Statistics and Probability Letters 116 (2016) 101–106

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Pitman closeness properties of point estimators and predictive densities with parametric constraints

Takeru Matsuda[a,*], William E. Strawderman[b]

[a] Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan

[b] Department of Statistics and Biostatistics, Rutgers University, NJ, USA

ARTICLE INFO

Article history:
Received 26 February 2016
Received in revised form 23 April 2016
Accepted 23 April 2016
Available online 28 April 2016

MSC:
62F10
62F15
62C15

Keywords:
Point estimation
Predictive density estimation
Restricted parameters
Uniform priors
Bayes estimation
Pitman closeness

ABSTRACT

We contrast Pitman closeness and risk evaluations for Bayes procedures in point estimation and predictive density estimation problems when the mean of the underlying normal distribution is restricted to be nonnegative. Interesting reversals in preferences arise.

1. Introduction

Let $X \sim \mathrm{N}(\theta, \sigma^2)$ where $\theta \geq 0$ and σ^2 is known. We consider point estimation of θ under squared error loss as well as predictive density estimation of the density of $Y \sim \mathrm{N}(\theta, \tau^2)$ with τ^2 known under Kullback–Leibler loss and a related log loss. In each case, we consider the Pitman closeness comparison (Pitman, 1937) of Bayes estimators with respect to the uniform prior on the unrestricted parameter space $(-\infty, \infty)$ and the uniform prior on the restricted parameter space $[0, \infty)$.

For point estimation comparisons, we consider squared error loss

$$L_S(\theta, d) = (d - \theta)^2. \tag{1}$$

For predictive density estimation comparisons, we consider Kullback–Leibler loss and log loss. Kullback–Leibler loss for estimating $p(y \mid \theta)$ by $\hat{p}(y \mid x)$ is

$$L_{\mathrm{KL}}(p(y \mid \theta), \hat{p}(y \mid x)) = \int p(y \mid \theta) \log \frac{p(y \mid \theta)}{\hat{p}(y \mid x)} \mathrm{d}y. \tag{2}$$

* Corresponding author.
E-mail address: Takeru_Matsuda@mist.i.u-tokyo.ac.jp (T. Matsuda).

http://dx.doi.org/10.1016/j.spl.2016.04.021
0167-7152/

Note that the difference of Kullback–Leibler loss for $\hat{p}_1(y \mid x)$ and $\hat{p}_2(y \mid x)$ is

$$D(p(y \mid \theta), \hat{p}_1(y \mid x)) - D(p(y \mid \theta), \hat{p}_2(y \mid x)) = \int p(y \mid \theta)(-\log \hat{p}_1(y \mid x))\mathrm{d}y - \int p(y \mid \theta)(-\log \hat{p}_2(y \mid x))\mathrm{d}y.$$

Therefore, Kullback–Leibler loss is essentially the same as the expectation of $-\log \hat{p}(y \mid x)$ with respect to $y \sim p(y \mid \theta)$. Log loss (Grünwald and Dawid, 2004), which has been considered in different settings including data compression, is

$$L_{\log}(\hat{p}(y \mid x)) = -\log \hat{p}(y \mid x). \tag{3}$$

Whereas Kullback–Leibler loss concerns the quality of prediction averaged over y, log loss involves each realization of y individually.

For a given loss function $L(\theta, d)$, the risk of an estimator $\delta(x)$ is given by

$$R(\theta, \delta) = \mathrm{E}_\theta L(\theta, \delta(x)),$$

and we say that $\delta_1(x)$ dominates $\delta_2(x)$ in risk if

$$R(\theta, \delta_1) \le R(\theta, \delta_2)$$

for every θ, with strict inequality for at least one value of θ. We say that $\delta_1(x)$ dominates $\delta_2(x)$ in terms of Pitman closeness under the loss $L(\theta, d)$ if

$$\Pr_\theta[L(\theta, \delta_1) < L(\theta, \delta_2)] > \frac{1}{2}$$

for every θ (Pitman, 1937; Robert et al., 1993). Also, we say that $\delta_1(x)$ dominates $\delta_2(x)$ in terms of modified Pitman closeness under the loss $L(\theta, d)$ if

$$\Pr_\theta[L(\theta, \delta_1) < L(\theta, \delta_2)] \ge \frac{1}{2}$$

for every θ (Nayak, 1990; Khatree, 1992).

Our findings contrast markedly with results of risk comparisons in these settings. For point estimation with squared error loss, it is well known (Katz, 1961) that the Bayes estimator with respect to the restricted uniform prior (sometimes called the Katz estimator) dominates the Bayes estimator with respect to the unrestricted uniform prior. Our results for Pitman closeness comparison under squared error loss are just the opposite, i.e., the unrestricted Bayes estimator (equivalently the unrestricted maximum likelihood estimator) dominates the Katz estimator. For predictive density estimation under log loss, the orderings under risk and under (modified) Pitman closeness are in agreement, i.e., the restricted Bayesian predictive density dominates the unrestricted Bayesian predictive density. Under Kullback–Leibler loss, although the restricted Bayesian predictive density dominates in risk, it does not dominate in terms of Pitman closeness. We note that the results of risk comparison and Pitman closeness comparison become reversed also in variance estimation (Khatree, 1992; Biau and Yatracos, 2012).

We consider point estimation comparisons in Section 2 while comparisons for predictive densities are considered in Section 3. We give some concluding remarks in Section 4.

2. Point estimation

Suppose that we have an observation $X \sim \mathrm{N}(\theta, \sigma^2)$ and estimate the parameter θ. The parameter θ is constrained to $\theta \ge 0$. In the following, we assume $\sigma^2 = 1$ without loss of generality. The unrestricted maximum likelihood estimator is

$$\hat{\theta}_{\mathrm{I}}(X) = X,$$

which is also the Bayes estimator with respect to the uniform prior on $(-\infty, \infty)$. The Katz estimator (Katz, 1961) is

$$\hat{\theta}_{\mathrm{K}}(X) = X + \gamma(X) = X + \frac{\phi(X)}{\Phi(X)}, \tag{4}$$

which is the Bayes estimator with respect to the uniform prior restricted to $[0, \infty)$. Here, ϕ is the probability density function and Φ is the cumulative distribution function of the standard normal distribution, $\mathrm{N}(0, 1)$. Katz (1961) showed that the Katz estimator is minimax and admissible under quadratic loss. Hence the Katz estimator dominates the unrestricted Bayes estimator in risk. It is somewhat surprising that the comparison in terms of Pitman closeness favors the unrestricted Bayes estimator over the restricted Bayes estimator, as is shown in Theorem 1. We use the following lemma in the proof.

Lemma 1. *The function $\gamma(u)$ in (4) satisfies $-1 < \gamma'(u) < 0$. In particular, $\gamma(u)$ is monotone decreasing.*

Proof. See Sampford (1953). □

Theorem 1. *$\hat{\theta}_{\mathrm{I}}$ dominates $\hat{\theta}_{\mathrm{K}}$ in terms of Pitman closeness under quadratic loss.*