JOURNAL OF APPROXIMATION THEORY **80**, 367–377 (1995)

Reverse Martingales and Approximation Operators

RASUL A. KHAN

Department of Mathematics, Cleveland State University, Cleveland, Ohio 44115, U.S.A.

Communicated by Paul L. Butzer

Received March 29, 1993; accepted in revised form February 11, 1994

Let $\{\xi_n, \mathscr{F}_n, n \geqslant m \geqslant 1\}$ be a reverse martingale such that the distribution of ξ_n depends on $x \in I \subset R = (-\infty, \infty)$ for each $n \geqslant m$, and $\xi_n \xrightarrow{\text{a.s.}} x$. For a continuous bounded function f on R let $L_n(f, x) = Ef(\xi_n)$ be the associated positive linear operator. The properties of ξ_n are used to obtain the convergence properties of $L_n(f, x)$, and some more details are given when ξ_n is a reverse martingale sequence of $\mathscr{U}$-statistics. Lipschitz properties for a subclass of these operators resulting from an exponential family of distributions are also given. It is further shown that this class of operators of convex functions preserves convexity also. An example of a reverse supermartingale related to the Bleimann–Butzer–Hahn operator is also discussed. © 1995 Academic Press, Inc.

1. INTRODUCTION

Probabilistic methods have proved quite useful in the theory of approximation operators. The object of this note is to exploit reverse martingales to unify some properties of approximation operators. A reverse martingale sequence is defined as follows. Let $(\Omega, \mathscr{F}, P)$ be a probability space and let $\{\mathscr{F}_n, n \geqslant m\}$ be a decreasing sequence of σ-fields of sets of $\mathscr{F}$. A sequence of random variables $\{\xi_n, n \geqslant m\}$ or more precisely $\{\xi_n, \mathscr{F}_n, n \geqslant m\}$ is said to be a reverse martingale if $\forall n \geqslant m$, ξ_n is $\mathscr{F}_n$-measurable, $E|\xi_n| < \infty$, and $E(\xi_n \mid \mathscr{F}_{n+1}) = \xi_{n+1}$ a.s. Suppose that the distribution of ξ_n depends on $x \in I \subset R = (-\infty, \infty)$ for each $n \geqslant m$. The interval I depends on the nature of the distribution. For example, in Bernoulli distribution with success probability x, $I = (0, 1)$, and in Gaussian distribution with expectation x, $I = (-\infty, \infty)$. It is well known (cf. [3, p. 18]) that $\{\xi_n, n \geqslant m\}$ necessarily converges to a finite limit which is assumed to be x. Let $C_B(R)$ be a class of continuous bounded functions f on R. For $f \in C_B(R)$ let $L_n(f, x) = Ef(\xi_n)$ be the associated positive linear operator. The properties of $L_n(f, x)$ such as convergence, rate of convergence, and monotonic convergence are obtained from the properties of ξ_n.

0021-9045/95 $6.00

A particularly interesting large class of reverse martingales is given by a sequence of $\mathscr{U}$-statistics. A special case is the well known sequence $\xi_n = (Y_1 + \cdots + Y_n)/n$ where $Y_1, Y_2, \ldots$ are iid random variables with a common mean x. If the distribution of Y_1 is a member of an exponential family of distributions (cf. Lehmann [10]), we obtain the Lipschitz property of the operator $L_n(f, x)$ when f satisfies a Lipschitz condition. This class of operators also preserves convexity for convex functions.

Section 2 gives the properties of $L_n(f, x)$ for a general reverse martingale sequence ξ_n. Moreover, these results are made more precise when ξ_n is a sequence of $\mathscr{U}$-statistics. In Section 3 we prove the Lipschitz property of $L_n(f, x)$ when $\xi_n = (Y_1 + \cdots + Y_n)/n$ and Y_1 has a general exponential distribution. Finally, in Section 4 we discuss an example of a reverse supermartingale related to the Bleimann–Butzer–Hahn operator.

2. Properties of $L_n(f, x)$

In what follows $\omega(f, \delta)$ denotes the usual modulus of continuity of f. The general properties of $L_n(f, x)$ are given by the following.

