APPLICATIONES MATHEMATICAE
36,1 (2009), pp. 1–12

Zbigniew Ciesielski (Sopot)
Ryszard Zieliński (Warszawa)

# POLYNOMIAL AND SPLINE ESTIMATORS OF THE DISTRIBUTION FUNCTION WITH PRESCRIBED ACCURACY

*Abstract.* Dvoretzky–Kiefer–Wolfowitz type inequalities for some polynomial and spline estimators of distribution functions are constructed. Moreover, hints on the corresponding algorithms are given as well.

**1. Introduction.** The family of all continuous probability distribution functions on the real line is denoted by $\mathcal{F}$. If the probability corresponding to $F \in \mathcal{F}$ is supported on the interval $I$ then we write $F \in \mathcal{F}(I)$. Let $X_1, \ldots, X_n$ be a simple sample corresponding to a distribution $F \in \mathcal{F}$ and let

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} 1_{(-\infty,x]}(X_j). \tag{1.1}$$

The Dvoretzky–Kiefer–Wolfowitz (DKW) inequality in its final form (see [6])

$$P\{\|F_n - F\|_\infty \geq \varepsilon\} \leq 2e^{-2n\varepsilon^2} \quad \text{for all } F \in \mathcal{F}, \tag{1.2}$$

where $\|F_n - F\|_\infty = \sup_x |F_n(x) - F(x)|$, gives us a powerful tool for statistical applications (testing hypotheses concerning unknown $F$, confidence intervals for $F$ etc.). It seems, however, unnatural to estimate a continuous $F \in \mathcal{F}$ by the step function $F_n$. In the abundant literature of the subject one can find different approaches to smoothing empirical distribution functions but it turns out that smoothing may spoil the estimator. For classical kernel estimators (see e.g. [8]) no inequality of DKW type exists (Zieliński [9]). Consequently, one cannot tell how many observations $X_1, \ldots, X_n$ are

2000 *Mathematics Subject Classification*: Primary 62G07; Secondary 62G30, 60E15.
*Key words and phrases*: nonparametric estimation, distribution estimation, Dvoretzky–Kiefer–Wolfowitz inequality, polynomial estimators, spline estimators, smoothing of empirical distribution function.

DOI: 10.4064/am36-1-1  

needed to guarantee the prescribed accuracy of the kernel estimator. Using the method from [9] one can prove a similar negative result for some polynomial estimators (see Section 4 below).

In what follows we discuss three estimators and the DKW type inequalities for them. The first one (see Section 3), to be denoted by $\Phi_{m,n}$, is a smooth modification of $F_n$. It is a piecewise polynomial estimator of the class $C^m \cap \mathcal{F}$ with knots placed at the sample points $X_{1:n}, \ldots, X_{n:n}$. The smoothness parameter $m = 1, 2, \ldots$ can be fixed by the statistician according to his a priori knowledge of the smoothness of the estimated $F \in \mathcal{F}$. In this case the DKW type inequality holds in the class $\mathcal{F}$ of all continuous probability distribution functions on the real line. The second one (see Section 4) is the polynomial estimator constructed in Ciesielski [3], and the third one (see Section 5) is a spline estimator with equally spaced knots (Ciesielski [2], [4]). In both cases subclasses of $\mathcal{F}$ are explicitly specified for which the DKW type inequalities are proved. We start in Section 2 with preliminaries on specific approximations by algebraic polynomials and by splines with equally spaced knots.

**2. Preliminaries from approximation theory.** We recall the necessary ingredients from approximation theory, in particular the basic properties of the Bernstein polynomials and of the related B-splines. The relation is direct: the Bernstein polynomials are simply the degenerate B-splines.

