B.SC.(H) STATISTICS
SEM. – IV
PAPER- STATISTICAL INFERENCE
UPC - 32371401
TOPIC- BAYESIAN ESTIMATION

The following is a general setup for a statistical inference problem: There is an unknown parameter that we would like to estimate. We get some data From the population, and we estimate the desired parameter. In the previous study, we discussed unknown parameter θ is assumed to be a fixed (non-random) quantity that is to be estimated by the observed data.

Here, we would like to discuss a different framework for inference, namely the Bayesian approach.

In the Bayesian framework, we treat the unknown parameter, θ, as a random variable. More specifically, we assume that we have some initial guess about the distribution of θ . This distribution is called the prior distribution. After observing some data, we update the distribution of θ (based on the observed data). This step is usually done using Bayes' Rule. That is why this approach is called the Bayesian approach. Here, to motivate the Bayesian approach, we will provide two examples of statistical problems that might be solved using the Bayesian approach.

Example:

Suppose that you would like to estimate the portion of voters in your town that plan to vote for Party A in an upcoming election. To do so, you take a random sample of size n from the likely voters in the town. Since you have a limited amount of time and resources, your sample is relatively small. Specifically, suppose that n=20. After doing your sampling, you find out that 6 people in your sample say they will vote for Party A.

Solution :

Let θ be the true portion of voters in your town who plan to vote for Party A. You might want to estimate θ as

$$\hat{\theta}=\frac{6}{20}=0.3$$

In fact, in absence of any other data, that seems to be a reasonable estimate. However, you might feel that n=20 is too small. Thus, your guess is that the error in your estimation might be too high. While thinking about this problem, you remember that the data from the previous election is available to you. You look at that data and find out that, in the previous election, 40% of the people in your town voted for Party A.

How can you use this data to possibly improve your estimate of θ? You might argue as follows:

Although the portion of votes for Party A changes from one election to another, the change is not usually very drastic. Therefore, given that in the previous election 40% of the voters voted for Party A, you might want to model the portion of votes for Party A in the next election as a random variable θ with a probability density function, $f(\theta)$, that is mostly concentrated around $\theta=0.4$.

For example, you might want to choose the density such that

$$E[\theta]=0.4$$

Such a distribution shows your prior belief about θ in the absence of any additional data. That is, before taking your random sample of size $n=20$, this is your guess about the distribution of θ. Therefore, you initially have the prior distribution $f(\theta)$.

Then you collect some data, shown by D. More specifically, here your data is a random sample of size n=20 voters, 6 of whom are voting for Party A. As we will discuss in more detail, you can then proceed to find an updated distribution for θ , called the posterior distribution. using Bayes' rule:

$$f(\theta|D)=\frac{P(D|\theta)f(\theta)}{P(D)}.$$

We can now use the posterior density, $f(\theta|D)$, to further draw inferences about θ .

Bayesian Statistical Inference

The goal is to draw inferences about an unknown variable X by observing a related random variable Y. The unknown variable is modeled as a random variable X, with prior distribution

f(x),if X is continuous, P(x), if X is discrete.

After observing the value of the random variable Y, we find the posterior distribution of X.

This is the conditional PDF (or PMF) of X given Y=y, f(x|y) or P(x|y).

The posterior distribution is usually found using Bayes' formula. Using the posterior distribution, we can then find point or interval estimates of X. The above equation, as we have seen before, is just one way of writing Bayes' rule. If either X or Y are continuous random variables, we can replace the corresponding PMF with PDF in the above formula.

Prior and Posterior

Let X be the random variable whose value we try to estimate. Let Y be the observed random variable. That is, we have observed Y=y, and we would like to estimate X. Assuming both X and Y are discrete, we can write

$$P(X=x|Y=y) = \frac{P(X=x,Y=y)}{P(Y=y)}$$

$$= \frac{P(Y=y|X=x)P(X=x)}{P(Y=y)}.$$

Using our notation for PMF and conditional PMF, the above equation can be rewritten as

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}.$$

For example, if X is a continuous random variable, while Y is discrete we can write

$$f(x|y) = \frac{P(y|x)f(x)}{P(y)}.$$

To find the denominator (P(y) or f(y)), we often use the law of total probability.

Example :

Let $X \sim$ Uniform(0,1).

Suppose that we know $Y|X = x \sim$ Geometric(x).

Find the posterior density of X given Y=2, f(x|2).

Solution: Using Bayes' rule we have

$$f(x|2) = \frac{P(2|x)f(x)}{P(2)}.$$

We know $Y|X = x \sim$ Geometric(x), so

$$P(y|x) = \frac{x(1-x)}{y-1}, \text{ for } y=1,2,\cdots.$$

Therefore, $P(2|x) = x(1-x)$.

To find $P_Y(2)$, we can use the law of total probability

$$P_Y(2) = \int_{-\infty}^{\infty} P(2|x)f(x)dx$$

$$= \int_0^1 x(1-x)\cdot 1dx = \frac{1}{6}.$$

Therefore, we obtain

$$f(x|2) = \frac{x(1-x)\cdot 1}{\frac{1}{6}}$$

$$= 6x(1-x), \text{ for } 0 \le x \le 1.$$

The main problem of Bayesian estimation is that of combining prior feelings about a parameter with direct sample evidence, and this is accomplished by determining $\phi(\theta|x)$, the conditional density of Θ given $X = x$. In contrast to the prior distribution of Θ, this conditional distribution (which also reflects the direct sample evidence) is called the posterior distribution of Θ.

