GLOBAL PROPERTIES OF KERNEL ESTIMATORS FOR MIXING DENSITIES IN EXPONENTIAL FAMILY MODELS FOR DISCRETE VARIABLES

by

Wei-Liem Loh and Cun-Hui Zhang

Purdue University — Rutgers University

Technical Report #93-47

Department of Statistics
Purdue University

August 1993

*

GLOBAL PROPERTIES OF KERNEL ESTIMATORS FOR MIXING DENSITIES IN EXPONENTIAL FAMILY MODELS FOR DISCRETE VARIABLES

Wei-Liem Loh and Cun-Hui Zhang

Purdue University and Rutgers University

Abstract: This paper concerns the global performance of the kernel estimators considered in Zhang (1992) for a mixing density function g based on a sample from $f(x) = \int f(x|\theta)g(\theta)d\theta$ under L^p loss, where $1 \leq p \leq \infty$ and $f(x|\theta)$ is a known exponential family of density functions with respect to the counting measure on the set of nonnegative integers. Fourier methods are used to derive upper bounds for the rate of convergence of the kernel estimators and lower bounds for the optimal convergence rate over various smoothness classes of mixing density functions. In particular under mild conditions, it is shown that these estimators achieve the optimal rate of convergence for the negative binomial mixture and are almost optimal for the Poisson mixture. Global estimation of the mixing distribution function under L^p loss is also considered.

Key words and phrases: Mixing density, kernel estimator, discrete exponential family, rate of convergence.

1 Introduction.

Let $X_1, \ldots, X_n$ be independent observations from a mixture distribution with probability law

$$f(x;g) = \int_0^{\theta^*} f(x|\theta)g(\theta)d\theta, \tag{1}$$

where $f(x|\theta)$ is a known parametric family of probability density functions with respect to a σ-finite measure μ, and g is a mixing density function on $(0, \theta^*)$. Suppose

$$f(x|\theta) = C(\theta)q(x)\theta^x, \quad \forall x = 0, 1, 2, \ldots, \tag{2}$$

where $0 \leq \theta \leq$ (or $<$) $\theta^* \leq \infty$, $q(x) > 0$ whenever $x = 0, 1, 2, \ldots$ and μ is the counting measure on the set of nonnegative integers.

Zhang (1992) considered a class of kernel estimators for the mixing density function g, its derivatives, and the mixing distribution function, and proved that the mean squared error at a fixed point achieves or almost achieves the optimal rate of convergence under mild conditions. In this paper we consider the L^p properties of these estimators and the optimal rate of convergence under these global criteria. In particular, Sections 2 and 3 give upper bounds for the convergence rates of these estimators for the mixing density and distribution respectively for the case where θ^* is finite and known. Section 4 supplies corresponding lower bounds for the optimal convergence rate. A consequence of the results in Sections 2, 3, and 4 is that, under mild assumptions, the kernel mixing density estimators are optimal for a negative binomial mixture and are almost optimal (except for an iterated logarithmic factor) for a Poisson mixture [Corollaries 1 and 2, and Theorem 3].

Section 5 considers the case where θ^* is infinite. Upper bounds for the convergence rates are obtained for an improved version of the kernel estimator. Unfortunately the results are less satisfactory here compared with the lower bounds stated in Section 4; the upper and lower bounds differ by some power of a logarithmic factor even for well behaved weights.

A key point of this paper is that in general, without further assumptions, global nonparametric estimation of a mixing density (or distribution) of a discrete exponential family is difficult in that the optimal rate of convergence is logarithmic (not polynomial).

Among related mixture problems, the deconvolution problem appears to be the best understood. Recent and important advances to the solution were made by Carroll and Hall (1988), Fan (1991a), (1991b), Zhang (1990) and many others using Fourier analysis. In particular kernel estimators for the mixing density (or distribution) have been obtained which achieve the optimal convergence rate.

Another problem that has been of much interest is the estimation of the mixing distribution of a Poisson mixture. Tucker (1963) approached this problem through the method of moments, and Lambert and Tierney (1984) and Simar (1976) considered the nonparametric maximum likelihood estimation for the mixing distribution. Loh (1992) and Zhang (1992) have independently obtained results on this problem as well as for mixtures of discrete exponential distributions via Fourier analysis. Walter and Hamedani (1991) successfully applied orthogonal polynomial techniques to mixtures of exponential families. Rolph (1968), Meeden (1972), and Datta (1991) have also used Bayesian methods to construct consistent estimators for the mixing distribution.

Throughout this paper we shall denote by $P = P_g$ and $E = E_g$ the probability and expectation corresponding to g respectively, by $\mathcal{R}\{z\}$ the real part of a complex number z, by $h^{(j)}$ the jth derivative (if it exists) of any function h with $h^{(0)} = h$, by h^* the Fourier transformation of any integrable function h, so that $h^*(t) = \int e^{ity}h(y)dy$ whenever $\int |h(y)|dy < \infty$, and the L^p-norm of any measurable function h by

$$||h||_p = \begin{cases} (\int_{-\infty}^{\infty} |h(y)|^p dy)^{1/p} & \text{if } 1 \le p < \infty, \\ \text{ess sup}_{-\infty<y<\infty}|h(y)| & \text{if } p = \infty. \end{cases}$$

We shall use the notation κ' and κ'' to denote the decomposition $\kappa = \kappa' + \kappa''$ such that κ' is an integer and $0 < \kappa'' \le 1$ for all real numbers κ.

2 Kernel estimators.

In this and the next two sections, we shall assume that θ^* is finite and known. The case of $\theta^* = \infty$ is considered in Section 5.

Let $k : R \to R$ be a symmetric function satisfying

$$\int_{-\infty}^{\infty} k(y)dy = 1, \qquad k^*(t) = 0, \quad \forall |t| > 1,$$

$$\int_{-\infty}^{\infty} y^j k(y)dy = 0, \quad \forall 1 \le j < \alpha_0, \tag{3}$$

and

$$\int_{-\infty}^{\infty} |y^{\alpha_0} k(y)|\, dy < \infty, \tag{4}$$

for some positive number α_0. Define

$$K_n(x,\theta) = \frac{I\{0 \le x \le m_n\}}{2\pi q(x)x!} \int_{-c_n}^{c_n} \mathcal{R}\{(it)^x e^{-it\theta}\}k^*(t/c_n)dt, \tag{5}$$

where m_n and c_n are positive constants tending to ∞ and $I\{.\}$ denotes the indicator function.

Given any probability density function g on $(0, \theta^*)$, we shall extend its domain to the whole real line by setting $g(y) = g(y)I\{0 < y < \theta^*\}$ for all $y \in R$. Let

$$h(y) = C(y)g(y), \quad \forall -\infty < y < \infty. \tag{6}$$

It follows from (1) and (2) that $f(x;g)/q(x) = \int_0^{\theta^*} \theta^x h(\theta)d\theta$. By the Fourier inversion formula, we have

$$
\begin{aligned}
& \int_{-\infty}^{\infty} h(\theta - y/c_n)k(y)dy \\
= & \int_0^{\theta^*} c_n k(c_n(\theta - y))h(y)dy \\
= & \sum_{x=0}^{\infty}(2\pi x!)^{-1} \int_0^{\theta^*} y^x h(y)dy \int_{-c_n}^{c_n} \mathcal{R}\{(it)^x e^{-it\theta}\}k^*(t/c_n)dt.
\end{aligned}
$$

Thus as in Zhang (1992), we observe that $K_n(x,\theta)$ can be used as a kernel for h in the sense that for $-\infty < \theta < \infty$,

$$E_g K_n(X_1,\theta) - h(\theta) = b_{1n}(\theta) + b_{2n}(\theta) \to 0, \tag{7}$$

as $(m_n, c_n) \to (\infty, \infty)$ along a suitable path, where

$$b_{1n}(\theta) = \int_{-\infty}^{\infty} [h(\theta - y/c_n) - h(\theta)]k(y)dy \tag{8}$$

and

$$b_{2n}(\theta) = -\sum_{x>m_n} \frac{\int_0^{\theta^*} y^x h(y)dy}{2\pi x!} \int_{-c_n}^{c_n} \mathcal{R}\{(it)^x e^{-it\theta}\}k^*(t/c_n)dt. \tag{9}$$

With this as motivation, we estimate $g(\theta)$ by

$$\hat{g}_n(\theta) = n^{-1} \sum_{j=1}^{n} \{K_n(X_j,\theta)/C(\theta)\} I\{0 \le \theta \le a_n\}. \tag{10}$$

The constants m_n, c_n, and a_n are chosen such that

$$c_n + \max_{1 \le x \le m_n} \log(1/q(x)) \le \beta_0 \log n, \tag{11}$$

$$c_n \le (\theta^* e)^{-1}(m_n - \beta_1 \log c_n), \tag{12}$$

and

$$a_n = \begin{cases} \theta^* & \text{if } C(\theta^*) > 0, \\ \theta^* - a^*/c_n & \text{if } C(\theta^*) = 0, \end{cases} \tag{13}$$

with some constants $0 < \beta_0 < 1/2$, $\beta_1 > 0$, and $0 < a^* < \infty$.