Let us start with the *Bernstein polynomials*. For a given integer $r \geq 1$ let $\Pi_m$ denote the space of real polynomials of degree at most $m = r - 1$ (i.e. of order $r$). For convenience, both parameters $r$ and $m$ will be used; $m$ is more natural for polynomials and $r$ for splines. The *basic Bernstein polynomials* of degree $m$ are linearly independent and given by the formula

$$N_{i,m}(x) = \binom{m}{i} x^i (1-x)^{m-i} \quad \text{with } i = 0, \ldots, m. \tag{2.1}$$

Consequently,

$$\Pi_m = \operatorname{span}[N_{i,m} : i = 0, \ldots, m], \tag{2.2}$$

and therefore each $w \in \Pi_m$ has a unique representation

$$w(x) = \sum_{i=0}^{m} w_i N_{i,m}(x). \tag{2.3}$$

Now, given the coefficients $w_i$ and the $x \in (0,1)$ we can calculate the value $w(x)$ using the *Casteljeau algorithm* based on the identity

$$N_{i,m}(x) = (1-x) N_{i,m-1}(x) + x N_{i-1,m-1}(x). \tag{2.4}$$

The Bernstein polynomials are positive on $(0,1)$ and they form a partition of unity,

$$\sum_{i=0}^{m} N_{i,m}(x) = 1 \quad \text{for } i = 0, \ldots, m \text{ and } x \in (0,1). \tag{2.5}$$

On the other hand, the *modified Bernstein polynomials*

$$M_{i,m}(x) = (m+1)N_{i,m}(x) \tag{2.6}$$

are normalized in $L^1(0,1)$, i.e.

$$\int_0^1 M_{i,m}(x)\,dx = 1. \tag{2.7}$$

In our construction of the polynomial estimator, the *Durrmeyer kernel*

$$R_m(x,y) = \sum_{i=0}^{m} M_{i,m}(x) N_{i,m}(y) \tag{2.8}$$

plays an essential role. The kernel is symmetric and it has the following spectral representation:

$$R_m(x,y) = \sum_{i=0}^{m} \lambda_{i,m}\, l_i(x) l_i(y), \tag{2.9}$$

where the eigenfunctions are the orthonormal Legendre polynomials $l_i$ and for the corresponding eigenvalues $\lambda_{i,m}$ we have (cf. e.g. [5])

$$\lambda_{0,m} = 1, \quad \lambda_{i,m} = \prod_{j=1}^{i} \frac{m+1-j}{m+1+j} \quad \text{for } i = 1, \ldots, m, \tag{2.10}$$

whence, with the notation $\lambda_k = k(k+1)$,

$$\frac{2}{5}\frac{\lambda_j}{\lambda_{k+1}} \le 1 - \lambda_{j,m} \le \frac{\lambda_j}{m+2} \quad \text{for } \lambda_k \le m < \lambda_{k+1},\ 1 \le j \le k. \tag{2.11}$$

Now, for a given function $F$ which is either of bounded variation or continuous on $[0,1]$, the value of the underlying smoothing operator is a polynomial of degree $m+1$ defined by the formula

$$T_m F(x) = F(0) + \int_0^x \Big[ \int_0^1 R_m(y,z)\,dF(y) \Big]\,dz. \tag{2.12}$$

Now, since

$$\int_0^x M_{i,m}(y)\,dy = \sum_{j=i+1}^{m+1} N_{j,m+1}(x) \quad \text{for } i = 0, 1, \ldots, m, \tag{2.13}$$

the Casteljeau algorithm, in case $F = F_n$, can be applied to calculate the value $T_m F(x)$. It also follows from (2.12) that the integrated Legendre polynomials

$$L_j(x) = \int_0^x l_j(y)\,dy \quad \text{for } j = 0, 1, \ldots \tag{2.14}$$

are the eigenfunctions for the operators $T_m$ with the corresponding eigenvalues (2.6). For more details on the eigenvalues $\lambda_{i,m}$ we refer to [3]. An elementary argument gives, for any probability distribution functions $F$ and $G$, both supported on $[0,1]$, the important inequality

$$\|T_m F - T_m G\|_\infty \le \|F - G\|_\infty. \tag{2.15}$$

Moreover, if $F$ is a continuous function on $[0,1]$, then

$$\|T_m F\|_\infty \le 3\|F\|_\infty. \tag{2.16}$$

PROPOSITION 2.1. *For each $F \in C[0,1]$ we have*

$$\|F - T_m F\|_\infty \to 0 \quad \textit{as } m \to \infty. \tag{2.17}$$

*Proof.* The family $\{T_m\}$ of operators on $C[0,1]$ is by (2.16) uniformly bounded and therefore by the Banach–Steinhaus theorem it is sufficient to check (2.17) for $F = L_j$ with any fixed $j$. However, in this case

$$F - T_m F = L_j - T_m L_j = (1 - \lambda_{j,m}) L_j,$$

and an application of (2.11) completes the proof.

