Krawtchouk polynomials and finite probability theory

Philip Feinsilver, René Schott

▶ To cite this version:

Philip Feinsilver, René Schott. Krawtchouk polynomials and finite probability theory. [Research Report] RR-1744, INRIA. 1992. <inria-00076984>

HAL Id: inria-00076984
https://hal.inria.fr/inria-00076984

Submitted on 29 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNITÉ DE RECHERCHE
INRIA-LORRAINE

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1992
25ème anniversaire

N° 1744

Programme 2
Calcul Symbolique, Programmation et Génie logiciel

KRAWTCHOUK POLYNOMIALS AND FINITE PROBABILITY THEORY

Philip FEINSILVER
René SCHOTT

Septembre 1992

RR_1744

Krawtchouk Polynomials and Finite Probability Theory

Polynômes de Krawtchouk et Probabilités sur un Univers fini

Philip Feinsilver*, René Schott †

Abstract

Some general remarks on random walks and martingales for finite probability distributions are presented. Orthogonal systems for the multinomial ditribution arise. In particular, a class of generalized Krawtchouk polynomials is determined by a random walk generated by roots of unity. Relations with hypergeometric functions and some limit theorems are discussed.

Résumé

Nous présentons quelques propriétés concernant les marches aléatoires et les martingales quand les distributions sont définies sur des univers finis. Des systèmes orthogonaux pour la loi multinomiale sont mis en évidence. En particulier, une classe de polynômes de Krawtchouk généralisés a pu être déterminée en considérant une marche aléatoire associée aux racines complexes de l'unité. Certains liens entre fonctions hypergéométriques et théorèmes limites sont également explicités.

*Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901 USA

†CRIN, INRIA-Lorraine, Université de Nancy 1, 54506 Vandoeuvre-lès-Nancy, France

Krawtchouk Polynomials and Finite Probability Theory

P. Feinsilver [1] *and R. Schott* [2]

Abstract. Some general remarks on random walks and martingales for finite probability distributions are presented. Orthogonal systems for the multinomial distribution arise. In particular, a class of generalized Krawtchouk polynomials is determined by a random walk generated by roots of unity. Relations with hypergeometric functions and some limit theorems are discussed.

I. Introduction

Krawtchouk polynomials $K_\alpha(x,N)$ are polynomials orthogonal with respect to a binomial distribution; for convenience we consider Bernoulli random variables taking values in ± 1 with equal probability. They may be defined by the generating function

$$G(v)=\sum_{\alpha=0}^{N} v^\alpha K_\alpha(x,N)=(1+v)^{(N+x)/2}(1-v)^{(N-x)/2} \tag{1.1}$$

The generating function

$$(1-v)^{y-a}(1-(1-R)v)^{-y}=\sum_{n=0}^{\infty}\frac{v^n}{n!}(a)_n\,{}_2F_1\left(\begin{matrix}-n,y\\a\end{matrix}\,\middle|\,R\right) \tag{1.2}$$

with $(a)_n=\Gamma(a+n)/\Gamma(a)$, yields the identification

$$K_\alpha(x,N)=\binom{N}{\alpha}\,{}_2F_1\left(\begin{matrix}-\alpha,(x-N)/2\\-N\end{matrix}\,\middle|\,2\right) \tag{1.3}$$

A theory for multivariate 'Bernoulli-type' polynomials is presented in [4]. Here we take a different look at generalized Krawtchouk polynomials.

One class arises from a random walk on the lattice $\mathbf{N}[1,\zeta,\zeta^2,\ldots,\zeta^{d-1}]$ where $\zeta=e^{2\pi i/d}$. There are evident connections with the finite Fourier transform, see, e.g., [2], but a deeper study is yet to come. Some observations of interest will be made in the discussion in section V. It is quite likely that the classes of polynomials discussed here will prove useful in image processing, among other possible applications.

In section II we give the probabilistic approach to the binomial case. Section III provides a general approach for finite probability distributions. In section IV some related

[1] Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901 USA

[2] CRIN, INRIA-Lorraine, Université de Nancy I, B.P. 239, 54506-Vandoeuvre-lès-Nancy, France

limit theorems are presented. Specialization to and special adjustments for the cyclic case comprise section V which completes the study.

