JOURNAL OF APPROXIMATION THEORY **54**, 1–66 (1988)

Urn Models, Approximations, and Splines

RONALD N. GOLDMAN

Control Data Corporation, 4201 Lexington Avenue North, Arden Hills, Minnesota 55112, U.S.A.

Communicated by Oved Shisha

Received April 2, 1984; revised November 25, 1985

Urn models are used to construct polynomials which share many of the shape preserving characteristics of the Bernstein polynomials. Some of these urn models generate splines and one special model is shown to generate the uniform B-splines. The approximation schemes engendered by these polynomials and polynomial splines are studied and their common properties are traced back to their probabilistic origins.

1. INTRODUCTION

One of the simplest and most elegant ways to prove the Weierstrass Approximation Theorem is to show that the Bernstein approximations of a continuous function actually converge to the original function [3]. Unfortunately the convergence of the Bernstein polynomials is sluggish and this slow convergence often precludes their use in practical applications. Nevertheless the Bernstein approximations not only converge to the original function, but they also approximate, in a general way, its shape [3, 11]. It is due to this ability to approximate shape rather than to their convergence properties that the Bernstein polynomials have recently been applied quite successfully in the field of computer aided geometric design [5]. Now the binomial distribution generates the Bernstein polynomials and many of the most important geometric properties of the Bernstein approximations can be derived directly from their probabilistic interpretation. Perhaps then if we wish to generalize the shape approximating characteristics of the Bernstein polynomials, we should look to ways to extend the probabilistic properties of the binomial distribution.

A simple classical way to generate discrete probability distributions is to construct urn models. Urn models extend the probabilistic properties of the

Current address: Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

0021-9045/88 $3.00

binomial distribution in a very natural way and therefore we would expect that the corresponding approximation schemes would also capture many of the geometric properties of the original function. This is indeed the case. Moreover there is a big bonus. Many urn models generate polynomial splines and one particular urn model actually generates B-splines. Therefore it is possible to use discrete urn models to study continuous polynomial splines. These stochastic models provide fresh insight into the discovery and proof of many algebraic results. For example, we can derive many well known properties of B-splines, including the Cox–de Boor recursion formula, by simple counting arguments; thus there is no need to resort to complicated divided difference techniques. It should not be too surprising that a generalization of the binomial distribution leads directly to B-splines. Afterall, B-splines were invented to extend the approximating characteristics of the Berstein polynomials and urn models were created to extend the probabilistic properties of the binomial distribution. What is remarkable is that simple discrete counting arguments can be used to derive sophisticated analytic results.

This paper is divided into four main parts. In Section 2 urn models are introduced and some of their basic properties—symmetry, recursion, moments, laws of signs, limits, and derivatives—are derived. We go on in Section 3 to develop approximation schemes based on these urn models. The fundamental properties of these approximation techniques—convexity, symmetry, recursion, uniqueness, variation dimunition, limits, and derivatives—are traced back to the basic properties of the urn models derived in Section 2. Sections 4 and 5 discuss splines. Section 4 deals with a class of continuous polynomial splines which can be constructed from a distinguished set of distributions introduced in Section 2. The properties of these splines are studied and traced back to the distributions from which they are derived. Section 5 discusses a particularly important special case of the splines constructed in Section 4, namely the uniform B-splines.

We believe that the connection between urn models, approximations and splines leads to a new unity which helps to simplify and generalize many algebraic and geometric properties. We hope that this coupling of probability theory and approximation theory will ultimately prove beneficial to both disciplines and we expect that it will continue to be a fertile area for many future investigations.

2. Urn Models

We begin with a very simple, very general, urn model first introduced by B. Friedman in [6].

Friedman's Urn Model

Consider an urn initially containing w white balls and b black balls. One ball at a time is drawn at random from the urn and its color is inspected. It is then returned to the urn and a constant number c_1 of balls of the same color and a constant number c_2 of balls of the opposite color are added to the urn.

We wish to study the discrete distributions generated by the probabilities of selecting exactly K white balls in the first N trials. When $c_1 = c_2 = 0$, this probability distribution is simply the binomial distribution (sampling with replacement); when $c_1 \neq 0$, $c_2 = 0$, this urn model reduces to the classical Polya–Eggenberger urn model [4]; and when $c_1 = 0$, $c_2 = w + b$, then this urn model generates the normalized uniform B-spline basis functions (see Section 5). Although, in general, there are 4 independent urn parameters—w, b, c_1, c_2—we shall show below that the probability of selecting exactly K white balls in the first N trials always depends on only the following 3 parameters:

$t = w/(w+b) =$ probability of selecting a white ball on the first trial;

$a_1 = c_1/(w+b) =$ percentage of balls of the same color added to the urn after the first trial;

$a_2 = c_2/(w+b) =$ percentage of balls of the opposite color added to the urn after the first trial.

Now we shall be particularly interested in investigating what happens when we hold a_1, a_2 fixed and allow t to very. Therefore we introduce the following notation.

$D_K^N(t) = D_K^N(a_1, a_2, t) =$ probability of selecting exactly K white balls in the first N trials given initial conditions a_1, a_2, t;

$D_N(t) = D_N(a_1, a_2, t) =$ probability distribution consisting of the functions $D_0^N(t), \ldots, D_N^N(t)$;

$s_K^N(t) = s_K^N(a_1, a_2, t) =$ probability of selecting a white ball after selecting exactly K white balls in the first N trials;

$f_K^N(t) = f_K^N(a_1, a_2, t) =$ probability of selecting a black ball after selecting exactly K white balls in the first N trials;

$S_N(t) = S_N(a_1, a_2, t) =$ a priori probability of selecting a white ball on the Nth trial;

$M_r^N(t) = M_r^N(a_1, a_2, t) = r$th moment of the probability distribution $D_N(t)$

$$= \sum K^r D_K^N(t).$$

2.1 *Probability Distributions*

For any fixed values of a_1, a_2, the functions $D_0^N(t), \ldots, D_N^N(t)$ form a discrete probability distribution because they represent the probabilities of

PROPOSITION 2.1.1.

$$\sum_K D_K^N(t) = 1 \qquad D_K^N(t) \geqslant 0, \qquad 0 \leqslant t \leqslant 1.$$

PROPOSITION 2.1.2.

$$f_K^N(t) + s_K^N(t) = 1 \qquad f_K^N(t), s_K^N(t) \geqslant 0, \qquad 0 \leqslant t \leqslant 1.$$

$N+1$ mutually exclusive events one of which must occur. Similarly $f_k^N(t)$, $s_K^N(t)$ represent the probabilities of two mutually exclusive events one of which must occur. Since by definition $0 \leqslant t = w/(w+b) \leqslant 1$, we have the following basic results.

2.2 *Symmetry*

There is symmetry in our urn model between white balls and black balls because whatever action we take when we select a white ball we take a symmetrical action when we select a black ball. Therefore if $D_K^N(t)$ represents the probability of selecting exactly K white balls in the first N trials, then by symmetry $D_K^N(1-t)$ must represent the probability of selecting exactly K black balls in the first N trials. Similarly if $s_K^N(t)$ $(f_K^N(t))$ represents the probability of selecting a white (black) ball after selecting exactly K white balls in the first N trials, then by symmetry $s_K^N(1-t)$ $(f_K^N(1-t))$ must represent the probability of selecting a black (white) ball after selecting exactly K black balls in the first N trials. These simple observations lead directly to the following important results.

PROPOSITION 2.2.1. $D_K^N(t) = D_{N-K}^N(1-t)$.

PROPOSITION 2.2.2. $s_K^N(t) = f_{N-K}^N(1-t), \qquad f_K^N(t) = s_{N-K}^N(1-t)$.

2.3 *Some Explicit Formulas*

It is easy to derive explicit formulas for the functions $s_K^N(t), f_K^N(t)$. Indeed, we have the following general results.

PROPOSITION 2.3.1.

$$s_K^N(t) = \frac{t + Ka_1 + (N-K)\,a_2}{1 + N(a_1 + a_2)}$$

$$f_K^N(t) = \frac{(1-t) + (N-K)\,a_1 + Ka_2}{1 + N(a_1 + a_2)}.$$

Proof. Consider the contents of the urn after selecting exactly K white balls in the first N trials. By definition

$$s_K^N(t) = \frac{\text{number of white balls in urn}}{\text{total number of all balls in urn}}$$

$$= \frac{w + Kc_1 + (N-K)\,c_2}{w + b + N(c_1 + c_2)}.$$

Dividing numerator and denominator by $w+b$, we obtain

$$s_K^N(t) = \frac{t + Ka_1 + (N-K)\,a_2}{1 + N(a_1 + a_2)}.$$

A similar argument (or simply the fact that $f_K^N(t) + s_K^N(t) = 1$) shows that

$$f_K^N(t) = \frac{(1-t) + (N-K)\,a_1 + Ka_2}{1 + N(a_1 + a_2)}. \qquad \text{Q.E.D.}$$

In some special cases it is also possible to obtain explicit formulas for $D_K^N(t)$. For example, if an urn initially contains only white (black) balls then a white (black) ball must be selected on the first trial. This simple observation leads to the following result.

PROPOSITION 2.3.2.

$$D_0^N(1) = 0 \qquad N > 0$$

$$D_N^N(0) = 0 \qquad N > 0.$$

More generally, we have the following formulas for $D_0^N(t)$, $D_N^N(t)$.

PROPOSITION 2.3.3.

$$D_0^N(t) = \prod_{K=0}^{N-1} f_0^K(t) = \prod_{K=0}^{N-1} \frac{(1 - t + Ka_1)}{(1 + K[a_1 + a_2])}$$

$$D_N^N(t) = \prod_{K=0}^{N-1} s_K^K(t) = \prod_{K=0}^{N-1} \frac{(t + Ka_1)}{(1 + K[a_1 + a_2])}.$$

Proof. To select exactly 0 white balls in the first N trials, we must select a black ball on every trial. Therefore

$$D_0^N(t) = \prod_{K=0}^{N-1} f_0^K(t) = \prod_{K=0}^{N-1} \frac{(1 - t + Ka_1)}{(1 + K[a_1 + a_2])}.$$

Similarly to select exactly N white balls in the first N trials, we must select a white ball on every trial. Hence

$$D_N^N(t)=\prod_{K=0}^{N-1} s_K^K(t)=\prod_{K=0}^{N-1}\frac{(t+Ka_1)}{(1+K[a_1+a_2])}. \qquad \text{Q.E.D.}$$

When $a_2=0$, we can calculate $D_K^N(t)$ explicitly for every K.

PROPOSITION 2.3.4. *If $a_2=0$, then*

$$D_K^N(t)=\binom{N}{K}\frac{t\cdots(t+[K-1]\,a_1)(1-t)\cdots(1-t+[N-K-1]\,a_1)}{(1+a_1)\cdots(1+[N-1]\,a_1)}.$$

Proof. By Proposition 2.3.1 when $a_2=0$, we have

$$s_J^L(t)=\frac{t+Ja_1}{1+La_1}$$

$$f_J^L(t)=\frac{(1-t)+(L-J)\,a_1}{1+La_1}.$$

Now there are $\binom{N}{K}$ ways of selecting exactly K white balls in the first N trials. To calculate the probability of any one particular way, we must multiply together K success factors of type $s_J^L(t)$ and $N-K$ failure factors of type $f_J^L(t)$ where for each L either $s_J^L(t)$ or $f_J^L(t)$ must appear but not both. Now $J=0, 1, \ldots, K$, $L=0, 1, \ldots, N-1$, and $L-J=0, 1, \ldots, N-K-1$. Therefore, collecting all these factors, we obtain

$$D_K^N(t)=\binom{N}{K}\frac{t\cdots(t+[K-1)\,a_1)(1-t)\cdots(1-t+[N-K-1]\,a_1)}{(1+a_1)\cdots(1+[N-1]\,a_1)}.$$

Q.E.D.

COROLLARY 2.3.5. *If $a_1=a_2=0$, then*

$$D_K^N(t)=\binom{N}{K}\,t^K(1-t)^{N-K}.$$

The case $a_2=0$ is the classical Polya–Eggenberger urn model, and the case $a_1=a_2=0$ is, of course, just the binomial distribution. When $a_2\neq 0$, it is not so easy to derive facile explicit formulas for the functions $D_K^N(t)$. In these cases we must resort to a simple recursion formula which we shall derive in the following section.

2.4 *The Recursion Formula*

The following simple recursion formula is fundamental to the investigation of urn models.

PROPOSITION 2.4.1. $D_K^{N+1}(t)=f_K^N(t)\,D_K^N(t)+s_{K-1}^N(t)\,D_{K-1}^N(t)$.

Proof. In order to select exactly K white balls in the first $N+1$ trials, we must select either exactly K or exactly $K-1$ white balls in the first N trials. Thus the probability of selecting exactly K white balls in the first $N+1$ trials $[D_K^{N+1}(t)]$ is equal to the sum of the probabilities of two mutually exclusive events.

1. The probability of selecting exactly K white balls in the first N trails $[D_K^N(t)]$ and then selecting a black ball on the $N+1$st trial $[f_K^N(t)]$.
2. The probability of selecting exactly $K-1$ white balls in the first N trials $[D_{K-1}^N(t)]$ and then selecting a white ball on the $N+1$st trial $[s_{K-1}^N(t)]$.

Translating English to Algebra yields our result. Q.E.D.

If $a_1=a_2=0$, then

$$f_K^N(t)=1-t$$

$$s_{K-1}^N(t)=t$$

$$D_K^{N+1}(t)=(1-t)\,D_K^N(t)+tD_{K-1}^N(t).$$

This recursion formula is, of course, just the standard recursion formula for the Bernstein polynomials (binomial distribution).

If $a_1=0$, $a_2=1$, then

$$f_K^N(t)=\frac{1-t+K}{1+N}$$

$$s_{K-1}^N(t)=\frac{t+N+1-K}{1+N}$$

$$D_K^{N+1}(t)=\frac{(1-t+K)}{(1+N)}\,D_K^N(t)+\frac{(t+N+1-K)}{(1+N)}\,D_{K-1}^N(t).$$

Now this recursion formula is actually just the Cox–de Boor recursion formula for *B*-splines in its simplest form (see Section 5).

Since we have explicit formulas for the functions $f_K^N(t)$, $s_K^N(t)$ (Proposition 2.3.1), and since by definition

$$D_0^1(t)=1-t,$$

$$D_1^1(t)=t,$$

we can use the recursion formula to calculate $D_K^N(t)$ for any values of K, N. Moreover it follows from the recursion formula and induction on N that $D_K^N(t)$ is a degree N polynomial in t which depends only on a_1, a_2, t, K, N. Thus we can state the following corollary.

COROLLARY 2.4.2. *The functions* $D_0^N(t), \ldots, D_N^N(t)$ *are degree* N *polynomials in* t.

In the next section we shall show that the $N+1$ polynomials $D_0^N(t), \ldots, D_N^N(t)$ actually form a basis for all the degree N polynomials in t.

2.5 *Expectation and Higher Order Moments*

In this section we shall study the $N+1$ moments $M_0^N(t), \ldots, M_N^N(t)$. Explicit formulas for these moments when $a_2 = 0$ are given in [13]. Here we shall examine the general case. To begin with $M_0^N(t)$, notice that we can restate Proposition 2.1.1 in the following manner.

PROPOSITION 2.5.1. $M_0^N(t) = 1$.

The first moment, $M_1^N(t) = \sum_{K=1}^N KD_K^N(t)$, is the classical expectation of the distribution $D_N(t)$. Therefore we have the following general result.

PROPOSITION 2.5.2. $M_1^N(t) = \sum_{K=1}^N S_K(t)$.

Proof. Simply observe that the expectation of several disjoint random events is just the sum of the expectations of each individual event and the expectation of a single random event is simply the probability of that event. Therefore, in our case, the expected number of white balls selected in the first N trials must be equal to the sum of the a priori probabilities of selecting a white ball in each of the first N trials. Q.E.D.

COROLLARY 2.5.3. $S_{N+1}(t) = M_1^{N+1}(t) - M_1^N(t)$.

Thus the problem of computing the moments $M_1^N(t)$ is equivalent to the problem of computing the a priori probabilities $S_K(t)$.

PROPOSITION 2.5.4. $S_{N+1} = \sum_K s_K^N(t)\, D_K^N(t)$.

Proof. The a priori probability of selecting a white ball on the $N+1$st trial is equal to the sum of the probabilities of all the possible, mutually exclusive, ways of selecting a white ball on the $N+1$st trial. But the only possible ways in which we can select a white ball on the $N+1$st trial are first to select N balls some number K of which are white $[D_K^N(t)]$ and then to select a white ball on the $N+1$st trial $[s_K^N(t)]$. Q.E.D.

PROPOSITION 2.5.5.

$$S_{N+1}(t)=\frac{t+Na_2+(a_1-a_2)\,M_1^N(t)}{1+N(a_1+a_2)}.$$

Proof. By Proposition 2.5.4 and 2.3.1,

$$\begin{aligned}
S_{N+1}(t)&=\sum_K s_K^N(t)\,D_K^N(t)\\
&=\sum_K\left[\frac{t+Ka_1+(N-K)\,a_2}{1+N(a_1+a_2)}\right]D_K^N(t)\\
&=\left[\frac{t+Na_2}{1+N(a_1+a_2)}\right]\sum_K D_K^N(t)+\left[\frac{a_1-a_2}{1+N(a_1+a_2)}\right]\sum_K KD_K^N(t)\\
&=\frac{t+Na_2+(a_1-a_2)\,M_1^N(t)}{1+N(a_1+a_2)}.
\end{aligned}$$

Q.E.D.

COROLLARY 2.5.6.

$$M_1^{N+1}(t)=\frac{t+Na_2+[1+(N+1)\,a_1+(N-1)\,a_2]\,M_1^N(t)}{1+N(a_1+a_2)}.$$

Proof. This result is an immediate consequence of Corollary 2.5.3 and Proposition 2.5.5.

