ESTIMATION OF DISTRIBUTION AND DENSITY FUNCTIONS BY GENERALIZED BERNSTEIN POLYNOMIALS

B. L. S. PRAKASA Rao

Department of Mathematics and Statistics, University of Hyderabad, Gachibowli, Hyderabad 500 046, India
E-mail: blsprao@gmail.com

(Received 22 September 2004; after final revision 17 February 2005)

As an application of generalized Bernstein polynomials, we suggest estimators for distribution and probability density functions with bounded support and obtain their asymptotic properties.

Key Words: Generalized Bernstein Polynomial; Density Estimation; Estimation of Distribution Function; Bounded Support

1. INTRODUCTION

In their recent paper, Babu et al. (2002) studied application of Bernstein polynomials for estimation of a distribution and density function with bounded support. Earlier work in this direction is due to Vitale (1975). We now study these problems as an application of generalized Bernstein polynomials developed recently by Cao (1997) and Ding (2004). These estimators have faster rates of convergence than those developed in Babu et al. (2002). Prakasa Rao (1983, 1999) give reviews of various methods of density estimation and their applications.

2. PRELIMINARIES

Let \( C [a, b] \) be the space of continuous functions on a bounded interval \([a, b]\) with the associated norm

\[ \| g \| = \max_{x \in [a, b]} |g(x)|. \]

Associated with an integer \( m \geq 1 \), the Bernstein polynomials \( B_m \) in \( C [0, 1] \) are defined by

\[ B_m g(x) = \sum_{k=0}^{m} g \left( \binom{k}{m} \right) b_k (m, x), \quad g \in C [0, 1], \quad x \in [0, 1] \]

where
\[ b_m(m, x) = \binom{m}{k} x^k (1-x)^{m-k}, \quad k = 0, \ldots, m. \]

Let \( s \) be a sequence of integers with \( s \geq 1 \). Cao (1997) introduced the generalized Bernstein polynomials \( C_m \) in \( C[0, 1] \) defined by

\[ C_m g(x) = \frac{1}{s} \sum_{k=0}^{m} \sum_{j=0}^{s-1} \binom{m}{k} \binom{s-1}{j} b_k(m, x), \quad g \in C[0, 1], \quad x \in [0, 1]. \]

It is easy to see that the generalized Bernstein polynomials reduce to the classical Bernstein polynomials if we choose \( s = 1 \) for all \( m \geq 1 \). It is also known that the generalized Bernstein polynomials are dense in \( C[0, 1] \) under the supremum norm if \( \frac{s_m}{m} \to 0 \) as \( m \to \infty \) (cf. Cao 1997) generalising the classical result of approximation of continuous function by Bernstein polynomials.

For any \( g \in C[0, 1] \), let \( \omega(g, t) \) be the first order modulus of continuity of the function, that is

\[ \omega(g, t) = \sup_{x, y \in [0, 1] : |x - y| \leq t} |g(x) - g(y)|. \]

For any \( g \in C[0, 1] \), the second order Ditzian-Totik modulus of continuity (also called modulus of smoothness) (cf. Ditzian and Totik (1987)) is defined by

\[ \omega_{g}^{(2)}(g, t) = \sup_{0 < h \leq t} \left\| \frac{\Delta_h^2 g}{h} \right\| \]

where

\[ \phi(x) = \sqrt{x(1-x)}, \quad x \in [0, 1] \]

and

\[ \Delta_h^2 g(x) = g(x + h) - 2g(x) + g(x - h), \quad x \pm h \in [0, 1] \]

... (2.1)

\[ \Delta^2_h g(x) = 0 \text{ otherwise.} \]

The following theorem is due to Cao (1997).

**Theorem 2.1** — Suppose that \( s \) is a sequence of integers greater than or equal to 1 such that \( \lim_{m \to \infty} \frac{s_m}{m} = 0 \). Let \( g \in C[0, 1] \). Then
\[ \lim_{m \to \infty} \| C_m g - g \| = 0. \]

Furthermore, for any integer \( m \geq 1 \) such that
\[ 0 < \frac{s_m - 1}{m} + \frac{1}{\sqrt{m}} \leq 1, \]
the following inequality holds:
\[ \| C_m g - g \| \leq 4 \omega \left( g, \frac{s_m - 1}{m} + \frac{1}{\sqrt{m}} \right). \]

The following results are due to Ding (2004).

\textbf{Theorem 2.2} — Suppose \( g \in C[0,1] \), and \( \lim_{m \to \infty} \frac{s_m}{m} = 0 \). Then there exists a positive constant \( C \) independent of \( g, x \) and \( m \) such that
\[ \| C_m g - g \| \leq C \left( \omega_\phi^{(2)} \left( g, \frac{1}{\sqrt{m}} \right) + \omega \left( g, \frac{s_m - 1}{m} \right) \right). \]

Furthermore
\[ \omega_\phi^{(2)} \left( g, \frac{1}{\sqrt{m}} \right) \leq C m^{-1} \sum_{k=1}^{m} \| C_k g - g \| \]
and
\[ \omega \left( g, \frac{1}{m} \right) \leq C m^{-1} \left( \sum_{k=1}^{m} \| C_k g - g \| + \| g \| \right). \]

As a Corollary to the above theorem, the following result can be proved.

\textbf{Corollary 2.3} — Suppose that a function \( g \in C[0,1] \) satisfies the conditions
\[ \omega_\phi^{(2)} (g, t) = O (r^{2\alpha}) \]
and
\[ \omega (g, t) = O (t^{\alpha}) \]
for some \( 0 < \alpha < 1 \). Further suppose that \( s_m \geq 1 \) and \( \lim_{m \to \infty} \frac{s_m}{m} = 0 \). Then
\[ \| C_m g - g \| = O \left( \left( \frac{s}{m} \right)^{\alpha} \right). \]

Suppose that \( g \in C^1 [0, 1] \). Ding (2004) obtained the following limit results.

**Theorem 2.4** — Let \( s = Cm^{\tau} \) where \( 0 < \tau < 1 \) and \( C \) is a positive constant. Then, for any \( g \in C^1 [0, 1] \),

\[ \lim_{m \to \infty} m^{1-\tau} (C_m g(x) - g(x)) = C(2x-1)g'(x). \]

Furthermore, if \( s = C \geq 1 \), where \( C \) is a positive constant, then for any \( g \in C^2 [0, 1] \),

\[ \lim_{m \to \infty} m(C_m g(x) - g(x)) = (1-C)(2x-1)g'(x) + \frac{1}{2}x(1-x)g''(x). \]

The problem of estimation of the distribution function of a random variable with bounded support has been studied by Babu et al. (2002) as an application of the classical Bernstein polynomials. Vitale (1975) investigated the problem of density estimation using the Bernstein polynomial approach and more recently, Petrone (1999) proposed a Bayesian nonparametric procedure for density estimation. All these works use the classical Bernstein polynomials as the building blocks for the proposed estimators. Prakasa Rao (1983, 1999) discusses other methods for a density estimation and estimation of a distribution function.