In what follows we use the symbol $W_p^m(I)$ for the Sobolev space of order of smoothness $m$ and with the integrability exponent $p$.

THEOREM 2.2. *Let $F \in W_2^2[0,1] \cap \mathcal{F}[0,1]$, $D = d/dx$ and $M = \|D^2 F\|_2$. Then*

$$\|F - T_m F\|_\infty \le \frac{M}{m^{1/4}} \quad \textit{for } m = 2, 3, \ldots. \tag{2.18}$$

*Proof.* For the density $f = DF$ we have

$$(f, 1) = 1 \tag{2.19}$$

with the usual scalar product $(f, g)$ in $L^2[0,1]$, and

$$F(x) - T_m F(x) = \int_0^x (f - R_m f)(z)\,dz, \tag{2.20}$$

where $R_m$ is the Durrmeyer operator corresponding to (2.8). It now follows from (2.20) that

$$\|F - T_m F\|_\infty \le \|f - R_m f\|_2. \tag{2.21}$$

Using the spectral representation of $R_m$ we find that

$$\|f - R_m f\|_2^2 = \sum_{i=1}^{\infty} (1 - \lambda_{i,m})^2 a_i^2, \tag{2.22}$$

where the $a_i$'s are the Fourier–Legendre coefficients and by definition each $\lambda_{i,m}$ for $i > m$ is assumed to be 0. Moreover,

$$a_i = (f, l_i). \tag{2.23}$$

Using the notation from (2.14) we get

$$|a_i| = |(D^2 F, L_i)| \le M\|L_i\|_2. \tag{2.24}$$

It remains to estimate $\|L_i\|_2$. To this end apply the identity

$$2L_{i+1} = \frac{1}{\sqrt{(2i+1)(2i+3)}}\, l_{i+2} - \frac{1}{\sqrt{(2i+1)(2i-1)}}\, l_i, \qquad i \ge 1, \tag{2.25}$$

to get

$$\|L_{i+1}\|_2^2 = \frac{1}{8\lambda_i\big(1 - \frac{3}{4\lambda_i}\big)} \le \frac{1}{8\lambda_i\big(1 - \frac{3}{4\lambda_1}\big)} = \frac{1}{5\lambda_i}, \qquad i \ge 1. \tag{2.26}$$

Thus, for the unique $k$ satisfying the inequalities $\lambda_k \le m < \lambda_{k+1}$ we obtain

$$\sum_{i=k+1}^{\infty} |a_i|^2 \le M^2 \sum_{i=k}^{\infty} \frac{1}{5\lambda_i} = \frac{M^2}{5}\,\frac{1}{k}. \tag{2.27}$$

By the monotonicity in $m$ of $\|f - R_m f\|_2$ and by (2.11) we get

$$\|f - R_m f\|_2^2 \le \sum_{i=1}^{k} \left(\frac{\lambda_i}{m+2}\right)^2 |a_i|^2 + \frac{M^2}{5}\,\frac{1}{k}. \tag{2.28}$$

Now, $\lambda_1 = 2$, $\|L_1\|_2^2 = 1/10$, $\sqrt{m}/2 \le k < \sqrt{m}$, and therefore

$$\sum_{i=1}^{k} \left(\frac{\lambda_i}{m+2}\right)^2 |a_i|^2 \le \frac{3M^2}{5\sqrt{m}}. \tag{2.29}$$

Combining (2.29) and (2.28) we get (2.18).