We shall investigate the global performance of the estimators $\hat{g}_n$ with respect to the following classes of mixing density functions. Let $1 \le p \le \infty$ and w be a measurable function on $(0, \theta^*)$ with $||w||_p$ finite. For $\alpha > 0$ we define $\mathcal{G}_{\alpha,\theta^*} = \mathcal{G}_{\alpha,\theta^*}(p, w, M)$ to be the set of all probability density functions g on $(0, \theta^*)$ such that

$$||w(\theta)\{g^{(\alpha')}(\theta) - g^{(\alpha')}(\theta + \delta)\}||_p < M|\delta|^{\alpha''}, \quad \forall \delta, \tag{14}$$

where α' is the integer with $0 < \alpha'' = \alpha - \alpha' \le 1$, and M is a constant such that $\mathcal{G}_{\alpha,\theta^*}(p, w, M)$ is nonempty.

Assume that there exist constants $\gamma \ge 0$, C_1^*, C_2^*, and C_3^* such that

$$\sup_{0<\theta<\theta^*} (\theta^* - \theta)^\gamma / C(\theta) < C_1^*, \tag{15}$$

$$\sup_{0<\theta<\theta^*} (\theta^* - \theta)^j |C^{(j)}(\theta)| / \{C(\theta) j!\} < C_2^*, \quad \forall 0 \le j \le \rho', \tag{16}$$

and

$$|C^{(\rho')}(\theta + \delta) - C^{(\rho')}(\theta)| < C_3^* \delta^{\rho''}, \; 0 < \theta < \theta + \delta < \theta^*, \tag{17}$$

where ρ' is a nonnegative integer with $0 < \rho'' = \rho - \rho' \le 1$.

REMARK. If $C(\theta^*) > 0$, we shall set $\gamma = 0$ although (15) holds for all $\gamma \ge 0$ and $\rho > 0$. Conditions (16) and (17) are satisfied for every $\rho > 0$, if $C(\theta)$ is an analytic function in a neighborhood of θ^* (e.g. $C(\theta) = (1 - \theta)^\nu$ with $\theta^* = 1$ for the negative binomial family).

Throughout the paper, we use $\mathrm{Rem}_h(x, \delta, m)$ to denote the remainder of the $(m + 1)$-term Taylor expansion of h, which can be written as

$$\begin{aligned} \mathrm{Rem}_h(x, \delta, m) &= h(x + \delta) - \sum_{j=0}^{m} h^{(j)}(x)\delta^j / j! \\ &= \int_0^\delta \frac{(\delta - y)^{m-1}}{(m-1)!} \{h^{(m)}(x + y) - h^{(m)}(x)\} dy. \end{aligned} \tag{18}$$

Theorem 1 *Suppose $\alpha > 0$, $1 \le p \le \infty$, and that (15)-(17) hold with $\gamma \ge 0$ and $\rho = \alpha + \gamma$. Let $\hat{g}_n$ be given by (10) with the kernel $K_n(x, \theta)$ in (5) such that $\alpha_0 \ge \alpha + \gamma$ in (4). Let (11)-(13) hold with $\beta_1 \ge \alpha + \gamma$. Then*

$$\sup\{E_g ||w(\hat{g}_n - g)||_p : g \in \mathcal{G}_{\alpha,\theta^*}(p, w, M)\} = O(c_n^{-\alpha}) .$$

Corollary 1 *Suppose the conditions of Theorem 1 are satisfied and*

$$q(x) B_0 B^x (x!)^\beta \ge 1, \quad \forall x \ge 0, \tag{19}$$

for some constants B_0, B, and β. If (11) and (12) hold with equality, then

$$\sup_{g\in\mathcal{G}_{\alpha,\theta^*}} E_g\|w(\hat{g}_n-g)\|_p = \begin{cases} O(1)(1/\log n)^\alpha & \text{if } \beta=0, \\ O(1)(\log\log n/\log n)^\alpha & \text{if } 0<\beta<\infty. \end{cases}$$

REMARK. Corollary 1 applies to negative binomial and Poisson mixtures for which $\beta = 0$ and 1 respectively.

PROOF OF THEOREM 1. Let $\chi_{[0,a]}(\theta) = I\{0 \le \theta \le a\}$. Since $g^{(j)}(\theta-a) = 0$ for $0 \le \theta < a$, by (18) and (14)

$$\begin{aligned} &\|\chi_{[0,a]}wg^{(j)}\|_p \\ =\ & \left\|\chi_{[0,a]}(\theta)w(\theta)\mathrm{Rem}_{g^{(j)}}(\theta-a,a,\alpha'-j)\right\|_p \\ \le\ & \int_0^a \frac{(a-y)^{\alpha'-j-1}}{(\alpha'-j-1)!}\left\|w(\theta)\{g^{(\alpha')}(\theta-a+y)-g^{(\alpha')}(\theta-a)\}\right\|_p dy \\ \le\ & \int_0^a \frac{(a-y)^{\alpha'-j-1}}{(\alpha'-j-1)!}M\{|-a+y|^{\alpha''}+a^{\alpha''}\}dy, \end{aligned}$$

so that for all $a > 0$ and $0 \le j \le \alpha'$

$$\|\chi_{[0,a]}wg^{(j)}\|_p \le 2Ma^{\alpha-j}/(\alpha'-j)!\,. \tag{20}$$

Taking the expansion at $\theta + \theta^* - a$, we obtain in the same manner

$$\|(1-\chi_{[0,a]})wg^{(j)}\|_p \le 2M(\theta^*-a)^{\alpha-j}/(\alpha'-j)!\,,\ a\le\theta^*, 0\le j\le\alpha'. \tag{21}$$

Let $\chi_n(\theta) = I\{0 \le \theta \le a_n\}$ and $\delta_n = \theta^* - a_n$. By (6) and (7),

$$\begin{aligned} E_g\|w(\hat{g}_n-g)\|_p \ \le\ & E_g\|w(\hat{g}_n-E\hat{g}_n)\|_p + \|\chi_n wb_{1n}/C\|_p \\ & +\|\chi_n wb_{2n}/C\|_p + \|(1-\chi_n)wg\|_p\,. \end{aligned}$$

By (21) and (13), $\|(1-\chi_n)wg\|_p \le 2M\delta_n^\alpha/\alpha'! = O(c_n^{-\alpha})$. This and Lemmas 1 and 2 below imply that

$$E_g\|w(\hat{g}_n-g)\|_p \le O(c_n^{-\alpha}) + O(c_n^{-\alpha-\gamma})\|w\|_p/C(a_n)\,.$$

This proves Theorem 1 by (15) and (13).

Lemma 1 *Suppose the conditions of Theorem 1 hold. Then,*

$$E_g \sup_{-\infty<\theta<\infty}\left|n^{-1}\sum_{j=1}^n K_n(X_j,\theta) - E_gK_n(X_1,\theta)\right| \le n^{-1/2+\beta_0}\|k\|_1/\pi$$

and

$$\sup_{-\infty<\theta<\infty} |b_{2n}(\theta)| \le \sum_{x>m_n} \frac{(\theta^*)^x C(0) c_n^{x+1} \|k\|_1}{\pi(x+1)!} \le c_n^{-\beta_1} \frac{C(0)\|k\|_1}{\pi\theta^*} .$$

Lemma 2 *Suppose the conditions of Theorem 1 hold. Then,*

$$\|\chi_n w b_{1n}/C\|_p \le O(c_n^{-\alpha}),$$

where the $O(1)$ is uniform over $\mathcal{G}_{\alpha,\theta^*}(p,w,M)$.

We refer the reader to Loh and Zhang (1993) for the proofs of Lemmas 1 and 2.

3 Estimating a mixing distribution.

Suppose the marginal density of X is

$$f(x;G) = \int_0^{\theta^*} f(x|\theta) dG(\theta), \tag{22}$$

where $f(x|\theta)$ is as in (2) with θ^* finite and known, and G is the mixing distribution. If the density $g = G'$ exists, then $f(x;G) = f(x;g)$. In this section we consider the estimation of the mixing distribution G. Our results here are parallel to those in Section 2 for the estimation of the mixing density. We denote by E_G the expectation when G is the true mixing distribution.