Theorem 1. *Let $\{\xi_n, \mathscr{F}_n, n \geqslant m\}$ be a square integrable reverse martingale such that the distribution of ξ_n depends on $x \in I \subset R$ $\forall n \geqslant m$ and $E\xi_1 = x$. For $f \in C_B(R)$ let $L_n(f, x) = Ef(\xi_n)$ and $\sigma_n^2(x) = E(\xi_n - x)^2$. Then*

(i) $\lim_{n \to \infty} L_n(f, x) = f(x)$ *for each* $x \in I$.

(ii) $|L_n(f, x) - f(x)| \leqslant 2\omega(f, \sigma_n(x))$.

(iii) $L_n(f, x) \geqslant L_{n+1}(f, x) \geqslant \cdots \geqslant f(x)$ *provided f is convex.*

Proof. A reverse martingale necessarily converges to a finite limit (cf. [3, p. 18]), and since $E\xi_1 = x$, hence $\xi_n \xrightarrow{\text{a.s.}} x$ as $n \to \infty$. Moreover, since f is continuous and bounded, (i) follows from the bounded convergence theorem. To prove (ii), let $\delta > 0$ and set $\lambda = [|\xi_n - x|/\delta]$ where $[y]$ denotes the greatest integer $\leqslant y$. Clearly,

$$|f(\xi_n) - f(x)| \leqslant (1 + \lambda)\, \omega(f, \delta),$$

and

$$|L_n(f, x) - f(x)| \leqslant E\,|f(\xi_n) - f(x)| \leqslant E(1 + \lambda)\, \omega(f, \delta).$$

Since λ is an integer-valued random variable, we have

$$E(1 + \lambda) \leqslant (1 + E\lambda^2) \leqslant \left(1 + \frac{E(\xi_n - x)^2}{\delta^2}\right) = \left(1 + \frac{\sigma_n^2(x)}{\delta^2}\right),$$

or the Cauchy–Schwarz inequality gives

$$E(1+\lambda) \leqslant (1+\sqrt{E\lambda^2}) \leqslant \left(1+\frac{\sigma_n(x)}{\delta}\right).$$

Hence (ii) follows by choosing $\delta=\sigma_n(x)$. To prove (iii) we note that

$$L_n(f, x) = Ef(\xi_n) = E(E(f(\xi_n) \mid \mathscr{F}_{n+1})),$$

and the conditional version of Jensen's inequality gives

$$L_n(f, x) \geqslant Ef(E(\xi_n \mid \mathscr{F}_{n+1})) = Ef(\xi_{n+1}) = L_{n+1}(f, x).$$

This completes the proof.

We now specialize ξ_n to an important and a rather large class of reverse martingales given by a sequence of $\mathscr{U}$-statistics. To define it let $Y_1, Y_2, ...$ be iid (independent and identically distributed) random variables with a distribution function $F(y)$, and let $\phi(y_1, ..., y_m)$ be a real symmetric measurable function on R^m such that $E|\phi(Y_1, ..., Y_m)| < \infty$ where $m(\geqslant 1)$ is a fixed integer. Then the sequence of $\mathscr{U}$-statistics is defined by

$$\mathscr{U}_n = \binom{n}{m}^{-1} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant n} \phi(Y_{i_1}, ..., Y_{i_m}), \qquad n \geqslant m. \tag{1}$$

Let $\mathscr{F}_n = \mathscr{B}(\mathscr{U}_n, \mathscr{U}_{n+1}, ...)$ be a sequence of σ-fields generated by $\mathscr{U}_n$, $\mathscr{U}_{n+1}, ...$. It is well known (cf. [2, p. 377]) that $\{\mathscr{U}_n, \mathscr{F}_n, n \geqslant m\}$ is a reverse martingale. Now let $x = E\phi(Y_1, ..., Y_m)$ be a real functional of F such that $x \in I \subset R$. For $f \in C_B(R)$ define the operator $L_n(f, x) = Ef(\mathscr{U}_n)$. Clearly, $E\mathscr{U}_n = x$ $\forall n \geqslant m$ and $\mathscr{U}_n \xrightarrow{\text{a.s.}} x$ as $n \to \infty$. Consequently, Theorem 1 holds for the sequence $\{\mathscr{U}_n, \mathscr{F}_n, n \geqslant m\}$ if ξ_n is replaced by $\mathscr{U}_n$ and ϕ is square integrable.