Note. 1. Throughout, N will denote 'time' — the number of steps taken in the random walk.

2. We use $\langle\,\rangle$ to denote expected value, e.g. $\langle X\rangle$.
3. For brevity, we call Krawtchouk polynomials simply K-polynomials.

II. Binomial Case

Consider a random walk on $\mathbf{Z}$, with equiprobable increments ± 1. We write X_j, $1 \le j \le N$, for the corresponding Bernoulli variables.

The generating function

$$G(v) = \prod(1 + vX_j) = (1+v)^{(N+x)/2}(1-v)^{(N-x)/2} \tag{2.1}$$

where $x = \sum X_j$ is the position after N steps. As noted above, $G(v) = \sum v^\alpha K_\alpha(x, N)$, with Krawtchouk polynomials K_α. One can take the viewpoint of 'quantum probability' and consider the X_j as the spectrum of an operator A, an $N \times N$ matrix. Then the condition on A is that $\lambda \in \operatorname{spectrum}(A) \Rightarrow \lambda \in \{-1, 1\}$. And the variables of interest are the multiplicities. We have

$$G(v) = \det(1 + vA)\;.$$

And the variables x, N are simply $\operatorname{tr} A$, $\operatorname{tr} A^2$. We thus have the principal observation that since the X_j take two values, two variables suffice to specify the K_α, which are seen to be elementary symmetric functions in the X_j. The variables x, N are the corresponding power sums: $x = \sum X_j$, $N = \sum X_j^2$.

The generating function $G(v)$ as a function of N is readily seen to be a martingale, as the X_j are independent with mean zero. It is, in fact, the prototype exponential martingale (see [1], e.g.). The calculation

$$\langle G(v)\, G(w)\rangle = \prod\langle 1 + (v+w)X_j + vwX_j^2\rangle = (1+vw)^N \tag{2.2}$$

exhibits the orthogonality of the K_α nicely. So the K_α are important mainly for these two features:

1. They are the iterated integrals (sums) of the Bernoulli process.
2. They are orthogonal polynomials with respect to the binomial distribution.

III. Finite Probability Distributions

The probabilistic approach may be carried out for general finite probability spaces. Each increment X takes d possible values $\{\xi_0, \ldots, \xi_\delta\}$ with $P(X = \xi_j) = p_j$, $0 \le j \le \delta$, where throughout we will use the convention $\delta = d - 1$. Denote the mean and variance by μ and σ^2 as usual.

Take N independent copies of X: X_j, $1 \le j \le N$. Define the martingale

$$G(v) = \prod_{j=1}^{N}(1 + v(X_j - \mu)) \tag{3.1}$$

We may switch to the multiplicities as variables. So set

$$n_j = \sum_{k=1}^{N} \mathbf{1}_{\{X_k=\xi_j\}} \tag{3.2}$$

the number of times the value ξ_j is taken. We have

$$G(v) = \prod_{j=0}^{\delta} (1 + v(\xi_j - \mu))^{n_j} = \sum_{\alpha=0}^{N} v^\alpha K_\alpha(n_0, \dots, n_\delta) \tag{3.3}$$

this last defining our generalized K-polynomials. One quickly gets

3.1 Proposition. *Denoting the multi-index* $\mathbf{n} = (n_0, \dots, n_\delta)$ *and by* $\mathbf{e_j}$ *the standard basis on* $\mathbf{Z}^d$, *K-polynomials satisfy the recurrence*

$$K_\alpha(\mathbf{n} + \mathbf{e_j}) = K_\alpha(\mathbf{n}) + (\xi_j - \mu)K_{\alpha-1}(\mathbf{n}) \tag{3.4}$$

We also find by binomial expansion

3.2 Proposition.

$$K_\alpha(n_0, \dots, n_\delta) = \sum_{|\mathbf{k}|=\alpha} \prod \binom{n_j}{k_j} (\xi_j - \mu)^{k_j} \tag{3.5}$$

where $|\mathbf{k}| = k_0 + \cdots + k_\delta$.

