Serdica

Bulgariacae mathematicae publicationes

Сердика

Българско математическо списание

A PROBABILISTIC INTERPRETATION OF MULTIVARIATE B-SPLINES AND SOME APPLICATIONS

Z. G. IGNATOV, V. K. KAISHEV*

The multivariate B-spline is shown to be the density (with respect to the Lebesgue measure) of a linear transformation of order statistics from the uniform disfribution on (0, 1). This interpretation makes it possible to establish exact and asymptotic results both for B-splines and linear transformations of uniform order statistics as well.

1. Introduction. B-splines first introduced as univariate functions by H. B. Curry and I. J. Schoenberg [10] played an important role both as a theoretical and practical tool in dealing with polynomial splines. The geometric interpretation of the B-spline due to Curry and Schoenberg led to a multivariate version of this function, proposed by C. de Boor [6]. Since then, an intensive study of the basic properties of multivariate B-splines and certain linear combinations of such functions was carried out, mostly by C. Micchelli [24], W. Dahmen [11], Dahmen and Micchelli [13], K. Hollig [20], H. Hakopian [16, 17], C. de Boor and K. Hollig [7]. For a more complete list of contributions to the subject we refer to the survey by W. Dahmen and C. Micchelli [13].

The purpose of this paper is to investigate the relation between B-splines and certain probability distributions. Thus, it is shown in Section 2, Lemmas 2.1 and 2.2, that the multivariate B-spline coincides with the density of a random vector, which is a linear transformation of order statistics from the uniform distribution on (0, 1). This establishes an important connection between the theory of spline functions and some aspects of probability theory. On its basis, useful representations of both B-splines and densities of linear transformations of uniform order statistics (LTUOS) are derived. For instance, applying Lemma 2.1 to the recurrence relation for multivariate B-splines due to Micchelli [24] and to that of C. de Boor and M. Cox [5, 9], in the univariate case, yields Lemmas 2.3 and 2.4 respectively. The latter can serve as numerically efficient means of dealing with densities of LTUOS. The importance of the results, given in Section 2 is better understood if we recall, that distributions of LTUOS arise in such branches of statistics as serial correlation, analysis of contingency tables, robust estimation of scale and location parameters. Application of the achievements of B-spline theory in the above mentioned statistical context proves to be promising, both from a theoretical and computational point of view. Probabilistic and statistical facts and techniques, related to order statistics, are shown to be of interest for the study of B-splines as well.

Section 3 represents a short review of what can be added to the asymptotic properties of B-splines and LTUOS, after we have Lemma 2.1.

---

* This work was initiated at the Institute of Mathematics, Bulgarian Academy of Sciences and completed while the second author was a participant in the scientific exchange program between the National Academy of Sciences of the United States of America and the Bulgarian Academy of Sciences.

SERDICA Bulgaricae mathematicae publicationes, Vol. 15, 1989, p. 91—99.

2. Multivariate B-splines and their probabilistic interpretation. The object of this section is to motivate stochastically the notion of the multivariate B-spline. To this end, we recall certain definitions and introduce some notation to be used throughout the sequel. Elements of the Euclidean space R^s, $s\geqq 1$, are denoted by $x, y, z, \dots$, i. e. $x=(x_1, \dots, x_s)$. Superscripts will be used to enumerate vectors x^j, $j=0, \dots, n$. For a given set $A\subset R^s$, $\chi_A(x)$, $[A]$, $\mathrm{vol}_s(A)$, $\dim(A)$ denotes the indicator function, the closed convex hull, the s-dimensional Lebesgue measure, and the dimension, respectively.

We'll now give two equivalent definitions of a multivariate B-spline.

Definition 2.1. *Let the convex hull* $[\{x^0, x^1, \dots, x^n\}]$ *of the set* $\{x^0, \dots, x^n\}\subset R^s$ *have dimension s. The linear functional*

$$I_{[\{x^0, \dots, x^n\}]}(\Phi)=n!\int_{S^n}\Phi(u_0x^0+\dots+u_nx^n)du_1\dots du_n,$$

where $S^n=\{(u_1, \dots, u_n): u_i\geqq 0,\ \Sigma_{i=1}^n u_i\leqq 1\}$, $u_0=1-\Sigma_{i=1}^n u_i$, has a density $M(t;\, x^0, x^1, \dots, x^n)$ i. e.

$$n!\int_{S^n}\Phi(u_0x^0+\dots+u_nx^n)\,du_1\dots du_n=\int_{R^s}\Phi(t)M(t;\, x^0, \dots, x^n)dt \tag{2.1}$$

for all integrable Φ ($\Phi\in L^1(R^s)$). The density $M(t;\, x^0, \dots, x^n)$ we call the multivariate B-spline of degree $n-s$ with knots $x^0, x^1, \dots x^n$.