In the construction below of the spline estimator with random knots, a particular role is played by the following polynomial density function on $[0,1]$:

$$\phi_m(x) := M_{m,2m}(x) = (2m+1)\binom{2m}{m} x^m (1-x)^m, \tag{2.30}$$

where the positive integer $m$ is the given parameter of smoothness. The corresponding polynomial distribution function is

$$\Phi_m(x) = \int_0^x \phi_m(y)\, dy, \tag{2.31}$$

or else (see (2.13))

$$\Phi_m(x) = \sum_{i=m+1}^{2m+1} N_{i,2m+1}(x), \tag{2.32}$$

which permits calculating $\Phi_m(x)$ with the help of the Casteljeau algorithm. It is also important that $\Phi_m(x)$ solves the two-point Hermite interpolation problem of order $m$, i.e.

$$D^k\Phi_m(0) = 0 \quad \text{and} \quad D^k\Phi_m(1) = \delta_{k,0} \quad \text{for } k = 0, \dots, m. \tag{2.33}$$

For later convenience we also introduce the transformed polynomial distribution

$$\Phi_m(x; [a, b]) = \Phi_m\left(\frac{x - a}{b - a}\right). \tag{2.34}$$

Clearly,

$$\Phi_m(a; [a, b]) = 0 \quad \text{and} \quad \Phi_m(b; [a, b]) = 1.$$

On the real line $\mathbb{R}$ similar roles to polynomials and Bernstein polynomials on $[0, 1]$ are played by the cardinal splines and cardinal B-splines, respectively. For a given integer $r \ge 1$ the space $\mathcal{S}^r$ of *cardinal splines* of order $r$ is the space of all functions on $\mathbb{R}$ of class $C^{r-2}$ whose restrictions to $(j, j + 1)$ are polynomials of order $r$, i.e. $D^r S = 0$ on each $(j, j + 1)$ for $j = 0, \pm 1, \dots$. It is known that there is a unique $B^{(r)} \in \mathcal{S}^r$ (up to a multiplicative constant) positive on $(0, r)$ and with support $[0, r]$. The function $B^{(r)}$ becomes unique if normalized e.g. so that

$$\int_{\mathbb{R}} B^{(r)}(y)\, dy = 1. \tag{2.35}$$

The density $B^{(r)}$ can as well be defined probabilistically by the formula

$$P\{U_1 + \dots + U_r < x\} = \int_{-\infty}^{x} B^{(r)}(y)\, dy \tag{2.36}$$

where $U_1, \dots, U_r$ are independent uniformly distributed r.v. on $[0, 1]$. We also need the rescaled cardinal B-splines

$$N_{i,h}^{(r)}(x) = B^{(r)}\left(\frac{x}{h} - i\right) \quad \text{and} \quad M_{i,h}^{(r)}(x) = \frac{1}{h} B^{(r)}\left(\frac{x}{h} - i\right) \quad \text{for } i \in \mathbb{Z},$$

where $h > 0$ is the so called *window parameter*. They share the nice properties

$$\sum_i N_{i,h}^{(r)}(x) = 1 \quad \text{and} \quad \int_{\mathbb{R}} M_{i,h}^{(r)}(y)\, dy = 1. \tag{2.37}$$

Clearly, $\operatorname{supp} N_{i,h}^{(r)} = \operatorname{supp} M_{i,h}^{(r)} = [ih, (i + r)h]$. Moreover, the recurrent formula

$$N_{i,h}^{(r)}(x) = \frac{x - ih}{(r - 1)h} N_{i,h}^{(r-1)}(x) + \frac{(i + r)h - x}{(r - 1)h} N_{i+1,h}^{(r-1)}(x) \tag{2.38}$$

supplies an algorithm for calculating at a given $x$ the value of the cardinal spline

$$\sum_i a_i N_{i,h}^{(r)}(x) \tag{2.39}$$

given the coefficients $(a_i)$. The same algorithm can be used to calculate the value at $x$ of the distribution function

$$\int_{-\infty}^{x} M_{i,h}^{(r)}(y)\,dy = \sum_{j=i}^{\infty} N_{i,h}^{(r+1)}(x). \tag{2.40}$$

