韓國數學教育學會誌<數學教育>
1982. 6. Vol. XX, No. 3

Journal of the Korea Society of Mathematical Education
Jun. 1982. Vol. XX, No. 3

Jackknifing the density estimator and its property

By Lee Seung-Ho
Ajou University, Suweon, Korea

1. Introduction

The jackknife technique, since its introduction by Quenouille(1956) and Tukey(1958), it has been demonstrated that, as s rough-and-ready statistical tool for reducing bias and producing approximate confidence intervals, the jackknife can be beneficially applied in ratio problems,. in maximum likelihood estimation, and in transformations of statistics.

In 1971, Schucany, Gray and Owen generalized the jackknife technique to handle more general forms of bias.

Indiscriminate universal application of the jackknife can be hazardous.

This is illustrated in the case of interval estimation for a truncation point (Miller, 1964).

This paper investigates the effectiveness of the jackknife as tools for bias reduction, when applied to the density estimator, which is known to be asymptotically unbiased estimator.

2. Kernel-type density estimator and its jackknife

Let $X_1, X_2, \dots, X_n$ be *i.i.d.* random variables with continuous density function $f(x)$ and $f_n(x; X_1, \dots, X_n)$ be an estimator of $f(x)$.

To be meaningful, it is assumed that $f_n(x; X_1, \dots, X_n)$ is nonnegative and jointly Borel measurable in $(x; X_1, \dots, X_n)$.

It can be shown that such an $f_n(x; X_1, \dots, X_n)$ is not an unbiased estimator (Rosenblatt, 1956), and hence a natural candidate for jackknifing.

An obvious estimate of $f(x)$ is the difference quotient,

$$\hat{f}_n(x) = \frac{\#\text{sample points in } (x-b, x+b)}{2nb} = \frac{F_n(x+b) - F_n(x-b)}{2b}$$

where $F_n(\cdot)$ is the sample *c.d.f.* of $X_1, \dots, X_n$ and b is a positive constant.

Then, using Taylor expansion of $F(x)$,

$$E[\hat{f}_n(x)] = \frac{1}{2b}[F(x+b) - F(x-b)] = f(x) + \frac{f''(x)}{6} b^2 + 0(b^4)$$

$$\mathrm{Var}[\hat{f}_n(x)] = \frac{1}{4nb^2}[F(x+b) - F(x-b) + \{F(x+b) - F(x-b)\}^2]$$

and $$\mathrm{MSE}[\hat{f}_n(x)] = \mathrm{Var}[\hat{f}_n(x)] + \mathrm{Bias}^2[\hat{f}_n(x)] = \frac{f(x)}{2nb} + \frac{[f''(x)]^2}{36} b^4 + 0\left(\frac{1}{nb} + b^4\right)$$

Thus $\hat{f}_n(x)$ is a biased estimator with bias;

$$\mathrm{Bias}\,[\hat{f}_n(x)] = \frac{f''(x)}{6} b^2 + 0(b^4)$$

The $\hat{f}_n(x)$ is asymptotically unbiased in the sense that the bias approaches to 0, provided that $b \to 0$ as $n \to \infty$.

Moreover, $\hat{f}_n(x)$ is consistent estimator provided that $nb \to \infty$ as $n \to \infty$.

One criteria for choosing the constant b is to choose b so as to minimize the mean squared error of $\hat{f}_n(x)$.

The optimal choice of $\hat{b}$ is then

$$\hat{b}=\left\{\frac{9}{2}\ \frac{f(x)}{[f''(x)]^2}\right\}^{\frac{1}{5}} n^{-\frac{1}{5}}$$

and the bias of $\hat{f}_n(x)$ with b is

$$\mathrm{Bias}[\hat{f}_n(x)]=\frac{f''(x)}{6}\left\{\frac{9}{2}\ \frac{f(x)}{[f''(x)]^2}\right\}^{\frac{2}{5}} n^{-\frac{2}{5}}+0(b^4)$$

Now we investigate whether the jackknife would reduce the bias or not.