Our aim in this paper is to develop estimators by using the generalized Bernstein polynomials defined above and study their asymptotic properties. *Inter alia*, we obtain the rates of convergence of these estimators.

### 3. ESTIMATION OF A DISTRIBUTION FUNCTION

Let \( X_i, 1 \leq i \leq n \) be independent and identically distributed (i.i.d.) observations with distribution function \( F \) with bounded support. Without loss of generality, we assume that the support of the distribution is the interval \([0, 1]\). Let \( F_n(\cdot) \) be the empirical distribution function based on the observations \( X_i, 1 \leq i \leq n \).

For any bounded continuous function \( g \in C[0, 1] \), let \( C_m g \) be the generalised Bernstein polynomial as discussed in Section 2. We consider the function \( \hat{F}_{n,m}(x) = C_m F_n(x) \) as an estimator of the distribution function \( F(x), x \in [0, 1] \). It is obvious that the estimator \( \hat{F}_{n,m}(x) = C_m F_n(x) \) is a polynomial in \( x \) and hence is continuous and infinitely differentiable. Furthermore the function is nondecreasing in \( x \) with \( 0 \leq \hat{F}_{n,m}(x) \leq 1 \) for all \( 0 \leq x \leq 1 \). Hence it is a genuine random distribution function.
**Theorem 3.1** — Let \( F \) be a continuous distribution function with support as the interval \([0, 1]\). Suppose that \( s_m \) is a sequence of integers greater than equal to one such that \( \lim_{m \to \infty} \frac{s_m}{m} = 0 \). If \( m, n \to \infty \), then \( \left\| \hat{F}_{n,m} - F \right\| \to 0 \) almost surely.

**Proof:** Observe that, for all \( x \in [0, 1] \),

\[
\left\| \hat{F}_{n,m} - F \right\| \leq \left\| \hat{F}_{n,m} - C_m F \right\| + \left\| C_m F - F \right\|. \tag{3.1}
\]

For any function \( g(\cdot) \) defined on \([0, 1]\), let

\[
\gamma_{j,m}^g = \frac{1}{s_m} \sum_{r=0}^{s_m-1} g \left( \frac{j+r}{m+s_m-1} \right), \quad j = 0, 1, \ldots, m. \tag{3.2}
\]

Since

\[
\hat{F}_{n,m}(x) - C_m F(x) = \sum_{k=0}^{m} \left( \gamma_{k,m}^F - \gamma_{k,m}^F \right) b_k(m,x), \tag{3.3}
\]

it follows that

\[
\left\| \hat{F}_{n,m} - C_m F \right\| \leq \max_{0 \leq k \leq m} \left| \gamma_{k,m}^F - \gamma_{k,m}^F \right| \tag{3.4}
\]

\[
\leq \max_{0 \leq k \leq m, 0 \leq r \leq s_m-1} \left| F \left( \frac{k+r}{m+s_m-1} \right) - F \left( \frac{k+r}{m+s_m-1} \right) \right|
\]

\[
\leq \left\| F_n - F \right\|.
\]

Hence

\[
\left\| \hat{F}_{n,m} - F \right\| \leq \left\| F_n - F \right\| + \left\| C_m F - F \right\|. \tag{3.5}
\]

Applying the Glivenko-Cantelli theorem, we obtain that \( \left\| F_n - F \right\| \to 0 \) as \( n \to \infty \).

Suppose that \( m \) and \( n \) tend to infinity such that \( \lim_{m \to \infty} \frac{s_m}{m} = 0 \). Now the result stated in Theorem 3.1 follows from the inequality (3.5) as an application of Theorem 2.1.

The following theorem gives the rate of convergence of the estimator \( \hat{F}_{n,m} \).
Theorem 3.2 — Let $F$ be a continuous distribution function with support as the interval $[0, 1]$. Suppose that $s_m$ is a sequence of integers greater than equal to one such that
\[
\lim_{m \to \infty} \frac{s_m}{m} = 0. \text{ Then}
\]
\[
\left\| \hat{F}_{n,m} - F \right\| = O\left( \left( \frac{\log \log n}{n} \right)^2 \right) + O\left( \frac{1}{\sqrt{m}} \right) + O\left( \frac{F_{s_m - 1}}{m} \right).
\]

In particular, if there exists $0 < \alpha < 1$ such that $\omega^{(2)}(F, t) = O(t^{2\alpha})$ and $\omega(F, t) = O(t^\alpha)$, then
\[
\left\| \hat{F}_{n,m} - F \right\| = O\left( \left( \frac{\log \log n}{n} \right)^{1/2} \right) + O\left( \frac{s_m}{m} \right) \alpha.
\]

Proof: The first result is a consequence of the inequality (3.5), Theorem 2.2 and the theorem of Smirnov (1944) and Chung (1949) (cf. Csorgo and Revesz (1981), p. 157). The second result follows now as an application of Corollary 2.3.

Remarks: If we choose $m = \lambda_n$ tending to infinity as $n \to \infty$ and the sequence $s_m$ such that
\[
s_m = m \left( \frac{\log \log n}{n} \right)^{1/2\alpha}
\]
in Theorem 3.2, then we obtain that
\[
\left\| \hat{F}_{n,\lambda_n} - F \right\| = O\left( \left( \frac{\log \log n}{n} \right)^{1/2} \right)
\]
which gives the optimum rate of convergence under the conditions stated in Theorem 3.1.

We now obtain bound on the difference between the estimator $\hat{F}_{n,m}$ and the empirical distribution function following the techniques in Babu et al. (2002). We first state a result due to Babu (1989) which will be used in the sequel.

Lemma 3.3 — Let $Z_i, 1 \leq i \leq n$ be independent random variables with mean zero and finite variances. Suppose that $\sup_{1 \leq i \leq n} |Z_i| \leq b$ for some $b > 0$. Further suppose that
\[
\sum_{i=1}^{n} E\left( Z_i^2 \right) \leq v.
\]
Then, for any $0 < s < 1$, and for all $0 \leq a \leq \frac{v}{s b^2}$,
$P \left( \sum_{i=1}^{n} Z_i \geq a \right) \leq \exp \left\{ -a^2 s (1-s)/v \right\}$.