Now, just as in the polynomial case (cf. (2.8)) we introduce the spline kernel

$$R_h^{(k,r)}(x,y) = \sum_{i=-\infty}^{\infty} M_{i+\nu,h}^{(k)}(x) N_{i,h}^{(r)}(y). \tag{2.41}$$

In what follows it is assumed that $1 \le k \le r$ and that $r-k$ is even, i.e. $r-k=2\nu$ with $\nu$ being a non-negative integer. It then follows that the supports of $M_{i+\nu,h}^{(k)}$ and of $N_{i,h}^{(r)}$ are concentric. Denote by $C_\pm(\mathbb{R})$ the linear space of all right-continuous functions on $\mathbb{R}$ with finite limits at $\pm\infty$, and by $BV(\mathbb{R})$ the linear space of all continuous functions of finite total variation on $\mathbb{R}$. Now, for $F \in C_\pm(\mathbb{R}) \cup BV(\mathbb{R})$ the value of the operator $T_h^{(k,r)}F$ is defined by the formula

$$\begin{aligned} T_h^{(k,r)}F(x) &= \int_{-\infty}^{x} \Big( \int_{\mathbb{R}} R_h^{(k,r)}(z,y)\,F(dz)\Big)\,dy \\ &= \sum_{i=-\infty}^{\infty} \int_{\mathbb{R}} M_{i+\nu,h}^{(k)}\,dF \int_{-\infty}^{x} N_{i,h}^{(r)}(y)\,dy. \end{aligned} \tag{2.42}$$

To recall from [4] the direct approximation theorem by the operators $T_h^{(k,r)}$, the notion of modulus of a given order $m \ge 1$ is needed:

$$\omega_m(F;\delta) = \sup_{|t|<\delta} \|\Delta_t^m F\|_\infty,$$

where $\Delta_t^m$ is the $m$th order progressive difference with step $t$, i.e.

$$\Delta_t^m f(x) = \sum_{j=0}^{m} (-1)^{j+m} \binom{m}{j} f(x+jt).$$

Now, the following theorem, important for spline estimation, was proved in [4]:

THEOREM 2.3. *Let $1 \le k \le r$ with $r-k$ even and let $h>0$. Then for $F \in C_\pm(\mathbb{R}) \cup BV(\mathbb{R})$,*

$$\|T_h^{(r,k)}F\|_\infty \le \|F\|_\infty, \tag{2.42}$$

*and for* $F \in C_{\pm}(\mathbb{R})$,

$$\|F - T_h^{(k,r)} F\|_\infty \le 2(4 + (r+k)^2)\omega_2(F; h), \tag{2.43}$$

*and in particular for* $k = 1$,

$$\|F - T_h^{(1,r)} F\|_\infty \le \omega_1\Big(F; \frac{r+1}{2}\, h\Big). \tag{2.44}$$

*Consequently, for each continuous probability distribution* $F$,

$$\|F - T_h^{(k,r)} F\|_\infty \to 0 \quad \textit{as } h \to 0_+. \tag{2.45}$$

**3. A smooth modification of $F_n$.** We start, for a given positive integer $m$, with the polynomial density (2.30) and its polynomial probability distribution (2.31). Let now $X_{1:n}, \ldots, X_{n:n}$ with $X_{1:n} \le \cdots \le X_{n:n}$ be the order statistics from the sample $X_1, \ldots, X_n$. For technical reasons we assume that $X_{1:n} < \cdots < X_{n:n}$; otherwise one should consider multiple knots. Moreover, define

$$X_{0:n} = \max\{0, X_{1:n} - (X_{2:n} - X_{1:n})\} = \max\{0, 2X_{1:n} - X_{2:n}\},$$
$$X_{n+1:n} = \min\{X_{n:n} + (X_{n:n} - X_{n-1:n}), 1\} = \min\{2X_{n:n} - X_{n-1:n}, 1\}.$$

