Dirichlet and quasi-Bernoulli laws for perpetuities

Pawel Hitczenko, Gérard Letac

Dirichlet and Quasi Bernoulli laws for perpetuities

Paweł Hitczenko*, Gérard Letac †

Abstract

Let X, B and Y be three Dirichlet, Bernoulli and beta independent random variables such that $X \sim D(a_0, \dots, a_d)$, such that $\Pr(B = (0, \dots, 0, 1, 0, \dots, 0) = a_i/a$ with $a = \sum_{i=0}^d a_i$ and such that $Y \sim \beta(1, a)$. We prove that $X \sim X(1-Y) + BY$. This gives the stationary distribution of a simple Markov chain on a tetrahedron.

1 Introduction

In a recent paper [1], Ambrus, Kevei and Vígh make the following interesting observation: If V, Y, W are independent random variables such that $V \sim \frac{1}{\pi}(\frac{1}{4} - v^2)^{-1/2}\mathbf{1}_{(-1/2,1/2)}(v)dv$, Y is uniform on $(0, 1)$ and $\Pr(W = 1) = \Pr(W = -1) = 1/2$ then $V \sim V(1 - Y) + \frac{W}{2}Y$. Random variables V satisfying $V \sim VM + Q$ where the pair (M, Q) is independent of V on the right-hand side are often called perpetuities. Thus, another way of stating the observation from [1] is that an arcsine random variable on $(-1/2, 1/2)$ is a solution of perpetuity equation with $(M, Q) \sim (1-Y, WY/2)$. Part of the reason we found it interesting is that there are relatively few examples of exact solutions to perpetuity equations in the literature. We will generalize this result, and our generalization will provide more example of perpetuities, including power semicircle distributions (see [2]).

To carry out the generalization, we reformulate this result on $(0, 1)$ by writing $X = V + \frac{1}{2}$ and $B = (1 + W)/2$. Clearly X, Y, B are independent with $X \sim \beta(1/2, 1/2)$, $B \sim \frac{1}{2}(\delta_0 + \delta_1)$ and $X \sim X(1 - Y) + BY$. In Theorem 1 below, we give a generalization of the result in [1] expressed in terms of X, Y and B.

We will need the following notation. The natural basis of $\mathbb{R}^{d+1}$ is denoted by $e_0, \dots e_d$. If $p_0, \dots, p_d$ are positive numbers with sum equal to one, the distribution $\sum_{i=0}^d p_i\delta_{e_i}$ of $B = (B_0, \dots, B_d) \in \mathbb{R}^{d+1}$ is called a Bernoulli distribution. It satisfies $\Pr(B = e_i) = p_i$ and $B_0 + \dots + B_d = 1$. If $a_0, \dots, a_d$ are positive numbers the Dirichlet distribution $D(a_0, \dots, a_d)$ of $X = (X_0, \dots, X_d)$ is such that $X_0 + \dots + X_d = 1$ and the law of $(X_1, \dots, X_d)$ is

$$\frac{\Gamma(a_0 + \dots + a_d)}{\Gamma(a_0) \dots \Gamma(a_d)}(1 - x_1 - \dots - x_d)^{a_0-1}x_1^{a_1-1} \dots x_d^{a_d-1}\mathbf{1}_{T_d}(x_1, \dots, x_d)dx_1, \dots dx_d$$

where T_d is the set of $(x_1, \dots, x_d)$ such that $x_i > 0$ for all $i = 0, 1, \dots, d$, with the convention $x_0 = 1 - x_1 \dots - x_d$. For instance, if the real random variable X_1 has the beta distribution $\beta(a_1, a_0)(dx) = \frac{1}{B(a_1,a_0)}x^{a_1-1}(1 - x)^{a_0-1}\mathbf{1}_{(0,1)}(x)dx$ then $(1 - X_1, X_1) \sim D(a_0, a_1)$.