Let us choose

$$a_m = [(s_m + m - 1)^{-1} \log (s_m + m - 1)]^{1/2}$$

and

$$b_{n,m} = (n^{-1} \log n)^{1/2} [(s_m + m - 1)^{-1} \log (s_m + m - 1)]^{1/4}$$

It is obvious that $b_{n,m} \leq a_m$. Divide the interval $[0,1]$ into subintervals of length $b_{n,m}$. Let

$$N_{x,m} = \left\{ 0 \leq k \leq m : \left| \frac{k+r}{m} - x \right| \leq a_m \text{ for some } 0 \leq r \leq s - 1 \right\}.$$

Note that

$$\hat{F}_{n,m}(x) - F(x) = \sum_{k=0}^{m} b_k(m,x) \left( \gamma_{k,m}^F - \gamma_{k,m}^F \gamma_{k,m}^F - F_n(x) + F(x) \right) + C_m F(x) - F(x).$$

Hence

$$\left\| \hat{F}_{n,m} - F_n \right\| \leq \sup_{0 \leq x \leq 1} \left| \sum_{k=0}^{m} b_k(m,x) \left( \gamma_{k,m}^F - \gamma_{k,m}^F \gamma_{k,m}^F - F_n(x) + F(x) \right) \right| + \left\| C_m F - F \right\| \ldots (3.6)$$

provided the distribution function $F$ satisfies the conditions in the Corollary 2.3. Suppose the distribution function $F$ satisfies the conditions in Corollary 2.3. Note that the distribution function $F$ is Lipschitzian of order $\alpha$. Let us now estimate the first term on the right side of the above inequality. Observe that

$$\left| \sum_{k=0}^{m} b_k(m,x) \left( \gamma_{k,m}^F - \gamma_{k,m}^F \gamma_{k,m}^F - F_n(x) + F(x) \right) \right|$$
\[
\sum_{k=0}^{m} b_k (m, x) \left\{ \frac{1}{s} \sum_{r=0}^{s-1} \left( F_n \left( \frac{k+r}{m+s_m-1} \right) \right) \right\} \\
- F \left( \frac{k+r}{m+s_m-1} \right) - F_n (x) + F (x) \right) \right| \\
\leq \sum_{k \in N_{x,m}} b_k (m, x) \sup_{0 \leq x \leq 1, 0 \leq r \leq s_m-1, k \in N_{x,m}} \left| F_n \left( \frac{k+r}{m+s_m-1} \right) - F \left( \frac{k+r}{m+s_m-1} \right) - F_n (x) + F (x) \right| \\
+ 4 \sum_{k \in N_{x,m}} b_k (m, x) \\
\leq \sum_{k \in N_{x,m}} b_k (m, x) \sup_{0 \leq x \leq 1, 1 \leq \lfloor b_{n,m} \rfloor \leq \frac{1}{m+s_m-1}} \left| F_n (j b_{n,m} - F_n (i b_{n,m}) - F (j b_{n,m}) + F (i b_{n,m})) \right| \\
+ \sum_{k \in N_{x,m}} b_k (m, x) \sup_{0 \leq x \leq 1, 1 \leq \lfloor b_{n,m} \rfloor \leq \frac{1}{m+s_m-1}} \left| F ((j+1) b_{n,m} - F (j b_{n,m}) + F ((i+1) b_{n,m} - F (i b_{n,m})) \right| \\
+ 4 \sum_{k \in N_{x,m}} b_k (m, x) \\
\leq \frac{D_{n,m} + D_{n,m,1} + 4 \sum_{k \in N_{x,m}} b_k (m, x)}{\text{(say).}} \\
\]

Note that \(D_{n,m,1} = O \left( b_{n,m}^\alpha \right)\) by the assumption that the distribution function \(F\) is Lipshitz of order \(\alpha\). For \(u \leq v\), let

\[
Z_{i, u, v} = I_{u < X_i < v} - (F (v) - F (u)), \; 1 \leq i \leq n.
\]

It is easy to see that \(Z_{i, u, v}, 1 \leq i \leq n\) are i.i.d. random variables with mean zero and

\[
\sum_{i=1}^{n} \text{Var} (Z_{i, u, v}) = n \text{Var} (Z_{1, u, v}) \leq cn \{ u - v \}^{\alpha}
\]

for some constant \(c \geq 2\). Applying Lemma 3.3 (cf. Babu (1989)), we obtain that

\[
P \left( \mid F_n (u) - F_n (v) + F (v) \mid > 4 c b_{n,m}^2 / a_m^\alpha \right) \leq 2 \exp \left\{ - cn b_{n,m}^2 / a_m^\alpha \right\} \quad \ldots (3.7)
\]
Hence

\[ P \left( D_{n,m} > 4cb_{n,m} \right) \leq 2n^2 \exp \left\{ -c n b_{n,m}^2 / a_m^\alpha \right\} \] \quad \ldots \ (3.8)

From the choice of the sequences \( b_{n,m} \) and \( a_m \), it follows that

\[ D_{n,m} = O \left( b_{n,m} \right) \quad \text{a.s.} \]

by an application of the Borel-Cantelli lemma. Let

\[ H_{n,m} = \sup_{0 \leq x \leq 1} \max_{k \in \mathbb{N}_{x,m}} \left| \gamma_{k,m}^F - \gamma_{k,m}^F - F_n(x) + F(x) \right|. \]

\[ \ldots \ (3.9) \]

It now follows by the inequalities obtained above that

\[ H_{n,m} \leq D_{n,m} + D_{n,m,1} = O \left( b_{n,m} \right). \]

\[ \ldots \ (3.10) \]

Under the conditions stated in Corollary 2.3, it follows that

\[ \| C_m F - F \| = \frac{1}{C_m} \left( \frac{N_{x,m}}{m} \right)^\alpha. \]

\[ \ldots \ (3.11) \]

Applying Lemma 3.3 (cf. Babu (1989)) again for a sequence of i.i.d. random variables \( W_i \),

\[ 0 \leq i \leq m \] with \( P (W_i = 1 - x) = x = 1 - P (W_i = -x), \)

\[ a = \left[ (s_m + m - 1) \log (s_m + m - 1) \right]^{1/2}, \]

\[ \nu = m/4 \] and \( s = 1/2, \) we get that

\[ \sum_{k=0, k \in \mathbb{N}_{x,m}} b_k (m, x) \leq \exp \left\{ -a^2 / m \right\}. \]

The above estimates lead to the following theorem giving a bound on the closeness of the estimator \( \hat{F}_{n,m} \) and the empirical distribution function \( F_n \).

