Group 5-3

Mathematical Algorithm for Brain and Neural System Analysis

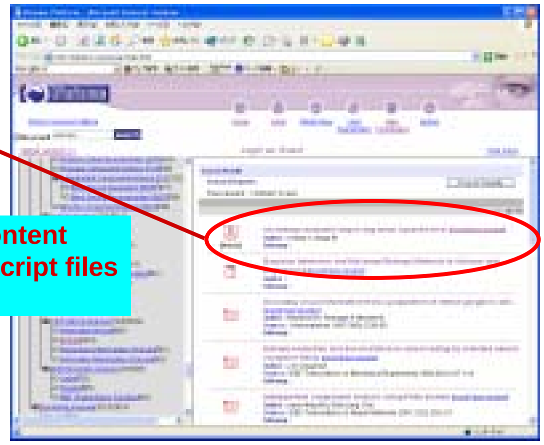
Yasunari YOKOTA
Gifu University

An entropy estimator improving mean squared error

Y. Yokota and M. Shiga, IEICE Trans, J86-A (2003) 936-944

This paper first derives an entropy estimator for a discrete random variable through Bayesian estimation of the occurrence probabilities of each value that the discrete random variable takes. It is shown that the derived entropy estimator achieves the least square error estimation. In the derived entropy estimator, if there is large number of sorts of values that the objective random variable takes, calculation for entropy estimation is very hard either analytically or by using numerical integration. Then, the present paper proposes a practical method for calculating entropy estimate in the derived entropy estimator; the practical method utilizes approximation by Taylor series of the entropy function. Numerical experiments demonstrate that the proposed entropy estimation method improves estimation precision of entropy in comparison to the typical conventional entropy estimation method.

Submitted content with Matlab script files (m-files)



Background

Suppose a discrete random variable X which takes M sorts of values $\boldsymbol{a}=(a_1,\ldots,a_M)$ with probabilities $\boldsymbol{r}=(r_1,\ldots,r_M)$, respectively.
Consider the case we try to estimate the entropy of such a random variable X using a sample set S of size N from the random variable X. Generally, the entropy of X has been estimated as follows:

(The conventional entropy estimator)

$H'_{\text{conv}}=\text{ent}(\boldsymbol{r}')$,
where
$\boldsymbol{r}'=\boldsymbol{n}/N=(n_1/N,\ldots,n_M/N)$, i.e., maximum likelihood estimate of $\boldsymbol{r}$,
$\boldsymbol{n}=(n_1,\ldots,{}_nM)$ means the numbers of each value $\boldsymbol{a}=(a1,\ldots,aM)$ included in the sample set S, and
ent() represents so-called M-ary entropy function.

Though the conventional entropy estimator H'_{conv} is a consistent estimator, but it is unfortunately neither an unbiased estimator nor a least square error estimator.
In brain science, it has been tried to estimate how much information related to an external stimuli a neuron transmits to following neurons. In such an estimation, entropy, i.e., average information amount, must be estimated as precisely as possible because an improper estimation, such as over- or under estimate, causes leading inappropriate conclusions.

Formulation

The entropy estimator H'_{Bayes} in which probabilities $\boldsymbol{r}$ are Bayesian estimated, that is, $\boldsymbol{r}$ is estimated as posterior probability density distribution $p(\boldsymbol{r}|\boldsymbol{n})$ after $\boldsymbol{n}$ was observed, is represented by

(Entropy estimator by Bayesian approach)

$$H'_{\text{Bayes}}=\frac{(M-1+N)!}{n_1!\cdots n_M!}\int_{\boldsymbol{r}\in R_r}\text{ent}(\boldsymbol{r})r_1^{n_1}\cdots r_M^{r_M}\,d\boldsymbol{r},$$

$$\text{where}\ \ R_r=\left\{\boldsymbol{r}\,\middle|\,\sum_{i=1}^{M}r_i=1,r_i\geq 0\right\}.$$

It is shown that the entropy estimator H'_{Bayes} minimizes mean squared error $\mathrm{E}[(H'_{\text{Bayes}}-H)^2]$, i.e., expectation of squared difference in the estimate and its true value, in other word, the entropy estimator H'_{Bayes} is the least square error estimator of entropy.

However, actual calculation for the entropy estimator H'_{Bayes} needs hard work because of multiple integral with respect to $\boldsymbol{r}$, especially when M is large.

For practical use, we may utilize an approximation version of H'_{Bayes} in which a truncated Taylor series of M-ary entropy function is used instead of M-ary entropy function itself:

(Approximation version of the proposed entropy estimator H'_{Bayes})

$$H'_{\text{Bayes}}=\sum_{i=1}^{M}s_i$$
$$-\frac{1}{N+M}\sum_{i=1}^{M}(1+\log s_i)(n_i+1)$$
$$-\sum_{k=2}^{K}\frac{1}{k(k-1)}\sum_{j=0}^{k}\binom{k}{j}\frac{(N+M-1)!}{(N+M+j-1)!}(-1)^j\sum_{i=1}^{M}s_i^{-j+1}\frac{(n_i+j)!}{n_i!},$$

where s_i means center in Taylor series of M-ary entropy function, and K is the truncation order in Taylor series. Sufficient approximation accuracy can be obtained when s_i=0.501, a small real number over 1/2, and K=20.

Simulation

As an example, suppose a discrete random variable X with M=3, $\boldsymbol{r}$=(0.1, 0.3, 0.6). A sample set of size N is generated by the random variable X, and entropy is estimated both by H'_{Bayes} and H'_{conv}. This entropy estimation process is repeated by 10,000 trials, and their mean and S.D. are calculated. The mean and (mean ±S.D.) curves are shown in Fig.1. Figure 1 reveals that the proposed entropy estimator H'_{Bayes} is superior to the conventional entropy estimator H'_{conv} in the meaning of both less bias error and less mean squared error.

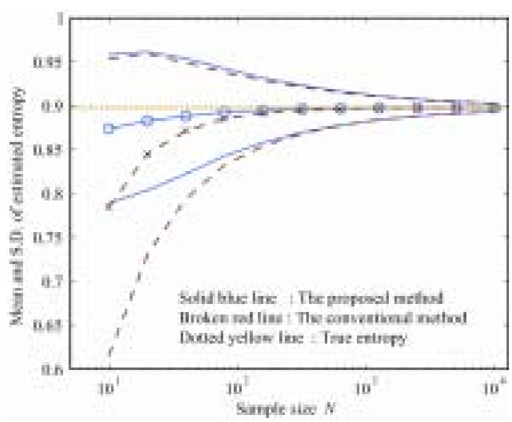


Fig.1 Comparison of entropy estimation precision by the proposed entropy estimator H'_{Bayes} and the conventional entropy estimator H'_{conv}.