Infinitely Divisible Approximations to I.I.D. Density Estimation Experiments

Harrison H. Zhou

Cornell University

June 21, 2003

Jointly with Michael Nussbaum

Overview

1. **Introduction**
   - **Le Cam distance**
   - **Two Pioneering Results**
2. **Infinitely Divisible Experiment**
3. **Main Results**
4. **Related Discussion**
5. **Conclusion**

1. Introduction

1.1 Introduction to Le Cam distance

Definition 1. An experiment $E = (\Omega, \mathcal{B}, (P_\theta, \theta \in \Theta))$ is an indexed set of Probability distributions $(P_\theta, \theta \in \Theta)$ on a measurable space $(\Omega, \mathcal{B})$.

Example. Suppose that we observe $x_1, x_2, \ldots, x_n$ i.i.d with density $f(x) \in \sum, x \in [0,1]$, then $E_n = \left([0,1]^n, \mathcal{B}^n_{[0,1]}, \left(P_f^{\otimes n}, f \in \sum\right)\right)$.

Definition 2. For experiments

$$E = (\Omega_1, \mathcal{B}_1, (P_\theta, \theta \in \Theta))$$
$$F = (\Omega_2, \mathcal{B}_2, (Q_\theta, \theta \in \Theta))$$

the deficiency is defined as

$$\delta(E, F) = \inf_K \sup_{\theta \in \Theta} \|KP_\theta - Q_\theta\|_{TV}$$

where the infimum is taken over the set of all Markov kernels from measures on $(\Omega_1, \mathcal{B}_1)$ to measures on $(\Omega_2, \mathcal{B}_2)$ (under weak assumptions).

The **Le Cam distance** is defined as

$$\Delta(E, F) = \max\{\delta(E, F),\ \delta(F, E)\}$$

Definition 3. Sequences E_n and F_n are **asymptotically equivalent** if $\Delta(E_n, F_n) \to 0$ as $n \to \infty$. The significance is

> **All asymptotically optimal statistical procedures can be carried over from one model to another**

The procedure can be estimation, testing or any asymptotic decision problem.

Definition 4. Sequences E_n and F_n are weakly asymptotically equivalent if for every finite set $S \in \Theta$, $\Delta\left(E_n^S, F_n^S\right) \to 0$, where E_n^S (F_n^S) denotes the subexperiment of E_n (or F_n) with restriction to the finite parameter set S.

1.2. Two Pioneering Results

Brown and Low (1996)

Regression experiment

$$E_n : y_i = f\left(\frac{i}{n}\right) + n^{-1/2} z_i,\ z_i \overset{\text{i.i.d.}}{\sim} N(0,1),\ i = 1, \ldots, n$$

Gaussian white noise experiment

$$F_n : dy(t) = f(t)\, dt + n^{-1/2} dW(t)$$

For given $\alpha > 1/2$ and $M > 0$ the parameter space $\sum$ is

$$\{f : |f(x) - f(y)| \le M\,|x - y|^{\alpha},\ x \in [0,1]\}$$

we have

$$\Delta(E_n, F_n) \to 0$$

Nussbaum (1996)

Density estimation experiment

$E_n: \quad x_1, x_2, ..., x_n$ i.i.d with density f

Gaussian white noise experiment

$$F_n : dy(t) = \sqrt{f(t)}dt + \frac{1}{2}n^{-1/2}dW(t)$$

For given $\alpha > 1/2$, $M > 0$ and $\epsilon > 0$ the parameter space $\sum$ is

$$\left\{ \begin{array}{c} f : |f(x) - f(y)| \leq M |x - y|^{\alpha}, \\ f \geq \epsilon, \ \int_0^1 f = 1, \ x \in [0,1] \end{array} \right\}$$

we have

$$\Delta(E_n, F_n) \to 0$$

Infinitely Divisible Experiments

Definition 5. The experiment $E = \{P_\theta, \theta \in \Theta\}$ indexed by Θ is called infinitely divisible if for any n there exists E_n such that E is equivalent to the direct product of n copies of E_n.

Proposition 1 (Le Cam (1986)).

(I) E is infinitely divisible iff it is a weak limit of Poisson experiments

(II) A pairwise imperfect experiment E is infinitely divisible iff it is a direct product of a Gaussian shift experiment and a Poisson experiment.