Definition 2.2. *Let* $\sigma=[v^0, \dots, v^n]$ *be any n-simplex in* R^n, *such that* $v^i|_{R^s}=x^i$, $i=0, 1, \dots, n$ *i. e. the first s coordinates of* v^i *agree with the vector* $x^i\in R^s$. *The multivariate B-spline* $M(t;\, x^0, \dots, x^n)$ is defined as

$$M(t;\, x^0, \dots, x^n)=\mathrm{vol}_{n-s}(\{u\in\sigma: u|_{R^s}=t\})/\mathrm{vol}_n(\sigma),\quad t\in R^s.$$

Before discussing Definitions 2.1 and 2.2 we remark, that the univariate B-spline (i. e. when $s=1$) was first introduced by H. Curry and I. Schoenberg [10] using the notion of a divided difference. Since we encounter divided differences recall

Definition 2.3. *Given any real numbers* $t_1, \dots, t_{r+1}$ *and any sufficiently smooth function* $f(y)$, *we define its r-th order divided difference over the points* $t_1, \dots, t_{r+1}$ by

$$[t_1, \dots, t_{r+1}]f=([t_2, \dots, t_{r+1}]f-[t_1, \dots, t_r]f)/(t_{r+1}-t_1),$$

provided $t_1\neq t_{r+1}$.

If $t_1=t_2=\dots=t_{r+1}$, *then* $[t_1, \dots, t_{r+1}]f=D^rf(t_1)/r!$, *where* $D^rf(t_1)$ *denotes the* r-th *derivative of* $f(y)$ *at* $y=t_1$, $r\geqq 0$, $(D^0f(t_1)=f(t_1))$.

Following Curry and Schoenberg [10] the univariate B-spline $M(t;\, x^0, \dots, x^n)$ of degree $n-1$ with knots $x^0, \dots, x^n\in R^1$ is the n-th order divided difference of the function $f(y)=n(y-t)_+^{n-1}$, i. e.

$$M(t;\, x^0, \dots, x^n)=[x^0, \dots, x^n]f(y), \tag{2.2}$$

where $(x)_+=\max\{0, x\}$ and $\dim[\{x^0, \dots, x^n\}]=1$.

It is known (see L. Schumaker [27], page 46) that when the knots $x^0, \dots, x^n$ are pairwise distinct

$$[x^0, \dots, x^n]f(y)=n\sum_{j=0}^{n}(x^j-t)_+^{n-1}/\prod_{\substack{i=0\\ i\neq j}}^{n}(x^j-x^i). \tag{2.3}$$

When some of the knots $x^0, \dots, x^n$ coincide, i. e.

$$x^0, \ldots, x^n = \underbrace{\tau^0, \ldots, \tau^0}_{\nu_0}, \underbrace{\tau^1, \ldots, \tau^1}_{\nu_1}, \ldots, \underbrace{\tau^l, \ldots, \tau^l}_{\nu_l}, \quad \tau^0 < \tau^1 < \ldots < \tau^l$$

$\nu_0 + \ldots + \nu_l = n+1$ we can apply the representation of a divided difference with multiple knots given in I. Berezin, N. Jidkov ([4], page 127) to the function $f(y) = n(y-t)_+^{n-1}$ to obtain

$$M(t; \underbrace{\tau^0, \ldots, \tau^0}_{\nu_0}, \ldots, \underbrace{\tau^l, \ldots, \tau^l}_{\nu_l}) = \sum_{i=0}^{l} D^{\nu_i - 1} g_i(\tau^i)/(\nu_i - 1)!, \tag{2.4}$$

where $g_i(y) = n(y-t)_+^{n-1} / \prod_{\substack{j=0 \\ j \neq i}}^{l} (y - \tau^j)^{\nu_j}$.

Let us now return to Definitions 2.1 and 2.2 and note that H. Curry, I. Schoenberg [10] used the Hermite—Genocchi formula for divided differences (see N. Nörlund [25], page 16) to establish relation (2.1) for univariate B-splines. The latter was extended to the multivariate case by W. Dahmen [11] and C. Micchelli [24]. Definition 2.2, proposed by C. de Boor [6], reveals an intrinsic geometrical feature of multivariate B-splines. It extends the geometrical interpretation of univariate B-splines due to H. Curry and I. Schoenberg [10], Definitions 2.1 and 2.2 are equivalent and this is proved and discussed by C. Micchelli [24], Lemma 1.

