Assignment 5: CS 215

Devansh Jain	Harshit Varma
190100044	190100055

November 17, 2020

Question 3

Prior:

$$p(\theta) = K \left(\frac{\theta_m}{\theta}\right)^{\alpha} I(\theta \geq \theta_m)$$

Likelihood:
Let $x_M = \max\{x_i\}_{i=1}^n$

$$L(\{x_i\}_{i=1}^n|\theta) = \frac{1}{\theta^n} I(\theta \geq x_M)$$

For maximizing/computing the mean, we would want Posterior to be non-zero.
This will be possible iff $\theta \geq \max\{x_M, \theta_m\}$
Let $M = \max\{x_M, \theta_m\}$
Thus, a constraint on θ is $\theta \geq M$

Evidence:

$$\begin{aligned} e(\{x_i\}_{i=1}^n) &= \int_M^{\infty} L(\{x_i\}_{i=1}^n|\theta)p(\theta)d\theta \\ &= \frac{K\theta_m^{\alpha}}{(n+\alpha-1)M^{n+\alpha-1}} \end{aligned}$$

Posterior:

$$\begin{aligned} g(\theta|\{x_i\}_{i=1}^n) &= \frac{p(\theta) \cdot L(\{x_i\}_{i=1}^n|\theta)}{e(\{x_i\}_{i=1}^n)} \\ &= \left(\frac{(n+\alpha-1)(M^{n+\alpha-1})}{K\theta_m^{\alpha}}\right) K \left(\frac{\theta_m}{\theta}\right)^{\alpha} I(\theta \geq \theta_m) \frac{1}{\theta^n} I(\theta \geq x_M) \\ &= (n+\alpha-1)(M^{n+\alpha-1}) \frac{1}{\theta^{n+\alpha}} I(\theta \geq M) \\ &= K' \left(\frac{M}{\theta}\right)^{n+\alpha} I(\theta \geq M) \end{aligned}$$

This exactly matches the form of Pareto$(M, n+\alpha)$.
Since both the posterior and Pareto$(M, n+\alpha)$ integrate to 1, the normalizing constant is same in both, thus the posterior exactly equals Pareto$(M, n+\alpha)$.

ML Estimate

Likelihood is maximized when θ takes the least value it is allowed to take, $= x_M$ (Here, we ignore the prior and thus ignore the constraints arising due to the prior)

$$\boxed{\theta_{ML} = x_M}$$

MAP Estimate

Mode of the Posterior Pareto$(M, n + \alpha)$ distribution occurs at M

$$\boxed{\theta_{MAP} = M}$$

Posterior Mean Estimate

This is equal to the mean of the Posterior Pareto$(M, n + \alpha)$ distribution

$$\boxed{\theta_{PME} = \left(\frac{\alpha + n - 1}{\alpha + n - 2}\right) M}$$

Neither the MAP estimate, nor the PME estimate tends to the ML estimate when $M \neq x_M$.
This is not desirable, as ideally more data should improve our estimates.
When $M = x_M$, then both MAP and PME estimates tends to the ML estimate as n increases.

The above result is justifiable - if we choose a prior of θ which only takes values greater than x_M, then we would never get θ such that it matches with ML Estimate which is x_M.
Had this not been the case and $\theta_m \leq x_M$, i.e., ML estimate lies in the domain of prior function of θ, then we see that both MAP and PME tend to ML estimate (MAP is equal to ML estimate in this case).