Theorem 1: Let $a_0, \dots, a_d$ be positive numbers. Denote $a = a_0 + \dots + a_d$. Let X, Y and B be three Dirichlet, beta and Bernoulli independent random variables such that $X \sim D(a_0, \dots, a_d)$ and $B \sim \sum_{i=0}^d \frac{a_i}{a}\delta_{e_i}$ are valued in $\mathbb{R}^{d+1}$ and such that $Y \sim \beta(1, a)$. Then $X \sim X(1 - Y) + BY$.

Comments. Considering each coordinate, Theorem 1 says that for all $i = 0, \dots, d$ we have $X_i \sim X_i(1 - Y) + B_iY$. Since $1 = \sum_{i=0}^d X_i = \sum_{i=0}^d B_i$ the statement for $i = 0$ is true if it is verified for $i = 1, \dots, d$. For instance for $d = 1$ a reformulation of Theorem 1 is

*Department of Mathematics, Drexel University, Philadelphia P.A. 19104, U.S.A. Partially supported by a grant from Simons Foundation (grant #208766 to Paweł Hitczenko).

†Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062 Toulouse, France.

Theorem 2: Let $a_0, a_1 > 0$. Let X_1, Y, B_1 be three independent random variables such that $X_1 \sim \beta(a_1, a_0)$, $Y \sim \beta(1, a_0 + a_1)$, $B_1 \sim \frac{a_0}{a_0+a_1}\delta_0 + \frac{a_1}{a_0+a_1}\delta_1$. Then $X_1 \sim X_1(1-Y) + B_1 Y$.

Therefore the initial remark contained in [1] is a particular case of Theorem 1 for $d = 1$ and $a_0 = a_1 = 1/2$. More generally, the case $d = 1$ and $a_0 = a_1$ covers the power semicircle distributions discussed in [2] (with $\theta = a_0 - 3/2$). In particular, $a_0 = a_1 = 3/2$ is the classical semicircle distribution.

One can prove Theorem 2 by showing $\mathbb{E}(X_1(1-Y)+B_1Y)^n) = \mathbb{E}(X_1^n)$ for all integers n. Our proof of the more general Theorem 1 is somewhat linked to this method of moments.

2 A Markov chain on the tetrahedron

This section gives an application of Theorem 1. Consider the homogeneous Markov chain $(X(n))_{n\geq 0}$ valued in the convex hull E_{d+1} of $(e_0, \ldots, e_d)$ with the following transition process: Given $X(n) \in E_{d+1}$ choose a vertex $B(n+1)$ randomly among $\{e_0, \ldots, e_d\}$ such that $\Pr(B(n+1) = e_i) = a_i/a$ and a random number $Y_{n+1} \sim \beta(1, a)$. Now draw the segment $(X(n), B(n+1))$ and take the point

$$X(n+1) = X(n)(1 - Y_{n+1}) + B(n)Y_{n+1}$$

on this segment. Theorem 1 says that the Dirichlet distribution $D(a_0, \ldots, a_d)$ is a stationary distribution for the Markov chain $(X(n))_{n\geq 0}$. Recall the following principle (see [3] Proposition 1):

Theorem 3: If E is a metric space and if C is the set of continuous maps $f : E \to E$ let us fix a probability $\nu(df)$ on C. Let $F_1, \ldots, F_n, \ldots$ be a sequence of independent random variables of C with the same distribution ν. Define $W_n = F_n \circ \ldots \circ F_2 \circ F_1$ and $Z_n = F_1 \circ \ldots \circ F_{n-1} \circ F_n$. Suppose that almost surely $Z = \lim_n Z_n(x)$ exists in E and does not depend on $x \in E$. Then

1. The distribution μ of Z is a stationary distribution of the Markov chain $(W_n(x))_{n\geq 0}$ on E;
2. if X and F_1 are independent and if $X \sim F_1(X)$ then $X \sim \mu$.

