CSIR UGC NET EXAM (Dec 2014), Question 113

Raghav Juyal

EP20BTECH11018

Some important concepts

1. The beta function, $B(x, y)$
2. Marginal distribution
3. Likelihood function and maximum likelihood estimator
4. Prior and posterior distributions
5. Posterior mean

The beta function, $B(x, y)$

The beta function, $B(x, y)$, is defined by the integral

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

where $Re(x) > 0$ and $Re(y) > 0$.

Marginal distribution

The marginal distribution of x is given by

$$f(x) = \int_y f(x, y) \, dy \tag{2}$$

Likelihood function

The likelihood function, represented by $f(X|\theta)$, refers to the joint probability of the data in the case of discrete distributions and joint probability density of the data in the case of continuous distributions, i.e.,

$$f(X|\theta) = \prod_{i=1}^{n} f(X_i|\theta) \tag{3}$$

Maximum Likelihood Estimator

The maximum likelihood estimator, represented by MLE, is the value which maximizes the likelihood function, i.e.,

$$\text{MLE} = \arg_{\theta} \max(f(X|\theta)) \tag{4}$$

Prior and Posterior Distribution

Prior distribution, where $f(\theta)$ is the prior density of θ, is the probability distribution that expresses established beliefs about an event before new evidence is taken into account.

When the new evidence is used to create a new distribution, that new distribution is called **posterior distribution**, where $f(\theta|X)$ is the posterior density of θ.From conditional probability we get,

$$f(\theta|X) = \frac{f(X,\theta)}{f(X)} \tag{5}$$

using (2) we get

$$f(\theta|X) = \frac{f(X,\theta)}{\int f(X,\theta)\, d\theta} \tag{6}$$

$$= \frac{f(X|\theta)\, f(\theta)}{\int f(X|\theta)\, f(\theta)\, d\theta} \tag{7}$$

Posterior mean

Posterior mean is the mean of the posterior distribution of θ, i.e.,

$$E(\theta|X) = \int \theta \, f(\theta|X) \, d\theta \tag{8}$$

Question

CSIR UGC NET EXAM (Dec 2014), Question 113

Let $X_1, X_2, ..., X_n$ 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

Let $f(\theta)$ be the prior density of θ.

$$f(\theta) \propto \frac{1}{\sqrt{\theta(1-\theta)}}$$
$$\implies f(\theta) = \frac{K}{\sqrt{\theta(1-\theta)}} \tag{9}$$

where K is the proportionality constant.

$$\int_0^1 f(\theta)\, d\theta = 1$$
$$\implies K \int_0^1 \frac{1}{\sqrt{\theta(1-\theta)}}\, d\theta = 1 \tag{10}$$

Solution Contd.

From (1) we get,

$$K \times B\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\implies K = \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \tag{11}$$

$$\therefore f(\theta) = \frac{\theta^{-\frac{1}{2}}\,(1-\theta)^{-\frac{1}{2}}}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \tag{12}$$

Solution Contd.

From (3) we get,

$$
\begin{aligned}
f\left(X|\theta\right) &= \prod_{i=1}^{n} \theta^{X_i} \left(1-\theta\right)^{1-X_i} \\
&= \theta^{S} \left(1-\theta\right)^{n-S} \qquad (13)
\end{aligned}
$$

Using log of likelihood function and differentiating we get,

$$
\begin{aligned}
\ln\left(f\left(X|\theta\right)\right) &= S\ln\left(\theta\right) + \left(n-S\right)\ln\left(1-\theta\right) \qquad (14) \\
\frac{\partial \ln\left(f\left(X|\theta\right)\right)}{\partial\theta} &= \frac{S}{\theta} + \frac{S-n}{1-\theta} = 0 \\
\therefore \text{MLE} &= \frac{S}{n} \qquad (15)
\end{aligned}
$$

Solution Contd.

Using (12) and (13) in (7) we get,

$$f(\theta|X) = \frac{\theta^{S-\frac{1}{2}}\ (1-\theta)^{n-S-\frac{1}{2}}}{\int_0^1 \theta^{S-\frac{1}{2}}\ (1-\theta)^{n-S-\frac{1}{2}}\ d\theta} \tag{16}$$

Using (1) we get,

$$f(\theta|X) = \frac{\theta^{S-\frac{1}{2}}\ (1-\theta)^{n-S-\frac{1}{2}}}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)} \tag{17}$$

Solution Contd.

From (8) we get,

$$
\begin{aligned}
E\left(\theta|X\right) &= \int_0^1 \theta\ f\left(\theta|X\right)\, d\theta \\
&= \int_0^1 \frac{\theta^{S+\frac{1}{2}}\ \left(1-\theta\right)^{n-S-\frac{1}{2}}}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)}\, d\theta \\
&= \frac{B\left(S+\frac{3}{2}, n-S+\frac{1}{2}\right)}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)} \qquad (18)
\end{aligned}
$$

Using (1) we get

$$
E\left(\theta|X\right) = \frac{S+\frac{1}{2}}{n+1} \qquad (19)
$$

Solution Contd.

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

$$\frac{S+\frac{1}{2}}{n+1} > \frac{S}{n}$$
$$\therefore n > 2S \tag{20}$$

Solution Contd.

We see that

$$E\left(\theta|X\right) = \frac{S + \frac{1}{2}}{n+1}$$

from (19) and

$$n > 1$$

given in the question $\implies E\left(\theta|X\right)$ exists.
From (20) we see that $E\left(\theta|X\right) >$ MLE for some values of S.
$\therefore$ Options 2 and 4 are correct.