A great deal of properties is known about the B-splines. Thus, $M(t; x^0, \ldots, x^n)$ is a non-negative piecewise polynomial of degree $n-1$ with $n-2$ continuous derivatives when $s=1$ and $x^0, \ldots, x^n$ are distinct real numbers. In general, $(s \geq 1)$, it has been shown that $M(t; x^0, \ldots, x^n)$ is a piecewise polynomial of total degree not exceeding $n-s$ with $n-s-1$ continuous derivatives when the convex hull of every subset of $s+1$ points of $\{x^0, \ldots, x^n\}$ forms an s-dimensional simplex. For a proof of this result, see, C. Micchelli [24] and also C. de Boor [8].

We'll now recall the definition of Dirichlet distributed random variables which will be needed to prove Lemma 2.1.

Definition 2.4. *The random variables $\theta_0, \theta_1 \ldots, \theta_n$ have the joint Dirichlet distribution $D(\nu_0, \nu_1, \ldots, \nu_n)$ with parameters $\nu_0 > 0, \ldots, \nu_n > 0$, $(\theta_0, \ldots, \theta_n) \in D(\nu_0, \ldots, \nu_n)$ if $\theta_0 = 1 - \theta_1 - \ldots - \theta_n$ and the joint probability density of $\theta_1, \ldots, \theta_n$ with respect to the Lebesgue measure on the simplex S^n, introduced in Definition 2.1 is*

$$\frac{\Gamma(\nu_0 + \ldots + \nu_n)}{\Gamma(\nu_0) \ldots \Gamma(\nu_n)} (1 - u_1 - u_2 - \ldots - u_n)^{\nu_0 - 1} u_1^{\nu_1 - 1} \ldots u_n^{\nu_n - 1}, \tag{2.5}$$

($\Gamma(.)$ *is the well-known Gamma function*).

Let $Z_1, Z_2, \ldots, Z_n$ be n independent uniformly distributed on (0, 1) random variables and let $Z_{1,n} \geq Z_{2,n} \geq \ldots \geq Z_{n,n}$ be their order statistics. Consider,

$$\theta_0 = 1 - Z_{1,n}, \ \theta_1 = Z_{1,n} - Z_{2,n}, \ldots, \theta_{n-1} = Z_{n-1,n} - Z_{n,n}, \ \theta_n = Z_{n,n} \tag{2.6}$$

known as the spacings. It is well known, that the random variables $\theta_0, \ldots, \theta_n$ have a Dirichlet distribution with parameters $(\underbrace{1, 1, \ldots, 1}_{n+1})$ (c.f. S. Karlin and M. Taylor [23], page 105).

A probabilistic interpretation of the multivariate B-spline is established by the following.

Lemma 2.1. *Let $x^0, x^1, \ldots, x^n$ be fixed vectors in R^s and let* $\dim([\{x^0, \ldots, x^n\}]) = s$. *Then the density $f_L(t)$ of the random vector*

$$L=\theta_0x^0+\theta_1x^1+\ldots+\theta_{n-1}x^{n-1}+\theta_nx^n$$

with respect to the s-dimensional Lebesgue measure is $M(t;\ x^0,\ x^1,\ldots,\ x^n)$.

Proof. We observe that the left-hand side of (2.1) can also be represented as

$$\int_{R^n} n!\,\chi_{S^n}\Phi(u_0x^0+\ldots+u_nx^n)du_1\ldots du_n. \tag{2.7}$$

Since $n!\,\chi_{S^n}$ is the joint probability density of $\theta_1,\ldots,\theta_n$, given by (2.5) with $\nu_0=\nu_1=\ldots=\nu_n=1$, (2.7) can be viewed as the mean value of the random variable $\Phi(\theta_0x^0+\ldots+\theta_nx^n)$. Hence, (2.1) can be written in the form

$$E\{\Phi(\theta_0x^0+\ldots+\theta_nx^n)\}=\int_{R^s} M(t;\ x^0,\ldots,\ x^n)\Phi(t)dt.$$

On the other hand, $E\{\Phi(\theta_0x^0+\ldots\theta_nx^n)\}=E\{\Phi(L)\}=\int_{R^s}\Phi(t)f_L(t)dt$, where $f_L(t)$ is the density of the random vector L with respect to the s-dimensional Lebesgue measure. So, for every function $\Phi\in L^1(R^s)$ we obtain

$$\int_{R^s} M(t;\ x^0,\ldots,\ x^n)\Phi(t)dt=\int_{R^s}\Phi(t)f_L(t)dt$$

and therefore, using the properties of the Lebesgue integral we conclude that

$$\mathrm{vol}_s(\{t:\ M(t;\ x^0,\ldots,\ x^n)\neq f_L(t)\})=0,$$

i.e. $M(t;\ x^0,\ldots,\ x^n)$ and $f_L(t)$ coincide almost everywhere with respect to the s-dimensional Lebesgue measure. This completes the proof of Lemma 2.1.

