JOURNAL OF APPROXIMATION THEORY 4, 347–356 (1971)

The L_1 Norm of the Approximation Error for Bernstein-Type Polynomials*

WASSILY HOEFFDING

Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514

Communicated by Oved Shisha

Received July 2, 1968; revised April 17, 1970

1. INTRODUCTION AND STATEMENT OF RESULTS

This paper is concerned with the estimation of the L_1 norm of the difference between a function of bounded variation and an associated Bernstein polynomial, and with the analogous problem for a Lebesgue integrable function of bounded variation inside (0, 1). A real-valued function defined in the open interval (0, 1) is said to be of bounded variation inside (0, 1) if it is of bounded variation in every closed subinterval of (0, 1). The class of these functions will be denoted by BV^*. To formulate some of the results, we state the following lemma, which is a simple consequence of the well-known canonical representation of a function of bounded variation.

LEMMA 1. *A function f is in BV^* if and only if it can be represented as $f = f_1 - f_2$, where f_1 and f_2 are nondecreasing real-valued functions on* (0, 1). *Moreover, if $f \in BV^*$, the functions f_1 and f_2 can be so chosen that, for $0 < x < y < 1$, the total variation of f on $[x, y]$ is the sum of the total variations of f_1 and f_2 on $[x, y]$:*

$$f = f_1 - f_2, \qquad \mathrm{var}_{[x,y]} f = \mathrm{var}_{[x,y]} f_1 + \mathrm{var}_{[x,y]} f_2 . \tag{1}$$

If f is finite in the closed interval [0, 1], the associated Bernstein polynomial of order n, denoted by $B_n f$, is defined by

$$B_n f(x) = \sum_{i=0}^{n} f(i/n)\, p_{n,i}(x), \tag{2}$$

where

$$p_{n,i}(x) = \binom{n}{i} x^i (1-x)^{n-i}. \tag{3}$$

* This research was supported in part by the Mathematics Division of the Air Force Office of Scientific Research under Grant No. AFOSR-68-1415.

For f Lebesgue integrable on (0, 1), we shall use the modified Bernstein polynomials $P_n f(x) = d/dx\, B_{n+1}F(x)$, where $F(x) = \int_0^x f(y)\, dy$. Explicitly (see Lorentz [1, Chap. II]),

$$P_n f(x) = \sum_{i=0}^{n} (n+1) \int_{i/(n+1)}^{(i+1)/(n+1)} f(y)\, dy\, p_{n,i}(x). \tag{4}$$

For $f \in BV^*$, let

$$J(f) = \int_0^1 x^{1/2}(1-x)^{1/2}\, |\, df(x)|. \tag{5}$$

If f is represented in the form (1), we have $J(f) = J(f_1) + J(f_2)$. If f is nondecreasing,

$$J(f) = \int_0^1 x^{1/2}(1-x)^{1/2}\, df(x) = \int_0^1 f(x)(x - \tfrac{1}{2})\, x^{-1/2}(1-x)^{-1/2}\, dx. \tag{6}$$

THEOREM 1. *If f is a Lebesgue integrable function of bounded variation inside* (0, 1). *then*

$$\int_0^1 |\, P_n f(x) - f(x)|\, dx \leqslant C_n J(f), \tag{7}$$

where

$$C_n = 2^{1/2}(n + \tfrac{1}{2})^{n+1/2}(n+1)^{-n-1} < (2/e)^{1/2} n^{-1/2}. \tag{8}$$

Equality holds in (7) *if and only if f is constant in each of the intervals* $(0, a)$ *and* $(a, 1)$, *where* $a = \frac{1}{2}(n+1)^{-1}$ *or* $a = 1 - \frac{1}{2}(n+1)^{-1}$.

THEOREM 2. *Let f be a step function with finitely many steps in every closed subinterval of* (0, 1), *and such that the functions f_1 and f_2 in the representation* (1) *are Lebesgue integrable. Then*

$$\lim_{n\to\infty} n^{1/2} \int_0^1 |\, P_n f(x) - f(x)|\, dx = (2/\pi)^{1/2}\, J(f), \tag{9}$$

irrespective of whether $J(f)$ is finite or infinite.

