CS215 Assignment 5
Estimation Transformations

Tathagat Verma 180050111
Neel Aryan Gupta 180050067

Question - 4

Pareto distribution $P(\theta) \propto \left(\frac{\theta_m}{\theta}\right)^\alpha$ for $\theta \geq \theta_m$

0 otherwise

First we find the normalization constant

$\therefore \int_{-\infty}^{\infty} P(\theta \mid \theta_m, \alpha)\, d\theta = 1$

$\Rightarrow k \int_{-\infty}^{\theta_m} 0 \, d\theta + k \int_{\theta_m}^{\infty} \left(\frac{\theta_m}{\theta}\right)^\alpha d\theta = 1$

$\Rightarrow k \cdot \theta_m^\alpha \int_{\theta_m}^{\infty} \frac{d\theta}{\theta^\alpha} = 1$

$\Rightarrow k \cdot \theta_m^\alpha \left[\frac{\theta^{1-\alpha}}{1-\alpha}\right]_{\theta_m}^{\infty} = 1$

$\Rightarrow k \cdot \theta_m^\alpha \cdot \frac{\theta_m^{1-\alpha}}{\alpha-1} = 1$

$\Rightarrow \boxed{k = \frac{\alpha-1}{\theta_m}}$

Now, we find the mean

$\therefore \int_{-\infty}^{\infty} \theta \cdot P(\theta \mid \theta_m, \alpha)\, d\theta \equiv \text{Mean}$

$\Rightarrow \text{Mean} = \int_{\theta_m}^{\infty} \theta \cdot \frac{\alpha-1}{\theta_m} \cdot \left(\frac{\theta_m}{\theta}\right)^\alpha d\theta = (\alpha-1)\theta_m^{\alpha-1} \int_{\theta_m}^{\infty} \frac{d\theta}{\theta^{\alpha-1}}$

$= (\alpha-1)\, \theta_m^{\alpha-1} \left[\frac{\theta^{2-\alpha}}{2-\alpha}\right]_{\theta_m}^{\infty}$

$\therefore$ If $\alpha < 2$ then mean $= \infty$ (absurd)

Assume $\alpha > 2$

PAGE NO.
DATE / /

$$\Rightarrow \text{Mean} = (\alpha-1)\,\theta_m^{\alpha-1}\left[\frac{\theta_m^{2-\alpha}}{2-\alpha}\right]_{\theta_m}^{\infty}$$

$$= (\alpha-1)\,\theta_m^{\alpha-1}\cdot\frac{\theta_m^{2-\alpha}}{\alpha-2}$$

$$= \left(\frac{\alpha-1}{\alpha-2}\right)\theta_m \quad \text{where } \alpha>2$$

Let the dataset be $\{x_i\}_{i=1}^{N}$

(j) Likelihood : $\prod_{i=1}^{N}\frac{1}{\theta} = \frac{1}{\theta^N}$

$\Rightarrow$ θ should be minimised

But $(0, \theta)$ should still contain all the data $\{x_i\}_{i=1}^{N}$

$\Rightarrow$ Lower bound for θ : $\max\{x_i\}_{i=1}^{N}$

Since θ has to be minimised

$\Rightarrow \theta = \max\{x_i\}_{i=1}^{N}$

Let $\tilde{\theta} = \max\{x_i\}_{i=1}^{N}$

$\therefore\ \hat{\theta}_{MLE} = \tilde{\theta}$

Posterior $\propto \prod_{i=1}^{N}\frac{1}{\theta}\cdot$ Prior $= \frac{1}{\theta^N}\cdot\left(\frac{\theta_m}{\theta}\right)^{\alpha}$ for $\theta \geq \theta_m$

$\therefore$ θ should be minimised

θ still has to contain the dataset but

Prior should also be maximised

$\Rightarrow \theta = \max(\theta_m, \tilde{\theta})$

If $\theta_m < \tilde{\theta}$ then likelihood would become zero at datapoints $> \theta_m$ thereby halting the maximality of Posterior. That's why $\theta = \max(\theta_m, \tilde{\theta})$ and NOT $\theta = \tilde{\theta}$.

$\Rightarrow \hat{\theta}^{MAP} = \max(\theta_m, \tilde{\theta})$

$\& \ \tilde{\theta}^{MLE} = \tilde{\theta}$

(ii) Note that if $\theta_m < \tilde{\theta}$ then $\hat{\theta}^{MAP} = \hat{\theta}^{MLE}$ always.
For large sample size, we see that
$P(\theta_m > \tilde{\theta}) \approx 0$

$\therefore P(\theta_m > \tilde{\theta}) = P(\theta_m > x_1, \theta_m > x_2, \ldots, \theta_m > x_N)$

$= \frac{\theta_m}{\theta} \cdot \frac{\theta_m}{\theta} \cdots$

$= \left(\frac{\theta_m}{\theta}\right)^N$

0 x_i x_j θ_m θ

But here $\theta_m < \theta \Rightarrow \theta_m/\theta < 1$

$\Rightarrow (\theta_m/\theta)^N \approx 0$ for large N

So $\hat{\theta}^{MAP}$ will tend to $\hat{\theta}^{MLE}$ as N increases.
This is desirable because we know that MLE estimators are the most effecient estimators for large sample sizes, using CRLB.

(iii) Note that posterior is zero if $\theta < \max(\tilde{\theta}, \theta_m)$
and posterior $\propto \frac{\theta_m^\alpha}{\theta^\alpha} \cdot \frac{1}{\theta^N}$

$\Rightarrow$ Posterior $\propto \frac{\theta_m^\alpha}{\theta^{\alpha+N}}$ for $\theta \geq \max(\hat{\theta}, \theta_m) = \theta_m'$

0 otherwise

This is also a Pareto distribution which has $\alpha' = \alpha + N$ & $\theta_m' = \max(\tilde{\theta}, \theta_m)$
We know the Pareto mean which is
$\left(\frac{\alpha-1}{\alpha-2}\right)\theta_m$

So we get

Posterior mean $= \left(\frac{\alpha+N-1}{\alpha+N-2}\right) \cdot \max(\tilde{\theta}, \theta_m)$

$\Rightarrow \hat{\theta}^{\text{Posterior mean}} = \left(\frac{\alpha+N-1}{\alpha+N-2}\right) \cdot \theta_m' \quad$ where $\theta_m' = \max(\tilde{\theta}, \theta_m)$

(iv) As $N \to \infty$

$$\lim_{N\to\infty} \left(\frac{\alpha+N-1}{\alpha+N-2}\right) \theta_m' = \theta_m' = \hat{\theta}^{MAP}$$

We know that $\hat{\theta}^{MAP}$ tends to $\hat{\theta}^{MLE}$ for large N

$\Rightarrow \hat{\theta}^{\text{Posterior Mean}} \to \hat{\theta}^{MLE}$ as $N \to \infty$

This is desirable due to the same reason as mentioned in (ii) that $\hat{\theta}^{MLE}$ is the most effecient estimator for $N \to \infty$.

This shows Bayes estimators gives correct results even for large N (in this case).