To gain some more insight we use a fundamental result due to Hoeffding (cf. [4, 11]). Let

$$\phi_j(y_1, ..., y_j) = E\phi(y_1, ..., y_j, Y_{j+1}, ..., Y_m), \qquad 1 \leqslant j \leqslant m-1,$$

and

$$\phi_m(y_1, ..., y_m) = \phi(y_1, ..., y_m).$$

Note that $E\phi_j(Y_1, ..., Y_j) = E\phi(Y_1, ..., Y_m) = x$, and set

$$\rho_j = \operatorname{var}(\phi_j(Y_1, ..., Y_j)) = E(\phi_j - x)^2, \qquad 1 \leqslant j \leqslant m.$$

The following properties of $\mathscr{U}$-statistics are due to Hoeffding (cf. [4, 11]).

THEOREM (Hoeffding). *Let $\mathscr{U}_n$ be defined by* (1) *and assume that ϕ is square integrable. Then*

(i) $\dfrac{m^2\rho_1}{n} \leqslant \operatorname{var}(\mathscr{U}_n) \leqslant \dfrac{m\rho_m}{n}$

(ii) $\operatorname{var}(\mathscr{U}_n) = n^{-1}m^2\rho_1 + O(n^{-2})$

(iii) *$n \operatorname{var}(\mathscr{U}_n)$ is a non-increasing function of n*

(iv) $\operatorname{var}(\mathscr{U}_m) = \rho_m$, $\lim_{n\to\infty} n \operatorname{var}(\mathscr{U}_n) = m^2\rho_1$

(v) $\lim_{n\to\infty} P\left(\dfrac{\sqrt{n}(\mathscr{U}_n - x)}{m\sqrt{\rho_1}} \leqslant y\right) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-t^2/2)\, dt.$

From these properties one easily obtains the following.

THEOREM 2. *Let $L_n(f, x) = Ef(\mathscr{U}_n)$ where $\mathscr{U}_n$ is defined by* (1) *and ϕ is square integrable. Then*

(a) $|L_n(f, x) - f(x)| \leqslant (1 + m\rho_m)\, \omega(f, 1/\sqrt{n})$.

(b) *Assume that $E\phi^4(Y_1, ..., Y_m) < \infty$. If f' and f'' are continuous and f'' is bounded, then*

$$\lim_{n\to\infty} n(L_n(f, x) - f(x)) = \frac{m^2\rho_1}{2} f''(x).$$

Moreover, the same conclusion holds if the first three derivatives are continuous and f''' is bounded.

Proof. Let $\lambda = [|\mathscr{U}_n - x|/\delta]$, $\delta > 0$, and recall that

$$|L_n(f, x) - f(x)| \leqslant (1 + E\lambda)\, \omega(f, \delta).$$

Since $E\lambda \leqslant E\lambda^2 \leqslant E(\mathscr{U}_n - x)^2/\delta^2 \leqslant m\rho_m/n\delta^2$, (a) follows by choosing $\delta = n^{-1/2}$. To prove (b) let $h_n = \mathscr{U}_n - x$ and obtain the Taylor expansion

$$f(\mathscr{U}_n) = f(x) + h_n f'(x) + \frac{h_n^2}{2} f''(x) + R_n,$$

where

$$R_n = \frac{h_n^2 v_n}{2}, \qquad v_n = (f''(x + \theta h_n) - f''(x)), \quad 0 < \theta < 1.$$

Consequently, taking expectation we have

$$L_n(f, x) = f(x) + \frac{Eh_n^2}{2} f''(x) + ER_n.$$

It should be noted that $h_n \xrightarrow{\text{a.s.}} 0$ and $v_n \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. Let $\varepsilon > 0$. Since f'' is continuous, there is $\delta > 0$ such that $|v_n| < \varepsilon$ whenever $|h_n| < \delta$. Also, $|v_n| \leqslant M$ due to the boundedness of f''. Using I as the indicator function we have

$$\begin{aligned} E\,|R_n| &= E\,|R_n|\, I(|h_n| < \delta) + E\,|R_n|\, I(|h_n| \geqslant \delta) \\ &\leqslant \frac{\varepsilon\, Eh_n^2}{2} + \frac{M}{2} Eh_n^2 I(|h_n| \geqslant \delta) \\ &\leqslant \frac{\varepsilon\, m\rho_m}{2n} + \frac{M}{2} Eh_n^2 I(|h_n| \geqslant \delta). \end{aligned}$$

