Problem 3

Part 1

Prior on Θ

$$PDF_{\Theta}(\theta) = \frac{(\alpha-1)\,\theta_m^{\alpha-1}}{\theta^{\alpha}} \quad \text{if } \theta \geq \theta_m$$

$$= 0 \quad \text{otherwise}$$

Likelihood

$$PDF(X_1 = x_1, \ldots, X_n = x_n \mid \Theta = \theta)$$

$$= \begin{cases} \frac{1}{\theta^n} & \text{if } \forall i = 1,2,\ldots n \;\; x_i \in [0,\theta] \\ 0 & \text{otherwise} \end{cases}$$

Evidence

$$PDF(X_1 = x_1, \ldots, X_n = x_n)$$

$$= \int_{\theta} P(X_1 = x_1, \ldots, X_n = x_n, \Theta = \theta)\, d\theta$$

$$= \int_{\theta} P(X_1 = x_1, \ldots, X_n = x_n \mid \Theta = \theta)\, P(\Theta = \theta)\, d\theta$$

$$= \int_{\theta} \frac{(\alpha-1)\,\theta_m^{\alpha-1}}{\theta^{\alpha+n}}\, d\theta \quad \text{if } \theta \geq \theta_m \text{ and } \forall i,\; x_i \in [0,\theta]$$

$$= \frac{(\alpha-1)\,\theta_m^{\alpha-1}}{(\alpha+n-1)\,\beta^{\alpha+n-1}} \quad \text{where } \beta = \max(x_1, x_2, \ldots, x_n, \theta_m)$$

Posterior

$$PDF(\Theta = \theta \mid X_1 = x_1, \ldots, X_n = x_n)$$

$$= \frac{PDF(X_1 = x_1, \ldots, X_n = x_n \mid \Theta = \theta)\, P(\Theta = \theta)}{P(X_1 = x_1, \ldots, X_n = x_n)}$$

$$= \frac{1}{\theta^n} \cdot \frac{(\alpha-1)\,\theta_m^{\alpha-1}}{\theta^{\alpha}} \cdot \frac{(\alpha+n-1)\,\beta^{\alpha+n-1}}{(\alpha-1)\,\theta_m^{\alpha-1}}$$

$$= \begin{cases} \frac{(\alpha+n-1)\,\beta^{\alpha+n-1}}{\theta^{\alpha+n}} & \text{if } x_i \in [0,\theta]\; \forall i \text{ and } \theta \geq \theta_m \text{ and } \beta = \max(x_1, x_2, \ldots, x_n, \theta_m) \\ 0 & \text{otherwise} \end{cases}$$

ML Estimate

PDF $(X_1=x_1, \ldots, X_n=x_n \mid \Theta=\theta)$

$$= \begin{cases} \frac{1}{\theta^n} & \text{iff } \forall i \in [n],\ x_i \in [0,\theta] \\ 0 & \text{otherwise} \end{cases}$$

For the condition to be satisfied, $\theta \geq \max(x_1, x_2, \ldots, x_n)$ and $x_i \geq 0\ \forall\ i \in [n]$. Since, $\frac{1}{\theta^n}$ is a decreasing function on positive θ,

$$\hat{\Theta}_{ML} = \max(x_1, x_2, \ldots, x_n) \quad \text{if } x_i \geq 0\ \forall\ i = 1, 2, \ldots, n$$

$$= 0 \quad \text{otherwise}$$

MAP Estimate

PDF $(\Theta=\theta \mid X_1=x_1, \ldots, X_n=x_n)$

$$= \begin{cases} \frac{(\alpha+n-1)\,\beta^{\alpha+n-1}}{\theta^{\alpha+n}} & \text{if } x_i \in [0,\theta]\ \forall\ i=1,2,\ldots,n \\ & \text{and } \theta \geq \theta_m \\ & \text{and } \beta = \max(x_1, x_2, \ldots, x_n, \theta_m) \\ 0 & \text{otherwise} \end{cases}$$

Using similar reasoning as above, $\theta \geq \max(x_1, \ldots, x_n, \theta_m)$. and $x_i \geq 0\ \forall\ i \in [n]$. Since $\frac{(\alpha+n-1)\,\beta^{\alpha+n-1}}{\theta^{\alpha+n}}$ is a decreasing function on positive θ,

$$\hat{\Theta}_{MAP} = \max(x_1, x_2, \ldots, x_n, \theta_m) \quad \text{if } x_i \geq 0\ \forall\ i=1, \ldots, n$$

$$= 0 \quad \text{otherwise}$$

Part 2

We now compare ML and MAP estimates for θ as $n \to \infty$

Define $\lambda = \max(x_1, x_2, \ldots, x_n)$.

