

Assignment 5

Raghav Juyal - EP20BTECH11018

Download latex-tikz codes from

<https://github.com/RaghavJuyal/AI1103/tree/main/Assignment5/Assignment5.tex>

QUESTION 113, CSIR UGC NET EXAM (DEC 2014)

Let X_1, X_2, \dots, X_n be independent and identically distributed Bernoulli(θ), where $0 < \theta < 1$ and $n > 1$. Let the prior density of θ be proportional to $\frac{1}{\sqrt{\theta(1-\theta)}}$, $0 < \theta < 1$. Define $S = \sum_{i=1}^n X_i$.

Then valid statements among the following are:

1. The posterior mean of θ does not exist;
2. The posterior mean of θ exists;
3. The posterior mean of θ exists and it is larger than the maximum likelihood estimator for all values of S .
4. The posterior mean of θ exists and it is larger than the maximum likelihood estimator for some values of S .

SOLUTION

Definition 1. Posterior mean is the mean of the posterior distribution of θ , i.e., $E(\theta|X) = \int \theta f(\theta|X) d\theta$

Let $f(\theta)$ be the prior density of θ and $f(X|\theta)$ be the likelihood function.

$$f(\theta) \propto \frac{1}{\sqrt{\theta(1-\theta)}} \quad (0.0.1)$$

$$\begin{aligned} f(X|\theta) &= \prod_{i=1}^n \theta^{X_i} (1-\theta)^{1-X_i} \\ &= \theta^S (1-\theta)^{n-S} \end{aligned} \quad (0.0.2)$$

Let MLE be the maximum likelihood estimator.

Definition 2. The maximum likelihood estimator is the value which maximizes the likelihood function, i.e., $MLE = \arg \max(f(X|\theta))$.

$$\ln f(X|\theta) = S \ln \theta + (n-S) \ln (1-\theta) \quad (0.0.3)$$

$$\frac{\partial \ln f(X|\theta)}{\partial \theta} = \frac{S}{\theta} + \frac{S-n}{1-\theta} = 0$$

$$\therefore MLE = \frac{S}{n} \quad (0.0.4)$$

$$\begin{aligned} f(\theta|X) &\propto f(X|\theta) f(\theta) \\ &\propto \theta^{S-\frac{1}{2}} (1-\theta)^{n-S-\frac{1}{2}} \end{aligned} \quad (0.0.5)$$

where $f(\theta|X)$ is the posterior density of θ .

$$\begin{aligned} \int_0^1 f(\theta|X) d\theta &= 1 \\ \therefore f(\theta|X) &= \frac{\theta^{S-\frac{1}{2}} (1-\theta)^{n-S-\frac{1}{2}}}{B(S+\frac{1}{2}, n-S+\frac{1}{2})} \end{aligned} \quad (0.0.6)$$

where $B(x, y)$ is the beta function. From definition of beta function we get

$$\begin{aligned} B(x, y) &= \int_0^1 t^{x-1} (1-t)^{y-1} dt \\ &= \frac{x+y}{xy} \times \frac{1}{x+y C_x} \end{aligned} \quad (0.0.7)$$

Let posterior mean be $E(\theta|X)$

$$\begin{aligned} E(\theta|X) &= \int_0^1 \theta f(\theta|X) d\theta \\ &= \int_0^1 \frac{\theta^{S+\frac{1}{2}} (1-\theta)^{n-S-\frac{1}{2}}}{B(S+\frac{1}{2}, n-S+\frac{1}{2})} d\theta \\ &= \frac{B(S+\frac{3}{2}, n-S+\frac{1}{2})}{B(S+\frac{1}{2}, n-S+\frac{1}{2})} \end{aligned} \quad (0.0.8)$$

Using (0.0.7) in (0.0.8) we get

$$E(\theta) = \frac{S+\frac{1}{2}}{n+1} \quad (0.0.9)$$

Since $n > 1$, $E(\theta|X)$ exists.

For $E(\theta|X)$ to be greater than MLE,

$$\frac{S + \frac{1}{2}}{n + 1} > \frac{S}{n}$$

$$n > 2S(S > 0) \text{ or } n < 2S(S < 0) \quad (0.0.10)$$

Since $n > 1$, for $E(\theta|X) > \text{MLE}$ we get $n > 2S$.

\therefore Option 2 and 4 are correct.