

## Dual roles of maximizing likelihood and Shannon entropy in Bayesian prediction

Toshio Ohnishi\*

Kyushu University, Fukuoka, Japan *ohnishi@econ.kyushu-u.ac.jp*

Takemi Yanagimoto

Chuo University, Tokyo, Japan *yanagmt@indsys.chuo-u.ac.jp*

The maximization of the likelihood and that of the Shannon entropy are the most famous principles in statistical inference. This paper reveals notable duality of these two important notions in the Bayesian prediction problems. We shed light on this duality through the dual Kullback-Leibler divergence losses. Under the  $e$ -divergence loss we find the following: 1) the minimization of the Bayesian risk is equivalent to the maximization of the Shannon entropy under a constraint, and 2) the maximization of the likelihood guarantees the minimum prediction in the sense that it derives the worst member of a class of nice predictors. An equality implying the balance of the log-likelihood ratio and the  $e$ -divergence plays an important role, which we call a saddlepoint equality. Dually, under the  $m$ -divergence loss the following findings are obtained: 1) the minimization of the Bayesian risk is equivalent to the maximization of the likelihood under a constraint, and 2) the maximization of the Shannon entropy guarantees the minimum prediction by deriving the worst member of a class of nice predictors. An equality showing a balance between the Shannon entropy difference and the  $m$ -divergence plays a key role.

Key Words: Kullback-Leibler divergence, Gateaux derivative, model averaging, saddlepoint equality.