Theorem 1 shows that the finiteness of $J(f)$ is a sufficient condition for the L_1 norm of the approximation error to be of order $n^{-1/2}$. Theorem 2 implies that the latter is guaranteed only if $J(f)$ is finite when no restrictions beyond $f \in BV^*$ are imposed. It also shows that the upper bound in (7), with the numerical constant $(2/e)^{1/2}$ reduced to $(2/\pi)^{1/2}$, is asymptotically attained for every fixed step function of the specified type.

If f is nondecreasing, the condition $J(f) < \infty$ is stronger, but not much stronger than square integrability of f. Explicitly (see [2, Appendix]), if f is nondecreasing, $J(f) < \infty$ implies $\int_0^1 f^2(x)\,dx < \infty$ (but not conversely), and $\int_0^1 f^2(x)\{\log(1 + |f(x)|)\}^{1+\delta}\,dx < \infty$ for some $\delta > 0$ implies $J(f) < \infty$. If f is nondecreasing and square integrable, we have for $n \geqslant 2$,

$$\int_0^1 | P_n f(x) - f(x)|\,dx \leqslant C(n^{-1}\log n)^{1/2} \left(\int_0^1 f^2(x)\,dx\right)^{1/2}, \tag{10}$$

where C is an absolute constant. The proof of (10) is sketched at the end of Section 2.

If f is convex, (10) is true with $\log n$ removed (as can be shown by means of Jensen's inequality).

Concerning the Bernstein polynomials (2), Theorem 1 immediately implies the following. If F is the difference of two convex, absolutely continuous functions on [0, 1] and if $J(F')$ is finite, then $\operatorname{var}_{[0,1]}(B_n F - F) = O(n^{-1/2})$. (I am indebted to Professor G. G. Lorentz for this observation.) We also have the following analogs of Theorems 1 and 2.

THEOREM 3. *Let f be of bounded variation in* [0, 1]. *Then*

$$\int_0^1 | B_n f(x) - f(x)|\,dx \leqslant C_n J(f) + (n+1)^{-1} \operatorname{var}_{[0,1]}(f), \tag{11}$$

where C_n is given by (8).

THEOREM 4. *Let f be a step function of bounded variation in* [0, 1] *with finitely many steps in every closed sub-interval of* (0, 1). *Then* (9) *holds, with P_n replaced by B_n.*

The upper bound in (11) can not be replaced by $Cn^{-1/2}J(f)$ with C an absolute constant, as the following example shows. Let $f(x) = b$ if $0 \leqslant x < a_n < 1$, $f(x) = c$ $(\neq b)$ if $a_n \leqslant x \leqslant 1$, where $a_n = o(n^{-1})$. By a simple calculation,

$$\int_0^1 | B_n f - f |\,dx = | b - c |\, n^{-1}(1 + o(1)), \quad J(f) = | b - c |\, a_n^{1/2}(1 + o(1)).$$

Hence $n^{1/2} \int_0^1 | B_n f - f |\,dx / J(f) \sim (na_n)^{-1/2} \to \infty$.

2. PROOF OF THEOREM 1

The modified Bernstein polynomial defined by (4) may be written in the form

$$P_n f(x) = \int_0^1 K_n(x, y) f(y)\,dy, \tag{12}$$

where

$$K_n(x, y) = (n + 1)\, p_{n,[(n+1)y]}(x) \tag{13}$$

and $[u]$ denotes the largest integer $\leqslant u$. We note that

$$\int_0^1 K_n(x, y)\, dy = 1, \qquad \int_0^1 K_n(x, y)\, dx = 1. \tag{14}$$

Let

$$H_n(x, u) = \int_u^1 K_n(x, y)\, dy. \tag{15}$$

A simple calculation shows that for $x, u \in [0, 1)$,

$$H_n(x, u) = \delta_n(u)\, G_{n,[(n+1)u]+1}(x) + (1 - \delta_n(u))\, G_{n,[(n+1)u]}(x), \tag{16}$$

where

$$\delta_n(u) = (n + 1)\, u - [(n + 1)\, u], \tag{17}$$

$$G_{n,k}(x) = \sum_{i=k}^{n} p_{n,i}(x) = n \int_0^x p_{n-1,k-1}(t)\, dt, \qquad k = 1,..., n, \tag{18}$$

and $G_{n,0}(x) = 1$, $G_{n,n+1}(x) = 0$.

