Cauchy-Stielties families with polynomial variance functions

Wlodek Bryc
Joint work with Wojtek Młotkowski (Wrocław, Poland) and Raouf Fakhfakh (Gurayat, Saudi Arabia)

October 21, 2019

Abstract

This is a third installment of the series of talks April 3, 2006 and Oct 27, 2011. This talk is based on a forthcoming paper http://arxiv.org/abs/1708.05343 with Raouf Fakhfakh and Wojciech Mlotkowski. The topic of this paper are properties of variance functions of Cauchy-Stieltjes Kernel families generated by a compactly supported (standardized) probability measure. After a brief introduction, I will describe some algebraic operations that can be used to construct additional variance functions from known variance functions. I will describe all quadratic and all cubic variance functions. I will also show how Cauchy-Stieltjes Kernel families with polynomial variance functions are related to generalized orthogonality of some families of polynomials.

Exponential family

Lehmann (1959), Barndorff-Nielsen (1978), Morris (1982), Letac, Mora (1990).
For a probability measure μ on $\mathbb{R}$ we define its *Laplace transform:*

$$L_\mu(\theta) := \int \exp(\theta x)\mu(dx),$$

cumulant function:

$$k_\mu(\theta) := \log L_\mu(\theta),$$

and define set

$$\Theta(\mu) := \mathrm{interior}\{\theta : L_\mu(\theta) < \infty\}.$$

We assume that μ is not concentrated at one point and that $\Theta(\mu)$ is nonempty.

For $\theta \in \Theta(\mu)$ we define family of probability distributions

$$P_\theta(dx) := \exp(\theta x - k_\mu(\theta))\mu(dx)$$

we call this the *natural exponential family* (NEF) generated by μ.
The mean function

$$\theta \mapsto m(\theta) = \int x P_\theta(dx)$$

is strictly increasing on $\Theta(\mu)$, so there is an inverse function

$$\psi : M_\mu \ni m \mapsto \theta \in \Theta(\mu).$$

The *variance function* of the NEF generated by μ is defined by

$$V_\mu(m) := \int (x-m)^2 P_{\psi(m)}(dx), \quad m \in M_\mu.$$

Question: describe variance functions

Morris (1982) characterized those polynomials of order at most 2 which are variance functions. They correspond to Gaussian, Poisson, binomial, negative binomial, gamma and hyperbolic cosine distributions.

Letac and Mora (1990) characterized variance functions which are polynomials of degree at most 3. Here in addition they obtained: Abel, Takacs, strict arcsine, large arcsine, Ressel, inverse Gaussian distribution.

Ismail, M. E. H., and May, C. P.
On a family of approximation operators.
J. Math. Anal. Appl. 63, 2 (1978), 446–462.

Morris, C. N.
Natural exponential families with quadratic variance functions.
Ann. Statist. 10, 1 (1982), 65–80.

Letac, G., and Mora, M.
Natural real exponential families with cubic variance functions.
Ann. Statist. 18, 1 (1990), 1–37.

Kernel family (Wesołowski, unpublished manuscript, 1999)

Kernel family generated by $k(x, \theta)$ and ν

For $\theta \in \Theta(\nu)$,

$$P_\theta(dx) = Z_\theta^{-1} k(x, \theta)\nu(dx), \qquad \text{where } Z_\theta = \int k(x, \theta)\nu(dx)$$

Exponential family:

$$k(x, \theta) = \exp(x\theta).$$

Cauchy-Stielties Kernel (CSK) family:

$$k(x, \theta) = \frac{1}{1 - x\theta}$$

Cauchy Stielties family generated by a distribution ν:

$$P_\theta(dx) := \frac{1}{Z_\theta(1-\theta x)}\nu(dx), \qquad \theta \in \Theta, \tag{1}$$

where

$$Z_\theta = \int \frac{\nu(dx)}{1-\theta x}.$$

We assume that ν is compactly supported in $[a, b]$, with mean 0 and nondegenerate, with variance, say 1. Then $a < 0 < b$ and $\Theta = (1/a, 1/b)$.

(Somewhat more generally, if ν has mean m_0 then one can find Θ that works, but it is quite convenient to work with the explicit Θ for $m_0 = 0$.)

Parametrization by the mean

The mean $m(\theta) := \int x P_\theta(dx)$ of P_θ is

$$m(\theta) = \frac{Z_\theta - 1}{\theta Z_\theta}.$$

The function $\theta \mapsto m(\theta)$ is strictly increasing on Θ.
Let $\psi(m)$ denote the inverse function, on a neighborhood of $\int x\nu(dx) = 0$.