Let us now allow the knots coalesce, i. e. assume

$$\{x^0,\ldots,\ x^n\}=\{x^{i_0}=x^{i_1}=\ldots=x^{i_{\nu_0-1}},\ x^{i_{\nu_0}}=x^{i_{\nu_0+1}}=\ldots=x^{i_{\nu_0+\nu_1-1}},\ldots,x^{i_{\nu_0+\ldots\nu_{l-1}}}$$

$$=x^{i_{\nu_0+\ldots+\nu_{l-1}+1}}=\ldots=x^{i_n}\}=\{\underbrace{\tau^0,\ldots,\ \tau^0}_{\nu_0},\ldots,\ \underbrace{\tau^l,\ldots,\ \tau^l}_{\nu_l}\},\ \nu_0+\ldots+\nu_l=n+1.\ \text{Then}$$

$L=\theta_0x^0+\theta_1x^1+\ldots+\theta_nx^n=\tilde\theta_0\tau^0+\tilde\theta_1\tau^1+\ldots+\tilde\theta_l\tau^l$, where we have set $\tilde\theta_0=\theta_{i_0}$

$$+\ldots+\theta_{i_{\nu_0-1}},\ \tilde\theta_1=\theta_{i_{\nu_0}}+\ldots+\theta_{i_{\nu_0+\nu_1-1}},\ldots,\ \tilde\theta_l=\theta_{i_{\nu_0+\ldots+\nu_{l-1}}}+\ldots+\theta_{i_n}.$$

By the properties of Dirichlet distribution (c.f. S. Wilks, S. [30], p. 177) we have $(\tilde\theta_0,\ldots,\ \tilde\theta_l)\in D(\nu_0,\ldots,\nu_l)$. So, we obtain the following more precise formulation of Lemma 2.1.

Lemma 2.1*. *If* $x^0,\ x^1,\ldots,\ x^l\in R^s$, $l\geq s$, *are pairwise distinct and* $\dim([\{x^0,\ldots,\ x^l\}])=s$, *then the B-spline with knots* $x^0,\ldots,\ x^l$ *having multiplicities, correspondingly* $\nu_0,\ldots,\ \nu_l$, $\nu_0+\ldots+\nu_l=n+1$, *i. e.*

$$M(t;\ \underbrace{x^0,\ldots,\ x_0}_{\nu_0},\ldots,\ \underbrace{x^l,\ldots,\ x^l}_{\nu_l})$$

is the density of the random vector $L=\theta_0x^0+\ldots+\theta_lx$ *with respect to the s-dimensional Lebesgue measure, where* $(\theta_0,\ldots,\ \theta_l)\in D(\nu_0,\ldots,\ \nu_l)$.

Throughout the sequel all the densities will be with respect to the Lebesgue measure.

If now in the expression of the random vector L of Lemma 2.1 we set $x^0=y^0$, $x^1-x^0=y^1$, $x^2-x^1=y^2, \ldots, x^n-x^{n-1}=y^n$ and use (2.6), we have

Lemma 2.2. *The density of* $L=y^0+Z_{1,n}y^1+\ldots+Z_{n,n}y^n$ *when* $\dim([\{y^0, y^0+y^1, \ldots, y^0+y^1+\ldots+y^n\}])=s$ *is*

$$F_L(t)=M(t;\, y^0, y^0+y^1, \ldots, y^0+y^1+\ldots+y^n). \tag{2.8}$$

The probabilistic interpretation of B-splines, given by Lemmas 2.1 and 2.2, can serve as a bridge between two different theories, that of spline-functions and order statistics. Along the next paragraphs we will illustrate, how new results for B-splines or linear transformations of order statistics can be established, by correspondingly comparing facts, known independently in each of the two areas.

Let us first mention some formulae, obtained for order statistics, which are seen to have corresponding analogues in the literature on B-splines. For instance, A. Dempster and R. Kleyle [15] have given the univariate $(s-1)$, distribution function $P(L<t)$ for L as in Lemma 2.2 and $y^i\neq 0$, $i=0, \ldots, n$.