Let $x \in (0, 1)$ be a continuity point of f. We have, from (12) and (14),

$$\begin{aligned} P_n f(x) - f(x) &= \int_0^1 K_n(x, y)\{f(y) - f(x)\}\, dy \\ &= -\int_0^x K_n(x, y) \int_y^x df(u)\, dy + \int_x^1 K_n(x, y) \int_x^y df(u)\, dy \\ &= -\int_0^x \int_0^u K_n(x, y)\, dy\, df(u) + \int_x^1 \int_u^1 K_n(x, y)\, dy\, df(u). \end{aligned}$$

Since $\int_0^u K_n(x, y)\, dy = 1 - H_n(x, u)$, we have

$$P_n f(x) - f(x) = -\int_0^x (1 - H_n(x, u))\, df(u) + \int_x^1 H_n(x, u)\, df(u). \tag{19}$$

Hence

$$\begin{aligned} \int_0^1 | P_n f(x) - f(x)|\, dx &\leqslant \int_0^1 \int_0^x (1 - H_n(x, u))|\, df(u)|\, dx \\ &\quad + \int_0^1 \int_x^1 H_n(x, u)|\, df(u)|\, dx \\ &= \int_0^1 D_n(u)|\, df(u)|, \end{aligned} \tag{20}$$

where

$$D_n(u) = \int_u^1 (1 - H_n(x, u))\, dx + \int_0^u H_n(x, u)\, dx. \tag{21}$$

Therefore,

$$\int_0^1 | P_n f(x) - f(x)|\, dx \leqslant C_n J(f), \tag{22}$$

where

$$C_n = \sup_{0 \leqslant u \leqslant 1} u^{-1/2}(1 - u)^{-1/2} D_n(u). \tag{23}$$

From (15) and (14), it is easily seen that

$$D_n(u) = 2 \int_0^u H_n(x, u)\, dx. \tag{24}$$

We now show that

$$D_n(u) = 2u(1 - u)\, p_{n,[(n+1)u]}(u). \tag{25}$$

For $k \leqslant (n + 1)\, u < k + 1$ $(k = 0, 1,..., n)$, we have $[(n + 1)\, u] = k$, $1 - \delta_n(u) = k + 1 - (n + 1)\, u$, and, by (16) and (18),

$$H_n(x, u) = G_{n,k+1}(x) + (k + 1 - (n + 1)\, u)\, p_{n,k}(x).$$

Hence it is sufficient to show that the function

$$g(u) = \int_0^u G_{n,k+1}(x)\, dx + (k + 1 - (n + 1)\, u) \int_0^u p_{n,k}(x)\, dx - u(1 - u)\, p_{n,k}(u)$$

is identically zero.

It is easy to verify the identities

$$u(1 - u)\, p'_{n,k}(u) = (k - nu)\, p_{n,k}(u),$$

$$up'_{n,k}(u) = np_{n,k}(u) - np_{n-1,k}(u).$$

Hence

$$\begin{aligned} g'(u) &= G_{n,k+1}(u) - (n + 1) \int_0^u p_{n,k}(x)\, dx + (k + 1 - (n + 1)u)\, p_{n,k}(u) \\ &\quad - (1 - 2u)\, p_{n,k}(u) - u(1 - u)\, p'_{n,k}(u) \\ &= G_{n,k+1}(u) - (n + 1) \int_0^u p_{n,k}(x)\, dx + up_{n,k}(u), \end{aligned}$$

$$g''(u) = np_{n-1,k}(u) - (n + 1)\, p_{n,k}(u) + p_{n,k}(u) + up'_{n,k}(u).$$

Thus $g''(u) = 0$, and since $g(0) = g'(0) = 0$, identity (25) is proved.

For $[(n+1)u]=k$ fixed, $u^{1/2}(1-u)^{1/2}\,p_{n,k}(u)=\binom{n}{k}\,u^{k+1/2}(1-u)^{n-k+1/2}$ attains its maximum at $u=(k+\frac{1}{2})/(n+1)$. Hence, by (25),

$$\begin{aligned} u^{-1/2}(1-u)^{-1/2}\,D_n(u) &= 2u^{1/2}(1-u)^{1/2}\,p_{n,k}(u) \\ &\leqslant 2\binom{n}{k}(k+\tfrac{1}{2})^{k+1/2}(n-k+\tfrac{1}{2})^{n-k+1/2}(n+1)^{-n-1} \\ &= c_n(k), \qquad \text{say.} \end{aligned} \tag{26}$$