The *variance function* of the Cauchy-Stielties kernel (CSK) family (1) generated by ν is defined by:

$$V(m) := \int (x - m)^2 P_{\psi(m)}(dx).$$

Theorem (Theorem (B-2009))

For m in some neighborhood of $m_0 = 0$ the variance function $V(m)$ is analytic, strictly positive and

$$P_{\psi(m)}(dx) = \frac{V(m)}{V(m) + m(m-x)}\nu(dx).$$

Conversely, if a function $V(m)$ is analytic, strictly positive in a neighborhood of 0 and if ν is a probability distribution, with mean $m_0 = 0$, such that for every m in a neighborhood of 0

$$\frac{V(m)}{V(m) + m(m-x)}\nu(dx)$$

is a probability measure, then ν is compactly supported, non degenerate and determined uniquely.

So we have correspondence:

$$\nu \leftrightarrow V(m),$$

where ν is a compactly supported nondegenerate probability distribution with mean 0 and $V(m)$ is an analytic function in a neighborhood of 0, with $V(0) > 0$.
If $\nu_1 \neq \nu_2$, with the same mean, then the corresponding variance functions $V_1(m)$, $V_2(m)$ are different.

Question:

Which functions $V(m)$ are variance functions?

New results from the paper with Fakhfakh and Mlotkowski

Notation: $\mathcal{V}$

The class of all variance functions corresponding to those compactly supported probability distributions ν which have mean 0 and variance 1.

So that if $V \in \mathcal{V}$ then V is an analytic function on a neighborhood of 0 and $V(0) = 1$.
Therefore the corresponding probability distribution ν has mean 0 and variance 1. We will sometimes write $V = V_\nu$.

Notation: $\mathcal{V}_\infty$

The class of those $V \in \mathcal{V}$ that the function $m \mapsto V(cm)$ belongs to $\mathcal{V}$ for all real c.

Quadratic variance functions I

Theorem (B-Ismail-2005)

Suppose $V(m) = 1 + am + bm^2$. *Then*

1. $V \in \mathcal{V}$ *iff* $b \geq -1$
2. $V \in \mathcal{V}_\infty$ *iff* $b \geq 0$

Idea of proof

Expand the density of $P_{\psi(m)}(dx) = \frac{V(m)}{V(m)+m(m-x)}\nu(dx)$ into the power series in m near $m = 0$:

$$\frac{V(m)}{V(m) + m(m-x)} = \sum_{k=0}^{\infty} p_k(x) m^k$$

Quadratic variance functions II

The coefficients $p_k(x)$ turn out to be monic polynomials given by the recursion:

$$xp_n(x) = p_{n-1} + \sum_{k=0}^{n} \frac{V^{(k)}(0)}{k!} p_{n+1-k}(x)$$

For $V(m) = 1 + am + bm^2$, this becomes a "constant" three step recursion

$$xp_n = p_{n+1}(x) + ap_n(x) + (b+1)p_{n-1}(x).$$

Since $P_{\psi(m)}(dx)$ is a probability measure, we have

$$1 = \int_{\mathbb{R}} P_{\psi(m)}(dx) = \int_{\mathbb{R}} \sum_{k=0}^{\infty} p_k(x) m^k \nu(dx) = \sum_{k=0}^{\infty} m^k \int_{\mathbb{R}} p_k(x)\nu(dx)$$

$$= 1 + \sum_{k=1}^{\infty} m^k \int_{\mathbb{R}} p_k(x)\nu(dx)$$

Quadratic variance functions III

Thus $\int p_k(x)\nu(dx) = 0$ for $k = 1, 2, \ldots$. Together with the three-step recursion

$$xp_n = p_{n+1}(x) + ap_n(x) + (b+1)p_{n-1}(x), \qquad (*)$$

this implies that $\{p_n : n \geq 0\}$ are orthogonal polynomials.
So by Favard's theorem, the last coefficient in the three step recursion (*) must be non-negative: $1 + b \geq 0$.
(Furthermore, $b = -1$ corresponds to discrete $\nu = (1-p)\delta_a + p\delta_b$.)

The main theorems I

Here we write $V(m)$ for the function not for the value!

Theorem

1. If $V(m) \in \mathcal{V}$ and $|c| \leq 1$ then $V(cm) \in \mathcal{V}$;
2. If $V(m) \in \mathcal{V}$ and $a \in \mathbb{R}$ then $V(m) + am \in \mathcal{V}$;
3. If $V(m) \in \mathcal{V}$ then $V(m) + m^2 \in \mathcal{V}_\infty$;
4. If $V(m) \in \mathcal{V}_\infty$ then $V(m) - m^2 \in \mathcal{V}$;
5. If $V(m) \in \mathcal{V}_\infty$ and $a \in \mathbb{R}$ then $V(m) + am \in \mathcal{V}_\infty$;
6. If $V(m) \in \mathcal{V}_\infty$ and $c \geq 1$ then $cV(m) - c + 1 \in \mathcal{V}_\infty$;
7. If $V_1(m), V_2(m) \in \mathcal{V}_\infty$ then $V_1(m) + V_2(m) - 1 \in \mathcal{V}_\infty$.