An interesting connection with the multivariate theory comes in when we consider equations (1.2), (1.3). The Lauricella polynomials F_B are defined by

$$F_B \left(\begin{matrix} -\mathbf{r}, \mathbf{b} \\ t \end{matrix} \middle| \, \mathbf{s} \right) = \sum_{\mathbf{k} \in \mathbf{N}^\delta} \frac{(-\mathbf{r})_\mathbf{k} (\mathbf{b})_\mathbf{k}}{(t)_{|\mathbf{k}|} \mathbf{k}!} \mathbf{s}^\mathbf{k} \tag{3.6}$$

with, e.g., $\mathbf{r} = (r_1, \dots, r_\delta)$, $(\mathbf{r})_\mathbf{k} = (r_1)_{k_1} (r_2)_{k_2} \cdots (r_\delta)_{k_\delta}$ for multi-index $\mathbf{k}$, also $\mathbf{s}^\mathbf{k} = s_1^{k_1} \cdots s_\delta^{k_\delta}$, and $\mathbf{k}! = k_1! \cdots k_\delta!$. Note that t is a single variable. The generating function of interest here is

$$(1 - \sum v_i)^{\sum b_j - t} \prod_j (1 - \sum v_i + s_j v_j)^{-b_j} = \sum \frac{\mathbf{v}^\mathbf{r} (t)_{|\mathbf{r}|}}{\mathbf{r}!} F_B \left(\begin{matrix} -\mathbf{r}, \mathbf{b} \\ t \end{matrix} \middle| \, \mathbf{s} \right) \tag{3.7}$$

a multivariate version of (1.2).

3.3 Proposition. *If* $\xi_0 = 0$, *then*

$$K_\alpha(\mathbf{n}) = (-N)_\alpha \sum_{|\mathbf{r}|=\alpha} \frac{\prod (p_j \xi_j)^{r_j}}{\mathbf{r}!} F_B \left(\begin{matrix} -\mathbf{r}, -\mathbf{n} \\ -N \end{matrix} \middle| \, \frac{1}{p_1}, \dots, \frac{1}{p_\delta} \right) \tag{3.8}$$

Proof: Let $v_j = vp_j\xi_j$, $b_j = -n_j$, $t = -N$, $s_j = p_j^{-1}$ in (3.7), for $1 \le j \le \delta$. Note that $\sum v_j = v\mu$, $\sum b_j - t = N - (n_1 + \cdots + n_\delta) = n_0$. ∎

As for the binomial case, we may use the power sum variables, e.g. for centered increments,

$$x_i = \sum_{j=1}^{N} (X_j - \mu)^i = \sum_{j=0}^{\delta} n_j(\xi_j - \mu)^i \tag{3.9}$$

to express the functions K_α. (This will be useful in the cyclic case, section V below.)

Orthogonality follows similar to the binomial case.

3.4 Proposition. *The K-polynomials $K_\alpha(n_0, \ldots, n_\delta)$ are orthogonal with respect to the induced multinomial distribution. In fact,*

$$\langle K_\alpha K_\beta \rangle = \delta_{\alpha\beta}\, \sigma^{2\alpha} \binom{N}{\alpha} \tag{3.10}$$

Proof:

$$\begin{aligned} \langle G(v)\, G(w) \rangle &= \sum \binom{N}{n_0, \ldots, n_\delta} p_0^{n_0} \cdots p_\delta^{n_\delta} \prod \left(1 + (v+w)(\xi_j - \mu) + vw(\xi_j - \mu)^2\right)^{n_j} \\ &= \left(\sum \left(p_j + (v+w)p_j(\xi_j - \mu) + vwp_j(\xi_j - \mu)^2\right)\right)^N \end{aligned} \tag{3.11}$$

Thus, $\langle G(v)\, G(w) \rangle = (1 + vw\sigma^2)^N$. This shows orthogonality and yields the squared norms as well. ∎

IV. Limit Theorems

Limit theorems for products $\prod(1 + vX_j)$ are given by [1] for X_j discrete increments of a process. In cases discussed there, the limits yield iterated integrals of the process, as may be expected (see [3, Ch. VII]).