Now

$$\frac{c_n(k+1)}{c_n(k)}=\frac{n-k}{k+1}\,\frac{(k+\frac{3}{2})^{k+3/2}(n-k-\frac{1}{2})^{n-k-1/2}}{(k+\frac{1}{2})^{k+1/2}(n-k+\frac{1}{2})^{n-k+1/2}}=\frac{F(k)}{F(n-k-1)},$$

where

$$F(k)=(k+\tfrac{3}{2})^{k+3/2}\,(k+1)^{-1}\,(k+\tfrac{1}{2})^{-k-1/2}.$$

It is readily seen that

$$\frac{d}{dk}\log F(k)=\log(k+\tfrac{3}{2})-\log(k+\tfrac{1}{2})-(k+1)^{-1}$$

is positive for $k\geqslant 0$. Hence $F(k)$ is strictly increasing. Therefore

$$\max_{0\leqslant k\leqslant n} c_n(k)=c_n(0)=c_n(n)=2^{1/2}(n+\tfrac{1}{2})^{n+1/2}(n+1)^{-n-1}. \tag{27}$$

Also, the left-hand side of (26) is equal to $c_n(0)$ if and only if $u=\frac{1}{2}(n+1)^{-1}$ or $u=1-\frac{1}{2}(n+1)^{-1}$. Thus C_n, as defined by (23), is equal to the expressions in (27). By (20), equality in (22) can hold only if f takes two values and the saltus is at $\frac{1}{2}(n+1)^{-1}$ or $1-\frac{1}{2}(n+1)^{-1}$. A direct calculation shows that equality does hold in this case.

The inequality in (8) is easily verified, completing the proof.

We now indicate the proof of inequality (10). It has been shown that $D_n(u)\leqslant(2/e)^{1/2}\,n^{-1/2}u^{1/2}(1-u)^{1/2}$. Hence, if f is nondecreasing,

$$\int_\epsilon^{1-\epsilon} D_n(u)\,d\,|f(u)|\leqslant(2e)^{1/2}\,n^{-1/2}\int_\epsilon^{1-\epsilon} u^{1/2}(1-u)^{1/2}\,df(u).$$

Integration by parts and application of Schwarz's inequality show that the right side does not exceed

$$Cn^{-1/2}\left(\log\frac{1-\epsilon}{\epsilon}\right)^{1/2}\left(\int_0^1 f^2(u)\,du\right)^{1/2}$$

for $0<\epsilon\leqslant 1/3$. If we set $\epsilon=(n+1)^{-1}$, the remaining contribution to $\int_0^1 D_n(u)\mid df(u)|$ is of smaller order of magnitude, and (10) follows from (20).

3. Proof of Theorem 2

For convenience of notation, the proof will be given for a step function f with finitely many steps in every interval $(0, \delta)$ with $\delta < 1$. For the general case, the proof requires only trivial modifications. It is irrelevant how f is defined at its points of discontinuity, and we may assume that

$$\begin{aligned} f(x) = b_j \quad &\text{if} \quad a_{j-1} < x \leqslant a_j, \quad j = 1, 2, \ldots, \\ 0 = a_0 < a_1 < \cdots, \quad &\lim_{j\to\infty} a_j = 1. \end{aligned} \tag{28}$$

Let

$$\begin{aligned} I_n(x, u) &= H_n(x, u) - 1 \quad &&\text{if} \quad 0 < u \leqslant x < 1, \\ I_n(x, u) &= H_n(x, u) &&\text{if} \quad 0 < x < u < 1, \end{aligned} \tag{29}$$

and let m be a fixed positive integer. By (19), if $x \in (0, 1)$ is a continuity point of f,

$$P_n f(x) - f(x) = \sum_{i=1}^{m} I_n(x, a_i)(b_{i+1} - b_i) + \int_{a_{m+1}-}^{1} I_n(x, u)\, df(u).$$

Hence

$$\int_0^{a_m} | P_n f(x) - f(x)|\, dx = A_n + \theta R_n, \qquad | \theta | \leqslant 1, \tag{30}$$

where

$$A_n = \int_0^{a_m} \left| \sum_{i=1}^{m} I_n(x, a_i)(b_{i+1} - b_i) \right| dx, \tag{31}$$

$$R_n = \int_{a_{m+1}-}^{1} \int_0^{a_m} | I_n(x, u)|\, dx\, | df(u)|. \tag{32}$$