Choose $E = E_{d+1}$. Apply Theorem 3 to the distribution ν of the random map F_1 on E_{d+1} defined by $F_1(x) = (1 - Y_1)x + Y_1 B(1)$ where $Y_1 \sim \beta(1, a)$ and $B(1) \sim \sum_{i=0}^{d} \frac{a_i}{a}\delta_{e_i}$ are independent. If the F_n defined by $F_n(x) = (1 - Y_n)x + Y_n B(n)$ are independent with distribution ν, clearly

$$Z_n(x) = (\prod_{j=1}^{n}(1 - Y_j))x + \sum_{k=1}^{n}\left(\prod_{j=1}^{k-1}(1 - Y_j)\right) Y_k B(k)$$

converges almost surely to the sum of the following converging series

$$Z = \sum_{k=1}^{\infty}\left(\prod_{j=1}^{k-1}(1 - Y_j)\right) Y_k B(k) \tag{1}$$

and therefore hypotheses of Theorem 3 are met. As a consequence the Dirichlet law $D(a_0, \ldots, a_d)$ is the unique stationary distribution of the Markov chain $(X(n))_{n\geq 0}$ and is the distribution of Z defined by (1). Finally recall the definition of a perpetuity [5] on an affine space $\mathcal{A}$. Let $\nu(df)$ be a probability on the space of affine transformations f mapping $\mathcal{A}$ into itself. We say that the probability μ on $\mathcal{A}$ is a perpetuity (relatively to ν) if $X \sim F(X)$ when $F \sim \nu$ and $X \sim \mu$ are independent. If the conditions of Theorem 3 are met for ν, there is exactly one perpetuity relative to ν. Theorem 1 says that the Dirichlet distribution is a perpetuity for the random affine map $F(x) = (1 - Y)x + YB$ on the affine hyperplane $\mathcal{A}$ of $\mathbb{R}^{d+1}$ containing $e_0, \ldots, e_d$.

3 Proof of Theorem 1

If $f = (f_0, \dots, f_d)$ and $x = (x_0, \dots, x_d)$ we write $\langle f, x\rangle = \sum_{i=0}^d f_i x_i$. We need a lemma:

Lemma 4. Suppose that $a_i > 0$ for all $i = 0, \dots, d$ and denote $a = \sum_{i=0}^d a_i$. Let X be valued in the convex hull E_{d+1} of $(e_0, \dots, e_d)$. The three following conditions are equivalent

1. The random variable X has the Dirichlet distribution $D(a_0, \dots, a_d)$;
2. for all $f = (f_0, \dots, f_d)$ such that $f_i > 0$ we have $\mathbb{E}(\langle f, X\rangle^{-a}) = \prod_{i=0}^d f_i^{-a_i}$;
3. for all $f = (f_0, \dots, f_d)$ such that $f_i > 0$ we have $\mathbb{E}(\langle f, X\rangle^{-a-1}) = \left(\prod_{i=0}^d f_i^{-a_i}\right) \sum_{i=0}^d \frac{a_i}{a f_i}$.

Proof of Lemma 4. $1 \Leftrightarrow 2$ is standard (see for instance Proposition 2.1 in [4]). Statement $3 \Rightarrow 1$ is proved in the same way as $2 \Rightarrow 1$ by observing that if X satisfies 3, then X has the moments of $D(a_0, \dots, a_n)$. A short way to prove $2 \Rightarrow 3$ is to differentiate the formula in 2 with respect to f_i, obtaining

$$-a\mathbb{E}(X_i \langle f, X\rangle^{-a-1}) = -\frac{a_i}{f_i}\left(\prod_{i=0}^d f_i^{-a_i}\right).$$

Summing these equalities for $i = 0, \dots, n$ and using $\sum_{i=0}^n X_i = 1$ gives 3. □

We now prove Theorem 1. Fixing $f_i > 0$ for $i = 0, \dots, d$we can write

$$\mathbb{E}(\langle f, B\rangle^{-1}) = \sum_{i=0}^d \frac{a_i}{a f_i}. \tag{2}$$

Observe also that for $-\infty < t < 1$ we have

$$\mathbb{E}((1 - tY)^{-a-1}) = (1 - t)^{-1}. \tag{3}$$

One way to see this is to apply formula 2 of the Lemma to the Dirichlet random variable $(1 - Y, Y) \sim D(a, 1)$ and to $f_0 = 1$ and $f_1 = 1 - t$. Denote $X' = X(1 - Y) + BY$. We now write