Let $K_n(x,\theta)$ be as in (5), and define

$$\hat{G}_n(\theta) = \begin{cases} n^{-1}\sum_{j=1}^n \int_{a_*}^\theta K_n(X_j,y)\{C(y)\}^{-1}dy & \text{if } 0<\theta\le a_n, \\ 1 & \text{if } \theta > a_n, \end{cases} \tag{23}$$

where a_n is as in (13), and a_* is a negative constant such that $1/C(y) = \sum_{x=0}^\infty q(x)y^x$ is an increasing analytic function for $a_* \le y < \theta^*$. Similar to (7)-(9) , we have

$$E_G\hat{G}_n(\theta) - G(\theta) = B_{1n}(\theta) + B_{2n}(\theta) \to 0 \tag{24}$$

for $0 < \theta < a_n$ under suitable conditions, where

$$B_{1n}(\theta) = \int_{-\infty}^\infty \int_{a_*}^\theta \left\{\frac{C(z-y/c_n)}{C(z)}\right\} d_z G(z-y/c_n) k(y) dy - G(\theta) \tag{25}$$

and

$$(26)\quad B_{2n}(\theta) = -\sum_{x>m_n} \frac{\int_0^{\theta^*} y^x h(y)dy}{2\pi x!} \int_{-c_n}^{c_n} \mathcal{R}\left\{(it)^x \int_{a_*}^{\theta} \frac{e^{-itz}}{C(z)} dz\right\} k^*(t/c_n)dt.$$

Let $\alpha > -1$ and $1 \le p \le \infty$. Define $\mathcal{G}^{cdf}_{\alpha,\theta^*} = \mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)$ to be the set of all probability distribution functions G on $(0,\theta^*)$ such that

$$(27)\qquad \|G^{(\alpha'+1)}(\theta) - G^{(\alpha'+1)}(\theta+\delta)\|_p < M|\delta|^{\alpha''}, \quad \forall\delta.$$

REMARK. If $0 < 1+\alpha \le 1/p$, and $M > 1$, then $\mathcal{G}^{cdf}_{\alpha,\theta^*}$ is the class of all distribution functions on $(0,\theta^*)$. If $\alpha > 0$, then $\mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M) = \mathcal{G}_{\alpha,\theta^*}(p,1,M)$.

Theorem 2 *Let $\alpha > -1$ and $1 \le p \le \infty$. Suppose (15)-(17) hold with $\rho \ge \alpha + 1 + \max(\gamma,1)$ if $\gamma \ne 1$ and $\rho > \alpha + 2$ if $\gamma = 1$. Let $\hat{G}_n$ be given by (23) with $\alpha_0 \ge \alpha + 1 + \gamma$ in (4) and $\beta_1 \ge \alpha + \gamma$ in (12). Then,*

$$\sup\{E_G\|\hat{G}_n - G\|_p : G \in \mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)\} = O(c_n^{-\alpha-1}).$$

Corollary 2 *Suppose the conditions of Theorem 2 are satisfied, and that (19) holds for for some $0 \le \beta < \infty$. If (11) and (12) hold with equality, then*

$$\sup_{G\in\mathcal{G}^{cdf}_{\alpha,\theta^*}} E_G\|(\hat{G}_n - G)\|_p = \begin{cases} O(1)(1/\log n)^{\alpha+1} & \text{if } \beta = 0, \\ O(1)(\log\log n/\log n)^{\alpha+1} & \text{if } 0 < \beta < \infty. \end{cases}$$

PROOF OF THEOREM 2. Let χ_n and δ_n be as in the proof of Theorem 1. By (24)-(26) and Lemma 1,

$$\begin{aligned} &E_G\|(\hat{G}_n - G)\|_p \\ \le\ & E_G\|(\hat{G}_n - E_G\hat{G}_n)\|_p + \|\chi_n(B_{1n} + B_{2n})\|_p + \|(1-\chi_n)(1-G)\|_p \\ \le\ & O(c_n^{-\alpha-1}) + \|\chi_n B_{1n}\|_p + \|(1-\chi_n)(1-G)\|_p. \end{aligned}$$

Note that $|\int_{a_*}^{\theta} e^{-itz}\{C(z)\}^{-1}dz| \le 2/\{|t|C(\theta)\}$ implies

$$c_n\|\chi_n B_{2n} C\|_\infty \le 2(1+1/m_n)\sum_{x>m_n} \frac{(\theta^*)^x C(0) c_n^{x+1}\|k\|_1}{\pi(x+1)!}.$$

By the proof of (21), we have $\|(1-\chi_n)(1-G)\|_p \le 3M\delta_n^{\alpha+1}/(\alpha'+1)!$. The conclusion follows from

Lemma 3 *Under the conditions of Theorem 2, $\|\chi_n B_{1n}\|_p = O(c_n^{-\alpha-1})$, where the $O(1)$ is uniform over $\mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)$.*

This lemma is proved in the Appendix.

4 Optimal rate of convergence.

In Sections 2 and 3, we obtained upper bounds for the maximum $\|\cdot\|_p$ risk of our kernel estimators over the classes $\mathcal{G}_{\alpha,\theta^*} = \mathcal{G}_{\alpha,\theta^*}(p,1,M)$ and $\mathcal{G}^{cdf}_{\alpha,\theta^*} = \mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)$. Here we derive corresponding lower bounds for the rate of the minimax risk

$$(28) \qquad r_{n,\alpha,\theta^*} = \inf_{\tilde{g}_n} \sup\{E_g\|\tilde{g}_n - g\|_p : g \in \mathcal{G}_{\alpha,\theta^*}(p,1,M)\}$$

and

$$(29) \qquad r^{cdf}_{n,\alpha,\theta^*} = \inf_{\tilde{G}_n} \sup\{E_G\|\tilde{G}_n - G\|_p : G \in \mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)\},$$

where the infimum runs over all statistics $\tilde{g}_n$ and $\tilde{G}_n$ based on $X_1,\ldots,X_n$, and $\mathcal{G}_{\alpha,\theta^*}(p,1,M)$ and $\mathcal{G}^{cdf}_{\alpha,\theta^*}(p,1,M)$ are given by (14) and (27) respectively with $w(\theta) = I\{0 \le \theta \le \theta^*\}$. The rates of (28) and (29) can be regarded as characterizations of the degree of difficulty for estimating the mixing density g and the mixing distribution G respectively.

Theorem 3 *Let r_{n,α,θ^*} and $r^{cdf}_{n,\alpha,\theta^*}$ be as in (28) and (29) respectively.*
(i) If $1 \le p \le \infty$ and $\alpha > 0$, then

$$\liminf_{n\to\infty} (\log n)^\alpha r_{n,\alpha,\theta^*} > 0.$$

(ii) If $1 \le p \le \infty$ and $\alpha \ge 0$, then

$$\liminf_{n\to\infty} (\log n)^{\alpha+1} r^{cdf}_{n,\alpha,\theta^*} > 0.$$

The basic idea behind the proof of Theorem 3 is to find mixing densities g_{0n}, g_{1n}, and g_{2n} in $\mathcal{G}_{\alpha,\theta^*}(p,1,M)$ such that $\max_{j=1,2}\|g_{jn} - g_{0n}\|_p$ tends to 0 at a much slower rate than $f(.;g_{jn}) - f(.;g_{0n})$, $j = 1,2$.

The densities g_{jn} are constructed in the following manner. Let a, θ_0, and θ_1 be fixed constants satisfying $0 < a < \theta_0 < \theta_1 < \theta^*$. Define

$$(30) \qquad h_{u,v}(\theta) = v^u\theta^{u-1}e^{-v\theta}/\Gamma(u)$$

and

$$(31) \qquad g_{u,v}(\theta) = \begin{cases} h_{u,v}(\theta)/C(\theta), & \text{if } \theta < \theta_0, \\ l_{u,v}(\theta)/C(\theta), & \text{if } \theta_0 \le \theta \le \theta_1, \\ 0, & \text{if } \theta_1 < \theta, \end{cases}$$

where $l_{u,v}$ is a polynomial of degree $(2\alpha' + 1)$ such that $g_{u,v}$ is α' times continuously differentiable. Let g_0 be a probability density in $\mathcal{G}_{\alpha,\theta^*}$, and define

$$(32) \qquad g_{0n}(\theta) = g_0(\theta) + \frac{3w_0}{u_n^{(p-1)/(2p)}} \left(\frac{\theta_0}{u_n}\right)^\alpha \{g_{u_n,v_n}(\theta) - w_{0n} g_0(\theta)\}$$

$$(33)\ g_{1n}(\theta) = g_{0n}(\theta) + \frac{w_0}{u_n^{(p-1)/(2p)}} \left(\frac{\theta_0}{u_n}\right)^\alpha \left[\sin\left(u_n \frac{\theta - a}{\theta_0}\right) - \frac{w_{1n}}{w_{0n}}\right] g_{u_n,v_n}(\theta),$$

$$(34)\ g_{2n}(\theta) = g_{0n}(\theta) + \frac{w_0}{u_n^{(p-1)/(2p)}} \left(\frac{\theta_0}{u_n}\right)^\alpha \left[\cos\left(u_n \frac{\theta - a}{\theta_0}\right) - \frac{w_{2n}}{w_{0n}}\right] g_{u_n,v_n}(\theta),$$

where the constants w_{jn} are given by $\int g_{jn}(\theta) d\theta = 1$, w_0 is a small positive constant, $u_n = \delta_0 \log n$, and $v_n = u_n/a$, with

$$\delta_0 = \max\left\{\frac{\theta_0/(\theta_1 - \theta_0)}{\log(\theta_1/\theta_0)}, \frac{1}{\theta_0/a - 1 - \log(\theta_0/a)}, \frac{2}{\log(1 + a^2/\theta_0^2)}\right\}.$$

Note that for $(p-1)/(2p) + \alpha \geq 0$ and small w_0, g_{jn} are all close to g_0.