Here (and in section V) we look at limit theorems based on the power sum functions. First, define the normalized values $\beta_j = (\xi_j - \mu)/\sigma$. Then the central limit theorem says that

$$\frac{1}{\sigma\sqrt{N}} \sum_{j=1}^{N} (X_j - \mu) = \frac{1}{\sqrt{N}} \sum_{j=0}^{\delta} n_j \beta_j \tag{4.1}$$

converges in distribution to a standard Gaussian as $N \to \infty$. We look at the power sums of the normalized variables:

$$\sum \left(\frac{X_j - \mu}{\sigma\sqrt{N}}\right)^i = N^{-i/2} \sum n_j \beta_j^i \tag{4.2}$$

4.1 Theorem. *Let Y_j satisfy $n_j = p_jN + Y_j\sqrt{N}$. Then, as $N \to \infty$, $\sum_{j=0}^{\delta} Y_j\beta_j$ converges to a standard Gaussian.*

Proof: We have, recalling that $\sum p_j\beta_j = 0$,

$$\frac{1}{\sqrt{N}} \sum n_j\beta_j = \frac{1}{\sqrt{N}} \sum (p_j + Y_j\sqrt{N})\beta_j = \sum Y_j\beta_j \tag{4.3}$$

and the result follows from the central limit theorem. ∎

4.2 Theorem. *The martingale for the normalized variables*

$$G(v) = \prod_{j=1}^{N} \left(1 + v\,\frac{X_j - \mu}{\sigma\sqrt{N}}\right) \tag{4.4}$$

converges to the Brownian martingale at time 1,

$$G(v) \to e^{vY - v^2/2} \tag{4.5}$$

where Y is the standard Gaussian denoting the limit of the normalized sums.

Proof: Write

$$\begin{aligned}\prod_{j=1}^{N} \left(1 + v\,\frac{X_j - \mu}{\sigma\sqrt{N}}\right) &= \prod_{j=0}^{\delta} (1 + v\beta_j/\sqrt{N})^{n_j} \\ &= \prod \exp(n_j \,\log(1 + v\beta_j/\sqrt{N})) \\ &= \exp\left(\sum_{i=1}^{\infty} (-1)^{i-1}\,\frac{v^i}{i}\,S_i\right)\end{aligned} \tag{4.6}$$

where S_i are the scaled power sums $N^{-i/2}\sum n_j\beta_j^i$. Theorem (4.1) deals with $i = 1$. For $i = 2$ we have, with $n_j = p_jN + Y_j\sqrt{N}$,

$$N^{-1}\sum n_j\beta_j^2 = \sum p_j\beta_j^2 + N^{-1/2}\sum Y_j\beta_j^2 = 1 + o(1), \tag{4.7}$$

since the β's are scaled to variance 1. For $i \geq 3$,we have

$$N^{1-i/2}\sum p_j\beta_j^i + N^{(1-i)/2}\sum Y_j\beta_j^i \to 0 \tag{4.8}$$

as $N \to \infty$. Denoting the limit of $\sum Y_j\beta_j$ by Y the result follows. ∎

We conclude that the corresponding K-polynomials converge to Hermite polynomials in the variable Y.

V. Cyclic K-polynomials

Now we consider a random walk in $\mathbf{C}$, with increments X taking values in the dth roots of unity: $1, \zeta, \zeta^2, \ldots, \zeta^\delta$, with $\zeta = e^{2\pi i/d}$, and $\delta = d - 1$ as above. For simplicity we discuss the isotropic case, all values occurring with equal probability $1/d$.

The distribution of x = position after N steps is an interesting problem in number theory. Namely, to count how many 'paths' lead to the same position x. Here we make some remarks concerning this problem. Again, let n_j denote the multiplicities of occurrences of the ζ^j. We have the power sum variables:

$$\begin{aligned}
x_0 &= n_0 + n_1 + \ldots = N \\
x_1 &= n_0 + n_1\zeta + \ldots = \sum n_j\zeta^j = x \\
x_2 &= n_0 + n_1\zeta^2 + \ldots = \sum n_j\zeta^{2j} \\
&\vdots \\
x_k &= n_0 + n_1\zeta^k + \ldots = \sum n_j\zeta^{kj}
\end{aligned} \tag{5.1}$$

Notice that the x_k are the finite Fourier transform of the variables $(n_0, \ldots, n_\delta)$. So here we can conveniently go back and forth between the two sets of variables.