From (16) and (18) we obtain by straightforward calculation,

$$\int_0^1 (x - u)^2\, d_n H_n(x, u) \leqslant 3u(1 - u)\, n^{-1}, \qquad 0 \leqslant u \leqslant 1.$$

Hence, if $0 < x < u$,

$$H_n(x, u) \leqslant (u - x)^{-2} \int_0^x (u - y)^2\, d_y H_n(y, u) \leqslant 3u(1 - u)(u - x)^{-2}\, n^{-1}.$$

For $u < x < 1$, we have the same upper bound for $1 - H_n(x, u)$, so that

$$| I_n(x, u)| \leqslant 3u(1 - u)(u - x)^{-2}\, n^{-1}, \qquad 0 < x, \quad u < 1. \tag{33}$$

From (32) and (33), we have

$$R_n \leqslant 3(a_{m+1} - a_m)^{-2} n^{-1} \int_{0+}^{1} (1 - u) \mid df(u)|.$$

The last integral is finite since f_1 and f_2 in (1) are Lebesgue integrable. Hence $R_n = O(n^{-1})$ and

$$\int_0^{a_m} | P_n f(x) - f(x)| \, dx = A_n + O(n^{-1}). \tag{34}$$

Let $a_{j-1} < x < (a_{j-1} + a_j)/2$. Then, by (33), $I_n(x, a_i) = O(n^{-1})$ if $i \neq j - 1$, uniformly in x for $i = 1,..., m$. Hence

$$\sum_{i=1}^{m} I_n(x, a_i)(b_{i+1} - b_i) = (H_n(x, a_{j-1}) - 1)(b_j - b_{j-1}) + O(n^{-1}) \quad \text{if} \quad a_{j-1} < x < \frac{a_{j-1} + a_j}{2}, \tag{35}$$

for $j = 1,..., m$, uniformly for $x \in (0, a_m)$. (For $j = 1$, the first term on the right is zero.)

In a similar way, it is seen that

$$\sum_{i=1}^{m} I_n(x, a_i)(b_{i+1} - b_i) = H_n(x, a_j)(b_{j+1} - b_j) + O(n^{-1}) \quad \text{if} \quad \frac{a_{j-1} + a_j}{2} < x < a_j, \tag{36}$$

for $j = 1,..., m$, uniformly for $x \in (0, a_m)$.

It follows that

$$A_n = \sum_{j=2}^{m} | b_j - b_{j-1} | \int_{a_{j+1}}^{(a_{j+1}+a_j)/2} (1 - H_n(x, a_{j-1})) \, dx + \sum_{j=1}^{m} | b_{j+1} - b_j | \int_{(a_{j-1}+a_j)/2}^{a_j} H_n(x, a_j) \, dx + O(n^{-1}). \tag{37}$$

Another application of (33) shows that if the upper limits of integration in the first sum in (37) are replaced by 1 and the lower limits in the second sum by 0, then a term of order n^{-1} is added. Hence

$$A_n = \sum_{j=1}^{m-1} | b_{j+1} - b_j | \left\{ \int_{a_j}^{1} (1 - H_n(x, a_j)) \, dx + \int_0^{a_j} H_n(x, a_j) \, dx \right\} + | b_{m+1} - b_m | \int_0^{a_m} H_n(x, a_m) \, dx + O(n^{-1}). \tag{38}$$

With (34), (21), and (24), we thus have

$$\int_0^{a_m} | P_n f(x) - f(x)| \, dx = \sum_{j=1}^{m-1} D_n(a_j) \, | b_{j+1} - b_j | + \tfrac{1}{2} D_n(a_m) \, | b_{m+1} - b_m | + O(n^{-1}). \quad (39)$$

For $u \in (0, 1)$ fixed, we have by (25) and Stirling's formula

$$D_n(u) = (2/\pi)^{1/2} \, n^{-1/2} u^{1/2} (1 - u)^{1/2} + o(n^{-1/2}).$$

Inserting this expression in (39) and recalling (28), we obtain

$$\int_0^{a_m} | P_n f(x) - f(x)| \, dx = (2/\pi)^{1/2} \, n^{-1/2} \left\{ \int_0^{a_m} x^{1/2} (1 - x)^{1/2} \, | df(x)| + \frac{1}{2} \int_{a_m}^{a_m^+} x^{1/2} (1 - x)^{1/2} \, | df(x)| \right\} + o(n^{-1/2}). \quad (40)$$

If $J(f) = \infty$, the first integral on the right side of (40) may be made as large as we please by choosing m sufficiently large, and (9) is proved in this case.