We shall show in the proof of Theorem 3 that

$$(35) \quad p_n = \inf_{\tilde{g}_n} \max\left(P_g\{\|\tilde{g}_n - g\|_p > \varepsilon_0 (\log n)^{-\alpha}\} : g = g_{0n}, g_{1n}, \text{ or } g_{2n}\right)$$

is bounded away from 0 as $n \to \infty$ for some $\varepsilon_0 > 0$, and that g_{jn}, $0 \leq j \leq 2$, are members of $\mathcal{G}_{\alpha,\theta^*}(p, 1, M)$ for small w_0 and suitable g_0. This will prove the theorem since $r_{n,\alpha,\theta^*} \geq \varepsilon_0 p_n (\log n)^{-\alpha}$.

Lemma 4 *Let $h_{u,v}$ be given by (30) with $u/v = a$. Then as $u \to \infty$, we have*

$$\|\theta^x h_{u,v}(\theta)\|_\infty \approx a^{x-1}\sqrt{u}/\sqrt{2\pi}, \quad \forall x,$$

and for $1 \leq p < \infty$,

$$\begin{aligned} \|h_{u,v}\|_p &= (v^u/\Gamma(u))\{\Gamma(pu - p + 1)/(pv)^{pu-p+1}\}^{1/p} \\ &\approx \{u/(2\pi a^2)\}^{(p-1)/(2p)} p^{-1/(2p)}. \end{aligned}$$

In addition, there exist constants c_j^* such that

$$|h_{u,v}^{(j)}(\theta)|/h_{u,v}(\theta) \leq c_j^* \theta^{-j}\{1 + |u - 1 - v\theta|^j + (v\theta)^{j/2}\}, \quad \forall j \geq 0,$$

and as $u \to \infty$

$$||h_{u,v}^{(j)}||_p/||h_{u,v}||_p \approx a^{-j}u^{j/2}\{E|Q_j(p^{-1/2}Z)|^p\}^{1/p}, \quad \forall j \geq 0,$$

where Z *is a N(0,1) random variable and* $Q_j(x)$ *are polynomials such that*

$$Q_{j+1}(x) = xQ_j(x) - \frac{d}{dx}Q_j(x), \quad Q_0(x) = 1.$$

If $u = u_n$ *and* $v = v_n$ *as in (33) and (34), then* $\left|h_{u,v}^{(j)}(\theta_0)\right| = O(n^{-1}u^{j+1/2})$ *and*

$$\left\|h_{u,v}(\theta)I\{\theta > \theta_0\}\right\|_p \leq (2\pi\theta_0^2/u)^{-(p-1)/(2p)}n^{-1}, \; \forall 1 \leq p \leq \infty.$$

REMARK. It can be proved by mathematical induction that

$$Q_j(x) = \sum_{0 \leq l \leq j/2} \frac{j!(-1)^l x^{j-2l}}{(j-2l)!l!2^l}.$$

It is also clear that $0 < E|Q_j(p^{-1/2}Z)|^p < \infty$.

Lemma 5 *There exists a constant* $C^* = C^*_{\alpha,\theta_0,\theta_1}$ *such that*

$$\|l_{u,v}^{(m)}(\theta)I\{\theta_0 \leq \theta \leq \theta_1\}\|_p \leq C^* \sum_{j=0}^{\alpha'} \left|h_{u,v}^{(j)}(\theta_0)\right|, \tag{36}$$

for all $u > 0$, $v > 0$, $1 \leq p \leq \infty$, *and* $m \geq 0$. *If* $u = u_n$ *and* $v = v_n$ *as in (33) and (34), then* $\|l_{u,v}^{(m)}\|_p = O(n^{-1}u_n^{\alpha'+1/2})$.

Lemmas 4 and 5 are proved in the Appendix. We also need the following result from Zhang (1992)

$$\sum_{x=0}^{\infty} q(x) \left| \int_0^{\theta_0} \cos(u(\theta - a)/\theta_0)\theta^x h_{u,v}(\theta)d\theta \right| \leq 3/(nC(\theta_1)). \tag{37}$$

PROOF OF THEOREM 3. For the sake of clarity, we shall break the proof down into 4 separate steps and drop the subscript n in u_n and v_n. Part (i) is proved in Steps 1-3, while Part (ii) is proved in Step 4. Also, we shall use the notation $\chi_0 = I\{0 < \theta < \theta_0\}$ and $\chi_1 = I\{\theta_0 \leq \theta < \theta_1\}$.

STEP 1. Verify the membership of g_{jn} in $\mathcal{G}_{\alpha,\theta^*}(p,1,M)$. Let $\epsilon_1 > 0$ and g_0 be a density function in $\mathcal{G}_{\alpha,\theta^*}(p,1,M-\epsilon_1)$ such that

$$g_0(\theta) \geq \epsilon_1 \chi_1(\theta),\ \forall\theta. \tag{38}$$

For small $\epsilon_1 > 0$, such g_0 exists if $\mathcal{G}_{\alpha,\theta^*}$ is nonempty.

By Lemma 5 $\|\chi_1 g_{u,v}\|_\infty = O(n^{-1}u^{\alpha'+1/2})$. Also, we have by (32)

$$1/C(\theta_0) + o(1) \geq w_{0n} = \int_0^{\theta_1} g_{u,v}(y)dy \geq 1/C(0) + o(1), \tag{39}$$

as $|w_{0n} - \|\chi_0 g_{u,v}\|_1| \leq \|\chi_1 g_{u,v}\|_1 = o(1)$ and $1/C(\theta_0) \geq \|\chi_0 g_{u,v}\|_1 \geq (1 - \|(1-\chi_0)h_{u,v}\|_1)/C(0)$. By (32), (38) and Lemma 5, $\chi_1 g_{0n} \geq 0$ for small w_0 and $(p-1)/(2p) + \alpha \geq 0$. It follows that g_{0n} is a density function, as $\chi_0 g_{u,v} \geq 0$. In the same manner, we find $|w_{jn}| \leq w_{0n} + o(1)$, so that by (33) and (34) g_{1n} and g_{2n} are all density functions.

It remains to verify (14). By the smoothness of $C(\theta)$ on $[0,\theta_1]$ and Lemmas 4 and 5, we have

$$u^{-(p-1)/(2p)}\|g_{u,v}^{(m)}\|_p = O(u^{m/2}) = O(u^m), \qquad m = \alpha', \alpha'+1,$$

which implies

$$u^{-(p-1)/(2p)-\alpha}\|g_{u,v}^{(\alpha')}(\theta) - g_{u,v}^{(\alpha')}(\theta+\delta)\|_p = O(1)\min(u^{-\alpha''}, u^{1-\alpha''}\delta).$$

Since $g_0 \in \mathcal{G}_{\alpha,\theta^*}(p,1,M-\epsilon_1)$ and $\min(u^{-\alpha''}, u^{1-\alpha''}\delta) \leq \delta^{\alpha''}$,

$$\|g_{0n}^{(\alpha')}(\theta) - g_{0n}^{(\alpha')}(\theta+\delta)\|_p \leq (M - \epsilon_1 + O(1)w_0)\delta^{\alpha''}$$

by (32). This implies (14) with $g = g_{0n}$ for small w_0. Since $\|(d/d\theta)^j h_0(u(\theta - a)/\theta_0)\|_\infty = (u/\theta_0)^j$ for $h_0(y) = \sin(y)$ and $h_0(y) = \cos(y)$, we also have

$$\begin{aligned}\left\|g_{jn}^{(m)} - g_{0n}^{(m)}\right\|_p &\leq O(1)\frac{w_0\theta_0^\alpha}{u^{(p-1)/(2p)+\alpha}}\sum_{l=0}^m \|g_{u,v}^{(l)}\|_p\left\{1 + (u/\theta_0)^{(m-l)}\right\} \\ &= w_0 O(u^{m-\alpha})\end{aligned}$$

for $m = \alpha', \alpha'+1$, and $j = 1, 2$, so that

$$\|(g_{jn} - g_{0n})^{(\alpha')}(\theta) - (g_{jn} - g_{0n})^{(\alpha')}(\theta+\delta)\|_p = O(1)w_0\delta^{\alpha''}.$$

Therefore, we also have (14) with $g = g_{1n}$ and g_{2n} for small w_0.