Now, for the given sample $X_1, \ldots, X_n$ and for given $m \ge 0$ a new estimator $\Phi_{m,n}(x)$ for $F$ is defined as follows: it equals 0 for $x < X_{0:n}$, 1 for $x \ge X_{n+1:n}$, and for $i = 1, \ldots, n+1$ and $X_{i-1:n} \le x < X_{i:n}$,

$$\Phi_{m,n}(x) = \frac{1}{n}\, \Phi(x; [X_{i-1:n}, X_{i:n}]) + F_n(X_{i-1:n}) - \frac{1}{2n}. \tag{3.1}$$

We consider $\Phi_{m,n}(x)$ as an estimator of the unknown distribution function $F \in \mathcal{F}$ which generates the sample $X_1, \ldots, X_n$. Given the sample, the estimator $\Phi_{m,n}(x)$ may be easily calculated by the Casteljeau algorithm or using the standard incomplete beta functions or distribution functions of beta random variable; the functions are available in numerical computer packages. Summarizing we get

PROPOSITION 3.1. *The estimator $\Phi_{m,n}$ has the following properties*:

1. *It is a piecewise polynomial probability distribution supported on $[X_{0:n}, X_{n+1:n}]$ with knots at the sample points.*
2. *The function $\Phi_{m,n}(x) + 1/2n$ interpolates $F_n$ at the sample points, i.e.*

$$\Phi_{m,n}(X_{i:n}) + \frac{1}{2n} = F_n(X_{i:n}) \quad \textit{for } i = 1, \ldots, n.$$

3. $\Phi_{m,n} \in C^m(\mathbb{R})$.
4. $D^k \Phi_{m,n}(X_{i:n}) = 0$ *for* $k = 1, \ldots, m$ *and* $i = 1, \ldots, n$.
5. $\|\Phi_{m,n} - F\|_\infty \le \|F_n - F\|_\infty + 1/2n$.

6. *Moreover, the following DKW type inequality holds:*

$$P\{\|\Phi_{m,n} - F\|_\infty \geq \varepsilon\} \leq 2\,e^{-2n(\varepsilon-1/2n)^2}, \quad n > \frac{1}{2\varepsilon},\ F \in \mathcal{F}. \tag{3.2}$$

**4. Estimating by polynomials on** $[0,1]$**.** A polynomial estimator on $[0,1]$ (see [3] and cf. (2.12)) can be defined by

$$F_{m,n}(x) = T_m F_n(x), \tag{4.1}$$

where $T_m$ transforms distributions on $[0,1]$, continuous or not, into distribution functions on $[0,1]$ which are polynomials of degree $m+1$. It turns out that for the family $\mathcal{F}[0,1]$ of all continuous distributions supported on $[0,1]$, the DKW type inequality for the estimator $F_{m,n}$ does not hold. More precisely, we have the following negative result

THEOREM 4.1. *There are $\varepsilon > 0$ and $\eta > 0$ such that for any $m$ and $n$ one can find a distribution function $F \in \mathcal{F}[0,1]$ such that*

$$P\{\|F_{m,n} - F\|_\infty > \varepsilon\} > \eta. \tag{4.2}$$

To prove (4.2) it is enough to demonstrate that for some $\varepsilon, \eta > 0$ and for every $n$ and for every odd $m$ the following inequality holds:

$$P\{F_{m,n}(1/2) > F(1/2) + \varepsilon > 0\} > \eta. \tag{4.3}$$

We start the proof by writing formula (4.1) in the form

$$F_{m,n}(x) = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=0}^{m} N_{i,m}(X_j)\int_0^x M_{i,m}(y)\,dy, \tag{4.4}$$

whence by (2.13),

$$F_{m,n}(x) = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=0}^{m} N_{i,m}(X_j)\sum_{k=i+1}^{m+1} N_{k,m+1}(x). \tag{4.5}$$

Let now $\{\varepsilon_k : k = 0,1,\ldots\}$ be a sequence of independent Bernoulli r.v. with $P\{\varepsilon_k = 0\} = 1/2 = P\{\varepsilon_k = 1\}$. Moreover, let $\zeta_{m+1} = \varepsilon_0 + \cdots + \varepsilon_{m+1}$ be the binomial r.v. Then for odd $m = 2\nu - 1$ and $i \leq \nu$,

$$\begin{aligned}\sum_{k=i+1}^{m+1} N_{k,m+1}(1/2) &= P\{\zeta_{m+1} > i\} \geq P\{\zeta_{m+1} > \nu\} \\ &= P\{\zeta_{m+1} \leq \nu\} = 1/2.\end{aligned} \tag{4.6}$$