The distribution of the circular serial correlation coefficient with lag A

$$_AR_N=\sum_{j=1}^{N}(U_j-\bar U)(U_{j+A}-\bar U)/\sum_{j=1}^{N}(U_j-\bar U)^2, \tag{2.9}$$

where $U_{N+j}=U_j$, $\bar U=\Sigma_{j=1}^N U_j/N$, U_j — independent normally $N(0, 1)$ distributed and $A<$ sample size N, was found by R. Anderson [3]. His expression for the marginal density of $_AR_N$, N-odd, is seen to represent the B-spline

$$M(t;\, {_AC_1}, {_AC_2}, \ldots, {_AC_{n+1}}), \tag{2.10}$$

where $_AC_1, \ldots, {_AC_{n+1}}$ are the distinct latent roots of the matrix of the quadratic form in the numerator of (2.9), $n+1=(N-1)/2$.

As Dempster and Kleyle have pointed out their formula for $P(L<t)$ gives also the distribution of $_AR_N$.

More general expressions for the distribution of L, covering the case of zero y^i-s were derived by M. Ali [1] and also by H. Weisberg [29]. The representation of Ali is a ratio of two determinants while that of Weisberg coincides, after differentiating it, with the right-hand side of (2.4).

In the general case, $(s\geq 1)$, the density of the random vector L from Lemma 2.1 was found by M. Ali and M. Mead [2] under the assumption that the knots $x^0, \ldots, x^n$ are in general position, i. e. every subset of $s+1$ points from $\{x^0, \ldots, x^n\}$ forms a simplex of dimension s. Note that Lemma 2.1 imposes only $\dim([\{x^0, \ldots, x^n\}])=s$.

In two earlier papers, Quenouille [26] and Watson [28] approached a distributional problem, closely related to that of Ali and Mead [2]. They considered the joint distribution of lag 1 to lag s serial correlation coefficients $_1R_N, \ldots, {_sR_N}$, given by (2.9). As Ali and Mead [2] point out, their formula is a special case of Watson's expression for the joint density of $_1R_N, \ldots, {_sR_N}$ and coincides with it, in the corresponding situation. In virtue of Lemma 2.1 we can directly apply the formulae of G. Watson and Ali & Mead to express the multivariate B-spline when the knots are in general position. Such formulae seems not to be known in the B-spline literature. However, on basis of the results of W. Dahmen [11] and Hakopian [17] on multivariate divided differences, an expression for the B-spline of a similar form could be derived. We shall note, that both formulae of Watson and Ali and Mead can hardly serve for direct computations, since they represent sums of ratios of determinants with rather cumbersome summation rule involved.

A numerically more convenient expression for the density of a linear transformation of uniform order statistics can be established by applying Lemma 2.1 to the recurrence relation for multivariate B-splines due to C. Micchelli [24]. The latter holds for arbitrary sets of knots not necessarily in general position and is therefore more general than the formulae of Watson and Ali and Mead. Specifically we have

Lemma 2.3. *For* $n>s+1$, $\dim([\{x^0,\ldots,x^n\}])=s$, and $x=\Sigma_{j=0}^n \lambda_j x^j$, $\Sigma_{j=0}^n \lambda_j=1$,

$$f_L(x)=\frac{n}{n-s}\sum_{j=0}^{n}\lambda_j f_{L^j}(x), \tag{2.11}$$

where $f_L(x)$ *and* $f_{L^j}(x)$ *are correspondingly the densities of the random vectors*

$$L=x^0+(x^1-x^0)Z_{1,n}+(x^2-x^1)Z_{2,n}+\ldots+(x^n-x^{n-1})Z_{n,n},$$

$$L^j=x^0+(x^1-x^0)Z_{1,n}+(x^2-x^1)Z_{2,n-1}+\ldots+(x^{j+1}-x^{j-1})Z_{j,n-1}$$

$$+(x^{j+2}-x^{j+1})Z_{j+1,n-1}+\ldots+(x^n-x^{n-1})Z_{n-1,n-1},$$

$j=0,1,\ldots,n$, (*for* $j=0$, $x^{-1}=0$, $Z_{0,n-1}=1$ *and for* $j=n-1,n$, $Z_{n,n-1}=Z_{n+1,n-1}=0$)