Since $E\phi^4(Y_1, ..., Y_m) < \infty$, it is well known (cf. [9, p. 21; 11, p. 9]) that

$$Eh_n^4 = E(\mathscr{U}_n - x)^4 = O(n^{-2}).$$

Now the Cauchy–Schwarz inequality gives

$$Eh_n^2 I(|h_n| \geqslant \delta) \leqslant \sqrt{Eh_n^4 P(|h_n| \geqslant \delta)},$$

and by the Markov inequality we have

$$P(|h_n| \geqslant \delta) \leqslant \frac{Eh_n^4}{\delta^4} = O(n^{-2}).$$

Thus

$$Eh_n^2 I(|h_n| \geqslant \delta) \leqslant O(n^{-2}),$$

and it follows from above that

$$nE\,|R_n| \leqslant \varepsilon_n \to 0 \qquad \text{as} \quad n \to \infty.$$

Hence the conclusion in (b) follows. Using the Taylor expansion up to the third derivative a similar argument applies if f''' is bounded.

Letting $\phi(y) = y$ in (1) we obtain an important special case of $\mathscr{U}_n$ given by $\xi_n = (Y_1 + \cdots + Y_n)/n$ where $Y_1, Y_2, ...$ are iid random variables with mean $EY_1 = x \in I$ and variance $\sigma^2(x)$. A number of well known operators such as Bernstein, Szász, Weierstrass, Baskakov, Gamma, etc., are covered by this special case (cf. [7]). Clearly, Theorems 1 and 2 hold for this reverse martingale. However, since $\rho_1 = \sigma^2(x)$, (a) in Theorem 2 can be restated as

$$|L_n(f, x) - f(x)| \leqslant (1 + \sigma^2(x))\, \omega\left(f, \frac{1}{\sqrt{n}}\right),$$

and

$$\max_{\alpha \leqslant x \leqslant \beta} |L_n(f, x) - f(x)| \leqslant (1 + \max_{\alpha \leqslant x \leqslant \beta} \sigma^2(x))\, \omega\left(f, \frac{1}{\sqrt{n}}\right),$$

where $[\alpha, \beta]$ is a fixed subinterval of I.

3. Lipschitz and Convexity Preserving Properties

We first prove the Lipschitz property of the operator $L_n(f, x)$ in the case when $\xi_n = (Y_1 + \cdots + Y_n)/n$ and the distribution of Y_i's is a member of an exponential family (cf. Lehmann [10]). Although the result is specialized to ξ_n, it still represents a large class and includes many well known operators in the literature.

Let $Y_1, Y_2, \ldots$ be iid random variables with a common exponential density

$$f(y, \theta) = \exp(\theta y - b(\theta)), \qquad \theta \in J, \tag{2}$$

relative to a σ-finite measure $\mu(y)$ where θ is a parameter with values in an open interval J (possibly infinite). Let $S_n = Y_1 + \cdots + Y_n$ and $\xi_n = S_n/n$, $n \geqslant 1$. It is well known (cf. [10]) that $b(\theta)$ is analytic and $EY_1 = b'(\theta)$ and $\operatorname{var}(Y_1) = \sigma^2(\theta) = b''(\theta) > 0$. Let $f \in C_B(R)$ and set $L_n(f, \theta) = Ef(\xi_n)$. The Lipschitz property of $L_n(f, x)$ (a version of $L_n(f, \theta)$) is a consequence of the following.

Theorem 3. *If $f(x) \in \mathrm{Lip}^{\alpha}(A)(0 \leqslant \alpha \leqslant 1)$, i.e., $|f(x_1) - f(x_2)| \leqslant A\, |x_1 - x_2|^{\alpha}$ $(0 \leqslant \alpha \leqslant 1)$, then*

$$|L_n(f, \theta_2) - L_n(f, \theta_1)| \leqslant A\, |b'(\theta_2) - b'(\theta_1)|^{\alpha}. \tag{3}$$