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{\langle f, X'\rangle^{a+1}}\right) &= \mathbb{E}\left((\langle f, X\rangle - Y\langle f, X - B\rangle)^{-a-1}\right) \\
&= \mathbb{E}\left(\langle f, X\rangle^{-a-1}\mathbb{E}\left([1 - Y\frac{\langle f, X - B\rangle}{\langle f, X\rangle}]^{-a-1}|X, B\right)\right) \qquad (4)\\
&= \mathbb{E}\left(\langle f, X\rangle^{-a-1}[1 - \frac{\langle X - B\rangle}{\langle f, X\rangle}]^{-1}\right) = \mathbb{E}\left(\langle f, X\rangle^{-a}\langle f, B\rangle^{-1}\right) \qquad (5)\\
&= \left(\prod_{i=0}^d f_i^{-a_i}\right)\sum_{i=0}^d \frac{a_i}{a f_i} = \mathbb{E}\left(\frac{1}{\langle f, X\rangle^{a+1}}\right). \qquad (6)
\end{aligned}$$

Equation (4) is obtained by conditioning using the independence of X, Y, B; Equality (5) comes from (3) applied to $t = \langle f, X - B\rangle / \langle f, X\rangle$ which is indeed < 1; Equality (6) comes from the independence of X and B, from (2)and from parts $1 \Rightarrow 2$, 3 of the Lemma. From part $3 \Rightarrow 1$ we get that $X' \sim X$. □

4 Extension to the quasi Bernoulli distributions

In this section, Theorem 6 generalizes Theorem 1. This choice of order of exposition comes from the fact that this extension of Theorem 1 is somewhat complicated. We need for this to define in Theorem 5 below the quasi Bernoulli distributions of order k on the tetrahedron E_{d+1} with parameters $a_0, \dots, a_d$.

On the open set $U_{d+1}=\{f=(f_0,\ldots,f_d)\ ;\ f_i>0\ \forall i\}$ we define the function $T(f)=\prod_{i=0}^d f_i^{-a_i}$ and the differential operator
$$H=\sum_{i=0}^d \frac{\partial}{\partial f_i}.$$
In the following statement $(a)_k$ is the Pochhammer symbol $(a)_0=1$ and $(a)_{k+1}=(a+k)(a)_k$.

Lemma 5. Let X be a random variable of E_{d+1} let $a_0,\ldots,a_d$ be a sequence of positive numbers and let k be a non negative integer. Then $X\sim D(a_0,\ldots,a_d)$ if and only if for all $f\in U_{d+1}$ we have
$$(-1)^k(a)_k\mathbb{E}(\langle f,X\rangle^{-a-k})=H^k(T)(f). \tag{7}$$

Proof. The if part is proved by induction. The case $k=0$ comes from [4] Proposition 2.1. If it is correct for k we apply H to both sides of (7) and we use the fact that $X_0+\ldots+X_d=1$ for claiming that $H\mathbb{E}(\langle f,X\rangle^{-a-k})=-(a+k)\mathbb{E}(\langle f,X\rangle^{-a-k-1})$. This extends the induction hypothesis to $k+1$. The only if part is proved by the same moment method used in [4] Proposition 2.1 in the particular case $k=0$. $\square$

