6.4 Prediction of Subsequent Data

Zongyi Liu

2023-06-14

6.4 Prediction of Subsequent Data

The second goal of Bayesian inference is predicting subsequent data affter incorporating an observed set of data. Our predicted probability for the next value of y is just the probability of that value happenin for each value of θ, weighted by the posterior believability of each θ

$$p(y|D) = \int d\theta p(y|\theta)p(\theta|D) = \sum_{\theta} p(y|\theta)p(\theta|D)$$

where $p(\theta|D)$ in the first line is a probability density, and $p(\theta|D)$ in the second line is a proabability mass at discrete values of θ. In particular, for $y = 1$, the equation above becomes

$$p(y = 1|D) \approx \sum_{\theta} p(y = 1|\theta)p(\theta|D) = \sum_{\theta} p(\theta|D)$$

which is just the average value of θ in the posterior distribution of θ.