Proof. It is well known (cf. [10]) that the distribution of S_n is again of the exponential type with density

$$g_n(y, \theta) = \exp(\theta y - nb(\theta)), \qquad \theta \in J, \tag{4}$$

relative to a σ-finite measure $\mu_n(y)$. Moreover, $g_n(y, \theta)$ is known to be stochastically ordered. Let $\theta_2 > \theta_1$ and let $F_i(y)$ $(i = 1, 2)$ denote the distribution functions with respective densities $g_n(y, \theta_i)$ $(i = 1, 2)$. Then

$$F_2(y) \leqslant F_1(y) \qquad \text{for all } y,$$

and it follows from Lemma 1 (cf. Lehmann [10, p. 73]) that there exist random variables $\mathcal{T}_n$ and V_n such that $\mathcal{T}_n \leqslant V_n$, and the distributions of $\mathcal{T}_n$ and V_n are F_1 and F_2, respectively. Let

$$V_n = W_n + \mathcal{T}_n, \qquad W_n = V_n - \mathcal{T}_n \geqslant 0,$$

and let $F(w, u)$ denote the joint distribution of W_n and $\mathcal{T}_n$. Then the marginal distribution of $\mathcal{T}_n$ is F_1, and the distribution of the sum $W_n + \mathcal{T}_n = V_n$ is F_2. Hence it follows that

$$L_n(f, \theta_1) = Ef\left(\frac{\mathcal{T}_n}{n}\right) \qquad \text{and} \qquad L_n(f, \theta_2) = Ef\left(\frac{\mathcal{T}_n + W_n}{n}\right).$$

Since $t^\alpha (0 \leqslant \alpha \leqslant 1)$, $t \geqslant 0$ is concave and $f \in \mathrm{Lip}^\alpha(A)$, it follows from Jensen's inequality that

$$\begin{aligned} |L_n(f, \theta_2) - L_n(f, \theta_1)| &\leqslant E\left|f\left(\frac{\mathcal{T}_n + W_n}{n}\right) - f\left(\frac{\mathcal{T}_n}{n}\right)\right| \\ &\leqslant AE\left(\frac{W_n}{n}\right)^\alpha \leqslant A\left(\frac{EW_n}{n}\right)^\alpha. \end{aligned}$$

Since $EW_n = E(V_n - \mathcal{T}_n) = n(b'(\theta_2) - b'(\theta_1)) > 0$, we have

$$|L_n(f, \theta_2) - L_n(f, \theta_1)| \leqslant A\, |b'(\theta_2) - b'(\theta_1)|^\alpha.$$

This completes the proof of (3).

To deduce the Lipschitz property for $L_n(f, x)$ we note that θ in (2) is merely a reparametrization of some other parameter. That is, let $x \in I$ be the parameter of the distribution of Y_1, then $\theta = \theta(x) \in J$ is a function of x. Here is an example to clarify the idea and the process of reparametrization.

Example. Let Y_1, $Y_2, \ldots$ be iid random variables having a Poisson distribution given by

$$P(Y_1 = k) = \frac{e^{-x} x^k}{k!}, \qquad k = 0, 1, 2, \ldots, x \in I = (0, \infty).$$

Clearly,

$$f(y, \theta) = \exp(\theta y - b(\theta)), \qquad d\mu(y) = \frac{1}{y!}, \; y = 0, 1, 2, \ldots,$$

where $\theta = \ln x$, $b(\theta) = e^\theta$, $b'(\theta) = e^\theta = x$. Thus θ is merely a reparametrization of x. The reader is referred to [7] for such reparametrizations in other examples of exponential family.

Now that $\theta=\theta(x)$ is a unique solution of $b'(\theta)=x$ in an exponential family, we can write $L_n(f, x)$ instead of $L_n(f, \theta)$. Thus we have the following corollary.

COROLLARY. *Let $b'(\theta)=x \in I$ and hence $\theta=\theta(x)$. Then*

$$|L_n(f, x_1)-L_n(f, x_2)| \leqslant A\,|x_1-x_2|^\alpha \tag{5}$$

if and only if $f \in \mathrm{Lip}^\alpha(A)$ $(0 \leqslant \alpha \leqslant 1)$.

Remarks. Since $b'(\theta)=x$ in the corollary is only a reparametrization, hence (5) holds true for the exponential family defined by (2). Included in (2) are such distributions as binomial, Poisson, geometric, and normal, etc., to name a few. Hence the corollary applies to such operators as Bernstein, Szász, Baskakov, and Weierstrass, etc. (see [7, p. 199]).

The general conditions for preserving convexity for positive linear operators of convex functions are given by [5]. However, in view of the fact that $\xi_n=(Y_1+\cdots+Y_n)/n=S_n/n$ and the family (2) cover many operators, the following special case is stated for the sake of completeness.