COROLLARY 2.5.7. *There exist constants p_N, q_N such that*

$$M_1^N(t)=p_Nt+q_N$$

(i) $p_N>0$

(ii) $q_N\geqslant 0$

(iii) $p_N+2q_N=N.$

Proof. By induction on N. Certainly this result is true for $N=1$ since $M_1^1(t)=t$. Now it follows easily from Corollary 2.5.6 and the inductive hypothesis that $M_1^N(t)$ is linear in t. Therefore there exist constants p_N, q_N such that

$$M_1^N(t)=p_Nt+q_N.$$

Moreover, again by Corollary 2.5.6, we have the recursion formulas

$$p_{N+1}=\frac{1+(1+(N+1)\,a_1+(N-1)\,a_2)\,p_N}{1+N(a_1+a_2)}$$

$$q_{N+1}=\frac{Na_2+(1+(N+1)\,a_1+(N-1)\,a_2)\,q_N}{1+N(a_1+a_2)}.$$

Therefore it follows easily by induction on N that

$$p_N > 0$$
$$q_N \geqslant 0.$$

Finally from the recursion formulas and the inductive hypothesis

$$\begin{aligned}
p_{N+1} + 2q_{N+1} &= \frac{1 + 2Na_2 + (1 + (N+1)\,a_1 + (N-1)\,a_2)(p_N + 2q_N)}{1 + N(a_1 + a_2)} \\
&= \frac{1 + 2Na_2 + (1 + (N+1)\,a_1 + (N-1)\,a_2)N}{1 + N(a_1 + a_2)} \\
&= \frac{(N+1) + (N+1)\,Na_1 + (N+1)\,Na_2}{1 + Na_1 + Na_2} \\
&= N + 1.
\end{aligned}$$

Q.E.D.

For the binomial distribution, and more generally for the Polya–Eggenberger urn model, we have the following more special results.

COROLLARY 2.5.8. *If* $a_2 = 0$, *then for all* N

(i) $s_N(t) = t$,

(ii) $M_1^N(t) = Nt$.

Proof. Again these results follows easily by induction on N. Indeed these results are clearly valid for $N = 1$. Moreover by the inductive hypothesis, Proposition 2.5.5, and Corollary 2.5.3

$$S_{N+1}(t) = t$$
$$M_1^{N+1}(t) = (N+1)\,t.$$

Q.E.D.

If $a_2 = 0$, then by Corollary 2.5.8 the a priori probability of selecting a white ball on any trial is the same as the probability of selecting a white ball on the first trial. This result is obvious when $a_1 = a_2 = 0$ (binomial distribution) since in this case the contents of the urn are the same for every trial. It is rather remarkable, though well known [2], that this result is still valid even when $a_1 \neq 0$ and the contents of the urn vary from trial to trial. Even more astonishing is the following result.

COROLLARY 2.5.9. *If* $a_2 = 1 + a_1$, *then*

(i) $S_N(t) = t \qquad N = 1$

$\qquad = 1/2 \qquad N \neq 1$,

(ii) $M_1^N(t) = t + (N-1)/2$.

Proof. Same as Corollary 2.5.8.

If $a_2 = 1 + a_1$, then by Corollary 2.5.9 the a priori probability of selecting a white ball on any trial after the first is always exactly 1/2. This extraordinary result suggests that these particular distributions may have other remarkable properties. We will return to study these special distributions further in Section 4.

By Proposition 2.5.1 the zeroth moment, $M_0^N(t)$, is simply a constant, and by Corollary 2.5.7 the first moment, $M_1^N(t)$, is always a linear function in t. We shall now generalize these results to higher order moments. We begin with a recursion formula which expresses the rth moment of $D_{N+1}(t)$ in terms of the first r moments of $D_N(t)$.

PROPOSITION 2.5.10 (Recursion Formula for Moments).

$$M_r^{N+1}(t) = \left[\frac{1+(N+r)\,a_1+(N-r)\,a_2}{1+N(a_1+a_2)}\right] M_r^N(t)$$
$$+ \sum_{i=1}^{r-1} \left[\frac{\binom{r}{i}t + \binom{r}{i-1}a_1 + \left[N\binom{r}{i} - \binom{r}{i-1}\right]a_2}{1+N(a_1+a_2)}\right] M_i^N(t)$$
$$+ \frac{t+Na_2}{1+N(a_1+a_2)}.$$

Proof. By Propositions 2.4.1, 2.3.1, 2.1.2,

$$\begin{aligned}
M_r^{N+1}(t) &= \sum_K K^r D_K^{N+1}(t)\\
&= \sum_K K^r f_K^N(t)\, D_K^N(t) + \sum_K K^r s_{K-1}^N(t)\, D_{K-1}^N(t)\\
&= \sum_K K^r D_K^N(t) + \sum_K \left[(K+1)^r - K^r\right] s_K^N(t)\, D_K^N(t)\\
&= M_r^N(t) + \sum_{i=0}^{r-1} \sum_K \binom{r}{i} \left[\frac{t+Ka_1+(N-K)\,a_2}{1+N(a_1+a_2)}\right] K^i D_K^N(t)\\
&= M_r^N(t) + \sum_{i=0}^{r-1} \frac{\binom{r}{i}(t+Na_2)}{1+N(a_1+a_2)} M_i^N(t)\\
&\quad + \sum_{i=0}^{r-1} \frac{\binom{r}{i}(a_1-a_2)}{1+N(a_1+a_2)} M_{i+1}^N(t)
\end{aligned}$$

$$= \left[\frac{1+(N+r)\,a_1+(N-r)\,a_2}{1+N(a_1+a_2)}\right] M_r^N(t)$$

$$+\sum_{i=1}^{r-1}\left[\frac{\binom{r}{i}t+\binom{r}{i-1}a_1+\left[N\binom{r}{i}-\binom{r}{i-1}\right]a_2}{1+N(a_1+a_2)}\right] M_i^N(t)$$

$$+\frac{t+Na_2}{1+N(a_1+a_2)}. \qquad \text{Q.E.D.}$$

Notice that Corollary 2.5.6, the recursion formula for expectation, is just a special case of the general recursion formula for moments. As additional consequences of this general recursion formula, we have the following results.

COROLLARY 2.5.11. *If* $0 \leqslant r \leqslant N$, *then there exist constants* $p_r^{N,r}, \ldots, p_r^{N,0}$ *such that*

$$M_r^N(t) = p_r^{N,r} t^r + \cdots + p_r^{N,0}$$

(i) $p_r^{N,r} > 0$

(ii) $p_N^{N,N} = \dfrac{N!}{\prod_{K=0}^{N-1}[1+K(a_1+a_2)]}$

(iii) $p_r^{N,i} \geqslant 0.$

Proof. By induction on N. Certainly by Proposition 2.5.1 and Corollary 2.5.7 this result is true for $N=1$. Moreover by the recursion formula and the inductive hypothesis if $0 \leqslant r \leqslant N$, then

$$p_r^{N+1,r} = \frac{[1+(N+r)\,a_1+(N-r)\,a_2]\,p_r^{N,r}+rp_{r-1}^{N,r-1}}{1+N(a_1+a_2)} > 0$$

$$p_{N+1}^{N+1,N+1} = \frac{(N+1)}{1+N(a_1+a_2)}\,p_N^{N,N}$$

$$= \frac{(N+1)!}{\prod_{K=0}^{N}[1+K(a_1+a_2)]}.$$

Similarly since each term in the recursion formula is non-negative

$$p_r^{N,i} \geqslant 0. \qquad \text{Q.E.D.}$$

COROLLARY 2.5.12. *The* $N+1$ *moments* $M_0^N(t), \ldots, M_N^N(t)$ *are a basis for the degree N polynomials in t.*

COROLLARY 2.5.13. *The $N+1$ distribution functions $D_0^N(t), \ldots, D_N^N(t)$ are a basis for the degree N polynomials in t.*

2.6 *Conjectures Concerning the Laws of Signs*

For any finite sequence of real numbers $C=(c_0, \ldots, c_N)$, let $z(C)$ denote the number of zeroes and $v(C)$ denote the number of sign changes ignoring zeroes in C. That is, set

$$z(c_0, \ldots, c_N) = \text{number of zeroes in } (c_0, \ldots, c_N)$$

$$v(c_0, \ldots, c_N) = \text{number of sign alternations in } (c_0, \ldots, c_N).$$

For a continuous real-valued function g, define the number of zeroes $z(g)$ and the number of sign changes $v(g)$ in the interval (a, b) by setting

$$z(g) = \sup z[g(c_0), \ldots, g(c_N)]$$

$$v(g) = \sup v[g(c_0), \ldots, g(c_N)],$$

where the supremums are taken over all finite sequences $a < c_0 < \cdots < c_N < b$. By continuity it follows that in any interval

$$v(g) \leqslant z(g).$$

An ordered collection of continuous functions $F_0(t), \ldots, F_N(t)$ is said to satisfy the Weak Law of Signs in the interval (a, b) iff for every sequence of constants $c_0, \ldots, c_N$

$$v\left[\sum c_K F_K(t)\right] \leqslant v(c_0, \ldots, c_N).$$

Similarly an ordered collection of continuous functions $F_0(t), \ldots, F_N(t)$ is said to satisfy the Strong Law of Signs, or Descartes' Law of Signs, in the interval (a, b) iff for every sequence of constants $c_0, \ldots, c_N$, c_K not all zero,

$$z\left[\sum c_K F_K(t)\right] \leqslant v(c_0, \ldots, c_N).$$

By continuity it again follows that in any interval

$$v\left[\sum c_K F_K(t)\right] \leqslant z\left[\sum c_K F_K(t)\right].$$

Therefore in any interval

$$\text{Strong Law of Signs} \Rightarrow \text{Weak Law of Signs}.$$

It is also obvious from our definitions that

$$\text{Strong Law of Signs} \Rightarrow \text{Linear Independence.}$$

Thus linear independence is a necessary, but not a sufficient, condition for a sequence of functions to satisfy the Strong Law of Signs. A necessary and sufficient condition for the functions $F_0(t), \ldots, F_N(t)$ to satisfy the Strong Law of Signs is that for any sequence $a < t_0 < t_1 < \cdots < t_N < b$ the $K \times K$ subdeterminants of

$$\begin{vmatrix} F_0(t_0) \cdots F_N(t_0) \\ \vdots \qquad \vdots \\ F_0(t_N) \cdots F_N(t_N) \end{vmatrix}$$

are of 1 strict sign [12].

It is easy to show that if $a_1 = a_2 = 0$ (binomial distribution), then the functions $D_0^N(t), \ldots, D_N^N(t)$ satisfy the Strong Law of Signs in the interval $(0, 1)$ [11], and this result remains valid if $a_1 \neq 0$, $a_2 = 0$ (the Polya–Eggenberger urn model) [7]. In addition, the functions $D_0^N(t), \ldots, D_N^N(t)$ are known to satisfy the Strong Law of Signs in the interval $(0, 1)$ when $a_1 = 0$, $a_2 = 1$ (uniform B-splines) and this result remains valid when $a_1 = 0$, $a_2 \neq 1$ (non-uniform B-splines) [12]. These special cases, together with the fact that by Corollary 2.5.13 the functions $D_0^N(t), \ldots, D_N^N(t)$ are linearly independent for all values of a_1, a_2, prompt us to propose the following conjectures.

Conjecture 2.6.1. For all positive finite values of a_1, a_2, the ordered set of functions $D_0^N(t), \ldots, D_N^N(t)$ satisfies the Weak Law of Signs in the interval $(0, 1)$.

Conjecture 2.6.2. For all positive finite values of a_1, a_2, the ordered set of functions $D_0^N(t), \ldots, D_N^N(t)$ satisfies the Strong Law of Signs in the interval $(0, 1)$.

Clearly

$$\text{Conjecture 2.6.2} \Rightarrow \text{Conjecture 2.6.1}$$

but, as yet, we know of no proof, probabilistic or otherwise, for either general conjecture. However, there is some numerical evidence for Conjecture 2.6.2. Heath has written a computer program to compute subdeterminants of matrices of the form

$$\begin{vmatrix} D_0^N(t_0) \cdots D_N^N(t_0) \\ \vdots \qquad \vdots \\ D_0^N(t_N) \cdots D_N^N(t_N) \end{vmatrix}.$$

He has tested several hundred random numerical examples and in all cases has found that the $K \times K$ subdeterminants were indeed of one strict sign [9]. We shall discuss the geometric significance of these conjectures further in Section 3.5.

2.7 *Some Simple Limits*

In this section we shall study the behavior of the probability distributions $D_N(a_1, a_2, t)$ as either a_1 or a_2 or both approach infinity. For each of our results we shall give both an intuitive and a rigorous argument. The reason for this apparent overkill is that our rigorous demonstrations are based on the recursion formula and follow by induction on N. Now induction is a fine technique for proof, but not for discovery. Our non-rigorous intuitive arguments provide the insight and motivation which are lacking in the inductive proofs.

To begin, consider what happens to the functions $f_K^N(a_1, a_2, t)$, $s_K^N(a_1, a_2, t)$ as a_1 approaches infinity. Let c_1, the number of balls of the same color added to the urn after each trial, be very large compared to the other urn parameters c_2, w, b. Then after selecting exactly K white balls in the first N trials, the urn will contain approximately Kc_1 white balls and $(N-K)\,c_1$ black balls. Therefore when a_1 is large.

$$f_K^N(a_1, a_2, t) \approx (N-K)/N$$

$$s_K^N(a_1, a_2, t) \approx K/N$$

and these approximations become more exact as a_1 (and hence c_1) approaches infinity. This argument suggests the following lemma.

LEMMA 2.7.1.

$$\lim_{a_1 \to \infty} f_K^N(a_1, a_2, t) = (N-K)/N$$

$$\lim_{a_1 \to \infty} s_K^N(a_1, a_2, t) = K/N.$$

Proof. By Proposition 2.3.1

$$\lim_{a_1 \to \infty} f_K^N(a_1, a_2, t) = \lim_{a_1 \to \infty} \frac{(1-t) + (N-K)\,a_1 + Ka_2}{1 + N(a_1 + a_2)}$$

$$= (N-K)/N$$

$$\lim_{a_1 \to \infty} s_K^N(a_1, a_2, t) = \lim_{a_1 \to \infty} \frac{t + Ka_1 + (N-K)\,a_2}{1 + N(a_1 + a_2)}$$

$$= K/N. \qquad \text{Q.E.D.}$$

Now consider what happens to the functions $f_K^N(a_1, a_2, t)$, $s_K^N(a_1, a_2, t)$ as a_2 approaches infinity. Let c_2, the number of balls of the opposite color added to the urn after each trial, be very large compared to the other urn parameters c_1, w, b. Then after selecting exactly K white balls in the first N trials, the urn will contain approaximately $(N-K)\,c_2$ white balls and Kc_2 black balls. Therefore when a_2 is large

$$f_K^N(a_1, a_2, t) \approx K/N$$
$$s_K^N(a_1, a_2, t) \approx (N-K)/N$$

and again these approximations become more exact as a_2 (and hence c_2) approaches infinity. This argument suggests the following lemma.

LEMMA 2.7.2.

$$\lim_{a_2 \to \infty} f_K^N(a_1, a_2, t) = K/N = f_K^{N-1}(0, 1, 1)$$

$$\lim_{a_2 \to \infty} s_K^N(a_1, a_2, t) = (N-K)/N = s_K^{N-1}(0, 1, 1).$$

Proof. By Proposition 2.3.1

$$\begin{aligned}\lim_{a_2 \to \infty} f_K^N(a_1, a_2, t) &= \lim_{a_2 \to \infty} \frac{(1-t) + (N-K)\,a_1 + Ka_2}{1 + N(a_1 + a_2)}\\ &= K/N\\ &= f_K^{N-1}(0, 1, 1)\end{aligned}$$

$$\begin{aligned}\lim_{a_2 \to \infty} s_K^N(a_1, a_2, t) &= \lim_{a_2 \to \infty} \frac{t + Ka_1 + (N-K)\,a_2}{1 + N(a_1 + a_2)}\\ &= (N-K)/N\\ &= s_K^{N-1}(0, 1, 1).\end{aligned}$$

Q.E.D.

Finally let us consider what happens to the functions $f_K^N(a_1, a_2, t)$, $s_K^N(a_1, a_2, t)$ when both a_1, a_2 approach infinity. Suppose $a_2 = a_1 + p$ for some fixed constant p. Let c_1, the number of balls of the same color added to the urn after each trial, be very large compared to w, b, p. Then since $a_2 = a_1 + p$, $c_2 \approx c_1$. Therefore after each pick an equal number of balls of each color must be added to the urn. Hence after N trials the urn will contain approximately Nc_1 white balls and Nc_2 black balls. Thus when a_1 is large

$$f_K^N(a_1, a_2, t) \approx \frac{Nc_1}{2Nc_1} = 1/2$$

$$s_K^N(a_1, a_2, t) \approx \frac{Nc_1}{2Nc_1} = 1/2$$

and as usual these approximations become more exact as a_1 (and hence c_1) approaches infinity. This argument suggests the following lemma.

LEMMA 2.7.3. *Let p be a fixed constant, and let $a_2 = a_1 + p$. Then*

$$\lim_{a_1 \to \infty} f_K^N(a_1, a_2, t) = 1/2 = f_K^N(1/2, 1/2, 1/2)$$

$$\lim_{a_1 \to \infty} s_K^N(a_1, a_2, t) = 1/2 = s_K^N(1/2, 1/2, 1/2).$$

Proof. By Proposition 2.3.1

$$\begin{aligned}
\lim_{a_1 \to \infty} f_K^N(a_1, a_2, t) &= \lim_{a_1 \to \infty} \frac{(1-t) + (N-K)\,a_1 + Ka_2}{1 + N(a_1 + a_2)} \\
&= \lim_{a_1 \to \infty} \frac{(1-t) + (N-K)\,a_1 + K(a_1 + p)}{1 + N(2a_1 + p)} \\
&= 1/2 \\
&= f_K^N(1/2, 1/2, 1/2)
\end{aligned}$$

$$\begin{aligned}
\lim_{a_1 \to \infty} s_K^N(a_1, a_2, t) &= \lim_{a_1 \to \infty} \frac{(1-t) + Ka_1 + (N-K)\,a_2}{1 + N(a_1 + a_2)} \\
&= \lim_{a_1 \to \infty} \frac{t + Ka_1 + (N-K)(a_1 + p)}{1 + N(2a_1 + p)} \\
&= 1/2 \\
&= s_K^N(1/2, 1/2, 1/2).
\end{aligned}$$

Q.E.D.

Now let us consider what happens to the probability distributions $D_N(a_1, a_2, t)$ as a_1 approaches infinity. Let $c_1 >> c_2$, w, b. Then after the first trial almost all the balls in the urn will be of the same color as the ball selected on the first trial. Therefore, with a probability approaching 1, all the balls selected on subsequent trials will be of the same color as the ball selected on the first trial. Hence when a_1 is large

$D_0^N(a_1, a_2, t) \approx$ probability that the ball selected on the first trial is black

$= 1 - t$

$$D_N^N(a_1, a_2, t) \approx \text{probability that the ball selected on the first trial is white}$$

$$= t$$

$$D_K^N(a_1, a_2, t) \approx 0, \qquad K \neq 0, N$$

and these approximations become more exact as a_1 (and hence c_1) approaches infinity. This argument suggests the following proposition.