From algebraic number theory, see [5, p.265ff.], it is known that for d = a prime power, $\phi(d)$ denoting Euler's function, the powers $1, \zeta, \zeta^2, \ldots, \zeta^{\phi(d)-1}$ form a basis for the $\mathbf{Z}$-module spanned by the ζ^j, $0 \le j \le \delta$. I.e. each sum of the form $\sum n_j\zeta^j$ has a *unique expression* as a sum involving ζ^j for $j < \phi(d)$. The problem is that we have to count how many general sums, involving *all* the ζ^j, reduce to the same canonical form, involving only ζ^j, $j < \phi(d)$. From the point of view of Fourier analysis and probablility theory this has the flavor of finite prediction theory and requires a separate study.

Here we note that if d is a prime, then $1, \zeta, \zeta^2, \ldots, \zeta^{d-2}$ form a $\mathbf{Z}$-basis. And we have the elementary relation $\sum \zeta^j = 0$. Thus,

5.1 Proposition. *Let d be prime. Given $x_0 = N$ and $x_1 = \sum n_j\zeta^j$, the n_j are uniquely determined.*

Proof: In x_1, substitute $\zeta^\delta = -1 - \zeta - \cdots - \zeta^{\delta-1}$ yielding

$$x_1 = (n_0 - n_\delta) + \zeta(n_1 - n_\delta) + \cdots + \zeta^{\delta-1}(n_{\delta-1} - n_\delta) \tag{5.2}$$

By the result quoted above, the numbers $n_j - n_\delta$, $0 \le j < \delta$ are uniquely determined. Thus, their sum, call it $\nu = n_0 + \cdots + n_{\delta-1} - \delta n_\delta$ is known, thence $dn_\delta = n_0 + \cdots + n_\delta - \nu = N - \nu$ is determined and hence the n_j, $0 \le j < \delta$ as well. ∎

Now we look at the K-polynomials for the cyclic case. The generating function is

$$G(v) = \prod_{j=0}^{\delta}(1 + v\zeta^j)^{n_j} = \sum_{\alpha=0}^{N} v^\alpha K_\alpha \tag{5.3}$$

They satisfy the recurrence, cf. Proposition (3.4),

$$K_\alpha(\mathbf{n} + \mathbf{e_j}) = K_\alpha(\mathbf{n}) + \zeta^j K_{\alpha-1}(\mathbf{n}) \tag{5.4}$$

The binomial expansion gives

$$K_\alpha(n_0,\ldots,n_\delta) = \sum_{|\mathbf{k}|=\alpha} \prod \binom{n_j}{k_j} \zeta^{jk_j} \tag{5.5}$$

The F_B representation is found as follows. In the generating function (3.7), let $v_j = v\zeta^j$, $1 \le j \le \delta$. With $b_j = -n_j$, $1 \le j \le \delta$, $t = -N$, we have

$$\sum_{j=1}^{\delta} v_j = -v \qquad \text{and} \qquad \sum_{j=1}^{\delta} b_j = n_0 - N \tag{5.6}$$

With $s_j = 1 - \bar{\zeta}^j$, writing $\bar{\zeta} = \zeta^{-1}$ as usual,

$$1 - \sum v_j + s_j v_j = 1 + v + (1 - \bar{\zeta}^j) v\zeta^j = 1 + v\zeta^j \tag{5.7}$$

for $1 \le j \le< d$. Thus, as in Proposition (3.3),

$$K_\alpha = (-N)_\alpha \sum_{|\mathbf{r}|=\alpha} \frac{\zeta^{\sum jr_j}}{\mathbf{r}!} F_B\left(\begin{matrix} -\mathbf{r}, -\mathbf{n} \\ -N \end{matrix} \,\middle|\, 1-\bar{\zeta},\ldots,1-\bar{\zeta}^\delta \right) \tag{5.8}$$

