6.8

Bayesian Estimation

$h(\theta) \rightarrow$ prior pdf.

If $h(\theta)$ is a constant, and thus θ has the uniform prior distribution, we say that the Bayesian has a non-informative prior.

If, in fact, some knowledge of θ exists in advance of experimentation, non-informative priors should be avoided if at all possible.

Consider a continuous random sample which is a good statistic, say, Y and for a parameter θ, the pdf of Y, say, $g(y;\theta)$, can be thought of as the conditional pdf of Y, given θ, that is $g(y;\theta) = g(y|\theta)$.

Thus, we get,

$$g(y|\theta)\, h(\theta) = k(y,\theta) \quad \text{--- (1)}$$

as the joint pdf of the statistic Y and the parameter.

Now, the marginal pdf of Y is

$$k_1(y) = \int_{-\infty}^{\infty} h(\theta)\, g(y|\theta)\, d\theta$$

$$\therefore \text{①} \Rightarrow \frac{k(y,\theta)}{k_1(y)} = \frac{g(y|\theta)\, h(\theta)}{k_1(y)} = k(\theta|y)$$

as the conditional pdf of the parameter, given that $Y=y$. This formula is essentially Bayes' theorem, and $k(\theta|y)$ is called the posterior pdf of θ, given that $Y=y$.

If we ~~pe~~ penalize by taking the square of the error between the guess, say, $w(y)$ and the real value of the parameter θ, we use the conditional mean,

$$w(y) = \int_{-\infty}^{\infty} \theta\, k(\theta|y)\, d\theta$$

as our Bayes estimate of θ.

Example :- 6.8.2 :-

The prior pdf of the parameter to be the beta pdf;

$$h(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, \theta^{\alpha-1} (1-\theta)^{\beta-1} \; ; \; 0 \leq \theta < 1$$

$\therefore$ The joint probability,

$$k(y,\theta) = g(y|\theta)\, h(\theta)$$

$$= {}^{n}C_{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \theta^{y+\alpha-1} (1-\theta)^{n-y+\beta-1}$$

for $y = 0, 1, 2, \ldots, n$
and $0 < \theta < 1$.

Now, $k_1(y) = \int_0^1 k(y,\theta)\, d\theta$

$$= {}^{n}C_{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \theta^{y+\alpha-1} (1-\theta)^{n-y+\beta-1} d\theta$$

$$= {}^{n}C_{y} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+y) \cdot \Gamma(n+\beta-y)}{\Gamma(n+\alpha+\beta)}$$

for $y = 0, 1, 2, \ldots, n$

$$\int_0^1 x^{m-1} (1-x)^{n-1} dx = B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

Now, $k(\theta|y) = \frac{k(y,\theta)}{k_1(y)}$

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

Example 6.8-1 :-

Here, the prior probabilities are:-

$P(\lambda=4) = 0.2$ and $P(\lambda=2) = 0.8$

The conditional probabilities are:-

$P(X=6 \mid \lambda=2) = \cancel{P(X\leq 6, \lambda=2) - P(X\leq 5, \lambda=2)}$

$= 0.995 - 0.983 = 0.012$

and, $P(X=6 \mid \lambda=4) = 0.889 - 0.785 = 0.104$

Here, $P(X=6) = P(\lambda=2) \cdot P(X=6 \mid \lambda=2) + P(\lambda=4)\, P(X=6 \mid \lambda=4)$

$= 0.8 \times 0.012 + 0.2 \times 0.104$

$= 0.0304$

$\therefore$ The posterior probabilities are,

$$P(\lambda=2 \mid X=6) = \frac{P(\lambda=2, X=6)}{P(X=6)} = \frac{P(\lambda=2)\, P(X=6 \mid \lambda=2)}{P(X=6)}$$

$$= \frac{0.8 \times .012}{0.0304} = 0.316$$

and, $$P(\lambda=4 \mid X=6) = \frac{P(\lambda=4, X=6)}{P(X=6)} = \frac{P(\lambda=4)\, P(X=6 \mid \lambda=4)}{P(X=6)}$$

$$= \frac{0.2 \times 0.104}{0.0304} = 0.684$$

$\boxed{P(A' \mid B) = 1 - P(A \mid B)}$