PROPOSITION 2.7.4.

$$\begin{aligned} \lim_{a_1 \to \infty} D_K^N(a_1, a_2, t) &= 1 - t & K &= 0 \\ &= 0 & K &\neq 0, N \\ &= t & K &= N. \end{aligned}$$

Proof. By induction on N. Certainly this result is true for $N = 1$ since

$$D_0^1(a_1, a_2, t) = 1 - t$$

$$D_1^1(a_1, a_2, t) = t$$

are independent of a_1, a_2. Now by Proposition 2.4.1 and Lemma 2.7.1

$$\begin{aligned} \lim_{a_1 \to \infty} D_K^{N+1}(t) &= \lim_{a_1 \to \infty} f_K^N(a_1, a_2, t)\, D_K^N(a_1, a_2, t) \\ &\quad \lim_{a_1 \to \infty} s_{K-1}^N(a_1, a_2, t)\, D_{K-1}^N(a_1, a_2, t) \\ &= \frac{(N-K)}{N} \lim_{a_1 \to \infty} D_K^N(a_1, a_2, t) \\ &\quad + \frac{(K-1)}{N} \lim_{a_1 \to \infty} D_{K-1}^N(a_1, a_2, t). \end{aligned}$$

Therefore by the inductive hypothesis

$$\begin{aligned} \lim_{a_1 \to \infty} D_K^{N+1}(t) &= 1 - t & K &= 0 \\ &= 0 & K &\neq 0, N \\ &= t & K &= N. \end{aligned}$$

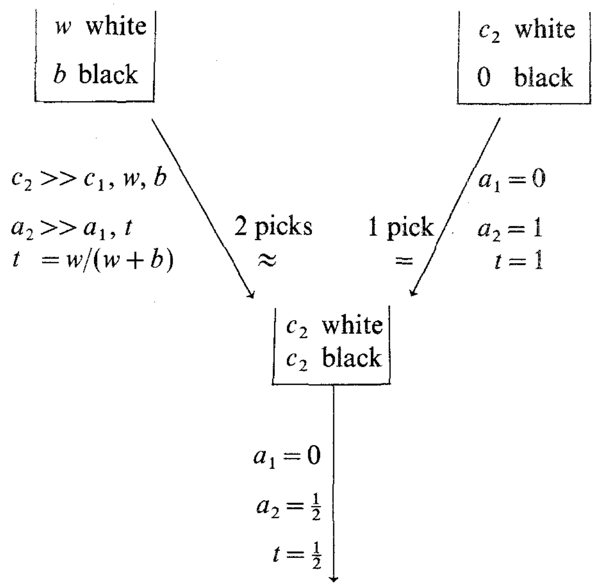
Q.E.D.

Now consider what happens to the probability distributions $D_N(a_1, a_2, t)$ as a_2 approaches infinity. Let $c_2 >> c_1$, w, b. Then after the first trial almost all the balls in the urn will be of the opposite color to the ball selected on the first trial. Therefore, with a probability approaching 1, the ball selected on the second trial will be of the opposite color to the ball selected on the first trial. Hence if $N>1$, it is not possible to select balls of only one color. Therefore when a_2 is large

$$D_K^N(a_1, a_2, t) \approx 0 \qquad K=0, N$$

and these approximations become more exact as a_2 (and hence c_2) approaches infinity.

To analyze the cases where $K \neq 0$, N, consider the contents of the urn after the first two trials. The fact that balls of the opposite color are necessarily selected on the first two trials leads to the following diagram.



Therefore, by inspection, when a_2 is large

$$\begin{aligned} D_K^N(a_1, a_2, t) &\approx D_{K-1}^{N-2}(0, 1/2, 1/2) \\ &= D_K^{N-1}(0, 1, 1) \end{aligned}$$

and this approximation becomes more exact as a_2 (and hence c_2) approaches infinity. These arguments suggest the following proposition.

PROPOSITION 2.7.5. *If $N>1$, then*

$$\begin{aligned} \lim_{a_2 \to \infty} D_K^N(a_1, a_2, t) &= D_K^{N-1}(0, 1, 1) \qquad && K \neq 0, N \\ &= 0 && K=0, N. \end{aligned}$$

Proof. By induction on N. It is easy to verify this result directly for

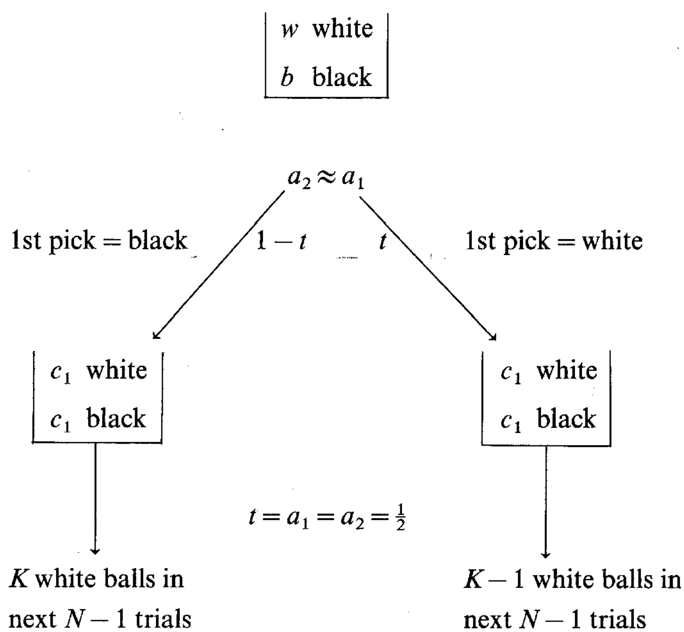
$N=2$. Now by the inductive hypothesis, Proposition 2.4.1, and Lemma 2.7.2

$$\begin{aligned}\lim_{a_2\to\infty} D_K^{N+1}(a_1, a_2, t) &= \lim_{a_2\to\infty} f_K^N(a_1, a_2, t)\, D_K^N(a_1, a_2, t)\\ &\quad + \lim_{a_2\to\infty} s_{K-1}^N(a_1, a_2, t)\, D_{K-1}^N(a_1, a_2, t)\\ &= f_K^{N-1}(0, 1, 1)\, D_K^{N-1}(0, 1, 1)\\ &\quad + s_{K-1}^{N-1}(0, 1, 1)\, D_{K-1}^{N-1}(0, 1, 1)\\ &= D_K^N(0, 1, 1).\end{aligned}$$

Q.E.D.

By Proposition 2.7.5 it follows that the functions $D_K^N(a_1, a_2, t)$ approach constant values independent of t as a_2 approaches infinity. Moreover, these constants are the values at $t=1$ of the distribution for which $a_1=0$, $a_2=1$. As we shall see in Section 5, these constants are particularly interesting; indeed they are actually the values at the knots of the uniform B-spline basis functions.

Finally let us consider what happens to the distributions $D_N(a_1, a_2, t)$ as both a_1, a_2 approach infinity. Suppose that $a_2=a_1+p$ for some fixed constant p. Let $c_1 >> w, b, p$. Then $c_2 \approx c_1$. Therefore after the first trial, the urn will contain approximately equal numbers of white balls and black balls (see diagram).



Therefore by inspection

$$D_K^N(a_1, a_2, t) \approx (1-t)\, D_K^{N-1}(1/2, 1/2, 1/2) + tD_{K-1}^{N-1}(1/2, 1/2, 1/2)$$

and this approximation becomes more exact as a_1 (and hence c_1) approaches infinity. This argument suggests the following proposition.

PROPOSITION 2.7.6. *Let p be a fixed constant, and let $a_2 = a_1 + p$. Then*

$$\lim_{a_1 \to \infty} D_K^N(a_1, a_2, t) = (1-t)\, D_K^{N-1}(1/2, 1/2, 1/2) + tD_{K-1}^{N-1}(1/2, 1/2, 1/2).$$

Proof. By induction on N. Certainly this result is true for $N=1$. Now by the inductive hypothesis, Proposition 2.4.1, and Lemma 2.7.3

$$\begin{aligned}
&\lim_{a_1 \to \infty} D_K^{N+1}(a_1, a_2, t)\\
&\quad = \lim_{a_1 \to \infty} f_K^N(a_1, a_2, t)\, D_K^N(a_1, a_2, t)\\
&\qquad + \lim_{a_1 \to \infty} s_{K-1}^N(a_1, a_2, t)\, D_{K-1}^N(a_1, a_2, t)\\
&\quad = 1/2[(1-t)\, D_K^{N-1}(1/2, 1/2, 1/2) + tD_{K-1}^{N-1}(1/2, 1/2, 1/2)\\
&\qquad + 1/2[(1-t)\, D_{K-1}^{N-1}(1/2, 1/2, 1/2) + tD_{K-2}^{N-1}(1/2, 1/2, 1/2)]\\
&\quad = (1-t)\, [f_K^{N-1}(1/2, 1/2, 1/2)\, D_K^{N-1}(1/2, 1/2, 1/2)\\
&\qquad + s_{K-1}^{N-1}(1/2, 1/2, 1/2)\, D_{K-1}^{N-1}(1/2, 1/2, 1/2)]\\
&\qquad + t[f_{K-1}^{N-1}(1/2, 1/2, 1/2)\, D_{K-1}^{N-1}(1/2, 1/2, 1/2)\\
&\qquad + s_{K-2}^{N-1}(1/2, 1/2, 1/2)\, D_{K-2}^{N-1}(1/2, 1/2, 1/2)]\\
&\quad = (1-t)\, D_K^N(1/2, 1/2, 1/2) + tD_{K-1}^N(1/2, 1/2, 1/2). \qquad \text{Q.E.D.}
\end{aligned}$$

Now we can compute $D_K^N(1/2, 1/2, 1/2)$ by the following intuitive argument. Consider an urn initially containing 1 white ball and 1 black ball. Then $t = 1/2$. Now if $a_1 = a_2 = 1/2$, then $c_1 = c_2 = 1$. Therefore after each trial 1 ball of each color will be added to the urn regardless of which color is selected. Thus the urn will always contain an equal number of white balls and black balls. Hence the probability of selecting a white (black) ball on any trial is always precisely 1/2. That is, this urn models the binomial distribution with $t = 1/2$. Therefore

$$\begin{aligned}
D_K^N(1/2, 1/2, 1/2) &= \binom{N}{K} (1/2)^K (1/2)^{N-K}\\
&= (1/2)^N \binom{N}{K}.
\end{aligned}$$

This argument suggests the following proposition.

PROPOSITION 2.7.7. $D_K^N(1/2, 1/2, 1/2) = (1/2)^N \binom{N}{K}$.

Proof. By induction on N. Certainly this result is true for $N=1$. Now by the inductive hypothesis, Proposition 2.4.1, and Lemma 2.7.3

$$\begin{aligned} D_K^{N+1}(1/2, 1/2, 1/2) &= f_K^N(1/2, 1/2, 1/2)\, D_K^N(1/2, 1/2, 1/2) \\ &\quad + s_{K-1}^N(1/2, 1/2, 1/2)\, D_{K-1}^N(1/2, 1/2, 1/2) \\ &= 1/2(1/2)^N \binom{N}{K} + 1/2(1/2)^N \binom{N}{K-1} \\ &= (1/2)^{N+1} \left[\binom{N}{K} + \binom{N}{K-1}\right] \\ &= (1/2)^{N+1} \binom{N+1}{K}. \end{aligned}$$

Q.E.D.

COROLLARY 2.7.8. *Let p be a fixed constant, and let $a_2 = a_1 + p$. Then*

$$\lim_{a_1 \to \infty} D_K^N(a_1, a_2, t) = (1/2)^{N-1} \left[\binom{N-1}{K}(1-t) + \binom{N-1}{K-1} t\right].$$

2.8 *Derivatives*

By Proposition 2.3.4 when $a_2 = 0$ we have explicit formulas for the functions $D_K^N(t)$. Therefore when $a_2 = 0$ it is no trouble at all to calculate the derivatives of these functions. This is not the case when $a_2 \neq 0$. In this section we shall develop formulas for the derivatives of the functions $D_K^N(t)$ when $a_1 = 0$, $a_2 \geqslant 0$.

Throughout this section we shall adopt the following notation

$$d_N = d_N(a_2) = 1 + Na_2.$$

Notice that by Proposition 2.3.1 if $a_1 = 0$, then d_N is simply the denominator of $f_K^N(t)$, $s_K^N(t)$. Therefore

$$\frac{df_K^N}{dt} = \frac{-1}{d_N}$$

$$\frac{ds_K^N}{dt} = \frac{1}{d_N}.$$

Notice too that

$$a_2 = 0 \Rightarrow d_N = 1$$

$$a_2 = 1 \Rightarrow d_N = N + 1.$$

LEMMA 2.8.1. *If* $a_1 = 0$, *then*

$$f_K^N(t) = \frac{d_{N-1}}{d_N} f_K^{N-1}(t)$$

$$s_K^N(t) = \frac{d_{N-1}}{d_N} s_{K-1}^{N-1}(t).$$

Proof. This result follows easily from Proposition 2.3.1.

PROPOSITION 2.8.2. *If* $a_1 = 0$, *then*

$$\frac{dD_K^N}{dt} = \frac{N}{d_{N-1}} [D_{K-1}^{N-1}(t) - D_K^{N-1}(t)].$$

Proof. By induction on N. This result is easy to verify directly for $N = 1$. Now by Proposition 2.1.2, Lemma 2.8.1, the recursion formula (Proposition 2.4.1), and the inductive hypothesis

$$\begin{aligned}
\frac{dD_K^{N+1}}{dt} &= \frac{df_K^N}{dt} D_K^N(t) + \frac{ds_{K-1}^N}{dt} D_{K-1}^N(t) \\
&\quad + f_K^N(t) \frac{dD_K^N}{dt} + s_{K-1}^N(t) \frac{dD_{K-1}^N}{dt} \\
&= \frac{D_{K-1}^N(t) - D_K^N(t)}{d_N} + f_K^N(t) \left[\frac{N}{d_{N-1}} (D_{K-1}^{N-1}(t) - D_K^{N-1}(t)) \right] \\
&\qquad + s_{K-1}^N(t) \left[\frac{N}{d_{N-1}} (D_{K-2}^{N-1}(t) - D_{K-1}^{N-1}(t)) \right] \\
&= \frac{D_{K-1}^N(t) - D_K^N(t)}{d_N} \\
&\quad + \frac{N}{d_{N-1}} [f_{K-1}^N(t) D_{K-1}^{N-1}(t) + s_{K-1}^N(t) D_{K-2}^{N-1}(t)] \\
&\quad - \frac{N}{d_{N-1}} [f_K^N(t) D_K^{N-1}(t) + s_K^N(t) D_{K-1}^{N-1}(t)] \\
&\quad + \frac{N}{d_{N-1}} [(f_K^N(t) + s_K^N(t)) - (f_{K-1}^N(t) + s_{K-1}^N(t))] D_{K-1}^{N-1}(t)
\end{aligned}$$

$$= \frac{D_{K-1}^N(t) - D_K^N(t)}{d_N} + \frac{N}{d_N}\,[f_{K-1}^{N-1}(t)\, D_{K-1}^{N-1}(t) + s_{K-2}^{N-1}(t)\, D_{K-2}^{N-1}(t)]$$

$$- \frac{N}{d_N}\,[f_K^{N-1}(t)\, D_K^{N-1}(t) + s_{K-1}^{N-1}(t)\, D_{K-1}^{N-1}(t)]$$

$$= \frac{D_{K-1}^N(t) - D_K^N(t)}{d_N} + \frac{N}{d_N}\,[D_{K-1}^N(t) - D_K^N(t)]$$

$$= \frac{(N+1)}{d_N}\,[D_{K-1}^N(t) - D_K^N(t)]. \qquad \text{Q.E.D.}$$

COROLLARY 2.8.3 (Bernstein Polynomials). *If* $a_1 = a_2 = 0$, *then*

$$\frac{dD_K^N}{dt} = N[D_{K-1}^{N-1}(t) - D_K^{N-1}(t)].$$

COROLLARY 2.8.4 (Uniform B-Splines). *If* $a_1 = 0$, $a_2 = 1$, *then*

$$\frac{dD_K^N}{dt} = D_{K-1}^{N-1}(t) - D_K^{N-1}(t).$$

When $a_1 = 0$, Proposition 2.8.2 gives us a simple formula for the derivative of the functions $D_K^N(t)$ in terms of the functions $D_K^{N-1}(t)$. This formula is not valid when $a_1 \neq 0$, $a_2 = 0$ as can be checked quite readily from Proposition 2.3.4. However, in this case the derivative can be computed explicitly. When $a_1 \neq 0$, $a_2 \neq 0$ we know of no simple method for computing the derivatives of $D_K^N(t)$.

Proposition 2.8.2 can be extended in the following manner.

PROPOSITION 2.8.5. *If* $a_1 = 0$, *then*

$$\frac{d^p D_K^N}{dt^p} = \frac{N(N-1)\cdots(N-p+1)}{d_{N-1}\, d_{N-1} \cdots d_{N-p}} \sum_j (-1)^{j+p} \binom{p}{j} D_{K-j}^{N-p}(t).$$

Proof. This result follows easily from Proposition 2.8.2 by induction on p.

COROLLARY 2.8.6 (Bernstein Polynomials). *If* $a_1 = a_2 = 0$, *then*

$$\frac{d^p D_K^N}{dt^p} = N(N-1)\cdots(N-p+1) \sum_j (-1)^{j+p} \binom{p}{j} D_{K-j}^{N-p}(t).$$

COROLLARY 2.8.7 (Uniform B-splines). *If* $a_1=0$, $a_2=1$, then

$$\frac{d^p D_K^N}{dt^p}=\sum_j (-1)\binom{p}{j} D_{K-j}^{N-p}(t).$$

The summation in Proposition 2.8.5 and its corollaries need not always be taken from $j=0$ to $j=p$. In fact since

$$D_K^N(t)=0 \qquad K<0 \text{ or } K>N,$$

the summation is really just from $j=\max(K+p-N, 0)$ to $j=\min(K, p)$.