THEOREM 4. *Let $Y_1, Y_2, \ldots$ be iid random variables with the exponential density (2). Let $\xi_n=(Y_1+\cdots+Y_n)/n$ and $L_n(f, x)=L_n(f, \theta)=Ef(\xi_n)$ where $b'(\theta)=x \in I$. If f is convex, then $L_n(f, x)$ is convex in x.*

Proof. Since the distribution of S_n is also exponential, it is enough to prove the theorem for $n=1$. The proof rests on the following facts. Let $\theta_1<\theta_2<\cdots<\theta_m$ and $y_1<y_2<\cdots<y_m$. The density function $f(y, \theta)=\exp(\theta y-b(\theta))$ is said to be TP_3 (totally positive of order 3) if the determinant $\Delta_m=|\exp(\theta_i y_j-b(\theta_i))|$, $i, j=1,2, \ldots, m$ is positive for $1 \leqslant m \leqslant 3$, and STP_3 if Δ_m is strictly positive for $1 \leqslant m \leqslant 3$. Clearly, the positivity of Δ_m is equivalent to the positivity of the determinant $\delta_m=|\exp(\theta_i y_j)|$, $i, j=1,2, \ldots, m$. It can be verified that $\delta_m>0$ for $1 \leqslant m \leqslant 3$ so that $f(y, \theta)$ is STP_3. In fact, $\delta_m>0$ for each $m=1,2, \ldots$ (see Lehmann [10, p. 115] and Karlin [5, pp. 11–12, 15–16]), and this is known as the STP property. Moreover, it is obvious that if $g(x)$ is an increasing function, then $k(y, x)=\exp(g(x) y)$ is also STP. Now that $b'(\theta)$ is increasing $(b''(\theta)>0)$, its inverse $b'^{-1}(\cdot)$ is also increasing. Consequently, the reparametrized density $f(y, \theta)=\exp(\theta y-b(\theta))$ by the equation $b'(\theta)=x$ $(\theta=b'^{-1}(x))$ is also STP and hence TP_3. Now recalling that $EY=\int y f(y, \theta)\, d\mu(y)=b'(\theta)=x$, the theorem follows from Proposition 3.2 of Karlin (cf. [5, p. 23]).

4. A Reverse Supermartingale and the Bleimann–Butzer–Hahn Operator

We now discuss an example of a reverse supermartingale which appears in the Bleimann–Butzer–Hahn operator. The primary purpose of this section is to cite one more example reflecting the basic spirit of the article. Secondly, this will also correct a minor error in [8]. Let $Y_1, Y_2, \ldots$ be iid Bernoulli random variables with probability $p = x/(1+x)(x>0)$, i.e., $P(Y_1 = 1) = 1 - P(Y_1 = 0) = p$. Let $S_n = Y_1 + \cdots + Y_n$ and $\xi_n = S_n/(n - S_n + 1)$. For $f \in C[0, \infty)$ the Bleimann–Butzer–Hahn operator is defined by

$$L_n(f, x) = Ef(\xi_n) = (1+x)^{-n} \sum_{k=0}^{n} f(k/(n-k+1)) \binom{n}{k} x^k.$$

In order to prove monotonic convergence it was claimed in [8] that $E(\xi_n \mid S_{n+1}) = \xi_{n+1}$. This is not quite correct. In fact the same argument as given in [8] shows that

$$\begin{aligned} E(\xi_n \mid S_{n+1}) &= \xi_{n+1} \qquad \text{if} \quad S_{n+1} \leqslant n \\ &= n \qquad\quad \text{if} \quad S_{n+1} = n+1. \end{aligned} \tag{6}$$

Since $\xi_{n+1} = n+1$ if $S_{n+1} = n+1$, we have

$$E(\xi_n \mid S_{n+1}) \leqslant \xi_{n+1} \qquad \text{and} \qquad E(\xi_n \mid \mathscr{F}_{n+1}) \leqslant \xi_{n+1}, \tag{7}$$

where $\mathscr{F}_n = \mathscr{B}(S_n, S_{n+1}, \ldots)$ is the σ-field generated by $S_n, S_{n+1}, \ldots$. Thus $\{\xi_n, \mathscr{F}_n, n \geqslant 1\}$ is a reverse supermartingale. Using (6) one also verifies that