Let $\hat{f}^i_{n-1}(x)$ denote the same type estimate of $f(x)$ obtained by deleting the i-th member in the sample $(X_1, \ldots, X_n)$ and estimating $f(x)$ based on the remaining $(n-1)$ observations; *i.e.*,

$$\hat{f}^i_{n-1}(x)=\frac{F^i_{n-1}(x+b')-F^i_{n-1}(x-b')}{2b'}$$

where $F^i_{n-1}(\cdot)$ is the sample *c.d.f.* of $X_1, \ldots, X_{i-1}, \ X_{i+1}, \ldots, X_n$.

Form the new estimates, called "*pseudo-values*" by Tukey;

$$J_i(\hat{f}_n(x))=n\hat{f}_n(x)-(n-1)\hat{f}^i_{n-1}(x).$$

The jackknife estimate $J(\hat{f}_n(x))$ of $f(x)$ is the average of the pseudovalues, $J_i(\hat{f}_n(x))$; *i.e.*,

$$J(\hat{f}_n(x))=\frac{1}{n}\sum_{i=1}^{n}J_i(\hat{f}_n(x))=n\hat{f}_n(x)-(n-1)\overline{\hat{f}^i_{n-1}(x)}$$

where $$\overline{\hat{f}^i_{n-1}(x)}=\frac{1}{n}\sum_{i=1}^{n}\hat{f}^i_{n-1}(x).$$

The jackknife exactly eliminates a n^{-1} bias term, and thus Quenouille conceived the jackknife to achieve this reduction in bias.

By analogous derivation as in $\hat{f}_n(x)$ we obtain

$$\begin{aligned}E[J(\hat{f}_n(x))]&=n\ E[\hat{f}_n(x)]-(n-1)\frac{1}{n}\sum_{i=1}^{n}E\hat{f}^i_{n-1}(x)]\\&=f(x)+\frac{f''(x)}{3!}[nb^2-(n-1)b'^2]+\cdots\cdots\\&\quad+\frac{f^{2k}(x)}{(2k+1)!}[nb^{2k}-(n-1)b'^{2k}]+0(b^{(2k+2)}+b'^{(2k+2)})\end{aligned}$$

provided that $f(x)$ is $2k$ times differentiable.

Therefore, the ratios of bias compared with $\hat{f}_n(x)$ are

$$\begin{aligned}R(n)&=\frac{\mathrm{Bias}[J(f_n(x))]_{2k}}{\mathrm{Bias}[f_n(x)]_{2k}}=n-(n-1)\left(\frac{b'}{b}\right)^{2k}\\&=\left(\frac{b'}{b}\right)^{2k}-n\left\{\left(\frac{b'}{b}\right)^{2k}-1\right\}>1, \text{ since } n>1.\end{aligned}$$

This means that the jackknife increases the bias compared with the original estimator.

3. Monte Carlo study

A random sample of size 300 has been generated from the bimodal mixture;

$$f(x)=0.75\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{(x+1.5)^2}{2}\right]+0.25\frac{3}{\sqrt{2\pi}}\exp\left[-\frac{(x-1.5)^2}{2/9}\right]$$

Compute $\hat{f}_n(x)$ and $J(\hat{f}_n(x))$ at points $x=-5.0\ (0.2)+5.0$, for $b=1.0$ and 0.4 respectively, $b'=\sqrt{n/(n-1)}$ and plot $f(x)$ by "*", $\hat{f}_n(x)$ by "0" and $J(\hat{f}_n(x))$ by "J" respectively.

The results show that the jackknife over-estimates than the original estimator and increases mean squared errors.

References

(1) Miller, R.G. (1964). A trustworthy jackknife. *Ann. Math. Statist.* 35, 1594-1605.

(2) Quenouille, M.H. (1956). Notes on bias in estimation. *Biometrika* 43, 353-360.

(3) Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* 27, 832-837.

(4) Schucany, W.R., Gray, H.L. & Owen, D.B. (1971). On bias reduction in estimation. *J. Am. Statist. Assoc.* 66, 524-533.

(5) Tukey, J.W. (1958). Bias and confidence in not-quite large samples. *Ann. Math. Statist.* 29, 614.