7.3 최대가능도추정법

정의 5 가능도함수

$x_1, x_2, \cdots, x_n$을 $f(x;\theta)$로부터의 관측값이라 할 때, 가능도함수(likelihood function) $L(\theta)$를

$$L(\theta; x_1, x_2, \cdots, x_n) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$$

으로 정의한다. 또한 로그가능도함수(log likelihood function) $l(\theta)$를

$$l(\theta) = \ln L(\theta)$$

으로 정의한다.

주어진 관측값 $x_1, x_2, \cdots, x_n$에 대해 최대가능도추정법은 가능도함수 $L(\theta)$ (또는 로그가능도함수 $l(\theta)$)를 최대로 하는 θ를 추정치로 정하는 방법이다. 이론 전개의 편의상 관측값 대신 확률표본 $X_1, X_2, \cdots, X_n$에 대해 위 과정을 적용하면 최대가능도추정량(Maximum Likelihood Estimator 또는 MLE)을 바로 얻게 된다.

예제 1 $X_1, X_2, \cdots, X_n$이 $B(1, \theta)$로부터의 확률표본일 때, 모수 θ에 대한 최대가능도추정량을 구하여라.

풀이 가능도함수

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

로부터

$$l(\theta) = \ln L(\theta) = \sum_{i=1}^{n} x_i \ln\theta + (n - \sum_{i=1}^{n} x_i) \ln(1-\theta)$$

이다. $l(\theta)$를 최대로 하는 θ는

$$l'(\theta) = \frac{1}{\theta}\sum x_i - \frac{n - \sum x_i}{1-\theta} \equiv 0$$

을 만족하는 근이 되므로

$$\hat{\theta} = \frac{1}{n}\sum X_i$$

임을 알 수 있다. ■

참고 이 과정에서 함수 $l(\theta)$는

$$l''(\theta) < 0$$

이므로 $l'(\theta) = 0$의 근이 최대가 된다.

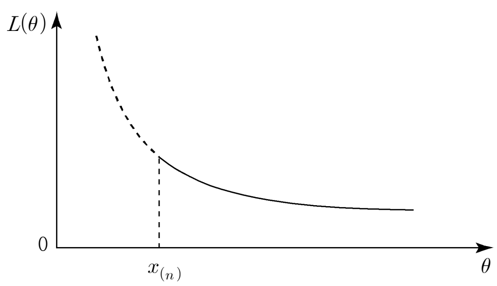
예제 2 $X_1, X_2, \cdots, X_n$이 $U(0, \theta)$로부터의 확률표본일 때, 모수 θ의 최대가능도추정량을 구하여라.

풀이 가능도함수는

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{[0,\theta]}(x_i)$$

$$= \left(\frac{1}{\theta}\right)^n I_{[0,\theta]}(x_{(n)}) I_{[0,\theta]}(x_{(1)})$$

이다. 이를 그림으로 나타내면



이다. 위 그림으로부터 $L(\theta)$를 최대로 하는 θ는 $x_{(n)}$임을 알 수 있다. 따라서

$$\hat{\theta} = X_{(n)}$$

이다. ■

예제 3 $X_1, X_2, \cdots, X_n$을 $N(\theta_1, \theta_2)$로부터의 확률표본이라 할 때, θ_1과 θ_2의 최대가능도추정량을 구하여라.

풀이 가능도함수

$$L(\theta_1, \theta_2) = \left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n \exp\left[-\frac{\sum_{i=1}^{n}(x_i - \theta_1)^2}{2\theta_2}\right]$$

로부터

$$l(\theta_1, \theta_2) = -\frac{n}{2}\ln(2\pi\theta_2) - \frac{\sum(x_i - \theta_1)^2}{2\theta_2}$$

이다. 이로부터

$$\frac{\partial l(\theta_1, \theta_2)}{\partial \theta_1} = \frac{\sum(x_i - \theta_1)}{\theta_2} \equiv 0$$

$$\frac{\partial l(\theta_1, \theta_2)}{\partial \theta_2} = -\frac{n}{2\theta_2} + \frac{\sum(x_i - \theta_1)^2}{2\theta_2^2} \equiv 0$$

을 풀면

$$\hat{\theta}_1 = \overline{X}, \quad \hat{\theta}_2 = S^2 = \frac{1}{n}\sum\left(X_i - \overline{X}\right)^2$$

임을 알 수 있다. ■