3. APPROXIMATIONS

Let $g(t)$ be a continuous real-valued function defined on some interval I and let $D_N(t)$ be any one of the distributions described in Section 2. Define a linear functional $D_N\colon C[I]\to C[0, 1]$ by setting

$$D_N[g](t)=\sum_K g(e_K^N)\, D_K^N(t)$$

for some constants e_K^N, $K=0, 1, \ldots, N$ in the domain of $g(t)$. Since we are free to choose the constants e_K^N any way we please, we shall choose them so that linear functions are exactly reproduced. Recall from Corollary 2.5.7 that there exist constants p_N, q_N such that

$$M_1^N(t)=\sum KD_K^N(t)=p_N t+q_N$$

(i) $p_N>0$

(ii) $q_N\geqslant 0$

(iii) $p_N+2q_N=N.$

Assume that $I\supseteq[-q_N/p_N, (N-q_N)/p_N]$ and set

$$e_K^N=(K-q_N)/p_N.$$

Notice in particular that by Corollaries 2.5.8, 2.5.9

$$a_2=0\Rightarrow e_K^N=K/N,$$
$$a_2=1+a_1\Rightarrow e_K^N=(2K+1-N)/2.$$

Now it follows immediately from Proposition 2.1.1 and Corollary 2.5.7 that

$$D_N[1]=1$$
$$D_N[t]=t.$$

By Corollary 2.4.2 the function $D_N[g](t)$ is a polynomial in t of degree less than or equal to N. Moreover by Corollary 2.5.11 if $g(t)$ is a polynomial of degree r, $0 \leqslant r \leqslant N$, then $D_N[g](t)$ is also a polynomial of degree r. For reasons which will soon become clear, we shall regard the functions $D_N[g](t)$ as polynomial approximations to the function $g(t)$. Indeed when $D_N(t)$ is the binomial distribution, the polynomials $D_N[g](t)$ are the usual Bernstein approximations to the function $g(t)$ on the interval $[0, 1]$. The approximations induced by the Polya–Eggenberger urn model $(a_2 = 0)$ have also been studied by several authors [7, 13, 14]. We now proceed to investigate the common properties of these rather special polynomial approximations in more detail.

3.1 *Convexity*

We begin with some simple consequences of results derived in Section 2.

PROPOSITION 3.1.1. *For any constant c, $D_N[c] = c$.*

Proof. This result is an immediate consequence of the fact that $D_N(t)$ is a probability distribution (Proposition 2.1.1).

PROPOSITION 3.1.2. $D_N[t] = t$.

Proof. This result is an immediate consequence of Corollary 2.5.7.

COROLLARY 3.1.3. *D_N is the identity on linear functions.*

PROPOSITION 3.1.4. *If $a_2 = 0$, then*

$$D_N[g](0) = g(0)$$
$$D_N[g](1) = g(1).$$

Proof. This result is an immediate consequence of Proposition 2.3.4.

The convex hull of a set S is the smallest convex set which contains S. Thus the convex hull of a finite set of points $\{P_0, ..., P_N\}$ is the set

$$\left\{\sum_K c_K P_K \,|\, c_K \geqslant 0 \text{ and } \sum_K c_K = 1\right\}.$$

PROPOSITION 3.1.5. $\text{graph}(D_N[g]) \subseteq \text{convex hull (graph } g).$

Proof. By Proposition 3.1.2

$$\begin{aligned}
\operatorname{graph}(D_N[g]) &= \{(t, D_N[g](t) \mid 0 \leqslant t \leqslant 1\} \\
&= \{(D_N[t](t), D_N[g](t)) \mid 0 \leqslant t \leqslant 1\} \\
&= \left\{\sum_K (e_K^N, g(e_K^N))\, D_K^N(t) \mid 0 \leqslant t \leqslant 1\right\} \\
&\subseteq \text{convex hull}\{(e_K^N, g(e_K^N)) \mid 0 \leqslant K \leqslant N\} \\
&\subseteq \text{convex hull (graph } g). \qquad \text{Q.E.D.}
\end{aligned}$$

By Proposition 3.1.5 the values of $D_N[g](t)$ necessarily lie in the general proximity of the values of $g(t)$. It is for this reason that we regard the functions $D_N[g](t)$ as approximations to the function $g(t)$. Later on we shall show that if $a_2 = 0$ collectively the approximations $D_N[g](t)$ uniquely determine the function $g(t)$. We shall also show that the graphs of the approximations $D_N[g](t)$ actually mimic the shape of the graph of $g(t)$.

3.2 *Symmetry*

We begin with a simple observation.

LEMMA 3.2.1. $e_{N-K}^N = 1 - e_K^N$.

Proof. By Corollary 2.5.7 and the definition of e_K^N

$$\begin{aligned}
1 - e_K^N &= 1 - (K - q_N)/p_N \\
&= \frac{p_N + q_N - K}{p_N} \\
&= \frac{N - q_N - K}{p_N} \\
&= e_{N-K}^N. \qquad \text{Q.E.D.}
\end{aligned}$$

Now let $\bar{g}(t) = g(1-t)$. Because of the symmetry of our urn models (Section 2.2), we have the following general result.

PROPOSITION 3.2.2. $D_N[\bar{g}](t) = \overline{D_N[g]}(t)$.

Proof. By Proposition 2.2.1 and Lemma 3.2.1

$$\begin{aligned}
D_N[\bar{g}](t) &= \sum \bar{g}(e_K^N)\, D_K^N(t) \\
&= \sum g(1 - e_K^N)\, D_K^N(t)
\end{aligned}$$

$$=\sum g(e^N_{N-K})\,D^N_{N-K}(1-t)$$

$$=\sum g(e^N_K)\,D^N_K(1-t)$$

$$=\overline{D[g](t)}. \qquad \text{Q.E.D.}$$

3.3 *Recursion*

The recursion formula

$$D^{N+1}_K(t)=f^N_K(t)\,D^N_K(t)+s^N_{K-1}(t)\,D^N_{K-1}(t)$$

for the distributions $D_N(t)$ (Proposition 2.4.1) engenders a recursive algorithm for the approximations $D_N[g](t)$. Define a triangular array $P^L_K[g](r)$, $0\leqslant K+L\leqslant N$, recursively by setting

$$P^0_K[g](r)=g(e^N_K)$$
$$P^L_K[g](r)=f^{N-L}_K(r)\,P^{L-1}_K[g](r)+s^{N-L}_K(r)\,P^{L-1}_{K+1}[g](r).$$

We shall show shortly that

$$P^N_0[g](r)=D_N[g](r).$$

This recursive construction algorithm is especially useful because it provides a simple, numerically stable technique for computing the value of $D_N[g](r)$ for any parameter r without the need to compute explicitly the values of $D^N_K(r)$, $0\leqslant K\leqslant N$.

To prove that $P^N_0[g](r)=D_N[g](r)$, we use a simple inductive argument.

LEMMA 3.3.1. $\sum P^1_K[g](r)\,D^{N-1}_K(r)=D_N[g](r)$.

Proof. By Proposition 2.4.1

$$\sum P^1_K[g](r)\,D^{N-1}_K(r)=\sum[f^{N-1}_K(r)P^0_K[g](r)+s^{N-1}_K P^0_{K+1}[g](r)]\,D^{N-1}_K(r)$$

$$=\sum[f^{N-1}_K(r)\,D^{N-1}_K(r)+s^{N-1}_{K-1}(r)\,D^{N-1}_{K-1}(r)]\;P^0_K[g](r)$$

$$=\sum g(e^N_K)\,D^N_K(r)$$

$$=D_N[g](r). \qquad \text{Q.E.D.}$$

LEMMA 3.3.2. $\sum P^L_K[g](r)\,D^{N-L}_K(r)=D_N[g](r)$.

Proof. This result follows from Lemma 3.3.1 by induction on L.

PROPOSITION 3.3.3. $P_0^N[g](r) = D_N[g](r)$.

Proof. This result follows immediately from Lemma 3.3.2 with $L = N$.

For the Polya–Eggenberger urn model ($a_2 = 0$), there exists a second recursion formula for the distributions $D_N(t)$ and this formula begets a second recursive algorithm for the approximations $D_N[g](t)$. Let $Q_K^L[g](r)$, $0 \leqslant K + L \leqslant N$, be the triangular array defined recursively by setting

$$Q_K^0[g](r) = g(e_K^N) = g(K/N)$$
$$Q_K^L[g](r) = f_K^{L-1}(r)\, Q_K^{L-1}[g](r) + s_K^{L-1}(r)\, Q_{K+1}^{L-1}[g](r).$$

We shall show that if $a_2 = 0$, then

$$Q_0^N[g](r) = D_N[g](r).$$

To proceed, we shall need to extend the definitions of the functions $f_K^N(t)$, $s_K^N(t)$, $D_K^N(t)$ to values of $t > 1$ and $K > N$. We do so simply by adopting the formulas of Propositions 2.3.1, 2.3.4 for arbitrary values of t, K, N.

PROPOSITION 3.3.4 (Recursions). *If* $a_2 = 0$, *then*

$$D_K^{N+1}(t) = f_0^N(t)\, D_K^N(t) + s_0^N(t)\, D_{K-1}^N(t + a_1).$$

Proof. By Proposition 2.3.1

$$f_0^N(t) = \frac{1 - t + Na_1}{1 + Na_1}$$

$$s_0^N(t) = \frac{t}{1 + Na_1}$$

and by Proposition 2.3.4

$$D_K^N(t) = \binom{N}{K} \frac{t \cdots (t + [K-1]\, a_1)(1 - t) \cdots (1 - t + [N - K - 1]\, a_1)}{(1 + a_1) \cdots (1 + [N-1]\, a_1)}$$

$$D_{K-1}^N(t + a_1) = \frac{K(1 - t - a_1)}{t(N + 1 - K)}\, D_K^N(t).$$

Combining all these terms, we obtain

$$
\begin{aligned}
&f_0^N(t)\, D_K^N(t) + s_0^N(t)\, D_{K-1}^N(t+a_1) \\
&\quad = \left[(1 - t + Na_1) + \frac{K(1-t-a_1)}{(N+1-K)} \right] \frac{D_K^N(t)}{(1+Na_1)} \\
&\quad = \left[\frac{(N+1)(1-t) + (N+1)(N-K)\, a_1}{(N+1-K)(1+Na_1)} \right] D_K^N(t) \\
&\quad = \frac{(N+1)(1-t+[N-K]\, a_1)}{(N+1-K)(1+Na_1)}\, D_K^N(t) \\
&\quad = D_K^{N+1}(t).
\end{aligned}
$$

Q.E.D.

LEMMA 3.3.5. *If* $a_2 = 0$, *then*

$$
\begin{aligned}
s_K^N(t) &= s_{K-1}^N(t+a_1) \\
f_K^N(t) &= f_{K-1}^N(t+a_1).
\end{aligned}
$$

Proof. These results follow immediately from Proposition 2.3.1.

PROPOSITION 3.3.6. *If* $a_2 = 0$, *then*

$$Q_0^N[g](r) = D_N[g](r).$$

Proof. By induction on N. Clearly this result is true for $N=1$. Now define two piecewise linear polynomials $f(t)$, $h(t)$ by setting

$$
\begin{aligned}
f(K/N) &= g[K/(N+1)] \qquad && K = 0, 1, ..., N \\
h(K/N) &= g[(K+1)/(N+1)] \qquad && K = 0, 1, ..., N.
\end{aligned}
$$

Then by construction and Lemma 3.3.5

$$
\begin{aligned}
Q_K^L[g](r) &= Q_K^L[f](r) \qquad && 0 \leqslant K+L \leqslant N \\
Q_K^L[g](r) &= Q_{K-1}^L[h](r+a_1) \qquad && 1 \leqslant K+L \leqslant N+1.
\end{aligned}
$$

Therefore by the inductive hypothesis

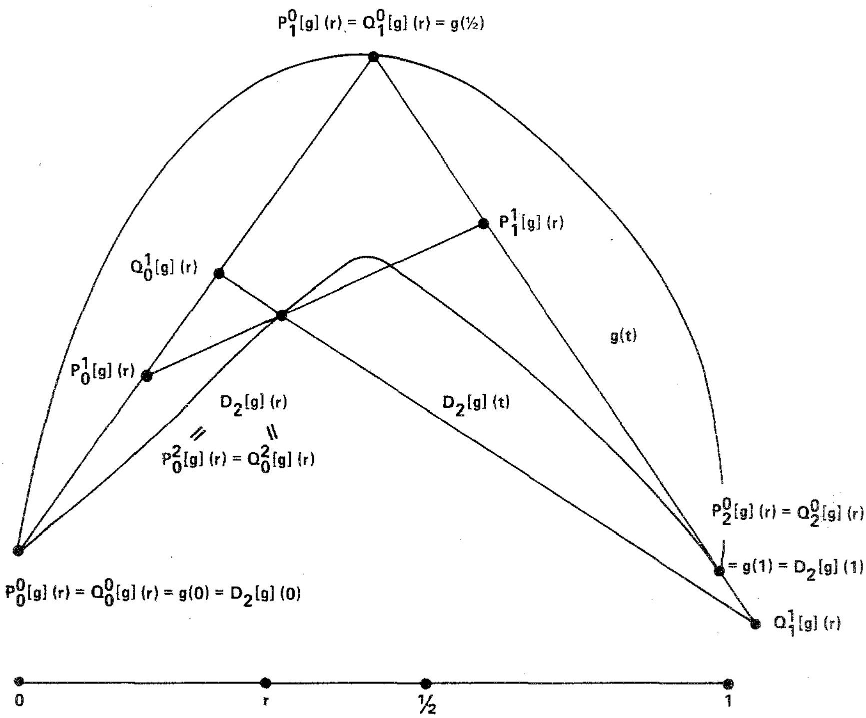
$$
\begin{aligned}
Q_0^N[g](r) &= Q_0^N[f](r) = D_N[f](r) \\
Q_1^N[g](r) &= Q_0^N[h](r+a_1) = D_N[h](r+a_1).
\end{aligned}
$$

Hence by Proposition 3.3.4

$$\begin{aligned} Q_0^{N+1}[g](r) &= f_0^N(r)\, Q_0^N[g](r) + s_0^N(r)\, Q_1^N[g](r) \\ &= f_0^N(r)\, D_N[f](r) + s_0^N(r)\, D_N[h](r+a_1) \\ &= \sum_K \left[f_0^N(r)\, D_K^N(r) + s_0^N(r)\, D_{K-1}^N(r+a_1)\right] g[K/(N+1)] \\ &= \sum_K g[K/(N+1)]\, D_K^{N+1}(r) \\ &= D_{N+1}[g](r). \end{aligned}$$

Q.E.D.

The second recursive algorithm for the approximations $D_N[g](t)$ goes beyond the bounds of probability theory. The construction employs functions $f_K^L(r)$, $s_K^L(r)$ for which $K > L$, and the proof resorts to values of $D_K^N(t+a_1)$ for which $t + a_1 > 1$. In neither case is there a clear probabilistic interpretation for these functions, and yet if we extend the formulas in Propositions 2.3.1, 2.3.4 beyond the realm in which they were originally derived, it all works. This is a somewhat bizarre and unexpected result.



Recursive Construction Algorithms for $D_N[g](t)$ $(N = 2, a_2 = 0)$

FIGURE 1

By Proposition 3.1.2 $D_N[t]=t$. Therefore for the Polya–Eggenberger distributions, we can illustrate the two recursive construction algorithms for $D_N[g](t)$ geometrically with the simple diagram in Figure 1.

For the binomial distribution

$$a_1 = a_2 = 0$$
$$f_K^{N-L}(r) = f_K^{L-1}(r) = 1 - r$$
$$s_K^{N-L}(r) = s_K^{L-1}(r) = r.$$

Therefore for the binomial distribution, the two recursive algorithms depicted above are identical. That is, for the binomial distribution

$$Q_K^L[g](r) = P_K^L[g](r) \qquad \text{for all } K, L.$$

3.4 *Uniqueness*

Corollary 2.5.13 states that the polynomials $D_0^N(t), ..., D_N^N(t)$ are linearly independent. This result has the following consequences.

PROPOSITION 3.4.1. $D_N[g]=D_N[h]$ *iff* $g(e_K^N)=h(e_K^N)$, $0 \leqslant K \leqslant N$.

PROPOSITION 3.4.2. *If* $a_2=0$, $D_N[g]=D_N[h]$ *for all* N *iff* $g(t)=h(t)$ *for* $0 \leqslant t \leqslant 1$.

Proof. Certainly if $g(t)=h(t)$ for all t, then $D_N[g]=D_N[h]$ for all N. Conversely if $a_2=0$, then $e_K^N=K/N$. Therefore by Proposition 3.4.1 if $D_N[g]=D_N[h]$ for all N, then $g(r)=h(r)$ for all rational fractions r. Hence it follows by continuity that $g(t)=h(t)$ for all values of $0 \leqslant t \leqslant 1$.

Q.E.D.

We can sharpen the preceding result somewhat as follows.

PROPOSITION 3.4.3. *If* $a_2=0$, *then for any integer* M $D_N[g]=D_N[h]$ *for all* $N \geqslant M$ *iff* $g(t)=h(t)$ *for* $0 \leqslant t \leqslant 1$.

Proof. Same as Proposition 3.4.2.

Thus if $a_2=0$ we can conclude that two continuous functions on $[0, 1]$ are identical iff their approximations are identical for all sufficiently large values of N. Hence collectively the approximations $D_N[g](t)$ uniquely characterize the function $g(t)$.

3.5. *The Variation Diminishing Property*

In Section 2.6 we introduced two conjectures concerning the Laws of Signs. In this section we shall derive some geometric consequences of these conjectures.

A linear functional $F: C[I] \to C[0, 1]$ is said to have the variation diminishing property iff for every function g

$$v(F[g]) \leqslant v(g).$$

Here v is the symbol defined in Section 2.6. Thus intuitively F is said to be variation diminishing iff for every function g the number of times the graph of $F[g]$ crosses the t-axis is less than or equal to the number of times the graph of g crosses the t-axis.

Now let $F(t) = (F_0(t), ..., F_N(t))$ be an ordered collection of continuous real-valued functions defined on the interval $[0, 1]$, and let $F[g]$ be the linear functional defined by setting

$$F[g](t) = \sum g(e_K^N)\, F_K(t).$$

Recalling the Laws of Signs from Section 2.6, we have the following general results.

PROPOSITION 3.5.1. *The linear functional $F[g]$ is variation diminishing iff the functions $F(t)$ satisfy the Weak Law of Signs in the interval* $(0, 1)$.

Proof. If the functions $F(t)$ satisfy the Weak Law of Signs in the interval $(0, 1)$, then by definition

$$\begin{aligned} v(F[g]) &= v\left[\sum g(e_K^N)\, F_K(t)\right] \\ &\leqslant v[g(e_K^N)] \\ &\leqslant v(g). \end{aligned}$$

Therefore $F[g]$ is variation diminishing. Conversely suppose that $F[g]$ is variation diminishing. Let $c_0, ..., c_N$ be a sequence of constants and let $g(t)$ be the piecewise linear function on I defined by setting

$$g(e_K^N) = c_K.$$

Then since $F[g]$ is variation diminishing

$$\begin{aligned} v\left[\sum c_K F_K(t)\right] &= v(F[g]) \\ &\leqslant v(g) \\ &= v[g(e_K^N)] \\ &= v(c_0, ..., c_N). \end{aligned}$$

Therefore the functions $F(t)$ satisfy the Weak Law of Signs in the interval $(0, 1)$. Q.E.D.

COROLLARY 3.5.2. *If the functions $F(t)$ satisfy the Strong Law of Signs in the interval* $(0, 1)$, *then the linear functional $F[g]$ is variation diminishing.*

Now Proposition 3.5.1 together with Conjecture 2.6.1 suggest the following general conjecture.