**Theorem 3.3** — Suppose that a distribution function \( F \in C [0, 1] \) satisfies the conditions

\[ \omega_{\theta} (g, t) = O (t^{2\alpha}) \]

and

\[ \omega (g, t) = O (t^\alpha) \]
for some $0 < \alpha < 1$. Further suppose that $s_m \geq 1$ and $\lim_{m \to \infty} \frac{s_m}{m} = 0$. Define the sequences $b_{n,m}$ and $a_m$ by the relations

$$a_m = [(s_m + m - 1)^{-1} \log (s_m + m - 1)]^{1/2\alpha}$$

and

$$b_{n,m} = (n^{-1} \log n)^{1/2} [(s_m + m - 1)^{-1} \log (s_m + m - 1)]^{1/4\alpha}.$$ 

Then

$$\left| \left| \hat{F}_{n,m} - F_n \right| \right| = O(b_{n,m}) + O\left(\exp\left(-\frac{\alpha^2}{m}\right)\right) + O\left(\left(\frac{s_m}{m}\right)^\alpha\right).$$

4. ESTIMATION OF DENSITY FUNCTION

Let $X_i, 1 \leq i \leq n$ be i.i.d. random variables with distribution function $F$ and density function $f$ with support as the interval $[0, 1]$. Let $F_n$ be the empirical distribution function based on the observations $X_i, 1 \leq i \leq n$ and define the estimator $\hat{F}_{n,m}$ as an estimator of the distribution function $F$ as given in the previous section. Note that

$$\hat{F}_{n,m}(x) = (C_m F_n)(x)$$

$$= \frac{1}{s_m} \sum_{k=0}^{m} \sum_{j=0}^{s_m - 1} F_{n}\left(\frac{k+j}{m+s_m - 1}\right) b_k(m,x)$$

$$= \sum_{k=0}^{m} \gamma_{k,m} F b_k(m,x)$$

$$= \sum_{k=1}^{m} \left\{\gamma_{k,m} F - \gamma_{k-1,m} F\right\} B_k(m,x) + \gamma_{0,m} F B_0(m,x)$$

$$= \sum_{k=0}^{m} f_{k,m}^{(n)} B_k(m,x)$$

where
\[ f_{0,m}^{(n)} = \gamma_{0,m}^n \]
\[ f_{k,m}^{(n)} = \gamma_{k,m}^n - \gamma_{k-1,m}^n, \]

and
\[ B_k (m, x) = \sum_{j = k}^{m} b_j (m, x), \quad k \geq 1. \]

It is easy to check that \( f_{k,m}^{(n)} \) is nonnegative for \( k = 0, 1, \ldots, m \) and \( B_0 (m, x) = 1 \) for \( 0 \leq x \leq 1 \). In addition
\[ B_k (m, x) = m \binom{m - 1}{k - 1} \int_0^x t^{k-1} (1-t)^{m-k} \, dt \]
for \( k \geq 1 \). Hence \( B_k (m, x) \) is nondecreasing in \( x \) for all \( k \geq 1 \). Let
\[ \hat{f}_{n,m} (x) = \sum_{k=1}^{m} f_{k,m}^{(n)} \frac{d}{dx} B_k (m, x). \]

Note that \( \hat{f}_{n,m} (x) \) is nonnegative and is the derivative of the function \( \hat{F}_{n,m} (x) \) which was considered as an estimator of the distribution function \( F (x) \). Hence we can consider the function \( \hat{f}_{n,m} (x) \) as an estimator for the probability density function \( f(x) \). We will now study the properties of the estimator \( \hat{f}_{n,m} (x) \). Observe that
\[ \hat{f}_{n,m} (x) = \sum_{k=1}^{m} f_{k,m}^{(n)} \frac{d}{dx} B_k (m, x) \]
\[ = m \sum_{k=1}^{m-1} f_{k+1,m}^{(n)} b_k (m-1, x) \]
\[ = m \sum_{k=0}^{m-1} \left( \gamma_{k+1,m}^n - \gamma_{k,m}^n \right) b_k (m-1, x). \]

We will now discuss asymptotic properties of the density estimator \( \hat{f}_{n,m} (x) \) for the probability density function \( f(x) \). Suppose the probability density function \( f \) is Lipschitz of order one.
Observe that

\[
\hat{f}_{n,m}(x) = m \sum_{k=0}^{m-1} \left( \frac{\gamma^n_{k+1,m} - \gamma^n_{k,m}}{\gamma^n_{k+1,m} - \gamma^n_{k,m}} \right) b_k (m-1, x)
\]

\[
= m \sum_{k=0}^{m-1} \left( \frac{\gamma^n_{k+1,m} - \gamma^n_{k,m}}{\gamma^n_{k+1,m} - \gamma^n_{k,m}} \right) b_k (m-1, x)
\]

\[
+ m \sum_{k=0}^{m-1} \left( \frac{\gamma^n_{k+1,m} - \gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}}{\gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}} \right) b_k (m-1, x)
\]

\[
= T_m (x) + T_{m,n} (x) \quad \text{(say)}.
\]

We will first estimate the term \( T_{m,n} (x) \). Let

\[
L_{m,n} = \sup_{0 \leq k \leq m-1} \left| \frac{\gamma^n_{k+1,m} - \gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}}{\gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}} \right|.
\]

It is easy to check that

\[
\frac{\gamma^n_{k+1,m} - \gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}}{\gamma^n_{k+1,m} - \gamma^n_{k,m} + \gamma^n_{k,m}}
\]

\[
= \frac{1}{s_m} \left( F_n \left( \frac{k + s_m}{m + s_m - 1} \right) - F_n \left( \frac{k}{m + s_m - 1} \right) \right)
\]

\[- \frac{1}{s_m} \left( F_n \left( \frac{k + s_m}{m + s_m - 1} \right) - F_n \left( \frac{k}{m + s_m - 1} \right) \right).
\]

Let \( Z_i, 1 \leq i \leq n \) be i.i.d. random variables defined by

\[
Z_i = \frac{1}{s_m} \int_{\frac{k}{m + s_m - 1} \leq X_i \leq \frac{k + s_m}{m + s_m - 1}} - \frac{1}{s_m} \left[ F_n \left( \frac{k + s_m}{m + s_m - 1} \right) - F_n \left( \frac{k}{m + s_m - 1} \right) \right].
\]

Note that \( E(Z_i) = 0 \) and

\[
Var(Z_i) \leq \frac{c_s}{s_m} \left[ F_n \left( \frac{k + s_m}{m + s_m - 1} \right) - F_n \left( \frac{k}{m + s_m - 1} \right) \right]
\]

\[
\leq \frac{c_s}{s_m} \left( \frac{s_m}{m + s_m - 1} \right)
\]
since the density function $f$ is bounded on the interval $[0,1]$. We assume that $c > 2$ without loss of generality.

Let

$$b = 1, \ s = \frac{1}{2}, \ a = 2 \ \text{cnd} n \left( \frac{s m}{m + s m - 1} \right)^{1/2}, \ \frac{1}{s m}$$

and

$$\nu = \text{cn} \left( \frac{s m}{m + s m - 1} \right)^{-2} s m$$

where

$$d_n = (n^{-1} \log n)^{1/2}.$$ 

An application of Lemma 3.3 shows that

$$P \left( L_{m, n} > 2c d_n m^{-1/2} \right) \leq 2m e^{-a^2/4\nu} \leq 2n^{1-c}.$$ 

Hence

$$L_{m, n} = O \left( m^{-1/2} d_n \right) \text{ a.s}$$

as $n \to \infty$ by the Borel-Cantelli lemma which in turn proves that

$$\|T_{m, n}\| = O \left( m^{1/2} d_n \right) \text{ a.s}$$

as $n \to \infty$.