The main theorems II

Corollary (from 3, 4)

The map $V(m) \mapsto V(m) + m^2$ is a bijection of $\mathcal{V}$ onto $\mathcal{V}_\infty$.

It turns our this is an old familiar Berkovici-Pata bijection which can be described by a composition of free-boolean and free-additive convolution powers: $\nu \mapsto (\nu^{\boxplus 2})^{\uplus 1/2}$. The later is $\boxplus$-infinitely divisible by a result of Belinschi-Nica (2008), or because $(\nu^{\boxplus 2})^{\uplus 1/2} = \lim_{\alpha\to\infty}(\nu^{\uplus 1/\alpha})^{\boxplus \alpha}$.

From formulas in a recent manuscript Raouf Fakfhfakh (2019) it follows that $V_{(\nu^{\boxplus 2})^{\uplus 1/2}}(m) = V_\nu(m) + m^2$.

Theorem

Consider cubic function $V(m) = 1 + am + bm^2 + cm^3$.

- *If $(b+1)^3 \geq 27c^2$ then $V(m) \in \mathcal{V}$.*
- *$V(m) \in \mathcal{V}_\infty$ if and only if $b^3 \geq 27c^2$.*

Generalized orthogonality I

There is a substantial literature on generalized orthogonality and finite-step recursions for polynomials. We introduce the following generalized orthogonality condition.

Definition

Fix $d \in \mathbb{N}$ and a non-degenerate probability measure ν with moments of all orders. We say that polynomials $\{p_n\}$ are $(\nu; d)$-*orthogonal* if

1. $\int p_n(x)\nu(dx) = 0$ for all $n \geq 1$;
2. $\int p_n(x)p_k(x)\nu(dx) = 0$ for all $n \geq 2 + (k-1)d, \quad k = 1, 2 \dots.$

Example ($d = 1$ is the usual orthogonality when ν has ∞ support)

First condition: $p_0 \perp p_1, p_2, \dots.$
Second condition: $p_n \perp p_k$ for $n \geq 2 + (k-1)d = k+1$, so $p_k \perp p_{k+1}, p_{k+2}, \dots$ for $k = 1, 2, \dots.$

Generalized orthogonality II

Example ($d = 2$)

For $d = 2$, the conditions are $\int p_n(x)p_0(x)\nu(dx) = 0$ for $n \geq 1$ and $\int p_n(x)p_k(x)\nu(dx) = 0$ for all $n \geq 2 + (k-1)d = 2k$. So $p_0 \perp p_1, p_2, \dots$ and $p_1 \perp p_2, p_3, \dots$. But $p_2 \perp p_4, p_5, \dots$ so there is no restriction on p_2, p_3. Similarly, $p_3 \perp p_6, p_7, \dots$ but there are no restrictions on p_3, p_4 and p_3, p_5.

Somewhat surprisingly, in the setting of Cauchy-Stieltjes kernel families, all the d-orthogonality conditions are implied by just two conditions

1. $\int P_n(x)P_0(x)\nu(dx) = 0$ for $n \geq 1$,
2. $\int P_n(x)P_2(x)\nu(dx) = 0$ for $n \geq 2 + d$,

Generalized orthogonality III

Question

Does $(\nu; d)$-orthogonality imply a $(d+2)$-step recursion?

The answer to the converse question seems to be quite complicated. Case $d = 2$ with a perturbation of a constant recursion will be mentioned below. The literature on $d+1$-step recursions (Maroni(1989), Iseghem(1987)) assumes non-degeneracy condition that is not satisfied on our case, and the conclusion is the existence of a d-tuple of functionals that need not to be positive.

Generalized orthogonality IV

Theorem

Suppose that $V \in \mathcal{V}$ corresponds to a probability measure ν. Consider

$$\frac{V(m)}{V(m)+m(m-x)} = \sum_{n=0}^{\infty} p_n(x) m^n. \tag{2}$$

Then the following statements are equivalent:

1. *$V(m)$ is a polynomial of degree at most $d+1$;*
2. *Polynomials $\{p_n\}$ satisfy $(d+2)$-step recursion*

$$x p_n(x) = p_{n+1}(x) + \sum_{k=1}^{d+1} b_k p_{n+1-k}(x), \quad n \geq 1 \tag{3}$$

   with initial conditions $p_0(x)=1$, $p_1(x)=x$, and $p_k \equiv 0$ for $k<0$.