Conjecture 3.5.3. For all positive finite values of a_1, a_2, the linear functionals $D_N[g]$ are variation diminishing.

By Proposition 3.5.1 it follows immediately that

$$\text{Conjecture 2.6.1} \Leftrightarrow \text{Conjecture 3.5.3.}$$

Now Conjecture 2.6.1 (the Weak Law of Signs) is known to be valid when $a_2 = 0$ (the Polya–Eggenberger urn model) and when $a_1 = 0$ (non-uniform B-splines). Thus Conjecture 3.5.3 (the variation diminishing property) must also be valid at least in these two special cases. We believe that the variation diminishing property is valid for all the distributions which arise from Friedman's urn model, but, like the Laws of Signs, we know of no proof, probabilistic or otherwise, for this general conjecture.

Conjecture 3.5.3 can be strengthened in the following manner.

Conjecture 3.5.4. For all positive finite values of a_1, a_2 and for all linear functions L

$$v(D_N[g] - L) \leqslant v(g - L).$$

Conjecture 3.5.4 is an immediate consequence of Conjecture 3.5.3 and Proposition 3.1.3 since by linearity

$$v(D_N[g] - L) = v(D_N[g - L]) \leqslant v(g - L).$$

Conjecture 3.5.4 says that for any straight line L the number of times the graph of $D_N[g]$ crosses L is less than or equal to the number of times the graph of g crosses L. Thus for any straight line L, $D_N[g]$ oscillates about L less than g oscillates about L. Hence globally the graph of $D_N[g]$ mimics the general shape of the graph of g.

3.6 *Limits*

In this section we shall investigate the behavior of the approximations $D_N[g]$ as a_1 or a_2 or N approaches infinity. We begin with some simple consequences of results derived in Section 2.7.

PROPOSITION 3.6.1. $\mathrm{Lim}_{a_1 \to \infty} D_N[g](t) = (1-t)\, g(e_0^N) + t g(e_N^N)$.

Proof. This result is an immediate consequence of Proposition 2.7.4.

PROPOSITION 3.6.2. $\mathrm{Lim}_{a_2 \to \infty} D_N[g](t) = \sum_{K=1}^{N-1} g(e_K^N)\, D_K^{N-1}(0, 1, 1)$.

Proof. This result is an immediate consequence of Proposition 2.7.5.

By Proposition 3.6.1 as a_1 approaches infinity the function $D_N[g](t)$ approaches the chord joining the end points of $g(t)$, and by Proposition 3.6.2 as a_2 approaches infinity the function $D_N[g](t)$ reduces to a constant. Thus when either a_1 or a_2 is very large, the functions $D_N[g](t)$ are not very good approximations to $g(t)$. But suppose we hold a_1, a_2 fixed and increase the value of N; what then can we say about the approximations $D_N[g](t)$? For the binomial distribution $(a_1 = a_2 = 0)$ we have the following well known result.

PROPOSITION 3.6.3 (The Weierstrass Approximation Theorem). *If $a_1 = a_2 = 0$, then as N approaches infinity the approximations $D_N[g](t)$ converge uniformly to the original function $g(t)$.*

Proof. See [3].

Thus increasing N makes the approximation better, but increasing a_1 or a_2 generally makes it worse. Now one might hope that for a_1, a_2 fixed the Weierstrass Approximation Theorem would remain valid; that is, that eventually N would dominate over a_1, a_2. However, we shall now show that this is not the case even for the Polya–Eggenberger distributions $(a_2 = 0)$.

LEMMA 3.6.4. *If $a_2 = 0$, then*

$$M_2^N(t) = \frac{N(N-1)\, t^2 + N(1 + Na_1) t}{1 + a_1}.$$

Proof. By induction on N. Certainly this result is true for $N = 1$. Moreover by the recursion formula for moments (Proposition 2.5.10), the formula for expectation when $a_2 = 0$ (Corollary 2.5.8), and the inductive hypothesis

$$M_2^{N+1}(t) = \frac{(1 + (N+2)\, a_1)}{(1 + Na_1)}\, M_2^N(t) + \frac{(2t + a_1)}{(1 + Na_1)}\, M_1^N(t) + \frac{t}{(1 + Na_1)}$$

$$=\frac{[1+(N+2)a_1][N(N-1)t^2+N(1+Na_1)t]}{(1+Na_1)(1+a_1)}$$

$$+\frac{2N(1+a_1)t^2+(1+Na_1)(1+a_1)t}{(1+Na_1)(1+a_1)}$$

$$=\frac{[(N^2+N-2)a_1+(N-1)+(2+2a_1)]Nt^2}{(1+Na_1)(1+a_1)}$$

$$+\frac{[(N^2+2N)a_1+N+(1+a_1)][1+Na_1]t}{(1+Na_1)(1+a_1)}$$

$$=\frac{(1+Na_1)(N+1)Nt^2+(N+1)([N+1]a_1+1)(1+Na_1)t}{(1+Na_1)(1+a_1)}$$

$$=\frac{(N+1)Nt^2+(N+1)([N+1]a_1+1)t}{1+a_1}. \qquad \text{Q.E.D.}$$

PROPOSITION 3.6.5. *If* $a_2=0$, *then*

$$D_N[t^2]=\frac{(N-1)t^2+(1+Na_1)t}{N(1+a_1)}.$$

Proof. This result is an immediate consequence of Lemma 3.6.4. It is also proved by an alternate method in [13].

PROPOSITION 3.6.6. *If* $a_2=0$, *then*

$$\lim_{N\to\infty} D_N[t^2]=\frac{t^2+a_1t}{1+a_1}.$$

By Proposition 3.6.6 the Weierstrass Approximation Theorem fails to hold for the Polya–Eggenberger distributions even for as simple a function as $g(t)=t^2$. Nevertheless by Proposition 3.4.3 the approximations $D_N[g](t)$, $N\geqslant M$, still uniquely determine the function $g(t)$ for $0\leqslant t\leqslant 1$.

3.7 *Derivatives*

For approximations induced by urn models with $a_2=0$ (Polya–Eggenberger urn models) we have explicit formulas for the functions $D_K^N(t)$ (Proposition 2.3.4). Therefore when $a_2=0$, we can compute the derivative of the approximation $D_N[g](t)$ explicitly. On the other hand when $a_1=0$, we can apply the results of Section 2.8 to compute the derivatives of $D_N[g](t)$. Let

$$d_N=d_N(a_2)=1+Na_2$$

as in Section 2.8. Then we have the following results.

PROPOSITION 3.7.1. *If* $a_1 = 0$, *then*

$$D'_N[g](t) = \frac{N}{d_{N-1}} \sum [g(e^N_{K+1}) - g(e^N_K)] D^{N-1}_K(t).$$

Proof. This result is an immediate consequence of Proposition 2.8.2.

COROLLARY 3.7.2 (Bernstein Approximations). *If* $a_1 = a_2 = 0$, *then*

$$D'_N[g](t) = N \sum \left[g\left(\frac{K+1}{N}\right) - g\left(\frac{K}{N}\right)\right] D^{N-1}_K(t).$$

COROLLARY 3.7.3 (Uniform *B*-Spline Approximations). *If* $a_1 = 0$, $a_2 = 1$, *then*

$$D'_N[g](t) = \sum [g(e^N_{K+1}) - g(e^N_K)] D^{N-1}_K(t).$$

PROPOSITION 3.7.4. *If* $a_1 = 0$, *then*

$$D^{(p)}_N[g](t) = \frac{N(N-1)\cdots(N-p+1)}{d_{N-1}\, d_{N-2} \cdots d_{N-p}} \times \sum_K \left[\sum_j (-1)^{j+p} \binom{p}{j} g(e^N_{K+j})\right] D^{N-p}_K(t).$$

Proof. This result is an immediate consequence of Proposition 2.8.5.

COROLLARY 3.7.5 (Bernstein Approximations). *If* $a_1 = a_2 = 0$, *then*

$$D^{(p)}_N[g](t) = N(N-1)\cdots(N-p+1) \times \sum_K \left[\sum_j (-1)^{j+p} \binom{p}{j} g\left(\frac{K+j}{N}\right)\right] D^{N-p}_K(t).$$

COROLLARY 3.7.6 (Uniform *B*-Spline Approximations). *If* $a_1 = 0$, $a_2 = 1$, *then*

$$D^{(p)}_N[g](t) = \sum_K \left[\sum_j (-1)^{j+p} \binom{p}{j} g(e^N_{K+j})\right] D^{N-p}_K(t).$$

As in Proposition 2.8.5 and its corollaries the summation in Proposition 3.7.4 and its corollaries need not always be taken from $j = 0$ to $j = p$. In fact the summation is really just from $j = \max(K + p - N, 0)$ to $j = \min(K, p)$.

4. Splines

In this section we shall restrict our attention to urn models for which $a_2 = 1 + a_1$. By Corollary 2.5.9 these urn models have the remarkable property that the a priori probability of selecting a white ball on any trial after the first is always exactly 1/2 regardless of the initial contents of the urn. We call such urn models spline models for reasons which will become clear shortly.

Let $x_0 < x_1 < \cdots < x_{M+a_1}$ be an increasing sequence of real numbers. A function $S(t)$ is said to be continuous polynomial spline of degree N, order $N+1$, with knots $(x_0, ..., x_{M+1})$ iff there are $M+1$ degree N polynomials $p_0(t), ..., p_M(t)$ such that

$$S(t) = p_K(t) \qquad x_K \leqslant t \leqslant x_{K+1}$$

$$p_K(x_{K+1}) = p_{K+1}(x_{K+1}) \qquad K = 0, 1, ..., M-1.$$

Conversely given $M+1$ degree N polynomials $p_0(t), ..., p_M(t)$ such that

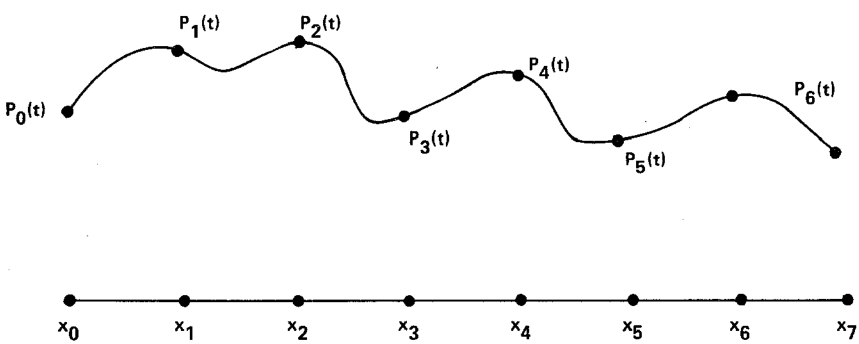
$$p_K(x_{K+1}) = p_{K+1}(x_{K+1}) \qquad K = 0, 1, ..., M-1$$

we can construct a continuous polynomial spline $S(t)$ by setting

$$S(t) = p_K(t) \qquad x_K \leqslant t \leqslant x_{K+1}$$

(see Figure 2).

We now proceed to make the connection between urn models and splines.



A continuous polynomial spline (M=6)

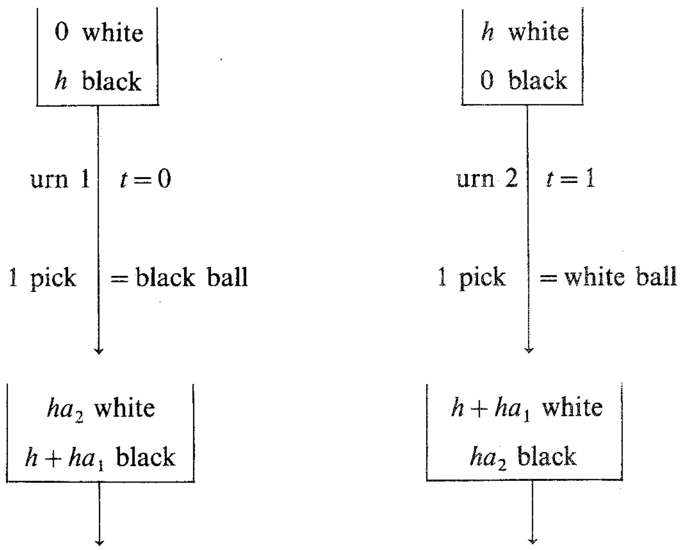
Figure 2

4.1 *Spline Distributions*

We begin with the following fundamental result.

PROPOSITION 4.1.1. *If $a_2 = 1 + a_1$, then $D_K^N(1) = D_{K-1}^N(0)$.*

Proof. Consider 2 urns—urn 1 representing the case $t=0$ and urn 2 the case $t=1$. Initially urn 1 contains no white balls and h black balls and urn 2 contains h white balls and no black balls. After 1 pick urn 1 will contain ha_2 white balls and $h + ha_1$ black balls. Similarly urn 2 will contain $h + ha_1$ white balls and ha_2 black balls (see diagram).



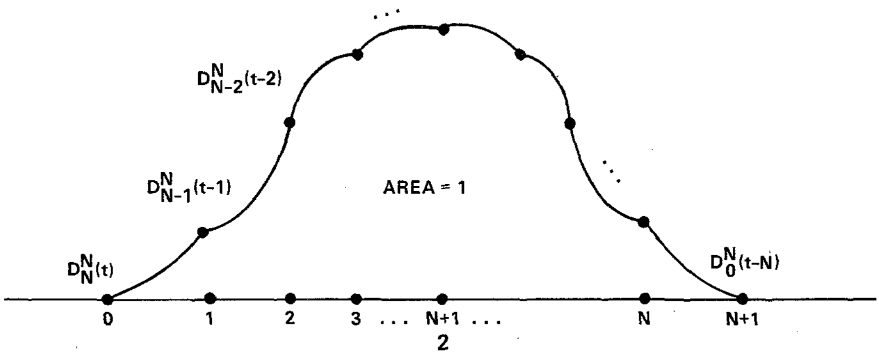
Hence if $a_2 = 1 + a_1$, then after 1 pick the contents of the 2 urns are identical. Therefore

$$a_2 = 1 + a_1 \Rightarrow D_K^N(1) = D_{K-1}^N(0). \qquad \text{Q.E.D.}$$

Define

$$\begin{aligned} S_{0N}(t) &= D_{N-K}^N(t-K) \qquad & 0 \leqslant K \leqslant t \leqslant K+1 \leqslant N+1 \\ &= 0 & t < 0 \text{ or } t > N+1. \end{aligned}$$

By Proposition 4.1.1 it follows that $S_{0N}(t)$ is a continuous polynomial spline of degree N with knots $(0, 1, \ldots, N+1)$ (see Figure 3).



The Polynomial Spline $S_{0N}(t)$

FIGURE 3

Thus the urn models for which $a_2 = 1 + a_1$ naturally generate polynomial splines. For this reason we call these special urn models spline models.

The splines $S_{0N}(t)$ are generally not differentiable at the knots. Indeed in the quadratic case it follows from the recursion formula (Proposition 2.3.1) that

$$D_0^2(t) = \frac{(1-t+a_1)(1-t)}{2(1+a_1)} = \frac{t^2-(2+a_1)t+(1+a_1)}{2(1+a_1)}$$

$$D_1^2(t) = \frac{(t+1+a_1)(1-t)+(2-t+a_1)t}{2(1+a_1)} = \frac{-2t^2+2t+(1+a_1)}{2(1+a_1)}$$

$$D_2^2(t) = \frac{(t+a_1)t}{2(1+a_1)} = \frac{t^2+a_1 t}{2(1+a_1)}.$$

Therefore if $a_1 \neq 0$, then by direct computation

$$0 \neq \frac{a_1}{2(1+a_1)} = \frac{dD_2^2}{dt}\Big|_{t=0}$$

$$\frac{dD_2^2}{dt}\Big|_{t=1} = \frac{2+a_1}{2(1+a_1)} \neq \frac{1}{1+a_1} = \frac{dD_1^2}{dt}\Big|_{t=0}$$

$$\frac{dD_1^2}{dt}\Big|_{t=1} = \frac{-1}{1+a_1} \neq \frac{-(2+a_1)}{2(1+a_1)} = \frac{dD_0^2}{dt}\Big|_{t=0}$$

$$\frac{\partial D_0^2}{dt}\Big|_{t=1} = \frac{-a_1}{2(1+a_1)} \neq 0.$$

Hence by construction

$$S'_{02}(K_-) \neq S'_{02}(K_+) \qquad K = 0, 1, 2, 3.$$

Notice however that we do obtain equality when $a_1 = 0$. We shall return to this very important special case (B-splines) again in Section 5.

Because the splines $S_{0N}(t)$ are generated from urn models, they inherit the following additional properties.

PROPOSITION 4.1.2 (Symmetry). $S_{0N}(t) = S_{0N}(N+1-t)$.

Proof. Let $K \leqslant t \leqslant K+1$. Then $N-K \leqslant N+1-t \leqslant N+1-K$. Therefore by Proposition 2.2.1

$$\begin{aligned} S_{0N}(t) &= D^N_{N-K}(t-K) \\ &= D^N_K(K+1-t) \\ &= S_{0N}(N+1-t). \end{aligned} \qquad \text{Q.E.D.}$$

COROLLARY 4.1.3 (Symmetry).

$$S_{0N}\left(\frac{N+1}{2} - t\right) = S_{0N}\left(\frac{N+1}{2} + t\right).$$

Proof. By Proposition 4.1.2

$$\begin{aligned} S_{0N}\left(\frac{N+1}{2} - t\right) &= S_{0N}\left(N+1 - \frac{N+1}{2} - t\right) \\ &= S_{0N}\left(\frac{N+1}{2} + t\right). \end{aligned} \qquad \text{Q.E.D.}$$

PROPOSITION 4.1.4 (End Points).

$$\begin{aligned} S_{0N}(0) &= 0 \qquad N > 0 \\ S_{0N}(N+1) &= 0 \qquad N > 0. \end{aligned}$$

Proof. By Proposition 2.3.2 if $N > 0$

$$\begin{aligned} S_{0N}(0) &= D^N_N(0) = 0 \\ S_{0N}(N+1) &= D^N_0(1) = 0. \end{aligned} \qquad \text{Q.E.D.}$$

PROPOSITION 4.1.5 (Explicit Formulas).

$$S_{0N}(t)=\prod_{K=0}^{N-1}\frac{(t+Ka_1)}{(1+K+2Ka_1)}\qquad 0\leqslant t\leqslant 1$$

$$S_{0N}(t)=\prod_{K=0}^{N-1}\frac{(N+1-t+Ka_1)}{(1+K+2Ka_1)}\qquad N\leqslant t\leqslant N+1.$$

Proof. These results follow immediately from Proposition 2.3.3.