We now estimate the bound on $\|T_m - f\|$. Note that, for $0 \leq x \leq 1$,

$$T_m (x) - f(x) = m \sum_{k=0}^{m-1} \left( \gamma_{k+1, m}^F - \gamma_{k, m}^F \right) b_k (m-1, x) - f(x)$$

$$= \frac{m}{s m} \sum_{k=0}^{m-1} \left[ \left( \frac{k + s_m}{m + s m - 1} \right) - F \left( \frac{k}{m + s m - 1} \right) \right] b_k (m-1, x)$$

$$- \sum_{k=0}^{m-1} f(x) b_k (m-1, x)$$
\[
= \frac{m}{s} \sum_{k=0}^{m-1} \left[ f\left( \frac{k}{m+s-1} \right) \frac{s^2}{m+s-1} + O\left( \left( \frac{s^2}{m+s-1} \right) \right) \right] b_k(m-1, x)
\]

\[+ \frac{m}{s} \sum_{k=0}^{m-1} \frac{s^2}{m+s-1} f(x) b_k(m-1, x) \]

\[
= \frac{m}{s} \sum_{k=0}^{m-1} \left[ f\left( \frac{k}{m+s-1} \right) \frac{s^2}{m+s-1} - \frac{s^2}{m} f(x) \right] b_k(m-1, x)
\]

\[+ \frac{m}{s} O\left( \frac{s^2}{(s+m-1)^2} \right) \]

\[
= \sum_{k=0}^{m-1} \left[ f\left( \frac{k}{m+s-1} \right) \frac{m}{m+s-1} - f(x) \right] b_k(m-1, x)
\]

\[+ \frac{m}{s} O\left( \frac{s^2}{(s+m-1)^2} \right) \]

\[
\frac{m}{m+s-1} \sum_{k=0}^{m-1} \left[ f\left( \frac{k}{m+s-1} \right) - f(x) \right] b_k(m-1, x) + O\left( \frac{s^2}{m+s-1} \right)
\]

\[+ O\left( \frac{ms}{(s+m-1)^2} \right) . \]

Hence

\[
\sup_{0 \leq x \leq 1} |T_m(x) - f(x)| \leq \frac{m}{m+s-1} \sum_{k=0}^{m-1} \left| \frac{k}{m+s-1} - x \right| b_k(m-1, x)
\]

\[+ O\left( \frac{s^2}{m+s-1} \right) + O\left( \frac{ms}{(s+m-1)^2} \right) \]

\[
\leq \frac{m}{m+s-1} \sum_{k=0}^{m-1} \left| \frac{k}{m} - x \right| b_k(m-1, x)
\]
\[ + \frac{m}{m + s} \left( \frac{s - 1}{m (m + s)} \right) \sum_{k=0}^{m-1} k b_k (m - 1, x) \]

\[ + O \left( \frac{s - 1}{m + s} \right) + O \left( \frac{m s}{(s + m - 1)^2} \right) \]

\[ \leq \frac{m}{m + s} \ O (m^{-1/2}) + \frac{m}{m + s} \frac{s - 1}{m (m + s)} \ O (m) \]

\[ + O \left( \frac{s - 1}{m + s} \right) + O \left( \frac{m s}{(s + m - 1)^2} \right) \]

As a consequence of the above estimates, we obtain the following theorem.

**Theorem 4.1** — Suppose the probability density function \( f \) is Lipschitz of order one. Define \( \hat{f}_{n,m}(x) \) as defined above as an estimator of the probability density function \( f(x) \). Then

\[ \left| \left| \hat{f}_{n,m} - f \right| \right| = O \left( m^{1/2} d_n \right) + O \left( \frac{m^{1/2} + s - 1}{m + s - 1} \right) + O \left( \frac{m s}{(m + s - 1)^2} \right). \]

**Remarks**: Note that the second and the third terms in the bound given above reduce to terms of the order \( O (m^{1/2}) \) and \( O (m^{-1}) \) if \( s_m = 1 \) for all \( m \geq 1 \) which lead to the bound for the classical Bernstein polynomial density estimator as discussed in Babu et al. (2002).

We now discuss about the limiting distribution of the estimator \( \hat{f}_{n,m}(x) \) after proper scaling. Recall that

\[ \hat{f}_{n,m}(x) - f(x) = T_m(x) - f(x) + T_{m,n}(x) \]
\begin{align*}
&= O \left( \frac{m^{1/2}}{m + s_m - 1} \right) + O \left( \frac{m (s_m - 1)}{(m + s_m - 1)^2} \right) \\
&+ O \left( \frac{s_m - 1}{m + s_m - 1} \right) + O \left( \frac{m s_m}{(m + s_m - 1)^2} \right) + T_{m, n}(x)
\end{align*}

Let
\[
a_{k, m} = F \left( \frac{k + s_m}{m + s_m - 1} \right) - F \left( \frac{k}{m + s_m - 1} \right),
\]
and
\[
Y_{i, m} = \sum_{k=0}^{m-1} \left( \frac{1}{s_m} \left\{ \frac{k}{m + s_m - 1} < X_i < \frac{k + s_m}{m + s_m - 1} \right\} - \frac{1}{s_m} a_{k, m} \right) b_k(m-1, x).
\]

It is easy to check that $Y_{i, m}, 1 \leq i \leq n$ are i.i.d. random variables for any $m \geq 1$ and
\[
T_{m, n}(x) = \frac{m}{n} \sum_{i=1}^{n} Y_{i, m}.
\]

Therefore
\[
\text{Var}(T_{m, n}(x)) = m^2 n^{-1} \text{Var}(Y_{1, m}).
\]

We will now compute $\text{Var}(Y_{1, m})$. Note that
\[
\text{Var}(Y_{1, m}) = s_m^{-2} \left[ \sum_{k=0}^{m-1} b_k(m-1, x) a_{k, m} - \left( \sum_{k=0}^{m-1} b_k(m-1, x) a_{k, m} \right)^2 \right].
\]

Note that
\[
\frac{m}{s_m} \sum_{k=0}^{m-1} b_k(m-1, x) a_{k, m} = T_m(x)
\]
\[
= f(x) + O \left( \frac{m^{1/2}}{m + s_m - 1} \right) + O \left( \frac{m (s_m - 1)}{(m + s_m - 1)^2} \right)
\]
\[ + O \left( \frac{s_m - 1}{m + s_m - 1} \right) + O \left( \frac{ms_m}{(s_m + m - 1)^2} \right). \]