Formula (7) is crucial for defining our quasi Bernoulli distributions. For simplicity denote $L=\log T$. Then
$$H(T)=TH(L),\ H^2(T)=TH(L)^2+TH^2(L),\ H^3(T)=TH(L)^3+3TH^2(L)H(L)+TH^3(L),$$
$$H^4(T)=T(H(L))^4+6H^2(L)(H(L))^2+4TH^3(L)H(L)+3T(H^2(L))^2+TH^4(L).$$
More generally an easy induction on k shows $T^{-1}H^k(T)$ is a polynomial P_k with respect to $H^j(L)$ with $j=1,\ldots,k$ with non negative integer coefficients $c_k(\alpha_1,\ldots,\alpha_k)$ such that $c_k(\alpha_1,\ldots,\alpha_k)\neq 0$ implies $\alpha_1+2\alpha_2+\cdots+k\alpha_k=k$:
$$T^{-1}H^k(T)=P_k(H(L),\ldots,H^k(L))=\sum_{\alpha_1,\ldots,\alpha_k}c_k(\alpha_1,\ldots,\alpha_k)H(L)^{\alpha_1}\cdots(H^k(L))^{\alpha_k}. \tag{8}$$
Now we observe that for $j>0$ we have
$$H^j(L)=(-1)^j(j-1)!\sum_{i=0}^d\frac{a_i}{f_i^j}. \tag{9}$$
The next theorem defines the quasi Bernoulli distribution $\mathcal{B}_k(a_0,\ldots,a_d)$.

Theorem 5. Let $a_0,\ldots,a_d>0$ with $a=\sum_{i=0}^d a_i$. Fix a positive integer k. Then there exists a unique probability distribution $\mathcal{B}_k(a_0,\ldots,a_d)$ for the random variable B on E_{d+1} such that for all $f_i>0$, $i=0\ldots,d$ we have
$$\begin{aligned}\mathbb{E}(\langle f,B\rangle^{-k}) &= \frac{1}{(a)_k}\sum_{\alpha_1,\ldots,\alpha_k}c_k(\alpha_1,\ldots,\alpha_k)\prod_{j=1}^k\left((j-1)!\sum_{i=0}^k\frac{a_i}{f_i^j}\right)^{\alpha_j} && (10)\\ &= \frac{1}{(a)_k}\sum_{b_0,\ldots,b_d\in\mathbb{N};b_0+\cdots+b_d=k}\frac{k!}{b_0!\ldots b_d!}\prod_{i=0}^d\frac{(a_i)_{b_i}}{f_i^{b_i}}. && (11)\end{aligned}$$
In particular, if $X\sim D(a_0,\ldots,a_d)$ we have
$$\mathbb{E}(\langle f,X\rangle^{-a-k})=\mathbb{E}(\langle f,X\rangle^{-a})\mathbb{E}(\langle f,B\rangle^{-k}) \tag{12}$$

Proof. The uniqueness is easily proved by moments. For proving the existence it is worthy to explain the idea first by considering the cases $k=1$ and $k=2$. The case $k=1$ gives the familiar Bernoulli distribution since for $B \sim \sum_{i=0}^d \frac{a_i}{a}\delta_{e_i}$ equality (2) holds. For $k=2$ formula (10) becomes

$$\mathbb{E}(\langle f,B\rangle^{-2}) = \frac{1}{a(a+1)}\left((\sum_{i=0}^d \frac{a_i}{f_i})^2 + \sum_{i=0}^d \frac{a_i}{f_i^2}\right) = \frac{1}{a(a+1)}\left(\sum_{i=0}^d \frac{a_i(a_i+1)}{f_i^2} + \sum_{0\leq i<j\leq d}\frac{2a_ia_j}{f_if_j}\right). \quad (13)$$

We now make the two observations

- If $0 \leq i < j \leq d$ and if $B \sim \lambda_{ij}$ which is the uniform probability on the edge (e_i, e_j) of the tetrahedron E_{d+1} we have $\mathbb{E}(\langle f,B\rangle^{-2}) = 1/f_if_j$.
- The second observation is that $\frac{1}{a(a+1)}\left(\sum_{i=0}^d a_i(a_i+1) + \sum_{0\leq i<j\leq d} 2a_ia_j\right) = 1.$

Therefore we see that $\mathcal{B}_2(a_0,\ldots,a_d)$ is the mixture

$$\frac{1}{a(a+1)}\left(\sum_{i=0}^d a_i(a_i+1)\delta_{e_i} + \sum_{0\leq i<j\leq d} 2a_ia_j\lambda_{ij}\right).$$