STEP 2. Next we shall show that

$$(40)\qquad \begin{aligned} &\sum_{x=0}^{\infty}|f(x;g_{jn})-f(x;g_{0n})| \\ =\ &\sum_{x=0}^{\infty}q(x)\left|\int_0^{\theta^*}\theta^x C(\theta)\{g_{jn}(\theta)-g_{0n}(\theta)\}d\theta\right| = o(n^{-1}). \end{aligned}$$

We shall only prove this for g_{2n}. Set $l_{u,v,x}(\theta)=\theta^x l_{u,v}(\theta)$. By the definition of $l_{u,v}$ and Lemmas 4 and 5 we have

$$|l_{u,v,x}^{(m)}(\theta_0)| = O(1)\sum_{j=0}^{m}(x+1)^j\theta_0^{x-j}\left|h_{u,v}^{(m-j)}(\theta_0)\right| = O(n^{-1}u^{m+1/2})(x+1)^m\theta_0^x,$$

for $m\geq 0$ and by Lemma 5

$$\int_{\theta_0}^{\theta_1}|l_{u,v,x}^{(\alpha'+1)}(\theta)|d\theta = O(n^{-1}u^{\alpha'+1/2})(x+1)^{\alpha'+1}\theta_1^x.$$

Integrating by parts $\alpha'+1$ times, we obtain

$$\begin{aligned} &\left|\int_{\theta_0}^{\theta_1}\cos(u(\theta-a)/\theta_0)\theta^x l_{u,v}(\theta)d\theta\right| \\ \leq\ &(\theta_0/u)^{\alpha'+1}\int_{\theta_0}^{\theta_1}|l_{u,v,x}^{(\alpha'+1)}(\theta)|d\theta+\sum_{j=0}^{\alpha'}(\theta_0/u)^{j+1}|l_{u,v,x}^{(j)}(\theta_0)| \\ \leq\ &O(n^{-1}u^{-1/2})(x+1)^{\alpha'+1}\theta_1^x, \end{aligned}$$

so that

$$\begin{aligned} &\sum_{x=0}^{\infty}q(x)\left|\int_{\theta_0}^{\theta_1}\cos(u(\theta-a)/\theta_0)\theta^x l_{u,v}(\theta)d\theta\right| \\ =\ &O(n^{-1}u^{-1/2})\sum_{x=0}^{\infty}q(x)(x+1)^{\alpha'+1}\theta_1^x = o(n^{-1}). \end{aligned}$$

It follows from (31) and (37) that

$$(41)\qquad \sum_{x=0}^{\infty}q(x)\left|\int_0^{\theta_1}\cos(u(\theta-a)/\theta_0)\theta^x C(\theta)g_{u,v}(\theta)d\theta\right| = O(n^{-1}).$$

Since $\sum_x q(x)\theta^x C(\theta) = 1$, (39) and (41) imply

$$(42)\qquad \|g_{u,v}\|_1 \left|\frac{w_{2n}}{w_{0n}}\right| = \frac{\|g_{u,v}\|_1}{|w_{0n}|}\left|\int_0^{\theta_1} \cos(u(\theta-a)/\theta_0)g_{u,v}(\theta)d\theta\right| = O(n^{-1}).$$

Thus, by (34) the left-hand side of (40) is $o(n^{-1})$ for $j = 2$, as it is bounded by the product of $w_0\theta_0^\alpha u^{-(p-1)/(2p)-\alpha}$ and the sum of (41) and (42).

STEP 3. Verify (35) and prove Part (i). In view of (40), we only need to show

$$(43)\qquad \liminf_{n\to\infty}\,(\log n)^\alpha \max_{j=1,2}\|g_{jn} - g_{0n}\|_p > 2\varepsilon_0,$$

for some positive constant ε_0. By Lemma 4 there exists a positive constant δ_1 such that for large n

$$\begin{aligned} C(0)\,\|g_{u,v}\|_p &\geq \|h_{u,v}\|_p - \|(1-\chi_0)h_{u,v}\|_p \\ &\geq 2\delta_1 u^{(p-1)/(2p)} - O(n^{-1}u^{(p-1)/(2p)}) \geq \delta_1 u^{(p-1)/(2p)}. \end{aligned}$$

Since $\max\{|\sin(x)|, |\cos(x)|\} \geq 1/\sqrt{2}$, by (33) and (34)

$$\begin{aligned} & w_0^{-1}(u/\theta_0)^\alpha\{\|g_{1n} - g_{0n}\|_p^p + \|g_{2n} - g_{0n}\|_p^p\}^{1/p} \\ \geq\ & u^{-(p-1)/(2p)}\left(1/\sqrt{2} - \max_{j=1,2}|w_{jn}/w_{0n}|\right)\|g_{u,v}\|_p \\ \geq\ & \{\delta_1/C(0)\}\left(1/\sqrt{2} - \max_{j=1,2}|w_{jn}/w_{0n}|\right), \end{aligned}$$

which implies (43), as $w_{jn}/w_{0n} \to 0$ and $u = \delta_0 \log n$.

STEP 4. Prove Part (ii). Let $G_{u,v}$ and G_{jn} be the integrations of $g_{u,v}$ and g_{jn} respectively. If $\alpha > 0$, then $g_{jn} \in \mathcal{G}_{\alpha,\theta^*} \subseteq \mathcal{G}^{cdf}_{\alpha,\theta^*}$ by Step 1. For $\alpha = 0$, we have

$$\|G_{u,v}(\theta) - G_{u,v}(\theta+\delta)\|_p \leq \delta\|g_{u,v}\|_p,$$

so that $g_{jn} \in \mathcal{G}^{cdf}_{0,\theta^*}$ for small w_0 by Lemma 4. By Step 2, Part (ii) holds if

$$(44)\qquad \liminf_{n\to\infty}\,(\log n)^{\alpha+1}\max_{j=1,2}\|G_{jn} - G_{0n}\|_p > 2\varepsilon_0.$$

Integrating by parts three times, we have

$$\begin{aligned} & \int_0^\theta \cos(u(y-a)/\theta_0)g_{u,v}(y)dy \\ =\ & \sum_{j=1}^3(-\theta_0/u)^j\cos(u(\theta-a)/\theta_0 - j\pi/2)g_{u,v}^{(j-1)}(\theta) \\ & +(\theta_0/u)^3\int_0^\theta \sin(u(y-a)/\theta_0)g_{u,v}^{(3)}(y)dy\ . \end{aligned}$$

It follows from (33), (34), (42), and Lemmas 4 and 5 that

$$\|G_{2n} - G_{0n}\|_p = (\theta/u)(1 + O(1/\sqrt{u}))\|g_{1n} - g_{0n}\|_p \ ,$$

and likewise

$$\|G_{1n} - G_{0n}\|_p = (\theta/u)(1 + O(1/\sqrt{u}))\|g_{2n} - g_{0n}\|_p \ .$$

Hence, (44) follows from (43).

5 The case of infinite θ^*.

The natural value of θ^* in (2) is $\theta_0^* = \sup\{\theta : \sum_x q(x)\theta^x < \infty\}$. If (19) holds with $\beta = 0$, then θ_0^* is finite and known, and we can set $\theta^* = \theta_0^*$ and the results of Sections 2 and 3 follow. However if (19) does not hold for $\beta = 0$, then $\theta_0^* = \infty$ and the condition $\theta^* < \infty$ becomes an assumption in addition to the knowledge of $q(\cdot)$. In this section, we consider the case of $\theta^* = \infty$. Upper bounds of the L^p risks of our kernel estimators are provided in Theorems 4 and 5 below, which are proved in the Appendix. The lower bounds of Theorem 3 still apply here.