Furthermore

\[
\frac{m}{s_m} a_{k,m} = \frac{m}{s_m} \left[ F \left( \frac{k+s_m}{m+s_m-1} \right) - F \left( \frac{k}{m+s_m-1} \right) \right]
\]

\[
= \frac{m}{s_m} \frac{s_m}{m+s_m-1} f \left( \frac{k+s_m}{m+s_m-1} \right) + \frac{m}{s_m} O \left( \frac{k}{m+s_m-1} \right)^2 \]

\[
= \frac{m}{m+s_m-1} \left[ f(c) + O \left( \left| \frac{k}{m+s_m-1} - x \right| \right) \right] + O \left( \frac{ms_m}{(m+s_m-1)^2} \right)
\]

\[
= \frac{m}{m+s_m-1} \left[ f(x) + O \left( \left| \frac{k}{m+s_m-1} - k \right| \right) + O \left( \left| \frac{k}{m} - x \right| \right) \right]
\]

\[
+ O \left( \frac{ms_m}{(m+s_m-1)^2} \right)
\]

\[
= \frac{m}{m+s_m-1} f(x) + \frac{m}{m+s_m-1} O \left( \left| \frac{k}{m} - x \right| \right)
\]

\[
+ k O \left( \left| \frac{s_m - 1}{(m+s_m-1)^2} \right| \right) + O \left( \frac{ms_m}{(m+s_m-1)^2} \right)
\]

\[
= f(x) + \sum_{k=0}^{m-1} \frac{s_m - 1}{m+s_m-1} a_{k,m} = \sum_{k=0}^{m-1} \frac{b_k^2 (m-1, x) a_{k,m}}{s_m} \sum_{k=0}^{m-1} b_k^2 (m-1, x) a_{k,m}
\]
\[
\begin{align*}
&= \frac{1}{s} \sum_{k=0}^{m-1} b_k^2(m-1, x) f(x) + O \left( \frac{s m - 1}{m + s m - 1} \right) \\
&+ \frac{m}{m + s m - 1} O \left( \frac{k}{m} - x \right) k O \left( \frac{s m - 1}{(m + s m - 1)^2} \right) + O \left( \frac{m s m}{(m + s m - 1)^2} \right) \\
&= \frac{1}{s} m^{-1/2} \gamma(x) + \left( \frac{s m - 1}{s m (m + s m - 1)} \right) \\
&+ O \left( \frac{m}{(m + s m - 1)^2} \right)
\end{align*}
\]

where \( \gamma(x) = f(x) (4\pi x (1-x))^{-1/2} \) by the estimates derived in Babu et al. (2002). Combining the bounds derived above, we get that

\[
Var(T_{m,n}(x)) = m^2 n^{-1} Var(Y_{1,m})
\]

\[
= m^2 n^{-1} \frac{1}{s} \sum_{k=0}^{m-1} b_k^2(m-1, x) a_{k,m} - \left( \sum_{k=0}^{m-1} b_k(m-1, x) a_{k,m} \right)^2
\]

\[
= mn^{-1} \left[ \frac{1}{s} m^{-1/2} \gamma(x) + O \left( \frac{s m - 1}{s m (m + s m - 1)^2} \right) \right. \\
+ \frac{m}{s m (m + s m - 1)} O(m^{-3/4}) + O \left( \frac{s m - 1}{s m (m + s m - 1)^2} \right) \\
+ O \left( \frac{m}{(m + s m - 1)^2} \right) \\
\left. - n^{-1} \left[ f(x) + O \left( \frac{m^{1/2}}{m + s m - 1} \right) + O \left( \frac{m(s m - 1)}{(m + s m - 1)^2} \right) \right] \right]
\]

\[
+ O \left( \frac{s m - 1}{m + s m - 1} \right) + O \left( \frac{m s m}{(m + s m - 1)^2} \right)^2
\]

Suppose the sequence \( 1 \leq s m \leq m \) is chosen such that \( \frac{s m}{\sqrt{m}} \to 0 \) as \( m \to \infty \). It can be checked that
\[ \text{Var}(T_{m,n}(x)) = \frac{m^{1/2} n^{-1}}{s_m} \left[ \gamma(x) + O(m^{-1/4}) + O(m^{-3/2}(s_m - 1)) \right] \\
- n^{-1} \left[ f(x) + O(m^{-1/2}) + O \left( \frac{s_m}{m} \right)^2 \right] \\
= \frac{m^{1/2} n^{-1}}{s_m} \left[ \gamma(x) + O(m^{-1/4}) + O(m^{-3/2}(s_m - 1)) + \frac{s_m}{m^{1/2}} O(1) \right] \]

Observe that

\[ n^{1/2} s_m^{1/2} m^{-1/4} T_{m,n}(x) = n^{1/2} s_m^{1/2} m^{-1/4} \frac{m}{n} \sum_{i=1}^{n} Y_{i,m} \]

Let

\[ W_{i,m} = n^{-1/2} s_m^{1/2} m^{3/4} Y_{i,m}, 1 \leq i \leq n. \]

Note that

\[ \left| W_{i,m} \right| \leq n^{-1/2} s_m^{1/2} m^{3/4} \left| Y_{i,m} \right| \]

\[ \leq n^{-1/2} s_m^{1/2} m^{3/4} \frac{1}{s_m} \max_{0 \leq k \leq m-1} (a_{k,m}, b_k(m-1, x)) \]

\[ = n^{-1/2} s_m^{1/2} m^{3/4} \frac{1}{s_m} \left[ O \left( \frac{s_m}{m+s_m-1} \right) + O(m^{-1/4}) \right] \]

\[ = n^{-1/2} s_m^{-1/2} m^{3/4} \left[ O \left( \frac{s_m}{m} \right) + O(m^{-1/4}) \right]. \]

Note that \( W_{i,m}, 1 \leq i \leq n \) are i.i.d. random variables with mean zero. Furthermore

\[ \mathbb{E} \left| W_{1,m} \right|^3 \leq n^{-3/2} s_m^{-3/2} m^{9/4} \left[ O \left( \frac{3}{m^3} \right) + O(m^{-3/4}) \right]. \]

Suppose that \( f(x) > 0 \). Then

\[ \text{Var}(W_{1,m}) = n^{-1} s_m^{3/2} \text{Var}(Y_{1,m}) \geq c_1 n^{-1} \]
for large $n$ and $m$ such that $s_m/\sqrt{m} \to 0$ as $m \to \infty$ for some positive constant $c_1$. Hence

$$\text{Var} \left( \sum_{i=1}^{n} W_{i,m} \right) = n \text{Var} (W_{1,m}) \geq c_1$$

for large $n$ and $m$ such that $s_m/\sqrt{m} \to 0$ as $m \to \infty$. Therefore

$$\sum_{i=1}^{n} \frac{E \{ W_{i,m} \}^3}{\text{Var} \left( \sum_{i=1}^{n} W_{i,m} \right)^{3/2}} \leq c_2 n^{-1/2} m^{3/2} s_m^{-3/2} m^{9/4} \left[ O \left( \frac{s_m^3}{m^3} \right) + O \left( m^{-3/4} \right) \right]$$

$$= o(1)$$

for some positive constant $c_2$ provided

$$n^{-1} \left( \frac{m}{s_m} \right)^{3/4} \to 0$$

as $m$ and $n$ tend to infinity such that $s_m/\sqrt{m} \to 0$. Hence the sequence of random variables $W_{i,m}, 1 \leq i \leq n$ satisfies the Lyapunov's condition (cf. Billingsley (1986), p. 371) provided

$$n^{-1} \left( \frac{m}{s_m} \right)^{3/4} \to 0$$

as $m$ and $n \to \infty$ such that $s_m/\sqrt{m} \to 0$. This shows that the sum $\sum_{i=1}^{n} W_{i,m}$ is asymptotically normal.