PROPOSITION 4.1.6 (Continuous Distribution).

$$S_{0N}(t)\geqslant 0$$

$$\int_0^{N+1} S_{0N}(t)\,dt=1.$$

Proof. Let $K\leqslant t\leqslant K+1$. Then by Proposition 2.1.1

$$S_{0N}(t)=D^N_{N-K}(t-K)\geqslant 0.$$

Moreover, again by Proposition 2.1.1,

$$\begin{aligned}\int_0^{N+1} S_{0N}(t)\,dt&=\sum_K\int_K^{K+1} D^N_{N-K}(t-K)\,dt\\&=\sum_K\int_0^1 D^N_{N-K}(u)\,du\\&=\int_0^1\left[\sum_K D^N_{N-K}(u)\right]du\\&=\int_0^1 du\\&=1.\end{aligned}$$

Q.E.D.

PROPOSITION 4.1.7 (Expectation).

$$\int_0^{N+1} tS_{0N}(t)\,dt=\frac{N+1}{2}.$$

Proof. By symmetry. From Corollary 4.1.3 the distribution $S_{0N}(t)$ is symmetric about the point $(N+1)/2$. Now by a standard argument the expectation of a symmetric distributions is the point of symmetry [10]. Hence

$$\int_0^{N+1} tS_{0N}(t)\,dt = \frac{N+1}{2}.$$

By integration.

$$\begin{aligned}\int_0^{N+1} tS_{0N}(t)\,dt &= \sum_K \int_K^{K+1} tD_{N-K}^N(t-K)\,dt \\ &= \int_0^1 \sum_K (u+K)\,D_{N-K}^N(u)\,du \\ &= \int_0^1 \sum_K (u+N-K)\,D_K^N(u)\,du \\ &= \int_0^1 (u+N)\sum_K D_K^N(u)\,du - \int_0^1 \sum_K KD_K^N(u)\,du.\end{aligned}$$

Now by Proposition 2.1.1 and Corollary 2.5.9

$$\sum_K D_K^N(u) = 1$$

$$\sum_K KD_K^N(u) = u + \frac{N-1}{2}.$$

Therefore

$$\begin{aligned}\int_0^{N+1} tS_{0N}(t)\,dt &= \int_0^1 u\,du + \int_0^1 N\,du - \int_0^1 u\,du - \int_0^1 \frac{(N-1)}{2}\,du \\ &= N - \frac{(N-1)}{2} \\ &= \frac{N+1}{2}.\end{aligned}$$

Q.E.D.

Thus many of the characteristic properties of the splines $S_{0N}(t)$ are simple consequences of the corresponding properties of the distributions $D_N(t)$. In particular, the splines $S_{0N}(t)$ are continuous distributions. We call such distributions spline distributions.

When $a_1 = 0$ we shall show shortly that the spline distribution $S_{0N}(t)$ is actually the normalized uniform B-spline basis function of degree N with knots $(0, 1, \ldots, N+1)$. At the other extreme, as a_1 approaches infinity, we have the following result.

PROPOSITION 4.1.8 (Limits). *If* $K \leqslant t \leqslant K+1$, *then*

$$\lim_{a_1 \to \infty} S_{0N}(t) = (1/2)^{N-1}\left[\binom{N-1}{K-1}(K+1-t) + \binom{N-1}{K}(t-K)\right].$$

Proof. Let $K \leqslant t \leqslant K+1$. Then by Corollary 2.7.8

$$\lim_{a_1 \to \infty} S_{0N}(t) = \lim_{a_1 \to \infty} D^N_{N-K}(t-K)$$

$$= (1/2)^{N-1}\left[\binom{N-1}{N-K}[1-(t-K)] + \binom{N-1}{N-K-1}(t-K)\right]$$

$$= (1/2)^{N-1}\left[\binom{N-1}{K-1}(K+1-t) + \binom{N-1}{K}(t-K)\right]. \quad \text{Q.E.D.}$$

By Proposition 4.1.8. the splines $S_{0N}(t)$ approach linear splines as a_1 approaches infinity. Thus, for example, in the limit we have the following diagram.

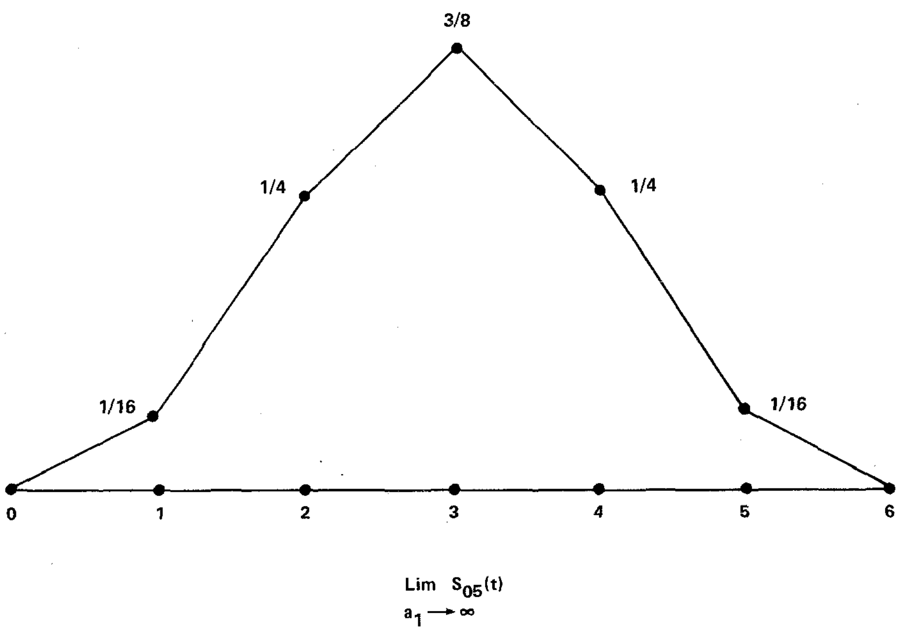


FIGURE 4

4.2 *Recursion*

By translating the functions $S_{0N}(t)$, we can define spline distributions over any sequence of consecutive integers $(J, J+1, ..., J+N+1)$. Let

$$S_{JN}(t) = S_{0N}(t-J)$$

or equivalently let

$$\begin{aligned} S_{JN}(t) &= D^N_{N-K}(t-J-K) \qquad && J \leqslant J+K \leqslant t \leqslant J+K+1 \leqslant J+N+1 \\ &= 0 && t < J \text{ or } t > J+N+1. \end{aligned}$$

Then $S_{JN}(t)$ is a spline distribution with knots $(J, ..., J+N+1)$.

We are going to derive a recursion formula for the splines $S_{JN}(t)$. We begin by recalling the standard recursion formula for $D^N_K(t)$.

LEMMA 4.2.1.

$$f^N_K(t) = \frac{K+1-t+Na_1}{1+N+2Na_1}$$

$$s^N_{K-1}(t) = \frac{t+N+1-K+Na_1}{1+N+2Na_1}.$$

Proof. These formulas follow immediately from Proposition 2.3.1.

PROPOSITION 4.2.2.

$$D^{N+1}_K(t) = \frac{(K+1-t)+Na_1}{(1+N)+2Na_1} D^N_K(t) + \frac{(t+N+1-K)+Na_1}{(1+N)+2Na_1} D^N_{K-1}(t).$$

Proof. This result follows immediately from Proposition 2.4.1 and Lemma 4.2.1.

LEMMA 4.2.3.

$$f^N_{N+1-K}(t-K) = \frac{N+2-t+Na_1}{1+N+2Na_1} \qquad K \leqslant t \leqslant K+1$$

$$s^N_{N-K}(t-K) = \frac{t+Na_1}{1+N+2Na_1} \qquad K \leqslant t \leqslant K+1.$$

Proof. These formulas follow immediately from Lemma 4.2.1.

PROPOSITION 4.2.4.

$$S_{J,N+1}(t)=\frac{(t-J)+Na_1}{(1+N)+2Na_1}S_{J,N}(t)+\frac{(N+2+J-t)+Na_1}{(1+N)+2Na_1}S_{J+1,N}(t).$$

Proof. Let $J+K\leqslant t\leqslant J+K+1$. Then by Proposition 2.4.1 and Lemma 4.2.3

$$\begin{aligned}
S_{J,N+1}(t)&=D^N_{N+1-K}(t-J-K)\\
&=f^N_{N+1-K}(t-J-K)\,D^N_{N+1-K}(t-J-K)\\
&\quad+s^N_{N-K}(t-J-K)\,D^N_{N-K}(t-J-K)\\
&=\frac{(N+2)-(t-J)+Na_1}{1+N+2Na_1}D^N_{N-(K-1)}[t-(J+1)-(K-1)]\\
&\quad+\frac{(t-J)+Na_1}{1+N+2Na_1}D^N_{N-K}(t-J-K)\\
&=\frac{(N+2+J-t)+Na_1}{1+N+2Na_1}S_{J+1,N}(t)\\
&\quad+\frac{(t-J)+Na_1}{1+N+2Na_1}S_{J,N}(t).
\end{aligned}$$

Q.E.D.

When $a_1=0$, Proposition 4.2.4 becomes

$$S_{J,N+1}(t)=\frac{(t-J)}{1+N)}S_{J,N}(t)+\frac{(N+2+J-t)}{(1+N)}S_{J+1,N}(t).$$

This recursion formula is identical to the Cox–de Boor recursion formula for B-splines with integral knots [1]. Therefore when $a_1=0$, $a_2=1$, the spline distributions $S_{JN}(t)$ are the normalized uniform B-spline basis functions. Thus Proposition 4.2.4 is a simple generalization of the Cox–de Boor recursion formula for B-splines. We shall return to the subject of B-splines again in Section 5.

We close this section with some additional observations about the functions $f^N_K(t)$, $s^N_K(t)$. These functions are defined probabilistically only for $0\leqslant t\leqslant 1$. However we can use the formulas of Lemma 4.2.1 to extend the definitions of $f^N_K(t)$, $s^N_K(t)$ outside the interval $[0, 1]$. We then have the following result.

LEMMA 4.2.5.

$$f^N_K(t)=f^N_{K-1}(t-1)$$
$$s^K_N(t)=s^N_{K-1}(t-1).$$

Proof. These results follow immediately from Lemma 4.2.1.

COROLLARY 4.2.6.

$$f_K^N(1) = f_{K-1}^N(0)$$
$$s_K^N(1) = s_{K-1}^N(0).$$

We can use Corollary 4.2.6 to give an alternate proof of Proposition 4.1.1.

PROPOSITION 4.1.1 (Revisited). $D_K^N(1) = D_{K-1}^N(0)$.

Proof. By induction on N. Certainly this result is true for $N = 1$ since

$$D_0^1(t) = 1 - t$$
$$D_1^1(t) = t.$$

Now by the recursion formula, Corollary 4.2.6, and the inductive hypothesis

$$\begin{aligned} D_K^{N+1}(1) &= f_K^N(1)\, D_K^N(1) + s_{K-1}^N(1)\, D_{K-1}^N(1) \\ &= f_{K-1}^N(0)\, D_{K-1}^N(0) + s_{K-2}^N(0)\, D_{K-2}^N(0) \\ &= D_{K-1}^{N+1}(0). \end{aligned}$$

Q.E.D.

4.3 *Additional Conjectures Concerning the Laws of Signs*

In Section 2.6 we presented two conjectures concerning the Laws of Signs for the polynomials $\{D_K^N(t)\}$. In this section we shall introduce two additional conjectures regarding the Laws of Signs for the splines $\{S_{JN}(t)\}$.

It is well known that the normalized uniform B-spline basis functions satisfy the Weak Law of Signs over any interval [12]. Thus when $a_1 = 0$, $a_2 = 1$, the splines $\{S_{JN}(t)\}$ satisfy the Weak Law of Signs. This one special case coupled with the basic similarity of all our spline distributions prompts us to propose the following general conjecture.

Conjecture 4.3.1. For all positive finite values of a_1, the splines $\{S_{JN}(t)\}$ satisfy the Weak Law of Signs in any interval.

If we restrict our attention to the unit interval (0, 1), then for $a_2 = 1 + a_1$

$$\text{Conjecture 4.3.1} \Rightarrow \text{Conjecture 2.6.1}$$

because

$$\sum c_K D_K^N(t) = \sum c_K S_{K-N,\,N}(t).$$

For arbitrary integral unit intervals, Conjecture 2.6.2 suggests the following general conjecture.

Conjecture 4.3.2. For all positive finite values of a_1, the splines $\{S_{JN}(t)\}$ satisfy the Strong Law of Signs over any unit interval $(K, K+1)$.

Notice that

$$\text{Conjecture 2.6.2} \Rightarrow \text{Conjecture 4.3.2}$$

since, by definition, over the interval $(K, K+1)$

$$\begin{aligned}\sum_J c_J S_{JN}(t) &= \sum_{J=0}^{N} c_{K-J} S_{K-J,N}(t) \\ &= \sum_{J=0}^{N} c_{K-J} D_{N-J}^{N}(t-K) \\ &= \sum_{J=0}^{N} c_{K+J-N} D_J^N(u),\end{aligned}$$

where $0 \leqslant u = t - K \leqslant 1$. Therefore if the polynomials $\{D_K^N(u)\}$ satisfy the Strong Law of Signs in the interval $(0, 1)$, then the splines $\{S_{JN}(t)\}$ satisfy the Strong Law of Signs in the interval $(K, K+1)$.

The Strong Law of Signs necessarily implies linear independence, and indeed because the polynomials $\{D_K^N(t)\}$ are linearly independent, we have the following result for the splines $\{S_{JN}(t)\}$.

PROPOSITION 4.3.3. *The splines $\{S_{JN}(t)\}$ are linearly independent.*

Proof. Suppose that there are constants $\{c_J\}$ such that for all t

$$\sum c_J S_{JN}(t) = 0.$$

For all t such that $K \leqslant t \leqslant K+1$, let $u = t - K$. Then

$$\sum c_{K+J-N} D_J^N(u) = \sum c_J S_{JN}(t) = 0.$$

But by Corollary 2.5.13 the functions $\{D_K^N(u)\}$ are linearly independent. Therefore

$$c_{K-N} = \cdots = c_K = 0.$$

Since this result is true for every K, it follows that

$$c_J = 0$$

for every J. Therefore the functions $\{S_{JN}(t)\}$ are linearly independent.

Q.E.D.

As yet we know of no proofs, probabilistic or otherwise, for Conjectures 4.3.1, 4.3.2. However, obviously any such proofs must be closely related to the proofs of Conjectures 2.6.1, 2.6.2. We shall discuss the geometric significance of these conjectures further in Section 4.4.

4.4 *Spline Approximations*

In Section 3 we constructed polynomial approximations $D_N[g](t)$ to continuous real-valued functions $g(t)$ defined on the interval $[e_0^N, e_N^N]$. Here we shall generalize this construction to spline approximations $S_N[f](t)$ of continuous real-valued functions $f(t)$ defined over the interval $(-\infty, \infty)$.

Recall from Section 3 that for $a_2 = 1 + a_1$

$$e_K^N = (2K + 1 - N)/2$$

$$e_{J+K}^N = J + e_K^N.$$

Now let $f(t)$ be a continuous real-valued function. Define the linear functionals $S_{JN}[f]$, $S_N[f]$ by setting

$$\begin{aligned} S_{JN}[f](t) &= \sum_K f(e_{J+K}^N)\, D_K^N(t-J) \qquad && J \leqslant t \leqslant J+1 \\ &= 0 && t < J \text{ or } t > J+1 \end{aligned}$$

and

$$S_N[f](t) = \sum_J S_{JN}[f](t).$$

For reasons which will soon become clear, we shall regard $S_{JN}[f]$ as a local approximation and $S_N[f]$ as a global approximation to f.

By construction

$$S_N[f](t) \begin{cases} \vdots \\ S_{0N}[f](t) \\ \vdots \\ S_{JN}[f](t) \\ \vdots \\ S_{NN}[f](t) \\ \vdots \end{cases}$$

$$\begin{array}{llllll} \vdots & \vdots & & & & \vdots \\ = f(e_0^N)\, D_0^N(t) + \cdots & + f(e_J^N)\, D_J^N(t) & + \cdots + f(e_N^N)\, D_N^N(t) & & & 0 \leqslant t \leqslant 1 \\ & \vdots & \vdots & & & \\ = & f(e_J^N)\, D_0^N(t-J) + \cdots & + f(e_N^N)\, D_{N-J}^N(t-J) + \cdots & & & J \leqslant t \leqslant J+1 \\ & & \vdots & & & \vdots \\ = & & f(e_N^N)\, D_0^N(t-N) & + \cdots & & N \leqslant t \leqslant N+1 \\ & & & \vdots & & \vdots \end{array}$$

Scanning these equations vertically rather than horizontally, we observe that in $S_N[f](t)$

$$\text{coefficient } f(e_N^N) = S_{0N}(t)$$
$$\text{coefficient } f(e_J^N) = S_{J-N,N}(t)$$
$$\text{coefficient } f(e_{J+N}^N) = S_{J,N}(t).$$

Therefore we have the following proposition.

PROPOSITION 4.4.1. $S_N[f](t) = \sum_J f(e_{J+N}^N)\, S_{JN}(t)$.

COROLLARY 4.4.2. *$S_N[f](t)$ is a polynomial spline.*

Locally the approximation $S_N[f](t)$ is given by $S_{JN}[f](t)$. Thus global properties of $S_{JN}[f](t)$ are local properties of $S_N[f](t)$. But the local approximations $S_{JN}[f](t)$ are essentially identical to the polynomial approximations $D_N[g](t)$ as we can see from the following lemma.

LEMMA 4.4.3. *Let $f_J(t) = f(t+J)$. Then*

$$S_{JN}[f](t) = D_N[f_J](t-J) \qquad J \leqslant t \leqslant J+1.$$

Proof. By definition

$$\begin{aligned} S_{JN}[f](t) &= \sum_K f(e_{J+K}^N)\, D_K^N(t-J) \\ &= \sum_K f_J(e_K^N)\, D_K^N(t-J) \\ &= D_N[f_J](t-J). \end{aligned}$$

Q.E.D.

Therefore all the standard global properties of the polynomial approximations $D_N[g](t)$ are local properties of the spline approximations $S_N[f](t)$. In particular, we have the following results.

PROPOSITION 4.4.4. *S_{JN} is the identity on linear functions.*

Proof. This result is an immediate consequence of Proposition 3.1.3 and Lemma 4.4.3. It also follows directly from the definition of $S_{JN}[f]$, Proposition 2.1.1, and Corollary 2.5.7.

PROPOSITION 4.4.5. $\text{graph}(S_{JN}[f]) \subseteq \text{convex hull (graph } f)$.