Let $\eta = \sqrt{\theta}$. Define

$$K_{s,n}(x, \eta) = [\pi q(x)(2x)!]^{-1}(-1)^x \int_{-c_n}^{c_n} \cos(t\eta)t^{2x}k^*(t/c_n)dt, \tag{45}$$

where k is as in Section 2 and c_n is a constant tending to ∞. Set

$$g_s(\eta) = g(\eta^2)I\{\eta \geq 0\}, \ C_s(\eta) = 2\eta C(\eta^2), \ h_s(\eta) = g_s(\eta)C_s(\eta). \tag{46}$$

Since $f(x; g) = \int_{-\infty}^{\infty} y^{2x}h_s(y)dy$ by (1) and (2), we have

$$\begin{aligned} & E_g K_{s,n}(X_1, \eta) - 2\eta C(\eta^2)g(\eta^2)I\{\eta \geq 0\} \\ = & \int_0^{\infty} c_n\{k(c_n(\eta - y)) + k(c_n(\eta + y))\}h_s(y)dy - h_s(\eta) \\ = & b_{3n}(\eta) + b_{4n}(\eta) \end{aligned}$$

by the Fourier inversion formula, where

$$b_{3n}(\eta) = \int_{-\infty}^{\infty} k(y)\{h_s(\eta - y/c_n) - h_s(\eta)\}dy, \tag{47}$$

and

$$b_{4n}(\eta) = \int_{-\infty}^{\infty} k(y)h_s(y/c_n - \eta)dy. \tag{48}$$

Zhang (1992) proposed to use $K_{s,n}(x,\sqrt{a})/[2\sqrt{a}C(a)]$ as a kernel for $g(a)$ in the case of $\theta^* = \infty$. Define

$$(49)\ \hat{g}_{n,\infty}(\theta) = \hat{g}_{s,n}(\sqrt{\theta})I\{a_{0n}^2 \le \theta \le a_{1n}^2\},\ \hat{g}_{s,n}(\eta) = n^{-1}\sum_{j=1}^{n}\frac{K_{s,n}(X_j,\eta)}{C_s(\eta)}$$

where a_{0n} and a_{1n} are positive constants tending to 0 and ∞ respectively. We shall study the global performance of this estimator under weighted L^p loss functions.

Let $1 \le p \le \infty$ and w be a measurable function on $(0,\infty)$ with $\|w\|_p < \infty$. For $\alpha > 0$ we define $\mathcal{G}_{\alpha,\infty} = \mathcal{G}_{\alpha,\infty}(p,w,\alpha_1,M,M_1)$ to be the set of all probability density functions g on $(0,\infty)$ such that

$$(50)\quad \|w_s(|\eta|)\{g_s^{(\alpha')}(\eta+\delta) - g_s^{(\alpha')}(\eta)\}\|_p < M|\delta|^{\alpha''},\quad \forall\delta,$$

$$(51)\quad \|w(\theta)g(\theta)I\{\theta > a\}\|_p < M_1[C(a)]^{\alpha_1},\quad \forall a > 0,$$

where $w_s(\eta) = (2\eta)^{1/p}w(\eta^2)$, and α_1, M, and M_1 are given constants. Note that with $\eta = \sqrt{\theta}$, $\|w(\theta)h_0(\sqrt{\theta})\|_p = \|w_s(\eta)h_0(\eta)\|_p$ for all Borel functions h_0.

We assume that for every $0 < \delta < 1$ there exists a finite constant C_δ^* such that

$$(52)\quad |C^{(j)}(\theta)|\{C(\theta)\}^{\delta-1} < C_\delta^*,\quad \forall\theta \ge 0,$$

for all $0 \le j < \rho$. This condition holds for $C(\theta) = e^{-\theta}$ of the Poisson mixture.

Theorem 4 *Let $\alpha > 0$ and $1 \le p \le \infty$. Suppose (17) and (52) hold with $\rho > \alpha(1 + 1/\alpha_1)$ and $\rho \ge \alpha + 1$, and that*

$$(53)\quad q(x)B_0B^{2x}[(2x)!]^{\beta/2} \ge 1,\quad \forall x \ge 0,$$

for some constants B_0, B, and $0 < \beta < 2$. Let $\hat{g}_{n,\infty}$ be given by (49) with $a_{0n} = a_*/c_n$, $C(a_{1n}^2) = c_n^{-\alpha/\alpha_1}$, $c_n = B^{-1}\{(\beta_0 \log n)/(1-\beta/2)\}^{1-\beta/2}$, and $K_{s,n}(x,\eta)$ as in (45) with $\alpha_0 \ge \rho$ in (4), where a_* and $\beta_0 < 1/2$ are positive constants. Then,

$$\sup\{E_g\|w(\hat{g}_{n,\infty} - g)\|_p : g \in \mathcal{G}_{\alpha,\infty}(p,w,\alpha_1,M,M_1)\} = O(1)(\log n)^{-\alpha(1-\beta/2)}.$$

REMARK. The β in (53) is the same as that in (19). The Poisson mixture satisfies the conditions of Theorem 4 with $\beta = 1$.

Now let us consider the estimation of the mixing distribution G for $\theta^* = \infty$. Define

$$(54)\quad \hat{G}_{n,\infty}(\theta) = \begin{cases} n^{-1}\sum_{j=1}^{n}\int_0^{\sqrt{\theta}} K_{s,n}(X_j,y)\{C(y^2)\}^{-1}dy & \text{if } 0<\theta\le a_n^2, \\ 1 & \text{if } \theta > a_n^2, \end{cases}$$

where $K_{s,n}(x,\eta)$ is given by (45), and $0 < a_n \to \infty$. Similar to (24)-(26) and (47)-(48) we have

$$(55)\qquad E_G\hat{G}_{n,\infty}(\theta) - G(\theta) = B_{3n}(\sqrt{\theta}) + B_{4n}(\sqrt{\theta})$$

for $0<\theta<a_n$, where with $G_s(y) = G(y^2)I\{y\ge 0\}$ and $C_{s,0}(y) = C(y^2)$,

$$B_{3n}(\eta) = \int_{-\infty}^{\infty}\int_0^{\eta}\left\{\frac{C_{s,0}(z-y/c_n)}{C_{s,0}(z)}\right\}d_zG_s(z-y/c_n)k(y)dy - G_s(\eta),$$

and

$$B_{4n}(\eta) = -\int_{-\infty}^{\infty}\int_0^{\eta}\left\{\frac{C_{s,0}(y/c_n-z)}{C_{s,0}(z)}\right\}d_zG_s(y/c_n-z)k(y)dy.$$

Let $1\le p\le\infty$ and unction on $(0,\infty)$ with $\|w\|_p<\infty$. For $\alpha>0$ we define $\mathcal{G}_{\alpha,\infty} = \mathcal{G}_{\alpha,\infty}(p,w,\alpha_1,M,M_1,M_2)$ to be the set of all probability density functions g on $(0,\infty)$ such that

$$\begin{aligned}(56)\quad \|w_s(|\eta|)\{G_s^{(\alpha'+1)}(\eta+\delta) - G_s^{(\alpha'+1)}(\eta)\}\|_p &< M|\delta|^{\alpha''}, \quad \forall\delta, \\ (57)\quad \|w(\theta)(1-G(\theta))I\{\theta>a\}\|_p &< M_1[C(a)]^{\alpha_1},\ \forall a>0,\end{aligned}$$

and

$$(58)\qquad G(\theta)\le M_2\theta^{(\alpha+1)/2},\quad \forall\theta>0,$$

where w_s is as in (50), and α_1, M, M_1, and M_2 are given constants.

REMARK. Although $b_{3n}(\eta)\to 0$ for $-\infty<\eta<\infty$ as $c_n\to\infty$, $b_{4n}(\eta)\to h_s(-\eta)$. Thus, the bias of $\hat{G}_{n,\infty}$ will not tend to 0 if we integrate from a negative number in (54) as we did in (23). This caused us to add condition (58).