3. *Polynomials $\{p_n(x)\}$ are $(\nu; d)$-orthogonal.*
4. *Polynomials $\{p_n\}$ satisfy conditions in definition of $(\nu; d)$-orthogonality for $k=1,2$ only.*

Theorem

Consider polynomials $\{p_n(x)\}$ given by the (eventually constant) 4-step recursion:

$$\begin{aligned} xp_1(x) &= p_2(x) + ap_1(x) + p_0(x), \\ xp_2(x) &= p_3(x) + ap_2(x) + bp_1(x), \\ xp_n(x) &= p_{n+1}(x) + ap_n(x) + bp_{n-1}(x) + cp_{n-2}(x), \;\; n \geq 3, \end{aligned}$$

with $p_0(x) = 1, p_1(x) = x$. Then the following conditions are equivalent.

1. $b^3 \geq 27c^2$.
2. *Polynomial $\{p_n\}$ are $(\nu; 2)$-orthogonal for some probability measure ν (which then necessarily has mean 0, variance 1, and compact support).*

Define $\Gamma_1(p)=\int_{\mathbb{R}} p(x)\nu(dx)$ and $\Gamma_2(p)=\int_{\mathbb{R}} xp(x)\nu(dx)$.

If polynomials $p_n(x)$ are $(\nu; d)$-orthogonal with $d=2$. Then $\Gamma_\alpha(p_np_k)=0$ if $n\geq 2k+\alpha$ for $\alpha=1,\dots,d$; however the non-singularity part of the requirement of $d=2$-orthogonality that $\Gamma_\alpha(p_np_k)\neq 0$ if $n=2k+\alpha-1$ fails.

Similarly, for arbitrary d, define $\Gamma_j(p)=\int x^{j-1}p(x)\nu(dx)$. Then $\Gamma_\alpha(p_np_k)=0$ if $n\geq dk+\alpha$, but $(\nu; d)$-orthogonality yields more zeros than what is allowed under d-orthogonality.

The End

Thank you!

Wesołowski, J.
Kernel families.
Unpublished manuscript, 1999.

W MINI NR FAKSU : 6257460 13 PAZ. 2005 12:35 STR. 2

Jacek Wesołowski 1999

KERNEL FAMILIES

Let μ be a positive measure on $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, and $k : \mathbf{R}^2 \to \mathbf{R}_+$ a measurable function, wkich will play the role of the kernel. Denote

$$G_\mu(\theta) = \int_{\mathbf{R}} k(x,\theta)\, \mu(dx),$$

for such θ that $\int_{\mathbf{R}} k(x,\theta)\, \mu dx < \infty$. Denote additionally

$$\Theta = Int\{\theta \in \mathbf{R} : \int_{\mathbf{R}} k(x,\theta)\, \mu(dx) < \infty\}.$$

The natural exponential family is an example of a kernel family with $k(x,\theta) = e^{\theta x}$. Consider two new special cases:

GAMMA KERNEL FAMILY

Belinschi, S. T. and Nica, A. (2008).
On a remarkable semigroup of homomorphisms with respect to free multiplicative convolution.
Indiana Univ. Math. J., 57(4):1679–1713.

Bryc, W., and Ismail, M.
Approximation operators, exponential, and q-exponential families.
arxiv.org/abs/math.ST/0512224, 2005.

Bryc, W.
Free exponential families as kernel families.
Demonstratio Mathematica XLII, 3 (2009), 657–672.

Bryc, W., and Hassairi, A.
One-sided Cauchy-Stieltjes kernel families.
Journal of Theoretical Probability 24, 2 (2011), 577–594.

Bryc, W., Fakhfakh, R., and Hassairi, A.
On Cauchy-Stieltjes kernel families.
Journal of Mutlivariate Analysis 124 (2014), 296–312.

BRYC, W., FAKHFAKH, R., AND MŁOTKOWSKI, W.
Cauchy-Stieltjes families with polynomial variance functions and generalized orthogonality.
Probability and Mathematical Statistics (Wroclaw) (to appear).

FAKHFAKH, R.
Characterization of quadratic CauchyStieltjes kernel families based on the orthogonality of polynomials
Journal of Mathematical Analysis and Applications 459 (2018), 577–589.

FAKHFAKH, R.
Variance function of boolean additive convolution.
umpublished manuscript, 2019.

HASSAIRI, A., ZARAI, M., ET AL.
Characterization of the cubic exponential families by orthogonality of polynomials.
The Annals of Probability 32, 3B (2004), 2463–2476.

Letac, G.
Lectures on natural exponential families and their variance functions, vol. 50 of *Monografías de Matemática [Mathematical Monographs]*. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1992.

Letac, G., and Mora, M.
Natural real exponential families with cubic variance functions. *Ann. Statist. 18*, 1 (1990), 1–37.

Morris, C. N.
Natural exponential families with quadratic variance functions. *Ann. Statist. 10*, 1 (1982), 65–80.