Then $\hat{\theta}_{ML} = \lambda$ if $x_i \geq 0 \ \forall i = 1, 2, \ldots, n$

$= 0$ otherwise

and $\hat{\theta}_{MAP} = \max(\lambda, \theta_m)$ if $x_i \geq 0 \ \forall i = 1, 2, \ldots, n$

$= 0$ otherwise

We see that $\hat{\theta}_{ML} = \hat{\theta}_{MAP}$ iff $\lambda \geq \theta_m$.

Let θ_{true} denote true value of θ.

Case 1 - $\theta_{true} < \theta_m$

Then all of $x_i \in [0, \theta_{true}] \subset [0, \theta_m]$

Hence, $\hat{\theta}_{ML} \neq \hat{\theta}_{MAP}$ for any n.

Case 2 - $\theta_{true} \geq \theta_m$

Since X is a uniform distribution on $[0, \theta]$,

$P(X \geq \theta_m) = 1 - \frac{\theta_m}{\theta}$ and $P(X < \theta_m) = \frac{\theta_m}{\theta}$

$$P(\lambda \geq \theta_m) = 1 - P(\lambda < \theta_m)$$

$$= 1 - P(x_1 < \theta_m) \ldots P(x_n < \theta_m)$$

$$= 1 - \left(\frac{\theta_m}{\theta}\right)^n$$

But as $\lambda \geq \theta_m$ implies $\hat{\theta}_{MAP} = \hat{\theta}_{ML}$,

$$P(\hat{\theta}_{MAP} = \hat{\theta}_{ML}) = 1 - \left(\frac{\theta_m}{\theta}\right)^n \to 1 \text{ as } n \to \infty$$

Hence, if $\theta_{true} < \theta_m$, they do not converge which isnt desirable as number of experiments $\to \infty$

If $\theta_{true} \geq \theta_m$, then they do converge, which is desirable as the real experiments dominates the prior belief.

Part 3

Posterior - Mean

$$\hat{\Theta}_{\text{Posterior Mean}} = E_{P(\Theta | x_1, \dots, x_n)} [\Theta]$$

$$= \int_{\Theta} \Theta \cdot P(\Theta | x_1, \dots, x_n) \, d\Theta$$

$$= \int_{\beta}^{\infty} \frac{(\alpha+n-1) \, \beta^{\alpha+n-1}}{\Theta^{\alpha+n-1}} \, d\Theta \quad \text{where } \beta = \max(x_1, x_2, \dots, x_n, \Theta_m)$$

$$= \frac{(\alpha+n-1)}{(\alpha+n-2)} \max(x_1, x_2, \dots, x_n, \Theta_m)$$

Part 4

We now compare $\hat{\Theta}_{ML}$ and $\hat{\Theta}_{\text{Posterior Mean}}$ as $n \to \infty$.

It is easy to see that as $n \to \infty$, $\frac{\alpha+n-1}{\alpha+n-2} \to 1$

Hence as $n \to \infty$, $\hat{\Theta}_{\text{Posterior Mean}} = \max(x_1, x_2, \dots, x_n, \Theta_m)$

Defining $\lambda = \max(x_1, \dots, x_n)$ as previously,

We get $\hat{\Theta}_{ML} = \lambda$ and $\hat{\Theta}_{\text{Posterior Mean}} = \max(\lambda, \Theta_m)$.

This is exactly like what we encountered before.

We use the same result. Hence,

if $\Theta_{true} < \Theta_m$, they do not converge, which is not desirable as $n \to \infty$.

if $\Theta_{true} \geq \Theta_m$, then they do converge, which is desirable as results of real experiments should dominate prior belief as $n \to \infty$.