Pitman Closeness Domination in Predictive Density Estimation for Two Ordered Normal Means Under α-Divergence Loss

Mejiro University Yuan-Tsung Chang
Keio University Nobuo Shinozaki
Rutgers University William, E. Strawderman

We consider Pitman closeness domination in predictive density estimation problems when the underlying loss metric is α-divergence $\{D(\alpha)\}$, a loss introduced by Csiszàr (1967) given by

$$D_\alpha\{\hat{p}(\tilde{y}|y), p(\tilde{y}|\psi)\} = \int f_\alpha\left(\frac{\hat{p}(\tilde{y}|y)}{p(\tilde{y}|\psi)}\right)p(\tilde{y}|\psi)d\tilde{y}, \tag{1}$$

where, for $-1 \le \alpha \le 1$

$$f_\alpha(z) = \begin{cases} \frac{6}{1-\alpha^2}(1 - z^{(1+\alpha)/2}), & |\alpha| < 1 \\ z\log z, & \alpha = 1 \\ -\log z, & \alpha = -1. \end{cases} \tag{2}$$

Here KL loss corresponds to $\alpha = -1$. The case $\alpha = 1$ is sometimes referred to as reverse KL loss.

If the true density function of Y is $N(\mu, \sigma^2)$ and the estimated predictive density of Y is $N(\hat{\mu}, \hat{\sigma}^2)$, Chang and Strawderman (2014) have derived the general form of D_α loss and have shown that it is a concave monotone function of quadratic loss and is also a function of the variances (observed, predicand, and plug-in). The general form is given as following:
a) for $-1 < \alpha < 1$,

$$D_\alpha(N(\tilde{y}|\hat{\mu}, \hat{\sigma}^2), N(\tilde{y}|\mu, \sigma^2)) = \frac{4}{1-\alpha^2}\left(1 - d(\sigma^2, \hat{\sigma}^2)e^{-A(\sigma^2,\hat{\sigma}^2)\frac{(\hat{\mu}-\mu)^2}{2}}\right), \tag{3}$$

where

$$d(\sigma^2, \hat{\sigma}^2) = \frac{\sigma^{(\alpha-1)/2}\tau}{\hat{\sigma}^{(\alpha+1)/2}}, \quad A(\sigma^2, \hat{\sigma}^2) = \left(\frac{1-\alpha}{2\sigma^2}\right)\left(1 - \frac{(1-\alpha)\tau^2}{2\sigma^2}\right) > 0, \quad \frac{1}{\tau^2} = \left(\frac{1+\alpha}{2\hat{\sigma}^2} + \frac{1-\alpha}{2\sigma^2}\right).$$

Further, $d(\sigma^2, \hat{\sigma}^2) < 1$ and $A(\sigma^2, \hat{\sigma}^2) > 0$.

b) for $\alpha = +1$, $D_{+1}(N(\tilde{y}|\hat{\mu}, \hat{\sigma}^2), N(\tilde{y}|\mu, \sigma^2)) = \frac{1}{2}\left[\left(\frac{\hat{\sigma}^2}{\sigma^2} - \log\frac{\hat{\sigma}^2}{\sigma^2} - 1\right) + \frac{(\hat{\mu}-\mu)^2}{\sigma^2}\right]. \quad (4)$

c) for $\alpha = -1$, $D_{-1}(N(\tilde{y}|\hat{\mu}, \hat{\sigma}^2), N(\tilde{y}|\mu, \sigma^2)) = \frac{1}{2}\left[\left(\frac{\sigma^2}{\hat{\sigma}^2} - \log\frac{\sigma^2}{\hat{\sigma}^2} - 1\right) + \frac{(\hat{\mu}-\mu)^2}{\hat{\sigma}^2}\right]. \quad (5)$

Note : The first part of the RHS of (4) and (5) is a form of Stein's loss for estimating variances and the second part is the squared error loss $(\hat{\mu} - \mu)^2$ divided by either the true or estimated variance. Also note that in each case, the $\{D(\alpha)\}$ loss is a concave monotone function of squared error loss $|\hat{\mu} - \mu|^2$ and is also a function of the variances.

The underlying distributions considered are normal, including the distribution of the observables, the distribution of the variable whose density is to be predicted, and the estimated predictive density which will be taken to be of the plug-in type. We demonstrate $\{D(\alpha)\}$ Pitman closeness domination of certain plug-in predictive densities over others for the entire class of metrics simultaneously when related Pitman's closeness domination holds in the problem of estimating the mean. We also consider $\{D(\alpha)\}$ Pitman domination of certain generalized Bayesian (best invariant) procedures suggested by Corcuera and Giummole (1999).

Examples of Pitman closeness domination presented relate to the problem of estimating the predictive density of the variable with the larger mean. More precisely, let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ be two independent random normal variables, where $\mu_1 \le \mu_2$. Under the above restriction we wish to predict a normal population with mean equal to the larger mean, μ_2, and variance equal to σ^2, $\tilde{Y} \sim N(\mu_2, \sigma^2)$. We consider different versions of this problem, depending on whether the $\sigma_i^2, i = 1, 2$ are known or are unknown but satisfy the additional order restriction, $\sigma_1^2 \le \sigma_2^2$. The case of two ordered normal means with known covariance matrix is also considered.