Proof. This result is an immediate consequence of Proposition 3.1.5 and

Lemma 4.4.3. It also follows directly from the definition of $S_{JN}[f]$, Proposition 2.1.1, and Corollary 2.5.7.

PROPOSITION 4.4.6 (Symmetry). *Let* $\bar{f}(t) = f(2J+1-t)$, *Then*

$$S_{JN}[\bar{f}](t) = \overline{S_{JN}[f]}(t).$$

Proof. This result is an immediate consequence of Lemmas 3.2.1 and 4.4.3.

PROPOSITION 4.4.7 (Recursion). *For* $J \leqslant r \leqslant J+1$ *and* $0 \leqslant K+L \leqslant N$, *define*

$$P^0_{J+K}[f](r) = f(e^N_{J+K})$$
$$P^L_{J+K}[f](r) = f^{N-L}_K(r-J)\, P^{L-1}_{J+K}[f](r) + s^{N-L}_K(r-J)\, P^{L-1}_{J+K+1}[f](r).$$

Then $S_{JN}[f](r) = P^N_J[f](r)$.

Proof. This result is an immediate consequence of Proposition 3.3.3 and Lemma 4.4.3.

COROLLARY 4.4.8 (Recursion). *For* $J \leqslant r \leqslant J+1$ *and* $L \leqslant I-J \leqslant N$, *define*

$$Q^0_K[f](r) = f(e^N_I)$$
$$Q^L_I[f](r) = f^{N-L}_{I-J-L}(r-J)\, Q^{L-1}_{I-1}[f](r) + s^{N-L}_{I-J-L}(r-J)\, Q^{L-1}_I[f](r).$$

Then $S_{JN}[f](r) = Q^N_{J+N}[f](r)$.

Proof. By construction $Q^L_I[f](r) = P^L_{I-L}[f](r)$. Therefore this result is an immediate consequence of Proposition 4.4.7.

PROPOSITION 4.4.9 (Uniqueness).

$$S_{JN}[g] = S_{JN}[h] \qquad \text{iff} \quad g(e^N_{J+K}) = h(e^N_{J+K}) \qquad \textit{for} \quad 0 \leqslant K \leqslant N.$$

Proof. This result is an immediate consequence of the definition of $S_{JN}[f]$ and Proposition 2.5.13.

PROPOSITION 4.4.10 (Uniqueness).

$$S_N[g] = S_N[h] \qquad \text{iff} \quad g(e^N_K) = h(e^N_K) \quad \text{for every integer } K.$$

Proof. This result is an immediate consequence of Propositions 4.3.3 and 4.4.1.

Finally Conjecture 3.5.3 concerning the variation diminishing property for the approximations $D_N[g]$ suggests the following conjectures for the approximations $S_{JN}[f]$, $S_N[f]$.

Conjecture 4.4.11. For all positive finite values of a_1, the linear functionals $S_{JN}[f]$ are variation diminishing in the interval $(J, J+1)$.

Conjecture 4.4.12. For all positive finite values of a_1, the linear functionals $S_N[f]$ are variation diminishing in any interval.

TABLE I

$D_K^N(t)$	$S_{JN}(t)$
1. Polynomial	1. Polynomial Spline
2. Discrete Probability Distribution a. $D_K^N(t) \geqslant 0 \quad 0 \leqslant t \leqslant 1$ b. $\sum_k D_K^N(t) = 1$	2. Continuous Probability Distribution a. $S_{0N}(t) \geqslant 0$ b. $\int_0^{N+1} S_{0N}(t)\,dt = 1$
3. A Priori Probability $S_N(t) = t \quad N = 1$ $= 1/2 \quad N > 1$	3. A Priori Probability ?
4. Expectation $\sum_{K=0}^{N} K D_K^N(t) = t + \frac{N+1}{2}$	4. Expectation $\int_0^{N+1} t S_{0N}(t)\,dt = \frac{N+1}{2}$
5. Symmetry $D_K^N(t) = D_{N-K}^N(1-t)$	5. Symmetry $S_{0N}(t) = S_{0N}(N+1-t)$
6. Explicit Formulas a. $D_0^N(t) = \prod \frac{(1-t+Ka_1)}{(1+K+2Ka_1)}$ b. $D_N^N(t) = \prod \frac{(t+Ka_1)}{(1+K+2Ka_1)}$	6. Explicit Formulas a. $S_{0N}(t) = \prod \frac{(t+Ka_1)}{(1+K+2Ka_1)} \quad 0 \leqslant t \leqslant 1$ b. $S_{0N}(t) = \prod \frac{(N+1-t+Ka_1)}{(1+K+2Ka_1)} \quad N \leqslant t \leqslant N+1$
7. Recursion Formula $D_K^{N+1}(t) = \frac{(K+1-t)+Na_1}{(1+N)+2Na_1} D_K^N(t) + \frac{(t+N+1-K)+Na_1}{(1+N)+2Na_1} D_{K-1}^N(t)$	7. Recursion Formula $S_{J,N+1}(t) = \frac{(t-J)+Na_1}{(1+N)+2Na_1} S_{JN}(t) + \frac{(N+2+J-t)+Na_1}{(1+N)+2Na_1} S_{J+1,N}(t)$
8. Limits $\lim_{a_1 \to \infty} D_K^N(t) =$ Linear Polynomial	8. Limits $\lim_{a_1 \to \infty} S_{0N}(t) =$ Linear Spline
9. Law of Signs Strong Law of Signs in the Interval (0, 1)?	9. Law of Signs Weak Law of Signs in any Interval?
10. Linear Independence	10. Linear Independence

By Lemma 4.4.3 and Proposition 3.5.1 for $a_2 = 1 + a_1$

$$\text{Conjecture 4.4.11} \Leftrightarrow \text{Conjecture 3.5.3} \Leftrightarrow \text{Conjecture 2.6.1}$$

and by Propositions 3.5.1 and 4.4.1

$$\text{Conjecture 4.4.12} \Leftrightarrow \text{Conjecture 4.3.1.}$$

Moreover since $S_N[f](t) = S_{JN}[f](t)$ for $J \leqslant t \leqslant J+1$

$$\text{Conjecture 4.4.12} \Rightarrow \text{Conjecture 4.4.11.}$$

Thus Conjecture 4.4.12 is the strongest of our conjectures concerning the variation diminishing property. Now the variation diminishing property is known to be valid for the normalized uniform B-spline basis functions [12]; thus Conjecture 4.4.12 is valid when $a_1 = 0$, $a_2 = 1$. We believe that Conjecture 4.4.12 (the variation diminishing property) is valid for all of our spline approximations—that is, for all values of a_1—but like Conjecture 4.3.1 (the Weak Law of Signs) we know of no proof, probabilistic or otherwise, for this general conjecture.

4.5 *Summary*

In Tables I and II we collect, compare, and contrast our results first for the functions $D_K^N(t)$ and $S_{JN}(t)$ and second for the approximations $D_N[g](t)$ and $S_N[f](t)$.

TABLE II

$D_N[g](t)$	$S_N[f](t)$
1. Polynomial Approximation	1. Spline Approximation
2. Linear Functions $D_N[L] = L$	2. Linear Functions $S_N[L] = L$
3. Convex Hull Property graph$(D_N[g]) \subseteq$ convex hill (graph g)	3. Convex Hull Property graph$(S_N[f]) \subseteq$ convex hull (graph f)
4. Symmetry $[\bar{g}(t) = g(1-t)]$ $D_N[\bar{g}](t) = \overline{D_N[g]}(t)$	4. Symmetry $[\bar{f}(t) = f(2J+1-t)]$ $S_{JN}[\bar{f}](t) = \overline{S_{JN}(f)}(t)$
5. Recursion $D_N[g](r) = P_0^N[g](r)$	5. Recursion $S_{JN}[f](r) = Q_{J+N}^N[f](r)$
6. Uniqueness $D_N[g] = D_N[h]$ iff $g(e_K^N) = h(e_K^N) \quad 0 \leqslant K \leqslant N$	6. Uniqueness $S_N[g](t) = S_N[h](t)$ iff $g(e_K^N) = h(e_K^N) \quad$ for all K.
7. Variation Diminishing $D_N[g]$ is Variation Diminishing in the Interval (0, 1)?	7. Variation Diminishing $S_N[f]$ is Variation Diminishing in Any Interval?

5. *B*-SPLINES

Let $x_0 < x_1 < \cdots < x_{N+1}$ be a sequence of increasing, evenly spaced values along the t-axis. A function $B(t)$ is said to be the normalized uniform B-spline basis function of degree N, order $N+1$, for the knot vector $(x_0, ..., x_{N+1})$ iff there are $N+1$ degree N polynomials $b_0(t), ..., b_N(t)$ such that

$$\begin{aligned} B(t) &= b_K(t) && x_K \leqslant t \leqslant x_{K+1} \\ &= 0 && t < x_0 \text{ or } t > x_{N+1} \end{aligned}$$

and the polynomials $b_0(t), ..., b_N(t)$ satisfy the following 4 conditions:

1. $b_0^{(p)}(x_0) = 0 \qquad p = 0, 1, ..., N-1$
2. $b_{K+1}^{(p)}(x_{K+1}) = b_K^{(p)}(x_{K+1}) \qquad p = 0, 1, ..., N-1$
3. $b_N^{(p)}(x_{N+1}) = 0 \qquad p = 0, 1, ..., N-1$
4. $\sum_k 1/\Delta x \int_{x_K}^{x_{K+1}} b_K(t)\, dt = 1$ (Normalization).

Thus a B-spline is a polynomial spline that has the maximum possible differentiability at the knots without collapsing 2 adjacent segments into a single polynomial.

To construct the normalized, uniform, degree N, B-spline basis function $B(t)$ for an arbitrary evenly spaced knot vector $(x_0, ..., x_{N+1})$, we need only construct the normalized, uniform, degree N, B-spline basis function $B_{0N}(t)$ for the canonical knot vector $(0, 1, ..., N+1)$. Indeed it is easy to verify that in general

$$B(t) = B_{0N}\left(\frac{t - x_0}{\Delta x}\right).$$

We shall now use an urn model to construct $B_{0N}(t)$.

5.1 *An Urn Model for B-Splines*

Consider an urn initially containing w white balls and b black balls. One ball at a time is drawn at random from the urn and its color is inspected. It is then returned to the urn and $w + b$ balls of the opposite color are added to the urn.

This urn model is just the special case of Friedman's urn model for which $a_1 = 0$, $a_2 = 1$. Moreover, it is the simplest spline model ($a_1 = 0$). For this special urn model we shall adopt the notation

$$B_K^N(t) = D_K^N(t)$$

$$B_{JN}(t) = S_{JN}(t).$$

By Proposition 4.2.4 the functions $B_{JN}(t)$ satisfy the Cox–de Boor recursion formula

$$B_{J,\,N+1}(t)=\frac{(t-J)}{(1+N)}\,B_{JN}(t)+\frac{(N+2+J-t)}{(1+N)}\,B_{J+1,N}(t).$$

Therefore it follows immediately, though somewhat obliquely, that the functions $B_{JN}(t)$ are the normalized uniform B-spline basis functions of degree N, order $N+1$, for the knot vectors $(J, ..., J+N+1)$. We shall now give a simpler more direct proof of this fact based on the simpler more primitive recursion formula

$$B_K^{N+1}(t)=\frac{(K+1-t)}{(N+1)}\,B_K^N(t)+\frac{(t+N+1-K)}{(N+1)}\,B_{K-1}^N(t)$$

of Proposition 4.2.2.

LEMMA 5.1.1.

$$B_K^N(t)=\frac{(-1)^{N-K}}{N!}\binom{N}{K}t^N+\cdots.$$

Proof. By induction on N. Certainly this result is true for $N=1$. Moreover by the recursion formula and the inductive hypothesis

$$\begin{aligned}B_K^{N+1}(t)&=\frac{(K+1-t)}{(N+1)}\,B_K^N(t)+\frac{(t+N+1-K)}{(N+1)}\,B_{K-1}^N(t)\\&=\frac{(-1)^{N+1-K}}{(N+1)!}\left[\binom{N}{K}+\binom{N}{K-1}\right]t^{N+1}+\cdots\\&=\frac{(-1)^{N+1-K}}{(N+1)!}\binom{N+1}{K}t^{N+1}+\cdots.\end{aligned}$$

Q.E.D.

COROLLARY 5.1.2.

$$\frac{d^N B_K^N}{dt^N}=(-1)^{N-K}\binom{N}{K}.$$

PROPOSITION 5.1.3. *The functions* $B_0^N(t), ..., B_N^N(t)$ *are degree* N *polynomials, and they satisfy the following* 4 *conditions*:

1. $\left.\dfrac{d^p B_N^N}{dt^p}\right|_{t=0} = 0 \qquad p = 0, 1, ..., N-1$

2. $\left.\dfrac{d^p B_K^N}{dt^p}\right|_{t=1} = \left.\dfrac{d^p B_{K-1}^N}{dt^p}\right|_{t=0} \qquad p = 0, 1, ..., N-1$

3. $\left.\dfrac{d^p B_0^N}{dt^p}\right|_{t=1} = 0 \qquad p = 0, 1, ..., N-1$

4. $\sum \int_0^1 B_K^N(t)\, dt = 1.$

Proof. When $p = 0$, parts 1, 2, 3 follow from Proposition 4.1.1. When $p \neq 0$, we proceed by induction on N.

1. By the recursion formula

$$B_{N+1}^{N+1}(t) = \frac{t}{(N+1)} B_N^N(t).$$

Therefore by Leibniz's Rule,

$$\frac{d^p B_{N+1}^{N+1}}{dt^p} = \frac{p}{(N+1)} \frac{d^{p-1} B_N^N}{dt^{p-1}} + \frac{t}{(N+1)} \frac{d^p B_N^N}{dt^p}.$$

Hence by the inductive hypothesis

$$\left.\frac{d^p B_{N+1}^{N+1}}{dt^p}\right|_{t=0} = \frac{p}{(N+1)} \left.\frac{d^{p-1} B_N^N}{dt^{p-1}}\right|_{t=0} = 0 \qquad p = 1, ..., N.$$

3. Again by the recursion formula

$$B_0^{N+1}(t) = \frac{(1-t)}{(N+1)} B_0^N(t).$$

Therefore by Leibniz's Rule,

$$\frac{d^p B_0^{N+1}}{dt^p} = \frac{-p}{(N+1)} \frac{d^{p-1} B_0^N}{dt^{p-1}} + \frac{(1-t)}{(N+1)} \frac{d^p B_0^N}{dt^p}.$$

Hence by the inductive hypothesis

$$\left.\frac{d^p B_0^{N+1}}{dt^p}\right|_{t=1} = \frac{-p}{(N+1)} \left.\frac{d^{p-1} B_0^N}{dt^{p-1}}\right|_{t=1} = 0 \qquad p = 1, ..., N.$$

2. Again by the recursion formula

$$B_K^{N+1}(t) = \frac{(1-t+K)}{(N+1)} B_K^N(t) + \frac{(t+N+1-K)}{(N+1)} B_{K-1}^N(t).$$

Therefore by Leibniz's Rule,

$$\frac{d^p B_K^{N+1}}{dt^p} = \frac{1}{(N+1)}\left[-p\frac{d^{p-1}B_K^N}{dt^{p-1}} + (1-t+K)\frac{d^p B_K^N}{dt^p}\right] + \frac{1}{(N+1)}\left[p\frac{d^{p-1}B_{K-1}^N}{dt^{p-1}} + (t+N+1-K)\frac{d^p B_{K-1}^N}{dt^p}\right].$$

Hence at $t=1$

$$\frac{d^p B_K^{N+1}}{dt^p}\Big|_{t=1} = \frac{p}{(N+1)}\left[\frac{d^{p-1}}{dt^{p-1}}[B_{K-1}^N(t) - B_K^N(t)]|_{t=1}\right] + \frac{1}{(N+1)}\left[\frac{d^p}{dt^p}[(N+2-K)B_{K-1}^N(t) + KB_K^N(t)]|_{t=1}\right].$$

Similarly for $K-1$ and $t=0$

$$\frac{d^p B_{K-1}^{N+1}}{dt^p}\Big|_{t=0} = \frac{p}{(N+1)}\left[\frac{d^{p-1}}{dt^{p-1}}B_{K-2}^N(t) - B_{K-1}^N(t)]|_{t=0}\right] + \frac{1}{(N+1)}\left[\frac{d^p}{dt^p}[(N+2-K)B_{K-2}^N(t) + KB_{K-1}^N(t)]|_{t=0}\right].$$

Therefore comparing term by term, it follows immediately from the inductive hypothesis that

$$\frac{d^p B_K^{N+1}}{dt^p}\Big|_{t=1} = \frac{d^p B_{K-1}^{N+1}}{dt^p}\Big|_{t=0} \qquad p=1, \ldots, N-1.$$

Moreover, to prove that this result is also true for $p=N$, we need only show that

$$(N+2-K)\frac{d^N B_{K-1}^N}{dt^N}\Big|_{t=1} + K\frac{d^N B_K^N}{dt^N}\Big|_{t=1} = (N+2-K)\frac{d^N B_{K-2}^N}{dt^N}\Big|_{t=0} + K\frac{d^N B_{K-1}^N}{dt_N}\Big|_{t=0}.$$

But by Corollary 5.1.2.

$$\frac{d^N B_K^N}{dt^N} = (-1)^{N-K}\binom{N}{K}.$$

Therefore

$$\begin{aligned}
&(N+2-K)\frac{d^N B_{K-1}^N}{dt^N} + K\frac{d^N B_K^N}{dt^N} \\
&\quad = (-1)^{N-K+1}(N+2-K)\binom{N}{K-1} + (-1)^{N-K}K\binom{N}{K} \\
&\quad = (-1)^{N+1-K}\binom{N}{K-1}[(N+2-K)-(N+1-K)] \\
&\quad = (-1)^{N+1-K}\binom{N}{K-1}.
\end{aligned}$$

Similarly

$$\begin{aligned}
&(N+2-K)\frac{d^N B_{K-2}^N}{dt^N} + K\frac{d^N B_{K-1}^N}{dt^N} \\
&\quad = (-1)^{N-K+2}(N+2-K)\binom{N}{K-2} + (-1)^{N-K+1}K\binom{N}{K-1} \\
&\quad = (-1)^{N+1-K}\binom{N}{K-1}[-(K-1)+K] \\
&\quad = (-1)^{N+1-K}\binom{N}{K-1}
\end{aligned}$$

so the result is true for $p = N$.

4. This is easy since by Proposition 2.1.1

$$\begin{aligned}
\sum \int_0^1 D_K^N(t)\,dt &= \int_0^1 \sum D_K^N(t)\,dt \\
&= \int_0^1 dt \\
&= 1.
\end{aligned}$$

Q.E.D.