Consequently, by (2.13) we obtain, for odd $m$,

$$F_{m,n}(1/2) \geq \frac{1}{2n}\sum_{j=1}^{n}\sum_{i=0}^{\nu} N_{i,m}(X_j) \geq \frac{1}{2n}\sum_{j=1}^{n}\int_{X_j}^{1} M_{\nu-1,m-1}(y)\,dy. \tag{4.7}$$

Now, the continuous function

$$M_m(x) = \int_x^1 M_{\nu-1,m-1}(y)\,dy$$

is decreasing on $[0,1]$, $M_m(0) = 1$, and $M_m(1/2) = 1/2$. Consequently, for $\varepsilon \in (0, 1/16)$ there is $\delta > 0$ such that $M_m(\delta + 1/2) > 4\varepsilon$. Choose now $F \in \mathcal{F}[0,1]$ such that

$$F(1/2) < \varepsilon \quad \text{and} \quad F(1/2+\delta) > \eta^{1/n}.$$

Thus, by our hypothesis we get, for $j = 1, \ldots, n$,

$$P_F\{X_j < 1/2 + \delta\} > \eta^{1/n} \quad \text{and} \quad P_F\{M_m(X_j) > 4\varepsilon\} > \eta^{1/n}.$$

Taking into account that

$$\bigcap_{j=1}^{n} \{M_m(X_j) > 4\varepsilon\} \subset \left\{\frac{1}{n}\sum_{j=1}^{n} M_m(X_j) > 4\varepsilon\right\}$$

one obtains

$$P_F\left\{\frac{1}{n}\sum_{j=1}^{n} M_m(X_j) > 4\varepsilon\right\} > \eta,$$

which proves (4.3).

To obtain a positive DKW type inequality for the above polynomial estimators we reduce the space $\mathcal{F}$ of all continuous distribution functions to a smaller subclass. We shall discuss, for a given constant $M$, subclass $W_M$ of $\mathcal{F}[0,1]$ such that $F \in W_M$ if and only if the density $f = F'$ is absolutely continuous and its derivative satisfies

$$\int_0^1 |f'(x)|^2\,dx \le M. \tag{4.8}$$

Theorem 4.2. *Given constants $M > 0$, $\varepsilon > 0$ and $\eta > 0$, one can find explicit values of $m$ and $n$ such that*

$$P\{\|F_{m,n} - F\|_\infty > \varepsilon\} < \eta \quad \textit{for all } F \in W_M. \tag{4.9}$$

*Specifically, it is enough to choose $m$ and $n$ so that*

$$\frac{2M}{m^{1/4}} < \varepsilon \quad \textit{and} \quad 2\exp\left(-2\,\frac{nM^2}{m^{1/2}}\right) < \eta.$$

*Proof.* Let us start with the identity

$$F - F_{m,n} = (F - T_m F) + T_m(F - F_n).$$

Now (see [3]) the triangle inequality and contraction property (2.15) imply

$$\|F - F_{m,n}\|_\infty \le \|F - F_n\|_\infty + \|F - T_m F\|_\infty, \tag{4.10}$$

whence by Theorem 2.2 and by the DKW inequality we get

$$\begin{aligned}
P\{\|F - F_{m,n}\|_\infty > \varepsilon\} &\le P\{\|F - F_{m,n}\|_\infty > 2M/m^{1/4}\} \\
&\le P\{\|F - F_n\|_\infty + \|F - T_m F\|_\infty > 2M/m^{1/4}\} \\
&\le P\{\|F - F_n\|_\infty + M/m^{1/4} > 2M/m^{1/4}\} \\
&= P\{\|F - F_n\|_\infty > M/m^{1/4}\} \le 2\exp\Big(-2\,\frac{nM^2}{m^{1/2}}\Big) < \eta,
\end{aligned}$$

and this completes the proof.