Finally we have the limit result

5.2 Theorem. *In $G(v)$, scale $v \to vn^{-1/d}$, $x_0 \to x_0 n$ and, for $1 \le k \le \delta$, $x_k \to x_k n^{(d,k)/d}$, where (d,k) = greatest common divisor of d and k. Then*

$$G(v) \to \exp\left(\sum_{k|d} (-1)^{k-1} v^k x_k / k \right) \tag{5.9}$$

as $n \to \infty$.

Proof: Fourier inversion says that $n_j = d^{-1} \sum x_k \bar{\zeta}^{jk}$. I.e.

$$G(v) = \prod_j (1 + v\zeta^j)^{\sum \bar{\zeta}^{jk} x_k / d} = \prod_{j,k} (1 + v\zeta^j)^{\bar{\zeta}^{jk} x_k / d} \tag{5.10}$$

For fixed k, let $(d,k) = g$. Then, for $0 \le j < d$, we can write $j = ld/g + r$, with $r < d/g$. Thus,

$$jk \equiv l(dk/g) + rk \equiv rk \pmod{d} \tag{5.11}$$

So

$$\begin{aligned} \prod_{j=0}^{\delta} (1 + v\zeta^j)^{\bar{\zeta}^{jk} x_k / d} &= \prod_{r=0}^{(d/g)-1} \prod_{l=0}^{g-1} \left(1 + v\zeta^{r + ld/g}\right)^{\bar{\zeta}^{rk} x_k / d} \\ &= \prod_{r=0}^{(d/g)-1} \left(1 + (v\zeta^r)^g (-1)^{g-1}\right)^{\bar{\zeta}^{rk} x_k / d} \end{aligned} \tag{5.12}$$

since

$$\prod_{l=0}^{g-1}(1+y\zeta^{ld/g}) = 1+y^g\zeta^{(d/g)(g(g-1)/2)} = 1+y^g(-1)^{g-1} \tag{5.13}$$

With the scalings indicated, $x_k \to x_k n^{g/d}$, $v \to vn^{-1/d}$, each factor of the product in (5.12) will converge to expressions of the form

$$\exp(v^g\zeta^{rg}(-1)^{g-1}\bar{\zeta}^{rk}x_k/d) \tag{5.14}$$

Taking the product over $0 \le r < d/g$, we have

$$\sum_{r=0}^{(d/g)-1} \zeta^{r(g-k)} = \frac{\bar{\zeta}^{kd/g}-1}{\zeta^{g-k}-1} = 0 \tag{5.15}$$

if $g \neq k$, and $= d/g$ for $g = k$, i.e. when $k \mid d$. ∎

It may be of interest to compare with [6].

VI. Concluding Remarks

It looks like an analysis for general abelian groups based on similar ideas should be feasible. It is interesting to recall that the K-polynomials in the binomial case provide representation spaces for $su(2)$ and are thus connected with spherical harmonics. Similar connections/interpretations for general K-polynomials would be quite interesting to find.

Acknowledgement. We are grateful to Jerzy Kocik for stimulating conversations.

References

1. F. Avram: *Weak convergence of the variations, iterated integrals and Doleans-Dade exponentials of sequences of semimartingales*, Ann. Prob., **16**, 1988, 246–250.

2. Ph. Feinsilver and R. Schott: *Orthogonal polynomial expansions via Fourier transform*, preprint, 1990.

3. Ph. Feinsilver: *Special functions, probability semigroups, and Hamiltonian flows*, Lecture Notes **696**, Springer-Verlag, 1978.

4. Ph. Feinsilver: *Bernoulli systems in several variables*, Lecture Notes **1064**, 1984, 86–98.

5. P. Ribenboim: *Algebraic numbers*, Wiley-Interscience, 1972.

6. W. von Waldenfels: *Noncommutative central limit theorems*, Lecture Notes **1210**, 1986, 174–202.

ISSN 0249-6399