The univariate version of (2.11) is a probabilistic analogue of the well-known de Boor-Cox [5, 9], efficient and stable recurrence relation. Namely, if $n\geq 2$, $s=1$, then

$$f_L(x)=\frac{n}{n-1}\left[\frac{x-x_0}{x_n-x_0}f_{L'}(x)+\frac{x_n-x}{x_n-x_0}f_{L''}(x)\right], \tag{2.12}$$

where

$$L=x_0+(x_1-x_0)Z_{1,n}+\ldots+(x_n-x_{n-1})Z_{n,n},$$

$$L'=x_0+(x_1-x_0)Z_{1,n-1}+\ldots+(x_{n-1}-x_{n-2})Z_{n-1,n-1},$$

$$L''=x_1+(x_2-x_1)Z_{1,n-1}+\ldots+(x_n-x_{n-1})Z_{n-1,n-1}.$$

Furthermore, let $t=\{t_0,t_1,\ldots,t_{n+1}\}$ be a real set obtained from $\{x_0,\ldots,x_n\}$ by addition of the point t_j satisfying $x_{j-1}\leq t_j\leq x_j$, $0<j\leq n$, $n>0$. Then we have

Lemma 2.4

$$f_L(x)=\frac{t_j-t_0}{t_{n-1}-t_0}f_{L'}(x)+\frac{t_{n+1}-t_j}{t_{n+1}-t_0}f_{L''}(x), \tag{2.13}$$

where

$$L=t_0+(t_1-t_0)Z_{1,n}+\ldots+(t_{j+1}-t_{j-1})Z_{j,n}+(t_{j+2}-t_{j+1})Z_{j+1,n}+\ldots+(t_{n+1}-t_n)Z_{n,n},$$

$$L'=t_0+(t_1-t_0)Z_{1,n}+\ldots+(t_n-t_{n-1})Z_{n,n}.\quad L''=t_1+(t_2-t_1)Z_{1,n}+\ldots+(t_{n+1}-t_n)Z_{n,n}.$$

Proof. Follows by Lemma 2.1 and the observation that a (univariate) B-spline. on the grid $\{x_0,\ldots,x_n\}$ is a non-negative linear combination of B-splines on the refined grid t (see C. de Boor [6]).

A formula, more general than (2.11), assuming more than one additional points in the refined grid and related to discrete B-splines is given in Z. Ignatov and V. Kaishev [21].

Relations (2.11), (2.12), (2.13) and some other recurrences of a similar kind can be helpful in computations involving linear combinations of uniform order statistics. For instance, the B-spline distribution function $\int_{-\infty}^x M(t;x^0,\ldots,x^n)dt$ can be expressed recurrently as a sum of B-splines of order $n+1$ (see de Boor [6], page 10). This relation proves useful in deriving efficient numerical algorithms for computing tables of significant points of ${}_AR_N$, $A=1,2\ldots$, whose density is the B-spline given by (2.10), (see Z. Ignatov and V. Kaishev [22]).

Finnally, we give another probabilistic interpretation of the B-spline. For the purpose, recall that the distribution $\mathcal{L}(Z_{1,n}, Z_{2,n}, \dots, Z_{n,n})$ of the order statistics $Z_{1,n} \geqq \dots \geqq Z_{n,n}$ obtained from a set of symmetrically dependent random variables $Z_1, \dots, Z_n$ coincides with the conditional distribution $\mathcal{L}(Z_1, Z_2, \dots, Z_n \mid Z_1 \geqq Z_2 \geqq \dots \geqq Z_n)$, given the probability

$$(2.14) \qquad P\{Z_i = Z_j\} = 0, \quad 1 \leqq i \neq j \leqq n.$$

The random variables $Z_1, \dots, Z_n$ from Lemma 2.1 are independent and identically distributed over the interval $(0, 1)$, hence, they are symmetrically dependent with equations (2.5) fulfilled, and we can write

$$(2.15) \qquad \mathcal{L}(Z_{1,n}, \dots, Z_{n,n}) = \mathcal{L}(Z_1, \dots, Z_n \mid Z_1 \geqq \dots \geqq Z_n).$$

Further, (2.15) implies $\mathcal{L}(y^0 + Z_{1,n} y^1 + \dots + Z_{n,n} y^n) = \mathcal{L}(y^0 + Z_1 y^1 + \dots + Z_n y^n \mid Z_1 \geqq \dots \geqq Z_n)$, and therefore (2.8), Lemma 2.2, can be rewritten as

$$(2.16) \qquad f_L(t) = f_{L'}(t \mid B) = M(t;\, y^0, y^0 + y^1, \dots, y^0 + y^1 \dots + y^n),$$

where $f_{L'}(t \mid B)$ denotes the conditional density of the random vector $L' = y^0 + Z_1 y^1 + \dots + Z_n y^n$ given $B = \{Z_1 \geqq Z_2 \geqq \dots \geqq Z_n\}$.