Density estimation experiment

$E_n: \quad x_1, x_2, ..., x_n$ i.i.d with density f

Poisson experiment

$F_n: \; x_1, x_2, ..., x_N$ i.i.d with density f, $N \sim Po(n)$

Proposition 2 (Le Cam (1986)).

The i.i.d density estimation experiment E_n and its accompanying Poisson experiment F_n are weakly asymptotically equivalent.

Main Results

E_n : $x_1, x_2, ..., x_n$ i.i.d with density f

F_n : $x_1, x_2, ..., x_N$ i.i.d with density f, $N \sim Po(n)$

Theorem 1 If we have (I) (Le Cam (1974))

$$\inf_{\hat{f}} \sup_{f \in \sum} E_f H^2 \left(\hat{f}, f\right) = o\left(n^{-1/2}\right)$$

or (II) (Low, Nussbaum and van de Geer(unpublished m.s.)) $\sum$ is compact in $C\,[0,1]$, then

$$\Delta\left(E_n, F_n\right) \to 0$$

For a given $\epsilon > 0$, $\alpha > 1/2, b > a$, we define the parameter space $\sum$ as follows

$$\left\{ \begin{array}{c} f(\cdot - \theta) : |f(x) - f(y)| \leq M |x-y|^{\alpha}, \; f \geq \epsilon \\ \int_0^1 f = 1, \; x \in [0,1], \; \theta \in [a,b] \end{array} \right\}$$

Theorem 2.

$$\begin{array}{lcl} E_{0,n} & : & x_1, x_2, ..., x_n \text{ i.i.d. with density } f(\cdot - \theta) \\ F_{0,n}^1 & : & X_n(\cdot), \text{ a Poisson process with intensity} \\ & & n \; \max\{f(0), f(1)\} \, I_{[\theta, \theta + 1/\sqrt{n}]} \end{array}$$

$$F_{0,n}^2 \; : \; dy(t) = \sqrt{f(t)} dt + \frac{1}{2} n^{-1/2} dW(t)$$

then $\Delta \left(E_{0,n}, F_{0,n}^1 \otimes F_{0,n}^2 \right) \to 0$

For a given $\epsilon > 0$, $\alpha > (1+k)/2$, $b > a, 1 > k > 0$, we define the parameter space $\sum$ as follows

$$\left\{ \begin{array}{c} f(\cdot - \theta) : f(x) = x^k (1-x)^k h(x), \\ |h(x) - h(y)| \le M |x-y|^\alpha, h(0) = h(1), \\ h \ge \epsilon, \ \int_0^1 h = 1, \ x \in [0,1], \ \theta \in [a,b] \end{array} \right\}$$

Theorem 3.

$$\begin{array}{lcl} E_{0,n} & : & x_1, x_2, ..., x_n \text{ i.i.d. with density } f(\cdot - \theta) \\ F^1_{0,n} & : & X_n(\cdot), \text{ a Poisson process with intensity} \\ & & n(x-\theta)^k \left(\theta + 1/\sqrt{n} - x\right)^k h(0) I_{[\theta, \theta + 1/\sqrt{n}]} \\ F^2_{0,n} & : & dy(t) = \sqrt{f(t)} dt + \frac{1}{2} n^{-1/2} dW(t) \end{array}$$

then $\Delta\left(E_{0,n}, F^1_{0,n} \otimes F^2_{0,n}\right) \to 0.$

Related Discussion

- Poisson type local limits: Ibragimov and Khasminskii (1981), Pflug (1983), Janssen, Milbrodt and Strasser (1985) etc.

- Nonparametric edge and endpoint estimation: Korostelev and Tsybakov (1993), Härdle, Park, Tsybakov (1995), Hall, Nussbaum, and Stern (1997), Donoho (1999), etc.

- Local limit experiments for the statistics of extreme values: Janssen and Marohn (1994), Drees (2001) etc.

- Quantile and weighted empirical processes: Csörgö and Mason (1989), Mason (2001)

Conclusion

We propose infinitely divisible approximations to i.i.d. experiments, within the framework of (nonparametric) asymptotic equivalence. The apparent connections to extreme value theory and edge estimation deserve further study.