We now pass to the general case of an arbitrary k. We slightly extend the definition of a Dirichlet distribution $D(a_0,\ldots,a_d)$ by allowing $a_i \geq 0$ instead of $a_i > 0$ while keeping $a = \sum_{i=0}^d a_i > 0$. For such a sequence $(a_0,\ldots,a_d)$ we define $T = \{i\ ;\ a_i > 0\}$ and $D(a_0,\ldots,a_d)$ has the Dirichlet distribution concentrated on the tetrahedron E_T generated by $(e_i)_{i\in T}$ with parameters $(a_i)_{i\in T}$. If $X \sim D(a_0,\ldots,a_d)$ the formula $\mathbb{E}(\langle f,X\rangle^{-a}) = \prod_{i=0}^d f_i^{-a_i}$ still holds. If T contains only one element i_0 then $D(a_0,\ldots,a_d)$ is simply $\delta_{e_{i_0}}$ and does not depend on a.

To finish the proof of the theorem, we observe that putting $f_i = 1$ for all $i = 0,\ldots,d$ in (8) gives $(-1)^k$ on the left hand side. As a consequence, putting $f_i = 1$ for all $i = 0,\ldots,d$ in (10) on the right hand side gives 1. The right hand side of (10) can be considered as a homogeneous polynomial of degree k with respect to the variables $1/f_i$, namely

$$\sum_{b_1,\ldots,b_d} w_k(b_0,\ldots,b_d)\prod_{i=0}^d \frac{1}{f_i^{b_i}}.$$

In fact, it will follow from (11) that $w_k(b_0,\ldots,b_d) = \frac{1}{(a)_k}\frac{k!}{b_0!\ldots b_d!}\prod_{i=0}^d (a_i)_{b_i}$ and hence they satisfy

$$\sum_{b_1,\ldots,b_d} w_k(b_0,\ldots,b_d) = 1.$$

Therefore, once we establish (11) we can claim that $\mathcal{B}_k(a_0,\ldots,a_d)$ exists and is the following mixture of extended Dirichlet distributions

$$\mathcal{B}_k(a_0,\ldots,a_d) = \sum_{b_1,\ldots,b_d} w_k(b_0,\ldots,b_d)D(b_0,\ldots,b_d).$$

To prove (11) recall that (8) and (9) imply that

$$H^k(T) = (-1)^k T\sum_{\alpha_1,\ldots,\alpha_k} c_k(\alpha_1,\ldots,\alpha_k)\prod_{j=1}^k\left((j-1)!\sum_{i=0}^k \frac{a_i}{f_i^j}\right)^{\alpha_j}.$$

Alternatively, applying $H^k(T)$ directly to the product $\prod_{i=0}^d f)i^{-a_i}$ we see that $H^k(T)$ is the sum over all possible choices of k applications of partial derivatives $\frac{\partial}{\partial f_i}$, $i = 0,\ldots,d$ to T. If $\frac{\partial}{\partial f_i}$ is applied b_i times, $i = 0,\ldots,d$ with $\sum b_i = k$ then the result is

$$(-1)^k\prod_{i=0}^d \frac{(a_i)_{b_i}}{f_i^{a_i+b_i}} = (-1)^k T\prod_{i=0}^d \frac{(a_i)_{b_i}}{f_i^{b_i}}.$$

Since there are $\binom{k}{b_0,\ldots,b_d}$ ways of applying $\frac{\partial}{\partial f_i}$ b_i times for $i = 0, \ldots, d$ it follows that

$$H^k(T) = (-1)^k T \sum_{b_0,\ldots,b_d} \frac{k!}{b_0! \ldots b_d!} \prod_{i=0}^{d} \frac{(a_i)_{b_i}}{f_i^{b_i}},$$

which proves (11).