In general, if $h(\theta)$ is the value of the prior distribution of Θ at θ and we want to combine the information that it conveys with direct sample evidence about Θ, for instance

The value of a statistic

$W = u(X_1, X_2, \ldots, X_n)$, we determine the posterior distribution of Θ by means of the formula

$$\phi(\theta|w) = f(\theta, w)\, g(w)$$

$$\phi(\theta|w) = h(\theta) \cdot f(w|\theta)\, g(w)$$

Here $f(w|\theta)$ is the value of the sampling distribution of W given $\Theta = \theta$ at w, $f(\theta, w)$ is the value of the joint distribution of Θ and W at θ and w, and g(w) is the value of the marginal distribution of W at w. Note that the preceding formula for $\phi(\theta|w)$ is, in fact, an extension of Bayes' theorem to the continuous case. Hence, the term "Bayesian estimation."

Once the posterior distribution of a parameter has been obtained, it can be used to make estimates.

Example: If $X_1, X_2, \ldots, X_n$ is a random sample from exponential distribution with pdf $f(x,\theta) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right)$ and the prior distribution of Θ is a uniform distribution with the pdf $\frac{1}{\theta}$, then find the posterior distribution of Θ given $X = x$. And find Bayes estimator for θ also.

For Θ = θ we have

$$f(x|\theta) = \frac{1}{\theta} e^{\left(-\frac{x}{\theta}\right)}, \ 0 < x < \infty.$$

$$h(\theta) = \frac{1}{\theta}, \ \theta > 0$$

Then the joint pdf of $X_1, X_2, \ldots, X_n$ and θ

$$f(\theta, x) = \prod_{i=1}^{n} [f(x|\theta)] . h(\theta)$$

$$f(\theta, x) = \prod_{i=1}^{n} \left[\frac{1}{\theta} e^{\left(-\frac{x}{\theta}\right)}\right] . \frac{1}{\theta}$$

$$f(\theta, x) = \left(\frac{1}{\theta}\right)^n e^{\left(-\frac{\sum x_i}{\theta}\right)} . \frac{1}{\theta}$$

$$g(x) = \int_0^\infty \frac{1}{\theta^n} e^{-\frac{\sum x_i}{\theta}} . \frac{1}{\theta} d\theta$$

$$= \int_0^\infty \frac{1}{\theta^{n+1}} . e^{-\frac{\sum x_i}{\theta}} d\theta$$

$$= \frac{\overline{|n}}{(\sum x_i)^n} \qquad \text{[by Gamma function]}$$

Then posterior distribution of θ

$$\phi(\theta|x) = \frac{f(\theta, x)}{g(x)} = \frac{\frac{1}{\theta^{n+1}} . e^{-\frac{\sum x_i}{\theta}}}{\frac{\overline{|n}}{(\sum x_i)^n}}$$

$$\phi(\theta|x) = \frac{(n\bar{x})^n}{\overline{|n}} \cdot \frac{1}{\theta^{n+1}} e^{-\frac{n\bar{x}}{\theta}}$$

$$\theta > 0$$

The above distribution is inverted gamma distribution

And the Bayes estimator for θ

$$E(\Theta|x) = \int_0^\infty \theta . \frac{(n\bar{x})^n}{\overline{|n}} \cdot \frac{1}{\theta^{n+1}} e^{-\frac{n\bar{x}}{\theta}} \, d\theta$$

$$= \frac{(n\bar{x})^n}{\overline{|n}} \int_0^\infty \frac{1}{\theta^n} e^{-\frac{n\bar{x}}{\theta}} \, d\theta$$

$$E(\Theta|x) = \frac{(n\bar{x})^n}{\overline{|n}} \quad \frac{\overline{|n-1}}{(n\bar{x})^{n-1}}$$

$$= \frac{n\bar{x}}{(n-1)}$$

is a value of an estimator of θ that minimize the Bayes risk when the loss function is quadratic.

<u>Question</u>: If $X_1, X_2, \ldots, X_n$ is a random sample from poisson distribution with unknown parameter λ, the prior distribution of its parameter is a gamma distribution with parameter α and β. find the posterior distribution and Bayes estimator for θ.

<u>Solution</u>: $E(\lambda|x) = \frac{(n+\beta)}{(\alpha+n\bar{x})}$ (try to solve)

Question: If X is a binomial random variable and the prior distribution of Θ is a beta distribution with the parameters α and β, then the posterior distribution of Θ given X = x is a beta distribution with the parameters x+α and n−x+β.

Solution:

Chapter -10 POINT ESTIMATION, Page No.-309.[Reference book 1]

REFERENCE BOOKS

1.John E. Freund's Mathematical Statistics with Applications Seventh Edition Irwin Miller and Marylees Miller, Pearson Education.

2. Hogg , R.V. and Tenis , E.A. (1998): Probability and Statistical Inference, Sixth Edition, Pearson Education.

3. Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics, Second Edition, John Wiley.