Hence we have the following result.

**Theorem 4.2** — Suppose the sequence $1 \leq s_m \leq m$ is chosen such that $s_m/\sqrt{m} \to 0$ as $m \to \infty$. Further suppose that

$$n^{-1} \left( \frac{m}{s_m} \right)^{3/4} \to 0$$

as $m$ and $n \to \infty$. Then

$$n^{1/2} \frac{1}{s_m} m^{-1/4} T_{m,n}(x) \overset{L}{\to} N(0, \gamma(x)).$$
Remarks: It is possible to weaken the above condition by using $2 + \delta$-th moment instead of the third moment in the above calculations for some $\delta > 0$. We will not go into this discussion here.

We have seen earlier that

$$\| T_m(x) - f(x) \| = O(m^{-1/2}) + O \left( \frac{s}{m} \right).$$

Hence

$$n^{1/2} s^{-1/2} m^{-1/4} (T_m(x) - f(x)) = n^{1/2} s^{-1/2} m^{-1/4} \left[ O(m^{-1/2}) + O \left( \frac{s}{m} \right) \right].$$

Suppose further that

$$ns_m m^{-3/2} \to 0$$

as $m$ and $n$ tend to infinity. It is easy to check that the term on the right side of the above equation goes to zero as $m$ and $n$ tend to infinity under the conditions stated above. Hence we have the following theorem proving the asymptotic normality of the estimator $\hat{f}_{n,m}(x)$.

Theorem 4.3 — Suppose the sequence $1 \leq s_m \leq m$ is chosen such that $\frac{s}{\sqrt{m}} \to 0$.

$$n^{-1} \left( \frac{m}{s_m} \right)^{3/4} = o(1) \quad \text{ ... (4.3)}$$

and

$$ns_m m^{-3/2} = o(1)$$

as $m$ and $n \to \infty$. Then

$$n^{1/2} s^{-1/2} m^{-1/4} \left( \frac{\hat{f}_{n,m}(x) - f(x)}{\sqrt{L}} \right) \to N(0, \gamma(x)).$$

Theorem 4.4 — Suppose that $s_m = m^\tau$ with $0 < \tau < \frac{1}{2}$ and $m$ and $n$ tend to infinity such that

$$n^{-1} m^{3/4 (1 - \tau)} = o(1).$$

Further suppose that the probability density function $f \in C^1 [0, 1]$, the derivative $f'$ is continuous in the interval $[0, 1]$ and
\[
\frac{1}{2} \left( \frac{3r}{2} - \frac{5}{4} \right) n^m \rightarrow \delta > 0
\]

as \( m \) and \( n \) tend to infinity. Then

\[
n^{1/2} \left( \frac{1}{s_m} \right)^{1/4} \left( \hat{f}_{n,m}(x) - f(x) \right) = \delta (2x - 1)f'(x)
\]

is asymptotically normal with mean zero and variance \( \gamma(x) \).

**Proof**: Let

\[
\beta_{m,n} = n^{1/2} \left( \frac{1}{s_m} \right)^{1/4} m^{-1/4}.
\]

Sharpening the inequality for \( \|T_m - f\| \) obtained earlier under the additional assumption on the existence of continuous derivative \( f' \) of \( f \) over the interval \([0, 1]\), observe that

\[
\beta_{m,n} \left( \hat{f}_{n,m}(x) - f(x) \right) = \beta_{m,n} T_{m,n}(x) + \beta_{m,n} (T_m(x) - f(x))
\]

\[
= \beta_{m,n} T_{m,n}(x) + \delta m^{1-\tau} (T_m(x) - f(x)) + o(1)
\]

\[
= \beta_{m,n} T_{m,n}(x) + \delta (2x - 1)f'(x) + o(1)
\]

as \( m \) and \( n \) tend to infinity by Theorem 2.4. As a Corollary to Theorem 4.2, we now obtain the result stated in the theorem by an application of Slutsky’s lemma.

**Remarks**: It can be seen from Theorem 4.3 that the rate of convergence of the density estimator \( \hat{f}_{n,m}(x) - f(x) \) is of the order \( n^{-1/2} s_m^{-1/2} m^{1/4} \) as opposed to the rate \( n^{-1/2} m^{1/4} \) for the estimator developed by Babu et al. (2002) using the classical Bernstein polynomials. If the sequence \( s_m \) is chosen such that \( s_m = 1 \) for all \( m \geq 1 \), then one obtains the estimator developed by Babu et al. (2002).

5. **Bayesian Density Estimation**

Following the methods developed by Petrone (1999), we now propose a Bayesian nonpara-metric procedure for density estimation for probability density functions with support in the interval \([0, 1]\) using a prior based on generalized Bernstein polynomials. As our methods are akin to those in Petrone (1999), we sketch the results.

Following the notation in Section 4, define

\[
w_{k,m} = \gamma_{k+1,m} - \gamma_{k,m}, 0 \leq k \leq m - 1
\]

and
\[ b(x; m, w_0, m, ..., w_{m-1}, m) = T_m(x) \]
\[ = m \sum_{k=0}^{m-1} \left( \gamma_{k+1, m}^F - \gamma_{k, m}^F \right) b_k(m-1, x). \]

We will call the above function as a generalized Bernstein density with parameters \( m \) and \( w_m = (w_0, m, ..., w_{m-1}, m). \) We have proved earlier that

\[ \| T_m(x) - f(x) \| = O(m^{-1/2}) + O\left(\frac{s}{m}\right). \]

In particular, it follows that

\[ \lim_{m \to \infty} b(x; m, w_0, m, ..., w_{m-1}, m) = f(x) \]

provided \( \frac{s}{m} \to 0 \) as \( m \to \infty \) for any \( x \in [0, 1]. \) In fact the convergence is uniform in \( x \in [0, 1]. \)

Suppose we choose the prior for \( M \) as a discrete distribution over the set of positive integers with \( P(M = m) = p_m > 0 \) for all \( m \geq 1. \) Further suppose that \( W_m = (W_0, m, ..., W_{m-1}, m) \) is a random vector such that the conditional distribution of \( W_m \) given \( M = m \) is a Dirichlet distribution with parameters \( \alpha_{0, m}, ..., \alpha_{m-1, m}. \) We assume that \( \alpha_{j, m} > 0 \) for all \( 0 \leq j \leq m-1 \) and \( m \geq 1. \) We will consider a special class of parameters \( \alpha_{j, m} > 0 \) for all \( 0 \leq j \leq m-1 \) and \( m \geq 1 \) given by

\[ \alpha_{k, m} = J_m \left( \gamma_{k+1, m}^F - \gamma_{k, m}^F \right), \quad 0 \leq k \leq m-1 \]

where \( J_m \) is a positive constant and \( F_0 \) is a distribution function with support as the interval \([0,1].\)