Let $J(f) < \infty$. Given a positive ϵ, choose $\eta = \eta(\epsilon) \in (0, 1)$ so that $\int_\eta^1 x^{1/2} (1 - x)^{1/2} \, | df(u)| < \epsilon$ and so that η is not a point of discontinuity of f. Let $\tilde{f}_1(x) = f(x)$ if $0 < x \leqslant \eta$, $\tilde{f}_1(x) = f(\eta)$ if $\eta < x < 1$, and let $\tilde{f}_2(x) = f(x) - \tilde{f}_1(x)$. Then

$$J(f) = J(\tilde{f}_1) + J(\tilde{f}_2), \qquad J(\tilde{f}_2) < \epsilon,$$

and $\int_0^1 | P_n f - f | \, dx$ differs from $\int_0^1 | P_n \tilde{f}_1 - \tilde{f}_1 | \, dx$ by at most $\int_0^1 | P_n \tilde{f}_2 - \tilde{f}_2 | \, dx$. Since $\tilde{f}_1$ has finitely many steps, (40) with $f = \tilde{f}_1$ implies

$$\lim_{n \to \infty} n^{1/2} \int_0^1 | P_n \tilde{f}_1 - \tilde{f}_1 | \, dx = \left(\frac{2}{\pi} \right)^{1/2} J(\tilde{f}_1).$$

By Theorem 1,

$$n^{1/2} \int_0^1 | P_n \tilde{f}_2 - \tilde{f}_2 | \, dx \leqslant (2/e)^{1/2} J(\tilde{f}_2) < (2/e)^{1/2} \epsilon.$$

Since ϵ is arbitrary, these facts imply (9). The proof is complete.

4. Proof of Theorems 3 and 4

Since a function of bounded variation can be represented in the form (1), and $\int | B_n f - f | \, dx \leqslant \sum_{i=1}^{2} \int | B_n f_i - f_i | \, dx$, we may assume in the proof of Theorem 3 that f is a nondecreasing function. By (4),

$$P_n f(x) = \sum_{i=0}^{n} c_{n,i} p_{n,i}(x), \qquad c_{n,i} = (n+1) \int_{i/(n+1)}^{(i+1)/(n+1)} f(y) \, dy.$$

Since f is nondecreasing,

$$f\left(\frac{i}{n+1}\right) \leqslant c_{n,i} \leqslant f\left(\frac{i+1}{n+1}\right), \qquad f\left(\frac{i}{n+1}\right) \leqslant f\left(\frac{i}{n}\right) \leqslant f\left(\frac{i+1}{n+1}\right),$$

$$i = 0, \ldots, n.$$

Hence $| c_{n,i} - f(i/n) | \leqslant f((i+1)/(n+1)) - f(i/(n+1))$ and therefore

$$\begin{aligned} \int_0^1 | P_n f(x) - B_n f(x) | \, dx &\leqslant \sum_{i=0}^{n} \int_0^1 \left| c_{n,i} - f\left(\frac{i}{n}\right) \right| p_{n,i}(x) \, dx \\ &\leqslant \sum_{i=0}^{n} \left\{ f\left(\frac{i+1}{n+1}\right) - f\left(\frac{i}{n}\right) \right\} (n+1)^{-1} \\ &= \operatorname{var}_{[0,1]}(f)(n+1)^{-1}. \end{aligned} \tag{41}$$

Inequality (11) now follows from

$$\int_0^1 | B_n f(x) - f(x) | \, dx \leqslant \int_0^1 | P_n f(x) - f(x) | \, dx + \int_0^1 | B_n f(x) - P_n f(x) | \, dx$$

and Theorem 1.

The conditions of Theorem 4 imply those of Theorem 2, and Theorem 4 follows from (9) and (41).

References

1. G. G. Lorentz, "Bernstein Polynomials," University of Toronto Press, Toronto, Canada, 1953.
2. Wassily Hoeffding, On the centering of a simple linear rank statistic. Submitted to *Ann. Math. Statist.*