Finally, formula (12) is a consequence of Lemma 4 and (10). $\square$

Comments. The complexity of $H^k(T)$ makes the general combinatorial description of $\mathcal{B}_k(a_0, \ldots, a_d)$ difficult. However it is worthwhile to write the details when $d = 1$ for small values of k. Observe first that if $B = (1 - B_1, B_1) \sim D(b_0, b_1)$, or $\mathbb{E}(\langle f, B\rangle^{-b_0-b_1}) = f_0^{-b_0} f_1^{-b_1}$ then $B_1 \sim \delta_0$ if $b_1 = 0$, $B_1 \sim \delta_1$ if $b_0 = 0$ and $B_1 \sim \beta(b_1, b_0)$ if $b_0, b_1 > 0$. As a consequence, if $B = (1 - B_1, B_1)$ has the quasi Bernoulli distribution $\mathcal{B}_k(a_0, a_1)$, the distribution of B_1 has the form

$$B_1 \sim \frac{(a_1)_k}{(a)_k}\delta_0(dx) + g_k(x)dx + \frac{(a_0)_k}{(a)_0}\delta_1(dx)$$

where the function g_k is zero outside of $(0, 1)$. The first values of $g_k(x)$ on $(0, 1)$ are the following

$$g_1(x) = 0,\ g_2(x) = \frac{2a_0a_1}{a(a+1)},\ g_3(x) = \frac{6a_0a_1}{(a)_3}(a_0 + 1 + (a_1 - a_0)x).$$

??????????????????????? *Pawel, here we should put more examples.*

Theorem 6. Let $a_0, \ldots, a_d > 0$ with $a = \sum_{i=0}^d a_i$.Fix a positive integer k. Consider three independent random variables $X \sim D(a_0, \ldots, a_d)$, $Y \sim \beta(k, a)$ and $B \sim \mathcal{B}_k(a_0, \ldots, a_d)$. Then $X \sim (1-Y)X + YB$.

Proof. It is quite similar to the proof of Theorem 1. Denote $X' = (1 - Y)X + YB$ and take $f_i > 0$. Then

$$\begin{aligned}
\mathbb{E}(\langle f, X'\rangle^{-a-k}) &= \mathbb{E}\left(\langle f, X\rangle^{-a-k}[1 - Y\frac{\langle f, X - B\rangle}{\langle f, X\rangle}]^{-a-k}\right) && (14)\\
&= \mathbb{E}\left(\langle f, X\rangle^{-a-k}\mathbb{E}\left([1 - Y\frac{\langle f, X - B\rangle}{\langle f, X\rangle}]^{-a-k}|X, B\right)\right) && (15)\\
&= \mathbb{E}\left(\langle f, X\rangle^{-a-k}[1 - \frac{\langle f, X - B\rangle}{\langle f, X\rangle}]^{-k}\right) = \mathbb{E}\left(\langle f, X\rangle^{-a}\langle f, B\rangle^{-k}\right) && (16)\\
&= \mathbb{E}(\langle f, X\rangle^{-a-k}). && (17)
\end{aligned}$$

Step (16) comes from the fact that $\mathbb{E}((1 - tY)^{-a-k}) = (1 - t)^{-k}$applied to $t = \frac{\langle f, X - B\rangle}{\langle f, X\rangle} < 1$. The proof comes from Lemma 4 part 2 applied to $d = 1$, $(1 - Y, Y) \sim D(a, k)$ and $f_0 = 1$ and $f_1 = 1 - t$.

5 References

[1] Gergely Ambrus, Péter Kevei and Viktor Vígh (2012) 'The diminishing segment process'. *Statist. Probab. Letters* **82** 191-195.

[2] Octavio Arzimendi and Víctor Pérez-Abreu (2010) 'On the non-classical infinite divisibilityof power semicircle distributions.' *Commun. Stoch. Anal.* **4** 161-178.

[3] Jean-François Chamayou and Gérard Letac (1991) 'Explicit stationary sequences for compositions of random functions and products of random matrices.' *J. Theoret. Probab.* **4** 3-36

[4] Jean-François Chamayou and Gérard Letac (1994) 'A transient random walk on stochastic matrices with Dirichlet distributions.' *Ann. Probab.* **22** 424-430.

[5] Charles M. Goldie and Ross A. Maller (2000) 'Stability of perpetuities.' *Ann. Probab.* **28** 1195-1218.