It is easy to see that the posterior distribution of the random vector \((M, W_m)\) given the i.i.d. observations \( x_i, 1 \leq i \leq n \) is proportional to

\[ p_m D(w_0, m, ..., w_{m-1}, m \mid m) \prod_{i=1}^n b(x; m, w_0, m, ..., w_{m-1}, m) \]

where \( D(w_0, m, ..., w_{m-1}, m \mid m) \) denotes the probability density of the Dirichlet distribution with parameters \( \alpha_{0, m}, ..., \alpha_{m-1, m} \) conditioned on the event \([M = m].\)

Following Petrone (1999), we can reformulate the model, by introducing auxiliary random variables \((Z_1, m, ..., Z_n, m),\) as follows:
(i) let $P (M = m) = p_m$ and suppose that the conditional distribution of $W_m$ given $M = m$ is $D (w_m; \alpha_{0,m}, ..., \alpha_{m-1,m})$;

(ii) let $Z_{1,m}, ..., Z_{n,m}$ be random variables conditionally i.i.d. given $(m, w_m)$ with

$$P (Z_{i,m} = j | m, w_m) = w_{j,m} \quad 0 \leq j \leq m-1;$$

and

(iii) let the conditional density of $(X_1, ..., X_n)$ given $m, w_m, z_{1,m}, ..., z_{n,m}$ be given by

$$m^n \prod_{i=1}^{n} b_{z_{i,m}} (m-1, x_i).$$

Combining the models described in (ii) and (iii), it follows that the $X_1, ..., X_n$, given $m, w_m$, are i.i.d. according to the mixture density

$$b (x; m, w_{0,m}, ..., w_{m-1,m}) = T_m (x)$$

$$= m \sum_{k=0}^{m-1} \left( \frac{\gamma_{k+1,m}}{\gamma_{k,m}} - \frac{\gamma_{k+1,m}}{\gamma_{k,m}} \right) b_k (m-1, x).$$

Let $r_{j,m}$ denote the number in $z_{1,m}, ..., z_{n,m}$ which are equal to $j$. Let $R_m = (R_{0,m}, ..., R_{m-1,m})$. It can be shown that the conditional distribution of $W_m$ given $m, x_1, ..., x_n$ is a mixture of Dirichlet distributions given by

$$\sum_{r_{0,m}, ..., r_{m-1,m}} D (\alpha_{0,m} + r_{0,m}, ..., \alpha_{m-1,m} + r_{m-1,m}) \frac{p (r_{0,m}, ..., r_{m-1,m}, m, k, x_1, ..., x_n)}{p (r_{0,m}, ..., r_{m-1,m})} \quad (5.1)$$

where $p (r_{0,m}, ..., r_{m-1,m}, m, x_1, ..., x_n)$ is the conditional probability function of the frequencies $(R_{0,m}, ..., R_{m-1,m})$. Furthermore the posterior probability function of $M$ is

$$p (m | x_1, ..., x_n) \propto \sum_{r_{0,m}, ..., r_{m-1,m}} E \left( W_{0,m}^{r_{0,m}} ... W_{m-1,m}^{r_{m-1,m}} \right) \prod_{i=1}^{n} b_{z_{i,m}} (m-1, x_i)$$

$$... (5.2)$$

The moments of the Dirichlet distribution in the above expression can be computed by using the formula
\[ E\left( W_{r_0}^{0, m} \ldots W_{r_{m-1}}^{m-1, m} \right) = \frac{\alpha_{r_0}^{(r_0)} \ldots \alpha_{r_{m-1}}^{(r_{m-1})}}{f^{(n)}_m} \]

where

\[ a^{(j)} = a(a + 1) \ldots (a + j - 1) \]

for any integer \( j \geq 1 \) and \( a^{(0)} = 1 \). (cf. Wilks (1962)).

We compute the density estimator conditionally on \( M = m \) and then the unconditional density estimator will be obtained by taking expectation with respect to the posterior distribution of \( M \).

A priori, conditionally on \( m \), the density of \( X_i \) is

\[ m \sum_{j=0}^{m-1} E(W_{j,m}) b_j(m-1, x) = b(x; m, E(W_m)). \]

Hence the conditional density of \( X_i \) given \( M = m \) is a generalized Bernstein density with parameters \( m \) and \( E(W_m) \). Following the computations in Section 3.3 of Petrone (1999), one can show that, given the i.i.d. sample \( x_1, \ldots, x_n \) and \( m \), the conditional density of \( X_{n+1} \) is given by

\[ h(x | m, x_1, \ldots, x_n) = \frac{J_m}{J_{m+n}} b(x; m, E(W_m | m)) + \frac{n}{J_{m+n}} b\left( x; m, E\left( \frac{R_m}{n} | m, x_1, \ldots, x_n \right) \right) \]

which can be considered as a generalized Bayesian density estimator. Observe that the unconditional density of \( X_i \) is given by

\[ \sum_{m=1}^{\infty} b(x; m, E(W_m)) p_m. \]

Given the i.i.d. observations \( x_1, \ldots, x_n \), the predictive density of \( X_{n+1} \) has an analogous expression given by

\[ h(x | x_1, \ldots, x_n) = \sum_{m=1}^{\infty} b(x; m, E(W_m | m, x_1, \ldots, x_n)) p(m | x_1, \ldots, x_n). \]

It is easy to check that the conditional density of \( X_{n+1} \) given \( x_1, \ldots, x_n \) is given by
\[ h(x | x_1, \ldots, x_n) = \sum_{m=1}^{\infty} \left[ \frac{J_m}{J_m + n} b(x; m, E(W_m | m)) p(m, x_1, \ldots, x_n) \right. \\
+ \left. \frac{1}{J_m + n} b\left(x; m, E\left(R_m | m, x_1, \ldots, x_n\right)\right) p(m, x_1, \ldots, x_n) \right]. \]

This estimator can be proposed as a generalized Bayesian density estimator.

Remarks: Rates of convergence, for the Bayesian density estimator proposed by Petrone (1999), were studied in Ghosal (2001). It would be interesting to obtain similar results for the generalized Bayesian density estimator discussed here. We conjecture that the rates of convergence for the generalized Bayesian density estimator based on the generalized Bernstein polynomials will be sharper than those for the Bayesian density estimator proposed by Petrone (1999) by suitable choice of the sequence \( \{s_m\} \) as in the case of density estimators discussed in Section 4. After completing this work, the author came to know the recent work of Kakizawa (2004) which suggests other types of estimators for density estimation still based on classical Bernstein polynomials. It is an open problem to find whether one can bring out estimators with faster rates of convergence by using generalized Bernstein polynomials and improve the results in Kakizawa (2004).

REFERENCES