

MATH 559: BAYESIAN THEORY AND METHODS

DECISION THEORY: TWO SPECIAL CASES

The Bayesian decision for loss function $L(d(y), \theta)$ and posterior distribution $\pi_n(\theta)$ is obtained as

$$\min_d \int L(d(y), \theta) \pi_n(\theta) d\theta.$$

We examine the solution to two specific decision problems.

1. **Credible intervals:** Suppose $0 < \kappa < 1$, and consider the loss function for $\theta \in \mathbb{R}$ defined as

$$L(d, \theta) = \kappa(d - \theta)\mathbb{1}(\theta < d) + (1 - \kappa)(\theta - d)\mathbb{1}(\theta \geq d) = L_{1-\kappa}(d, \theta)$$

say. The expected posterior loss for possible decision t is then

$$\begin{aligned} \int L(t, \theta) \pi_n(\theta) d\theta &= \int \kappa(t - \theta)\mathbb{1}(\theta < t) \pi_n(\theta) d\theta + \int (1 - \kappa)(\theta - t)\mathbb{1}(\theta \geq t) \pi_n(\theta) d\theta \\ &= \kappa \int_{-\infty}^t (t - \theta) \pi_n(\theta) d\theta + (1 - \kappa) \int_t^{\infty} (\theta - t) \pi_n(\theta) d\theta \end{aligned}$$

Differentiating with respect to t we obtain by the chain rule and fundamental rule of calculus

- First term:

$$\kappa \left[\int_{-\infty}^t \pi_n(\theta) d\theta + t\pi_n(t) - t\pi_n(t) \right] = \kappa \int_{-\infty}^t \pi_n(\theta) d\theta$$

- Second term:

$$(1 - \kappa) \left[-t\pi_n(t) - \int_t^{\infty} \pi_n(\theta) d\theta + t\pi_n(t) \right] = -(1 - \kappa) \int_t^{\infty} \pi_n(\theta) d\theta$$

and so the minimizing value of t solves

$$\kappa \int_{-\infty}^t \pi_n(\theta) d\theta = (1 - \kappa) \int_t^{\infty} \pi_n(\theta) d\theta$$

so that

$$\kappa = \int_t^{\infty} \pi_n(\theta) d\theta.$$

Thus the solution to the loss minimization problem is the $1 - \kappa$ quantile of $\pi_n(\theta)$. Now consider the decision $d = (d_1, d_2)$ that comprises two components, with loss function

$$L((d_1, d_2), \theta) = L_{1-\kappa}(d_1, \theta) + L_{\kappa}(d_2, \theta).$$

By the above calculation, we obtain the solution

$$(d_1(y), d_2(y))$$

as the κ and $1 - \kappa$ posterior quantiles respectively, so that the interval between the two values contains posterior probability $1 - 2\kappa$, and is the $100(1 - 2\kappa)\%$ equal-tailed credible interval.

The other versions of the Bayesian credible interval can also be obtained as loss minimization decisions.

2. **Predictive Distribution:** Suppose our task is to estimate the unknown conditional density $f_Y(y; \theta_0)$. We aim to construct a Bayesian estimate of $f_Y(y; \theta_0)$ using the Kullback-Leibler loss, that is, we aim to minimize

$$\int \left\{ \int \log \frac{f_Y(y; \theta)}{f(y)} f_Y(y; \theta) dy \right\} \pi_n(\theta) d\theta$$

with respect to the function $f(\cdot)$, that is, to maximize

$$\iint \log f(y) f_Y(y; \theta) \pi_n(\theta) d\theta dy.$$

Re-writing the double integral

$$\int \log f(y) \left\{ \int f_Y(y; \theta) \pi_n(\theta) d\theta \right\} dy = \int \log f(y) p_n(y) dy$$

and maximizing this quantity is the same as minimizing

$$\int \log \frac{p_n(y)}{f(y)} p_n(y) dy.$$

Thus by properties of the KL divergence, the optimal f is $p_n(y)$

$$p_n(y) = \int f_Y(y; \theta) \pi_n(\theta) d\theta.$$

that is, the posterior predictive distribution.