Theorem 5 *Let $\alpha>-1$ and $1\le p\le\infty$. Suppose (17) and (52) hold with $\rho>(\alpha+1)(1+1/\alpha_1)$, and that (53) holds for some $0<\beta<2$. Let $\hat{G}_{n,\infty}$ be given by (54) with c_n as in Theorem 4, $C(a_n^2) = c_n^{-(\alpha+1)/\alpha_1}$, and $\alpha_0\ge\rho$ in (4). Then*

$$\begin{aligned}&\sup\{E_G\|w(\hat{G}_{n,\infty}-G)\|_p : G\in\mathcal{G}^{cdf}_{\alpha,\infty}(p,w,\alpha_1,M,M_1,M_2)\} \\ =\ & O(1)(\log n)^{-(\alpha+1)(1-\beta/2)}.\end{aligned}$$

6 Appendix.

We shall extend the domain of C such that (17) holds for all real numbers θ and δ. By (18) we have the expansion

$$C(\theta - y/c_n)g(\theta - y/c_n) = \sum_{j=1}^{4} \xi_j(\theta, -y/c_n) \tag{59}$$

where $\xi_j(\theta,\delta) = \xi_j(\theta,\delta; g, C, \alpha, \gamma, \rho)$ are given by

$$\begin{aligned}
\xi_1(\theta,\delta) &= \sum_{j=0}^{\alpha'} g^{(j)}(\theta)\{\delta^j/j!\}\mathrm{Rem}_C(\theta,\delta,\rho'-j),\\
\xi_2(\theta,\delta) &= \mathrm{Rem}_g(\theta,\delta,\alpha')\sum_{j=0}^{\gamma'} C^{(j)}(\theta)\{\delta^j/j!\},\\
\xi_3(\theta,\delta) &= \mathrm{Rem}_g(\theta,\delta,\alpha')\mathrm{Rem}_C(\theta,\delta,\gamma'),
\end{aligned}$$

and

$$\xi_4(\theta,\delta) = \sum_{j=0}^{\alpha'} g^{(j)}(\theta)\{\delta^j/j!\}\sum_{l=0}^{\rho'-j} C^{(l)}(\theta)\{\delta^l/l!\}.$$

PROOF OF LEMMA 3. Let $\xi_j(\theta,\delta; g, C, \alpha, \gamma, \rho)$ be as in (59). For $j = 0, 1$, define

$$\xi_{jn} = \xi_{jn}(\theta) = \int_{-\infty}^{\infty} \sum_{l=1}^{3} \xi_l(\theta, -y/c_n; G, C^{(j)}, \alpha+1, \gamma_j, \rho_j)k(y)dy,$$

where $\rho_j = \alpha + 1 + \gamma_j$, $\min(\gamma_0, \gamma_1) \geq 0$, and $\max(\rho_0, \rho_1 + 1) \leq \rho$. Note that the pair (g, C) in (59) is replaced by $(G, C^{(j)})$ here. Integrating by parts in (25), we find

$$B_{1n} = \frac{\xi_{0n}(\theta)}{C(\theta)} - \frac{\xi_{0n}(a_*)}{C(a_*)} + \int_{a_*}^{\theta} \left\{ \frac{C^{(1)}(z)}{C(z)}\frac{\xi_{0n}(z)}{C(z)} - \frac{\xi_{1n}(z)}{C(z)} \right\} dz. \tag{60}$$

Since the proof of Lemma 2 depends only on the smoothness and boundedness of $g^{(j)}$, $C^{(j)}$, and χ_n/C, it also applies to the components of ξ_{0n} and ξ_{1n}. It follows that $\|\chi_n w \xi_{0n}/C\|_p = O(c_n^{-\alpha-1})$ and that there exist functions ζ_{jn} with $\|\chi_n w \zeta_{jn}\|_p = O(c_n^{-\alpha-1})$ such that

$$\frac{C^{(1)}(z)}{C(z)}\frac{\xi_{0n}(z)}{C(z)} - \frac{\xi_{1n}(z)}{C(z)} = \sum_{j=0}^{2} \zeta_{jn}(z)h_{jn}(z),$$

where $h_{0n} = C^{(1)}/[c_n^{\gamma_0}C^2]$, $h_{1n} = 1/[c_n^{\gamma_1}C]$, and $h_{2n} = 1/[c_n(\theta^* - \theta)^2]$. Notice the cancellation of the term $\{C^{(1)}/C\}\text{Rem}_g$ here. By the Hölder inequality and the monotonicity of w, we have for fixed $\theta \leq a_n$

$$\left| w(\theta) \int_{a_*}^{\theta} \left\{ \frac{C^{(1)}(z)}{C(z)} \frac{\xi_{0n}(z)}{C(z)} - \frac{\xi_{1n}(z)}{C(z)} \right\} dz \right| \leq O(c_n^{-\alpha-1}) \sum_{j=0}^{2} \bar{h}_{jn}(\theta),$$

where $\bar{h}_{jn}(\theta) = \|\chi_{[0,\theta]} h_{jn}\|_{p/(p-1)}$. This and (60) imply

$$\text{(61)} \quad \|\chi_n w B_{1n}\|_p \leq \|w\|_p \left| \frac{\xi_{0n}(a_*)}{C(a_*)} \right| + O(c_n^{-\alpha-1}) \left(1 + \sum_{j=0}^{2} \|\chi_n \bar{h}_{jn}\|_p \right).$$

Set $\gamma_0 = \gamma$, $\gamma_1 = \max(\gamma - 1, 0)$ if $\gamma \neq 1$, and $1 \geq \gamma_1 > 0$ if $\gamma = 1$. Let $h_0(z) = (\theta^* - z)^{-\kappa-1}$ and $\bar{h}_0(\theta) = \|\chi_{[0,\theta]} h_0\|_{p/(p-1)}$. Then for $\kappa > 0$ and $1 \leq p \leq \infty$,

$$\|\chi_n(\theta)\bar{h}_0(\theta)\|_p \leq (\theta^* - a_n)^{-\kappa}\{(p\kappa + 1)/(p-1)\}^{-(p-1)/p}\{p\kappa\}^{-1/p}.$$

This gives $\|\chi_n \bar{h}_{2n}\|_p = O(1)$ by (13). This also gives $\|\chi_n \bar{h}_{0n}\|_p = O(1)$ if $\gamma > 0$, while $\|\bar{h}_{0n}\|_\infty = O(1)$ by (17) if $\gamma = 0$. Since $C(\theta) = \sum_{x=0}^{\infty} q(x)\theta^x$,

$$\begin{aligned} \|\chi_n \bar{h}_{1n}\|_p &\leq c_n^{-\gamma_1} \sum_{x=0}^{\infty} a_n^{x+1}(xp/(p-1) + 1)^{-(p-1)/p}(px + p)^{-1/p} \\ &\leq c_n^{-\gamma_1} \|\chi_n / C\|_1 = O(1) \end{aligned}$$

by (15) and the choice of γ_1. Therefore, $\|\chi_n \bar{h}_{jn}\|_p = O(1)$ in all the cases. The proof is completed by (61) and

$$\begin{aligned} |\xi_{0n}(a_*)/C(a_*)| &= \left| \int C(a_* - y/c_n)G(a_* - y/c_n)k(y)dy/C(a_*) \right| \\ &\leq \int_{-\infty}^{a_* c_n} \left| \frac{y}{c_n a_*} \right|^{\alpha+1} |k(y)|dy = O(c_n^{-\alpha-1}). \end{aligned}$$

PROOF OF LEMMA 4. The approximations for $\|\theta^x h_{u,v}(\theta)\|_\infty$ and $\|h_{u,v}\|_p$ follow from the Stirling formula. Let $Q_j^*(x, y)$ be a function such that

$$h_{u,v}^{(j)}(\theta) = h_{u,v}(\theta)\theta^{-j} Q_j^*(u - 1 - v\theta, \sqrt{v\theta}).$$

Clearly $Q_0^* = 1$. Since $(\partial/\partial\theta) \log h_{u,v}(\theta) = (u-1)/\theta - v$,

$$\begin{aligned} &\theta^{-j-1} Q_{j+1}^*(u - 1 - v\theta, \sqrt{v\theta}) \\ = \ &[(u-1)/\theta - v]\theta^{-j} Q_j^* - j\theta^{-j-1} Q_j^* + \theta^{-j}\left[-vQ_{j,1}^* + \sqrt{v\theta} Q_{j,2}^*/(2\theta)\right], \end{aligned}$$

where $Q^*_{j,1}(x,y) = (\partial/\partial x)Q^*_j(x,y)$ and $Q^*_{j,2}(x,y) = (\partial/\partial y)Q^*_j(x,y)$. It follows that

$$Q^*_{j+1}(x,y) = xQ^*_j(x,y) - jQ^*_j(x,y) - y^2Q^*_{j,1}(x,y) + (y/2)Q^*_{j,2}(x,y),$$

so that $Q^*_j(x,y)$ is a polynomial of degree j. This gives the inequality for $|h^{(j)}_{u,v}(\theta)|$.