COROLLARY 5.1.4. *The spline $B_{0N}(t)$ is the normalized uniform B-spline basis function for the canonical knot vector* $(0, 1, ..., N+1)$.

Proof. Let $b_K(t) = D^N_{N-K}(t-K)$. Then by construction

$$B_{0N}(t) = b_K(t) \qquad K \leqslant t \leqslant K+1.$$

We must show that the polynomials $b_K(t)$ satisfy the 4 conditions which define the B-spline basis function for the knot vector $(0, 1, ..., N+1)$. Now by Proposition 5.1.3 for $p = 0, 1, ..., N-1$

1. $$b_0^{(p)}(0) = \frac{d^p D^N_N}{dt^p}\Big|_{t=0} = 0$$

2. $$b_{K+1}^{(p)}(K+1) = \frac{d^p D^N_{N-K-1}}{dt^p}\Big|_{t=0} = \frac{d^p D^N_{N-K}}{dt^p}\Big|_{t=1} = b_K^{(p)}(K+1)$$

3. $$b_N^{(p)}(N+1) = \frac{d^p D^N_0}{dt^p}\Big|_{t=1} = 0$$

4. $$\sum_K \int_K^{K+1} b_k(t)\, dt = \sum_K \int_K^{K+1} D^N_{N-K}(t-K)\, dt = \sum_K \int_0^1 D^N_{N-K}(u)\, du = 1.$$

Q.E.D.

COROLLARY 5.1.5. *The spline $B_{JN}(t)$ is the normalized uniform B-spline basis function for the knot vector* $(J, ..., J+N+1)$.

Proof. This result is an immediate consequence of Corollary 5.1.4 since by construction

$$B_{JN}(t) = B_{0N}(t-J).$$

We can also use the recursion formula to derive explicit expressions for $B^N_K(t)$, $B_{0N}(t)$.

PROPOSITION 5.1.6. $B^N_K(t) = 1/N! \sum_{J=0}^{N-K} (-1)^J \binom{N+1}{J} (t+N-K-J)^N$.

Proof. By induction on N. For $N=1$, we get

$$B_0^1(t)=(t+1)-2t=1-t$$
$$B_1^1(t)=t$$

as required. Now by the recursion formula and the inductive hypothesis

$$\begin{aligned}
B_K^{N+1}(t)&=\frac{(1-t+K)}{(N+1)}B_K^N(t)+\frac{(t+N+1-K)}{(N+1)}B_{K-1}^N(t)\\
&=\frac{1}{(N+1)!}\sum_{J=1}^{N+1-K}(-1)^{J-1}\binom{N+1}{J-1}\\
&\quad\times(t+N+1-K-J)^N(1-t+K)\\
&\quad+\frac{1}{(N+1)!}\sum_{J=0}^{N+1-K}(-1)^J\binom{N+1}{J}\\
&\quad\times(t+N+1-K-J)^N(t+N+1-K)\\
&=\frac{1}{(N+1)!}\sum_{J=0}^{N+1-K}(-1)^J(t+N+1-K-J)^NF(t)
\end{aligned}$$

$$\begin{aligned}
F(t)&=\frac{(N+1)!}{J!(N+2-J)!}[(N+2-J)(t+N+1-K)-J(1-t+K)]\\
&=\frac{(N+1)!}{J!(N+2-J)!}[(N+2)(t+N+1-K-J)].
\end{aligned}$$

Thus

$$F(t)=\binom{N+2}{J}(t+N+1-K-J)$$

so

$$D_K^{N+1}(t)=\frac{1}{(N+1)!}\sum_{J=0}^{N+1-K}(-1)^J\binom{N+2}{J}(t+N+1-K-J)^{N+1}$$

as required. Q.E.D.

COROLLARY 5.1.7.

$$\begin{aligned}
B_{0N}(t)&=\frac{1}{N!}\sum_{J=0}^{N}(-1)^K\binom{N+1}{J}(t-J)_+^N \qquad & 0\leqslant t\leqslant N+1\\
&=0 & \textit{otherwise.}
\end{aligned}$$

Proof. Let $0 \leqslant K \leqslant t \leqslant K+1 \leqslant N+1$. Then by Proposition 5.1.6

$$\begin{aligned} B_{0N}(t) &= B_{N-K}^{N}(t-K) \\ &= \frac{1}{N!} \sum_{J=0}^{K} (-1)^{J} \binom{N+1}{J} (t-J)^{N} \\ &= \frac{1}{N!} \sum_{J=0}^{N} (-1)^{J} \binom{N+1}{J} (t-J)_{+}^{N}. \end{aligned} \qquad \text{Q.E.D.}$$

We close this section with a result relating the limits of urn distributions to the values of the normalized uniform B-spline basis functions at the knots.

PROPOSITION 5.1.8. *If $N > 1$, then*

$$\lim_{a_2 \to \infty} D_K^N(a_1, a_2, t) = B_{0,\,N-1}(N-K).$$

Proof. By Proposition 2.7.5 if $N > 1$, then

$$\begin{aligned} \lim_{a_2 \to \infty} D_K^N(a_1, a_2, t) &= D_K^{N-1}(0, 1, 1) \\ &= B_K^{N-1}(1) \\ &= B_{0,\,N-1}(N-K). \end{aligned} \qquad \text{Q.E.D.}$$

By Proposition 2.7.5 as a_2 approaches infinity the urn distributions approach constant values and by Proposition 5.1.8 these values are just the values at the knots of the normalized uniform B-spline basis functions.

5.2 *Derivatives Revisited*

In Proposition 5.1.3 we proved that the functions $B_0^N(t), \ldots, B_N^N(t)$ can be joined together smoothly up to order $N-1$. However, this proof provides little or no insight into why the particular spline model $a_1 = 0$, $a_2 = 1$ is the correct model for B-splines rather than one of the other spline models. To rectify this situation, we now provide an alternate proof of Proposition 5.1.3 based on Propositions 2.8.5. and 4.1.1.

PROPOSITION 5.1.3 (Revisited). *The functions $B_0^N(t), \ldots, B_N^N(t)$ are degree N polynomials, and they satisfy the following* 4 *conditions:*

1. $\left.\dfrac{d^p B_N^N}{dt^p}\right|_{t=0} = 0 \qquad p = 0, 1, \ldots, N-1$

2. $\left.\dfrac{d^p B_K^N}{dt^p}\right|_{t=1} = \left.\dfrac{d^p B_{K-1}^N}{dt^p}\right|_{t=0} \qquad p = 0, 1, \ldots, N-1$

3. $\left.\dfrac{d^p B_0^N}{dt^p}\right|_{t=1} = 0 \qquad p = 0, 1, \ldots, N-1$

4. $\sum \int_0^1 B_K^N(t)\, dt = 1.$

Proof. The main facts are these: by Proposition 4.1.1 any urn model for which $a_2 = 1 + a_1$ satisfies

$$D_K^N(1) = D_{K-1}^N(0). \tag{$*$}$$

On the other hand if $a_1 = 0$, then by Proposition 2.8.5 we know the derivatives of the functions $D_K^N(t)$ in terms of the functions $D_K^{N-1}(t)$. Indeed if $a_1 = 0$, we have

$$\begin{aligned} \left.\frac{d^p D_K^N}{dt^p}\right|_{t=1} &= \frac{N(N-1)\cdots(N-p+1)}{d_{N-1}\, d_{N-2} \cdots d_{N-p}} \sum (-1)^{j+p} \binom{p}{j} D_{K-j}^{N-p}(1) \\ \left.\frac{d^p D_{K-1}^N}{dt^p}\right|_{t=0} &= \frac{N(N-1)\cdots(N-p+1)}{d_{N-1}\, d_{N-1} \cdots d_{N-p}} \sum (-1)^{j+p} \binom{p}{j} D_{K-j-1}^{N-p}(0). \end{aligned} \tag{$**$}$$

Now if $a_2 = 1 + a_1$ and $a_1 = 0$, then $(*)$ and $(**)$ together imply

$$\left.\frac{d^p B_K^N}{dt^p}\right|_{t=1} = \left.\frac{d^p B_{K-1}^N}{dt^p}\right|_{t=0} \qquad p = 0, 1, \ldots, N-1.$$

But this is exactly what we needed to prove for part 2. Moreover parts 1, 3 follow easily since by Proposition 2.3.2 and $(**)$ we have

$$\left.\frac{d^p B_N^N}{dt^p}\right|_{t=0} = \pm B_{N-p}^{N-p}(0) = 0 \qquad p = 0, 1, \ldots, N-1$$

$$\left.\frac{d^p B_0^N}{dt^p}\right|_{t=1} = \pm B_0^{N-p}(1) = 0 \qquad p = 0, 1, \ldots, N-1.$$

Finally part 4 is true for every urn model since by Proposition 2.2.1

$$\sum D_K^N(t) = 1.$$

Q.E.D.

The fact that we have formulas for the derivatives of $D_K^N(t)$ when $a_1 = 0$, $a_2 \neq 1$ is not an accident. Indeed these urn models actually generate certain special non-uniform B-splines [8]

We close this section by noting that we can also use the results of Section 2.8 to calculate the derivatives of the B-splines $B_{JN}(t)$ in terms of lower order B-splines. Indeed we have the following results.

PROPOSITION 5.2.1. $dB_{JN}/dt = B_{J,\,N-1}(t) - B_{J+1,\,N-1}(t)$.

Proof. This result follows directly from Corollary 2.8.7.

COROLLARY 5.2.2. $d^pB_{JN}/dt^p = \sum (-1)^i \binom{p}{i} B_{J+i,\,N-p}(t)$.

Proof. This result follows easily from Proposition 5.2.1 by induction on p.

5.3 *Summary*

Since the normalized uniform B-spline basis functions can be generated from an urn model, many of the special geometric features of these splines are simply reflections of the distinctive stochastic characteristics of the urn model. Thus symmetry, recursion, and normalization can all be derived by simple, discrete, counting arguments. In particular, the recursion formula

$$B_K^{N+1}(t) = \frac{(K+1-t)}{(N+1)} B_K^N(t) + \frac{(t+N+1-K)}{(N+1)} B_{K-1}^N(t)$$

is the Cox–de Boor recursion formula in its simplest, most primitive form. Thus the standard Cox–de Boor recursion formula

$$B_{J,\,N+1}(t) = \frac{(t-J)}{(N+1)} B_{JN}(t) + \frac{(N+2+J-t)}{(N+1)} B_{J+1,\,N}$$

is just a special case of the general recursion formula which is a characteristic feature of all Friedman urn models (see Section 2.4).

We summarize our results for B-splines in Table III. Except for items 2, 3, 8c, all of these properties follow from Table I in Section 4.5 by setting $a_1 = 0$. Item 2 is, of course, just Proposition 5.1.3 and Corollary 5.1.4; item 3 is just Corollaries 2.8.7 and 5.2.2; and item 8c is just Proposition 5.1.6 and Corollary 5.1.7.

All the results in Table II of Section 4.5 are also valid for B-splines, but since all these results are independent of the value of a_1 we shall not repeat them here.

TABLE III

$B_K^N(t)$	$B_{JN}(t)$
1. Polynomial	1. Polynomial Spline
2. Differentiability Conditions	2. Differentiability Conditions
$\left.\frac{d^p B_K^N}{dt^p}\right\|_{t=1} = \left.\frac{d^p B_{K-1}^N}{dt^p}\right\|_{t=0} \quad p=0, 1, ..., N-1$	$\left.\frac{d^p B_{0N}}{dt^p}\right\|_{t=K^-} = \left.\frac{d^p B_{0N}}{dt^p}\right\|_{t=K^+} \quad p=0, 1, ..., N-1$
3. Derivatives	3. Derivatives
$\frac{d^p B_K^N}{dt^p} = \sum (-1)^{i+p} \binom{p}{i} B_{K-i}^{N-p}(t)$	$\frac{d^p B_{JN}}{dt^p} = \sum (-1)^i \binom{p}{i} B_{J+i, N-p}(t)$
4. Discrete Probability Distribution	4. Continuous Probability Distribution
a. $B_K^N(t) \geqslant 0 \quad 0 \leqslant t \leqslant 1$	a. $B_{0N}(t) \geqslant 0$
b. $\sum_K B_K^N(t) = 1$	b. $\int_0^{N+1} B_{0N}(t)\, dt = 1$
5. A Priori Probability	5. A Priori Probability
$S_N(t) = t \quad N=1$ $= 1/2 \quad N>1$	?
6. Expectation	6. Expectation
$\sum_{K=0}^{N} KB_K^N(t) = t + \frac{N-1}{2}$	$\int_0^{N+1} tB_{0N}(t)\, dt = \frac{N+1}{2}$
7. Symmetry	7. Symmetry
$B_K^N(t) = B_{N-K}^N(1-t)$	$B_{0N}(t) = B_{0N}(N+1-t)$
8. Explicit Formulas	8. Explicit Formulas
a. $B_0^N(t) = \frac{(1-t)^N}{N!}$	a. $B_{0N}(t) = \frac{t^N}{N!} \quad 0 \leqslant t \leqslant 1$
b. $B_N^N(t) = \frac{t^N}{N!}$	b. $B_{0N}(t) = \frac{(N+1-t)^n}{N!} \quad N \leqslant t \leqslant N+1$
c. $B_K^N(t) = \frac{1}{N!} \sum_{J=0}^{N-K} (-1)^J \binom{N+1}{J} (t+N-K-J)^N$	c. $B_{0N}(t) = \frac{1}{N!} \sum_{J=0}^{N} (-1)^J \binom{N+1}{J} (t-J)_+^N$
9. Recursion Formula	9. Recursion Formula
$B_K^{N+1}(t) = \frac{(K+1-t)}{(N+1)} B_K^N(t) + \frac{(t+N+1-K)}{(N+1)} B_{K-1}^N(t)$	$B_{J,N+1}(t) = \frac{(t-J)}{(N+1)} B_{JN}(t) + \frac{(N+2+J-t)}{(N+1)} B_{J+1,N}(t)$
10. Law of Signs	10. Law of Signs
Strong Law of Signs in the Interval (0, 1)[12]	Weak Law of Signs in any Interval [12].
11. Polynomial Basis	11. Spline Basis

Finally we note that there also exist urn models which generate non-uniform B-splines and B-splines with multiple knots, but these are a subject for another paper [8].

6. Conclusions and Questions

Probability theory and approximation theory are intimately related. Many of the classical geometric properties of standard approximation techniques are just reflections of the simple stochastic properties of corresponding urn models. Thus rather than derive these geometric properties from explicit algebraic expressions, we have tried, whenever possible, to give probabilistic arguments. These arguments are simpler, more general, more natural, and more elegant. By adopting this high-level perspective, we have realized a deeper level of unity and understanding.

Still, many questions remain. Are both Laws of Signs indeed valid for all Friedman urn models? Do there exist simple probabilistic proofs for these Laws? It is not easy to see any obvious connection between probability theory and the Laws of Signs. Yet after all we have said and done, it is hard to believe that no link exists.

The Laws of Signs imply the variation diminishing property. Are all the approximation schemes derived from Friedman's urn model variation diminishing? Do the spline distributions generated from urn models all satisfy the Weak Law of Signs? Are the corresponding approximation schemes also always variation diminishing?

Spline distributions generalize the notion of normalized, uniform, B-spline functions. Do these spline distributions have any practical applications?

Differential conditions still elude direct probabilistic interpretations. Is there any insight that probability theory can provide about these critical conditions?

From Friedman's urn model we have singled out three fundamental sequences: the Polya–Eggenberger models ($a_2 = 0$) whose most prominent representative is the binomial distribution ($a_1 = 0$), the spline models ($a_2 = 1 + a_1$) whose most important representatives are the uniform B-splines ($a_1 = 0$), and certain very special non-uniform B-splines ($a_1 = 0$) whose most distinguished representatives are again the uniform B-splines ($a_2 = 1$). Are there any other interesting useful sequences within Friedman's urn model? Do they also have applications in approximation theory?

Friedman's urn model can be generalized in two ways: by adding different number of balls of each color after each distinct trial or by considering urns containing balls of three or more distinct colors. The first method can be used to generate many new types of splines including all univariate non-uniform B-splines [8]; the second method may lead to novel types of splines in two or more variables. In relation to splines neither of these techniques has been explored in any detail. Exactly what splines do they generate? What are their applications to approximation theory? Is there any relationship between urns with multiple colors and multivariate B-splines?

Other discrete probabilistic models—for example, the Poisson models—are important in probability theory. Can these models also be applied to solve problems in approximation theory?

Continuous probability distributions are barely touched upon in this paper. What precisely is the role of continuous distributions in approximation theory?

Finally, Laplace and Fourier transforms play a fundamental role in probability theory. Do they also have a central role in approximation theory?

ACKNOWLEDGMENTS

I thank Dr. Charles Micchelli for pointing out that the approximations $D_N[g](t)$ could always be defined so that D_N is the identity on linear functions simply by evaluation g at e_K^N rather than at K/N. I also thank the referee for many valuable suggestions concerning the style and organization of this paper.

REFERENCES

1. C. DEBOOR, "A Practical Guide to Splines," Springer-Verlag, New York/Berlin, 1978.
2. K. L. CHUNG, "Elementary Probability Theory with Stochastic Processes," Springer-Verlag, New York/Berlin, 1975.
3. P. J. DAVIS, "Interpolation and Approximation," Dover, New York, 1975.
4. F. EGGENBERGER AND G. POLYA, Uber die Statistik Verketterer Vorgange, *Z. Angew. Math. Mech.* **1** (1923), 279–289.
5. A. R. FORREST, Interactive interpolation and approximation by Bezier polynomials, *Comput. J.* **15** (1972), 71–79.
6. B. FRIEDMAN, A simple urn model, *Comm. Pure Appl. Math.* **2** (1949), 59–70.
7. R. N. GOLDMAN, Polya's urn model and computer aided geometric design, *SIAM J. Algebraic Discrete Methods* **6**, No. 1 (1985), 1–28.
8. R. N. GOLDMAN, Urn models and B-splines, *Constructive Approximation*, to appear.
9. D. C. HEATH, Private communication.
10. B. W. LINDGREN, AND G. W. MCELRATH, "Introduction to Probability and Statistics," Macmillan Co., New York, 1959.
11. G. POLYA AND I. J. SCHOENBERG, Remarks on De La Vallée Poussin means and convex conformal maps of the circle, *Pacific J. Math.* **8** (1958), 295–334.
12. L. L. SCHUMAKER, "Spline Functions: Basic Theory," Wiley, New York, 1981.
13. D. D. STANCU, Approximation of functions by a new class of linear polynomial operators, *Rev. Roumaine Math. Pures Appl.* **13** (1968), 1173—1194.
14. D. D. STANCU, Use of probabilistic methods in the theory of uniform approximation of continuous functions, *Rev. Roumaine Math. Pures Appl.* **14** (1969), 673–691.