$$\begin{aligned} E\xi_n &= \sum_{k=0}^{n+1} E(\xi_n \mid S_{n+1} = k)\, P(S_{n+1} = k) \\ &= \sum_{k=0}^{n} E(\xi_n \mid S_{n+1} = k)\, P(S_{n+1} = k) + nP(S_{n+1} = n+1) \\ &= \sum_{k=0}^{n} \xi_{n+1}(k)\, P(S_{n+1} = k) + nP(S_{n+1} = n+1), \end{aligned}$$

where $\xi_{n+1}(k) = k/(n-k+2) = \xi_{n+1}$ when $S_{n+1} = k$. Hence

$$E\xi_n = E\xi_{n+1} + np^{n+1} - (n+1)\, p^{n+1} = E\xi_{n+1} - p^{n+1} \leqslant E\xi_{n+1}.$$

The property of monotonic convergence for $L_n(f, x)$ can now be proved and a minor error in [8] is corrected as follows. Let $f(\cdot)$ be a decreasing convex function on $[0, \infty)$. Then by Jensen's inequality we have

$$L_n(f, x) = Ef(\xi_n) = EE(f(\xi_n) \mid S_{n+1}) \geqslant Ef(E(\xi_n \mid S_{n+1})). \tag{8}$$

It follows from (6) that

$$\begin{aligned}
Ef(E(\xi_n \mid S_{n+1})) &= \sum_{k=0}^{n+1} f(E(\xi_n \mid S_{n+1}=k))\ P(S_{n+1}=k) \\
&= \sum_{k=0}^{n} f(\xi_{n+1}(k))\ P(S_{n+1}=k) + f(n)\ P(S_{n+1}=n+1) \\
&= L_{n+1}(f, x) + (f(n) - f(n+1))\ P(S_{n+1}=n+1) \\
&\geqslant L_{n+1}(f, x).
\end{aligned}$$

Consequently from (8) we have

$$L_n(f, x) \geqslant L_{n+1}(f, x) \geqslant \cdots \geqslant f(x).$$

Moreover, if f is an increasing concave function, then by the preceding argument applied to $-f$ or directly from (7) and Jensen's inequality we have

$$\begin{aligned}
L_n(f, x) = Ef(\xi_n) = EE(f(\xi_n) \mid S_{n+1}) &\leqslant Ef(E(\xi_n \mid S_{n+1})) \\
&\leqslant Ef(\xi_{n+1}) = L_{n+1}(f, x).
\end{aligned}$$

ACKNOWLEDGMENT

I am thankful to the referees for their valuable comments.

REFERENCES

1. G. BLEIMANN, P. L. BUTZER, AND L. HAHN, A Bernstein-type operator approximating continuous functions on the semi-axis, *Nederl. Akad. Wetensch. Indag. Math.* **42** (1980), 255–262.
2. Y. S. CHOW AND H. TEICHER, "Probability Theory," Springer-Verlag, New York, 1978.
3. P. HALL AND C. C. HEYDE, "Martingale Limit Theory and Its Applications," Academic Press, New York, 1980.
4. W. HOEFFDING, A class of statistics with asymptotically normal distribution, *Ann. Math. Statist.* **19** (1948), 293–325.
5. S. KARLIN, "Total Positivity," Vol. I, Stanford Univ. Press, Stanford, CA 1968.
6. M. K. KHAN AND M. A. PETERS, Lipschitz constants for some approximation operators of a Lipschitz continuous function, *J. Approx. Theory* **59** (1989), 307–315.
7. R. A. KHAN, Some probabilistic methods in the theory of approximation operators, *Acta Math. Acad. Sci. Hungar.* **35** (1980), 193–203.
8. R. A. KHAN, Some properties of a Bernstein-type operator of Bleimann, Butzer and Hahn, *Progr. Approx. Theory* (1991), 497–504.

9. A. J. LEE, "$\mathcal{U}$-Statistics: Theory and Practice," Dekker, New York, 1990.
10. E. L. LEHMANN, "Testing Statistical Hypotheses," Wiley, New York, 1959.
11. R. N. SPROULE, A sequential fixed-width confidence interval for the mean of a $\mathcal{U}$-statistic, Inst. Statist. Mimeo Series No. 636, University of North Carolina, Chapel Hill, 1969.
12. R. N. SPROULE, Asymptotic properties of $\mathcal{U}$-statistics, *Trans. Amer. Math. Soc.* **199** (1974), 55–64.