In the univariate case $(s=1)$, $y^0, y^1, \dots, y^n$ are real numbers. Assume all of them are non zero, then relation (2.16) yields the following

Lemma 2.5. *Let $\eta_1, \eta_2, \dots, \eta_n$ be independent random variables and let η_i be uniformly distributed over the interval* $(\min(y^i, 0), \max(y^i, 0))$, $i = 1, 2, \dots, n$. *Then we have*

$$(2.17) \qquad f_{y^0 + \eta_1 + \dots + \eta_n}\left(t \mid \frac{\eta_1}{y^1} \geqq \frac{\eta_2}{y^2} \geqq \dots \geqq \frac{\eta_n}{y^n}\right) = M(t;\, y^0, y^0 + y^1, \dots, y^0 + y^1 + \dots + y^n).$$

By relation (2.16) and the theorem of total probability we obtain

$$(2.18) \qquad f_{L'}(t) = \sum_{(i_1, \dots, i_n) \in Q_n} P(B(i_1, \dots, i_n)) f_{L'}(t \mid B(i_1, \dots, i_n))$$

$$= \frac{1}{n!} \sum_{(i_1, \dots, i_n) \in Q_n} M(t;\, y^0, y^0 + y^{i_1}, y^0 + y^{i_1} + y^{i_2}, \dots, y^0 + y^{i_1} + \dots + y^{i_n}),$$

where $B(i_1, \dots, i_n) = \{Z_{i_1} \geqq \dots \geqq Z_{i_n}\}$ and the summation is over the set Q_n of all the permutations of the indexes $(1, 2, \dots, n)$. In the univariate case, $(s=1)$, another version of (2.18), based on the notion of discrete B-splines as introduced by de Boor [6] is available, (see Z. Ignatov and V. Kaishev [21]).

3. Summary of asymptotic results for B-splines and linear transformations of order statistics. In this section we shall briefly summarize some interesting applications of Lemma 2.1 to asymptotic results derived independently for B-splines and LTUOS and thus establish new facts for either one of the subjects.

Limits of univariate B-splines were investigated by Curry and Schoenberg [10] and in the multivariate case, $(s \geqq 1)$, by Dahmen and Micchelli [12]. Combining Lemma 2.1 with the corresponding limit theorems for B-splines (Theorem 3, Corollary 4 and Theorem 4, Corollary 5 of Dahmen & Micchelli [12]), one can restate them in terms of linear transformations. For example, by Theorem 3 of Dahmen and Micchelli [12] and Lemma 2.1 we obtain.

Theorem 3.1. *The sequence of random vectors*

$$L_n = y^{0,n} + Z_{1,n} y^{1,n} + \dots + Z_{n,n} y^{n,n}, \quad n \geqq 1 \text{ with}$$
$$\dim([\{y^{0,n}, y^{0,n} + y^{1,n}, \dots, y^{0,n} + \dots - y^{n,n}\}]) = s,$$

$y^{i,n} \in R^s$, *tends weakly to the random vector L if and only if its distribution is a Polya distribution (for a definition of a Polya distribution see Dahmen and Micchelli* [12]).

Further, we'll only indicate results while leaving technical details of their exact reformulation to the interested reader.

Applying Lemma 2.1 to Corollary 4 of Dahmen & Micchelli [12] yields a necessary and sufficient condition for the sequence L_n to converge to a normally distributed random vector L. It can be directly verified that the latter is equivalent to a sufficient condition of Ali and Mead [2] about the asymptotic normality of linear transformations of uniform order statistics and, in the univariate case, to a necessary and sufficient condition for convergence of linear combinations of ordered spacings to the standard normal distribution $N(0, 1)$. Under the condition of asymptotic normality of the sequence of random vectors L_n, an important expansion for the univariate B-spline distribution function

$$\int_{-\infty}^{x} M(t;\ x^{0,n},\ x^{1,n}, \ldots,\ x^{n,n})dt, \tag{3.1}$$

where $x^{j,n} = \Sigma_{i=0}^{j} y^{i,n}$, $j=0, 1, \ldots, n$, easily follows applying Lemma 2.1 to a theorem of van Zwet [31] which establishes the Edgeworth expansion for the distribution function of L_n (for $s=1$). A precise formulation of the expansion for (3.1) is given in Ignatov and Kaishev [21].

