1 Chapter 13 Combining Inference with Utility

Problem 13.3 *We build on our discussion of cost functions in continuous estimation in section 13.5.2. Suppose we use absolute error instead of squared error as the cost function:*

$$U(s, \hat{s}) = -|s - \hat{s}|.$$

Prove that the posterior **median** rather than the posterior mean maximizes expected utility.

Answer

So the gist is that we aim to get the estimate $\hat{s}$ that maximizes the expected utility:

$$\mathbb{E}[U(s, \hat{s})] = \mathbb{E}[-|s - \hat{s}|] = -\int |s - \hat{s}|\, p(s \mid x)\, ds.$$

Maximizing expected utility is equivalent to minimizing expected loss, so we need to minimize:

$$\int |s - \hat{s}|\, p(s \mid x)\, ds.$$

This is a classic result from Bayesian decision theory. The function

$$L(\hat{s}) = \int |s - \hat{s}|\, p(s \mid x)\, ds$$

is minimized when $\hat{s}$ is the **median** of the posterior distribution $p(s \mid x)$.

If we want to do a little proofing:
Let $F(\hat{s}) = \int_{-\infty}^{\hat{s}} p(s \mid x)\, ds$ be the cumulative distribution function of the posterior.
Taking derivative of $L(\hat{s})$ with respect to $\hat{s}$:

$$\frac{d}{d\hat{s}} L(\hat{s}) = \frac{d}{d\hat{s}} \int_{-\infty}^{\hat{s}} (\hat{s} - s) p(s \mid x) ds + \int_{\hat{s}}^{\infty} (s - \hat{s}) p(s \mid x) ds$$

Differentiate:

$$\frac{d}{d\hat{s}} L(\hat{s}) = \int_{-\infty}^{\hat{s}} p(s \mid x) ds - \int_{\hat{s}}^{\infty} p(s \mid x) ds = 2F(\hat{s}) - 1.$$

Set this derivative to 0:

$$2F(\hat{s}) - 1 = 0 \Rightarrow F(\hat{s}) = 0.5.$$

Thus, the minimum of the expected absolute error occurs at the median of the posterior distribution.
■