Bayesian Statistics - Point Estimation

모수 random variable $\ominus$가 주어진 상황에서 모수에 대한 예측값, 즉 decision function $\delta(x)$를 찾는 문제이다. 해는 $\delta(x) = \underset{\delta(x)}{\operatorname{argmin}} \int_{-\infty}^{\infty} L(\theta, \delta(x)) k(\theta|x) d\theta$. 여기서 $L(\cdot)$는 loss function, $k(\cdot)$는 posterior probability.

만약 $L(\theta, \delta(x)) = [\theta - \delta(x)]^2$로 주어진다면, $\delta(x) = E(\ominus|x)$로 정해진다. 즉, $\ominus$의 conditional distribution의 mean. 이 경우 $k(\theta|x) \propto g(x|\theta) h(\theta)$를 구한 다음, 여기서 mean을 구하면 된다.

Full Bayesian Learner

데이터 D로 X를 예측하고자하면,

$P(X|D) = \sum P(X|H_i, D) P(H_i|D) = \sum P(X|H_i) P(H_i|D)$.

MAP

모든 H_i를 보는 Full Bayesian Learner는 intractable하므로 $P(H_i|D)$가 최대인 H_i를 쓰기로 하자. 그러면 $P(X|D) \approx P(X|H_{MAP})$. 이 때, $P(H_i|D) = \dfrac{P(D|H_i) P(H_i)}{P(D)} \propto P(D|H_i) P(H_i)$.

ML

만약 모든 H_i가 equally probable 하다면 MAP에서 $P(D|H_i)$가 최대인 H_i를 찾으면 된다. 그 뒤, statistics라면 likelihood function $L(\theta|x) = \Pi f(x;\theta)$를 통해 (log취하고 미분하고...) θ를 결정.

References

1. Introduction to mathematical statistics by Hogg, McKean and Craig.

2. Artificial Intelligence: Modern Approach by S. J. Russel and P. Norvig.