REFERENCES

1. M. M. Ali. Distribution of linear combination of order statistics from rectangular population. *Bull. Inst. Statist. Res. Tr.*, **3**, 1969, No 1, 1-21.
2. M. M. Ali, E. R. Mead. On the distribution of several linear combinations of order statistics from the uniform distribution. *Bull. Inst. Statist. Res. Tr.*, **3**, 1969, No 1, 22-41.
3. R. L. Anderson. Distribution of the serial correlation coefficient. *Ann. Math. Statist.*, **13** 1942, 1-13.
4. И. С. Березин, Н. П. Жидков. Методы вычисления. Т. 1. М., 1959.
5. C. de Boor. On calculating with *B*-splines. *J. Approx. Theory*, **6**, 1972, 50-62.
6. C. de Boor. Splines as linear combinations of *B*-splines. — In: Approximation Theory, II. New, York, 1976. 1-47.
7. C. de Boor, K. Hollig. Recurrence relations for multivariate *B*-splines. *Proc. Amer. Math. Soc*, **85**, 1981, 397-400.
8. C. de Boor. Topics in multivariate approximation theory. — In: Topics in Numerical Analysis. (*Lecture Notes in Math., vol. 965*) Berlin, 1982, 39-78.
9. M. G. Cox. The numerical evaluation of *B*-splines. *J. Inst. Math. and Appl.*, **10, 1972**, No 2, 139-149.
10. H. B. Curry, I. J. Schoenberg. On Polya frequency functions IV The fundamental spline functions and their limits. *J. d'Analyse Math.*, **17**, 1966, 71-107.
11. W. Dahmen. On multivariate *B*-splines. *SIAM J. Numer. Anal.*, **17**, 1980, 179-191.
12. W. Dahmen, C., A. Micchelli. On limits of multivariate *B*-splines. *J. d'Analyse Math.*, **39**, 1981, 256-278.
13. W. Dahmen, C., A. Micchelli. Recent progress in multivariate splines. — In: Approximation Theory, IV. New York, 1983.
14. H. A. David. Order Statistics. Sec. ed., New York, 1981.
15. A. P. Dempster, R. M. Kleyle. Distributions determined by cutting a simplex with hyperplanes. *Ann. Math. Statist.*, **39**, 1968, 1473-1478.
16. H. Hakopian. On multivariate *B*-splines. *SIAM. J. Numer. Anal.*, **19**, 1982, 510-517.
17. H. Hakopian. Multivariate divided differences and multivariate interpolation of Lagrange and Hermite type. *J. Approx. Theory*, **34**, 1982, 286-305.
18. H. Hecker. A characterization of the asymptotic normality of linear combinations of order statistics from the uniform distribution. *Ann. Statistics*, **4**, 1976, 1244-1246.

19. L. Holst. On the lengths of the pieces of a stick broken at random. *J. Appl. Prob.*, **17**, 1980, 623-634.
20. K. Hollig. Multivariate splines. *SIAM J. Numer. Anal.*, **19**, 1982, 1013-1031.
21. Z. G. Ignatov, V. K. Kaishev. *B*-splines and linear combinations of uniform order statistics. Math. Research Cent., TSR, 2817, 1985, Univ. of Wisconsin, Madison.
22. Z. G. Ignatov, V. K. Kaishev. *B*-splines and serial correlation. — In: Proceedings of the IV international conference on statistical methods in experimental research and quality control., Varna, Bulgaria, 1986, 138-143.
23. S. Karlin, H. M. Taylor. A second course in stochastic processes. New York, 1981.
24. C. A. Micchelli. A constructive approach to Kergin interpolation in R: Multivariate *B*-splines and Lagrange interpolation. *Rocky Mountain J. Math.*, **10**, 1980, 485-497.
25. N. Nörlund. Vorlesungen über Differenzenrechnung. Berlin, 1924.
26. M. H. Quenouille. The joint distribution of serial correlation coefficients. *Ann. Math. Statist.*, **20**, 1949, 561-71.
27. L. L. Schumaker. Spline functions: Basic Theory. New York, 1981.
28. G. S. Watson. On the joint distribution of the circular serial correlation coefficients. *Biometrika*, **43**, 1956, 161-68.
29. H. Weisberg. The distribution of linear combinations of order statistics from the uniform distribution. *Ann. Math. Statist.*, **42**, 1971, 704-709.
30. S. S. Wilks. Mathematical Statistics. New York, 1962.
31. W. R. van Zwet. The Edgeworth Expansion for linear combinations of uniform order statistics. — In: Proceedings of the Second Prague Symposium on Asymptotic Statistics, Prague, 1979, 93-101.

Institute of Mathematics,
Bulgarian Academy of Sciences,
P. O. Box 373, Sofia,
Bulgaria

Received on 31. 05. 1986