P(x =< 1,0,0 >)} = \frac{(0.125)0.5}{(0.47 + 0.125)0.5}$$

$$\Sigma_{\theta\in\Theta}\ P(\theta) = 1$$

$$\Sigma_{\theta\in\Theta}\ P(\theta|x) = 1$$

$$\Sigma_{\theta\in\Theta}\ P(x|\theta) = \ any\ positive\ number$$

Lets take a look at the following equation,

$$P(\theta|x) = \frac{P(x,\theta)}{P(x)} \alpha P(x.\theta) \alpha P(x|\theta) P(\theta) \alpha P(x|\theta)$$

$P(\theta)$ may be considered a constant, which is why it disapears in the final expression, this is a product of Laplace's idea, that is the principle of indifference. So let us look at $\hat{\theta}_{MAP} = argmax_{\theta\in\Theta_0}\{P(\theta|x)\}$ You cannot change the value of theta if you divide the resulting value by a constant, theta stays the same. So let us look at the two equations side by side.

$$P(\theta|x) = \frac{P(x,\theta)}{P(x)} \alpha P(x.\theta) \alpha P(x|\theta) P(\theta) \alpha P(x|\theta)$$

$$\hat{\theta}_{MAP} = argmax_{\theta\in\Theta_0}\{P(\theta|x)$$

$$= argmax_{\theta\in\Theta_0}\{P(\theta|x)P(\theta) = argmax_{\theta\in\Theta_0}\{P(x|\theta) =?\hat{\theta}_{MLE}$$

Note that $\hat{\theta}_{MAP}$ is only equal to $\hat{\theta}_{MLE}$ is the MLE is a parameter set you specify. Let $\Theta_0 = \{0.1, 0.25, 0.5, 0.75, 0.9\}$. Calulcations get us the following,

$$x =< 1,1,0 >$$

$$P(x|\theta = 0.1) = 0.1^2 \times 0.9 = 0.009$$

$$P(x|\theta = 0.25) = 0.25^2 \times 0.75 = 0.047$$

$$P(x|\theta = 0.5) = 0.125$$

$$P(x|\theta = 0.75) = 0.141$$

$$P(x|\theta = 0.9) = 0.081$$

The largest probability that x is the given value is when $\theta = 0.75$, so we say that $\hat{\theta}_{MAP}$=0.75.

$$P(\theta = 0.75|x =< 1,1,0 >) = \frac{P(x|\theta = 0.75)}{\Sigma_{\theta \in \Theta} P(x|\theta)}$$

Lets examine Laplace's prior under many different parameter spaces approaching the full space.

$$\mathcal{F} iidBern(\theta), x =< 1,1,0 >, P(\theta) = \cup(0,1)$$

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)} = \frac{P(x|\theta)1}{P(x)} = \frac{P(x|\theta)}{\int_{\Theta} P(x,\theta)d\theta} = \frac{\theta^2(1-\theta)}{\int_0^1 \theta^2(1-\theta)d\theta} = \frac{\theta^2(1-\theta)}{[\frac{\theta^3}{3} - \frac{\theta^4}{4}]}$$