**5. Estimating by splines with equally spaced knots.** In this section we exhibit a rich family of subclasses of continuous probability distributions on $\mathbb{R}$ for which the DKW type inequality holds. The general scheme is very much like that in Section 4 (cf. the proof of Theorem 4.2). We start by recalling the generalized Hölder classes. A function $\omega(h)$ on $\mathbb{R}_+$ is said to be a *modulus of smoothness* if it is bounded, continuous, vanishing at 0, non-decreasing and subadditive, e.g. concave. Suppose we are given integers $(k, r)$ satisfying the conditions formulated just below formula (2.41). Taking into account what we already know it is reasonable to consider the following Hölder classes of probability distributions:

$$H_{\omega,2}^{(k,r)} = \{F \in \mathcal{F} : 2(4 + (r+k)^2)\omega_2(F; h) \le \omega(h) \text{ for all } h > 0\}, \tag{5.1}$$

$$H_{\omega,1}^{(k,r)} = \Big\{F \in \mathcal{F} : \omega_1\Big(F; \frac{r+k}{2}\,h\Big) \le \omega(h) \text{ for all } h > 0\Big\}. \tag{5.2}$$

According to our scheme, the spline estimator with window parameter $h$ and with the given $(k, r)$ is defined by the formula

$$F_{h,n} = F_{h,n}^{(k,r)} = T_h^{(k,r)} F_n. \tag{5.3}$$

For these estimators we also have the basic inequality

$$\|F - F_{h,n}\|_\infty \le \|F - F_n\|_\infty + \|F - T_h^{(k,r)} F\|_\infty. \tag{5.4}$$

We are now in a position to state the DKW type inequality for the generalized Hölder classes of continuous distributions.

THEOREM 5.1. *Let $i = 1, 2$ and $1 \le k \le r$ with $r - k$ even. Then for all $\varepsilon > 0$ and $\eta > 0$ there are $h > 0$ and $n \ge 1$ such that*

$$P\{\|F_{h,n} - F\|_\infty > \varepsilon\} < \eta \quad \textit{for all } F \in H_{\omega,i}^{(k,r)}. \tag{5.5}$$

*It suffices to choose $h$ and $n$ such that*

$$\omega(h) < \varepsilon/2 \quad \textit{and} \quad 2\exp(-n\varepsilon^2/2) < \eta.$$

The proof is immediate from the DKW inequality, (5.4) and Theorem 2.3.

## References

[1] C. de Boor, *Linear combinations of B-splines*, in: Approximation Theory II, Academic Press, 1976, 1–47.
[2] Z. Ciesielski, *Local spline approximation and nonparametric density estimation*, in: Constructive Theory of Functions '87, Sofia, 1988, 79–84.
[3] —, *Nonparametric polynomial density estimation*, Probab. Math. Statist. 9 (1988), 1–10.
[4] —, *Asymptotic nonparametric spline density estimation*, ibid. 12 (1991), 1–24.
[5] Z. Ciesielski and J. Domsta, *The degenerate B-splines as basis in the space of algebraic polynomials*, Ann. Polon. Math. 46 (1985), 71–79.
[6] P. Massart, *The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality*, Ann. Probab. 18 (1990), 1269–1283.
[7] I. J. Schoenberg, *Cardinal interpolation and spline functions*, J. Approx. Theory 2 (1969), 167–206.
[8] E. J. Wegman, *Kernel estimators*, in: Encyclopedia of Statistical Sciences, 2nd ed., Vol. 6, Wiley-Interscience, 2006.
[9] R. Zieliński, *Kernel estimators and the Dvoretzky–Kiefer–Wolfowitz inequality*, Appl. Math. (Warsaw) 34 (2007), 401–404.

Instytut Matematyczny PAN
Abrahama 18
81-825 Sopot, Poland
E-mail: Z.Ciesielski@impan.gda.pl

Instytut Matematyczny PAN
Śniadeckich 8
00-956 Warszawa, Poland
E-mail: R.Zielinski@impan.pl

*Received on 11.7.2008* (1945)