For the $\|\cdot\|_p$ norm, we have $|h^{(j)}_{u,v}|^p/\|h_{u,v}\|^p_p = |\theta^{-j}Q^*_j|^p h_{p(u-1)+1,pv}$. Since $h_{u,v}$ has mean u/v and variance u/v^2, by the central limit theorem

$$(v/\sqrt{u})(\theta - u/v) = (v\theta - u)/\sqrt{u} \to Z$$

in distribution under the density $h_{u,v}$ as $u \to \infty$. Since gamma distributions have finite moments,

$$\int_0^\infty \left|Q\left((u - v\theta)/\sqrt{u}, \sqrt{v\theta/u}\right)\right|^p h_{u,v}(\theta)d\theta \to E|Q(Z,1)|^p$$

for all polynomials Q and $p \geq 0$. Therefore, as $u \to \infty$

$$\begin{aligned} \|h^{(j)}_{u,v}\|^p_p/\|h_{u,v}\|^p_p &= \int |\theta^{-j}Q^*_j(u - 1 - v\theta, \sqrt{v\theta})|^p h_{p(u-1)+1,pv}(\theta)d\theta \\ &\approx E|(u/v)^{-j}Q^*_j(Z\sqrt{pu}/p, \sqrt{u})|^p, \\ &\approx a^{-j}u^{j/2}E|Q_j(Z/\sqrt{p})|^p, \end{aligned}$$

where $Q_j(x) = Q_j(x,1)$ and $Q_j(x,y)$ is the sum of all terms of degree j in $Q^*_j(x,y)$. The recursion of Q_j follows from that of Q^*_j.

If $u = u_n$ and $v = v_n$ as in (33) and (34), then $\|h_{u,v}(\theta)I\{\theta > \theta_0\}\|_1 \leq 1/n$ by Zhang (1992, Lemma 2). The rest follows, since by the expression for δ_0 and the Stirling formula we have $h_{u,v}(\theta)I\{\theta \geq \theta_0\} \leq \theta_0^{-1}(2\pi/u)^{-1/2}n^{-1}$.

PROOF OF LEMMA 5. Define

$$\|Q\|_0 = \sum_{j=0}^{\alpha'} \left\{|Q^{(j)}(\theta_0)| + |Q^{(j)}(\theta_1)|\right\}.$$

Since $\|\cdot\|_0$ is a norm for the $(2\alpha'+2)$-dimensional space of all polynomials Q of degree $2\alpha'+1$ on $[\theta_0, \theta_1]$, it is equivalent to all other norms on this linear space. This implies (36) as $l^{(j)}_{u,v}(\theta_1) = 0$ for $0 \leq j \leq \alpha'$. The rest follows from Lemma 4.

Proof of Theorem 4. Clearly, $\|w(\theta)g(\theta)I\{\theta > a_{1n}^2\}\|_p = O(c_n^{-\alpha})$ by (51) and the choice of a_{1n}. Also, as in Zhang (1992)

$$\|K_{s,n}(x,\cdot)\|_\infty \le \frac{2c_n^{2x+1}\|k\|_1}{\pi q(x)(2x+1)!} \le 2n^{\beta_0}B_0\|k\|_1/(\pi B).$$

It follows from (49) and the argument in the proof of Theorem 1 that

$$\begin{aligned} E_g\|w(\hat{g}_{n,\infty} - E_g\hat{g}_{n,\infty})\|_p &= O(1)n^{\beta_0-1/2} \max_{a_{0n}\le\eta\le a_{1n}} 1/C_s(\eta) \\ &= O(1)n^{\beta_0-1/2}c_n^{\max(1,\alpha/\alpha_1)}, \end{aligned}$$

so that

$$E_g\|w(\hat{g}_{n,\infty} - g)\|_p \le \|I\{0 \le \theta \le a_{1n}^2\}w(E_g\hat{g}_{n,\infty} - g)\|_p + O(c_n^{-\alpha}).$$

For the rest of the proof and unless otherwise specified we shall write everything as functions of $\eta = \sqrt{\theta}$, for which w_s is the actual weight function. Set $\chi_{0n} = I\{0 \le \eta < a_{0n}\}$, $\chi_{1n} = I\{a_{0n} \le \eta < 1\}$, and $\chi_{2n} = I\{1 \le \eta \le a_{1n}\}$. By (20) and (50) we have $\|w_s g_s \chi_{0n}\|_p = O(c_n^{-\alpha})$, so that

$$E_g\|w(\theta)(\hat{g}_{n,\infty}(\theta) - g(\theta))\|_p \le \|w_s(\chi_{1n} + \chi_{2n})(b_{3n} + b_{4n})/C_s\|_p + O(c_n^{-\alpha}),$$

where $b_{jn}(\eta)$ are given by (47) and (48).

Let ξ_j be as in (59). For $j = 1, 2$, define

$$\xi_{jn}(\eta, y) = \sum_{l=1}^{3} \xi_l(\eta, -y/c_n; g_s, C_s, \alpha, \gamma_j, \rho_j),$$

where $\rho_j = \alpha + \gamma_j$, $\gamma_1 = 1$, and $\alpha/\alpha_1 < \gamma_2 \le \rho - \alpha$. Then, by (47) $b_{3n}(\eta) = \int \xi_{jn}(\eta, y)k(y)dy$, $j = 1, 2$. As in the proof of Lemma 2, we have $\|w_s\chi_{jn}b_{3n}/C_s\|_p = O(c_n^{-\alpha})$, $j = 1, 2$. Notice that on the set $[0, 1]$ (15) is replaced by $\eta/C_s(\eta) \le 1/C(1)$ with $\gamma = \gamma_1 = 1$. Also notice that $\|\chi_{2n}C_s^{(j)}/C_s\|_\infty = o(c_n^{\epsilon_1})$ and $a_{1n} = o(c_n^{\epsilon_1})$ for all small $\epsilon_1 > 0$, while $\gamma = \gamma_2 > \alpha/\alpha_1$.

The proof of $\|w_s\chi_{jn}b_{4n}/C_s\|_p = O(c_n^{-\alpha})$ is similar and omitted. Note that (50) implies

$$\|I\{\eta > 0\}w_s(\eta)g_s^{(\alpha')}(\delta - \eta)\|_p < M|\delta|^{\alpha''}, \quad \forall\delta.$$

Proof of Theorem 5. We shall combine the methods in the proofs of Theorem 4 and Lemma 3 with the (g, C) in (59) replaced by $(G_s(y), C_{s,0}^{(j)}(y))$

in the case of B_{3n} and by $(G_s(-y), C_{s,0}^{(j)}(-y))$ in the case of B_{4n}, $j = 0, 1$. This gives

$$E_G\|w(\hat{G}_{n,\infty} - G)\|_p \leq O(c_n^{-\alpha}) + O(1)\left|\int_0^\infty C((y/c_n)^2)G((y/c_n)^2)k(y)dy\right|.$$

The proof is completed, as the integration on the right-hand side is bounded in absolute value by $C(0)M_2\int_0^\infty (y/c_n)^{\alpha+1}|k(y)|dy = O(c_n^{-\alpha-1})$.

Ackowledgement.

Research of Wei-Liem Loh was supported in part by the National Security Agency. Research of Cun-Hui Zhang was supported in part by the National Science Foundation and the Army Research Office.

References

[1] CARROLL, R. J. and HALL, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83** 1184-1186.

[2] DATTA, S. (1991). On the consistency of posterior mixtures and its applications. *Ann. Statist.* **19** 338-353.

[3] FAN, J. (1991a). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19** 1257-1272.

[4] FAN, J. (1991b). Global behavior of deconvolution kernel estimates. *Statist. Sinica* **1** 541-551.

[5] LAMBERT, D. and TIERNEY, L. (1984). Asymptotic properties of maximum likelihood estimates in the mixed Poisson model. *Ann. Statist.* **12** 1388-1399.

[6] LOH, W. L. (1992). Estimating the mixing density of a mixture of power series distributions. Manuscript.

[7] LOH, W. L. and ZHANG, C.-H. (1993). Global properties of kernel estimators for mixing densities in exponential family models for discrete variables. Technical Report, Statist. Dept., Purdue University.

[8] MEEDEN, G. (1972). Bayes estimation of the mixing distribution, the discrete case. *Ann. Math. Statist.* **43** 1993-1999.

[9] Rolph, J. E. (1968). Bayesian estimation of mixing distributions. *Ann. Math. Statist.* **39** 1289-1302.

[10] Simar, L. (1976). Maximum likelihood estimation of a compound Poisson process. *Ann. Statist.* **4** 1200-1209.

[11] Tucker, H. G. (1963). An estimate of the compounding distribution of a compound Poisson distribution. *Theor. Probab. Appl.* **8** 195-200.

[12] Walter, G. G. and Hamedani, G. G. (1991). Bayes empirical Bayes estimation for natural exponential families with quadratic variance functions. *Ann. Statist.* **19** 1191-1224.

[13] Zhang, C.-H. (1990). Fourier methods for estimating mixing densities and distributions. *Ann. Statist.* **18** 806-831.

[14] Zhang, C.-H. (1992). On estimating mixing densities in exponential family models for discrete variables. Tech. Rep. 055-92, Mathematical Sciences Research Institute, Berkeley, California.

Department of Statistics, Purdue University, West Lafayette, IN 47907, U.S.A.
Department of Statistics, Hill Center, Busch Campus, Rutgers University, New Brunswick, NJ 08903, U.S.A.