SMAI Class 12

September 12, 2022

MLE

Assume $\hat{\theta}$ to be constant but unknown.

What density will we assume given the mean of the distribution (example mean weight of apples)?

Normal density seems the best approach.

But the issue with normal density is that it has infinite support. (Negative weights will not make sense).

So, Rayleigh density seems more appropriate here. We know by central limit theorem tho, that everything will be normal at EOD.

If we don't know which density to pick **Normal is always the best to pick**

Revisited

We assume that we know $p(x|\theta)$.

$p(D|\mu\sigma)$: Likelihood of data given μ and σ.

This is equal to $\prod_{k=1}^{n} p(x_k|\mu\sigma)$.

Bayes' theorem intuition

We have a prior distribution (before observing) of what things could be in the world. And then it changes after observation.

Here,

$$P(\theta|D) = \frac{p(D|\theta)P(\theta)}{p(D)}$$

—

Assume now that variance is known, but mean is not known.

Now we are saying that the mean itself is a distribution.

In bayesian : we assume parameters are distributions.

In MLE : we assume parameters to be constant (which might be unkown to us).

Prior of μ is a distribution where let's say θ_o is the best estimate for it. (before measuring real apples).

MAP: Maximum A Posteriori (MAP) Estimation

$$\hat{\theta}_{MAP} = argmax p(D|\theta)p(\theta)$$

$p(x|D) = \int_\theta p(x|\theta)p(\theta|D)d\theta$

$$p(\mu|D) = \frac{p(D|\mu)p(\mu)}{\int p(D|\mu)p(\mu)d\mu}$$
$$= \alpha \left(\prod_{k=1}^{n} p(x_k|\mu)\right) p(\mu)$$

where α is the denominator value.

$$p(\mu|D) = \alpha \prod_k \frac{1}{\sqrt{2\pi}\sigma} exp\{-1/2(x_k - \mu)^2/\sigma^2\} \times \frac{1}{\sqrt{2\pi}\sigma_o} exp\{-1/2(\mu - \mu_o)^2/\sigma_o^2\}$$

product of density for x_k then the density of μ.

The above product can be simplified to:

$$\alpha \, exp\left\{-\frac{1}{2}\left[\left(\frac{n}{\sigma^2} + \frac{1}{\sigma_o^2}\right)\mu^2 - 2\left(\frac{1}{\sigma^2}\sum_k x_k + \frac{\mu_o}{\sigma_o^2}\right)\mu\right]\right\}$$