Asymptotic Equidistribution of Congruence Classes with respect to the Convolution Iterates of a Probability Vector

Gilles Gnacadja

Research and Development Information Systems, Amgen, Inc.
One Amgen Center Drive, Thousand Oaks, California 91320-1799, USA
gilles.gnacadja@gmail.com

Published in *Statistics and Probability Letters*

DOI 10.1016/j.spl.2012.05.025

This is an Enhanced Postprint of the published refereed version.
The main result has been augmented and the proof is different.

Revision C.4 30 July 2012

Abstract

Consider a positive integer d and a positive probability vector f over the numbers $0, \ldots, \ell$. The n-fold convolution f^{*n} of f is a probability vector over the numbers $0, \ldots, n\ell$, and these can be partitioned into congruence classes modulo d. The main result of this paper is that, asymptotically in n, these d congruence classes have equiprobability $1/d$. In the motivating application, one has N containers of capacity d and repeatedly retrieves one item from each of M randomly selected containers ($0 < M < N$); containers are replenished to full capacity when emptied. The result implies that, over the long term, the number of containers requiring replenishment is M/d. This finding is relevant wherever one would be interested in the steady-state pace of replenishing fixed-capacity containers.

Keywords. Equidistribution; Congruence Classes in Probability Convolution; Circulant Matrix; Doubly Stochastic Matrix; Inventory Replenishment.

Mathematics Subject Classification (2010): 60C05, 60G10, 05A16.

1 Introduction

This paper stems from a combinatorics problem that seems not to have been considered before. Suppose that we have N containers of capacity d, and that from each of M randomly selected of these ($0 < M < N$) we retrieve one item. This retrieval process is repeated indefinitely, and every container that becomes empty is replenished to full capacity before the retrieval process continues. What is the number of containers that need replenishment when the retrieval process has occurred n times? Of course, this number cannot be known deterministically (if $d \geqslant 2$) because of the randomness involved. But as is explained next, it will follow from the main result of this paper that, as n grows, the number converges to M/d.

A container needs replenishment after n repetitions of the retrieval process if and only if the number of times it was selected has just become a multiple of d. In more elaborate form, this condition says that

Event (1): the container was selected in the nth instance of the retrieval process, and

Event (2): the number of times the container was selected in the prior $n-1$ instances of the retrieval process is congruent to $d-1$ modulo d.

The probability of Event (1) is M/N, and it follows from Theorem 2.1 that, asymptotically in n, the probability of Event (2) is $1/d$. Consequently, asymptotically in n, the probability that a container needs replenishment is $M/(Nd)$, and the number of such containers is M/d.

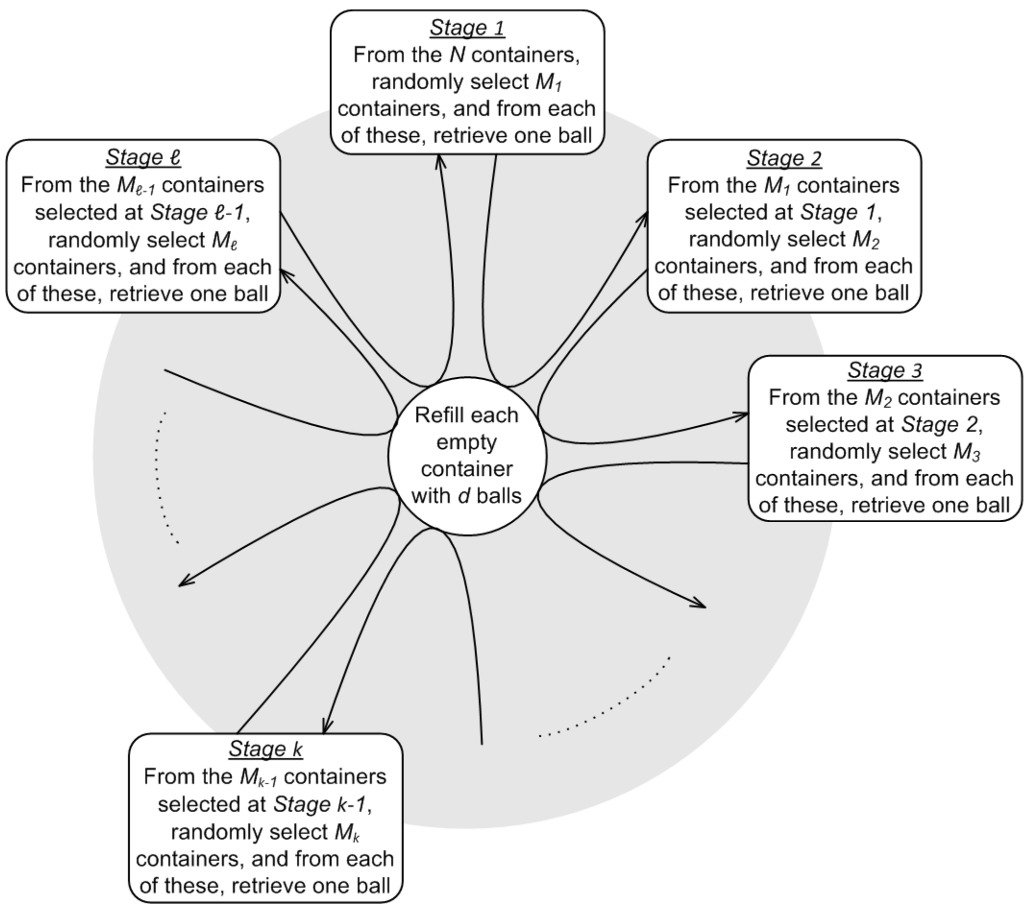

Figure 1.1: The multi-stage retrieval and replenishment process

The generality of Theorem 2.1 actually makes it applicable more broadly. Namely, the retrieval process may consist of ℓ subprocesses whereby items are retrieved from $M_1, \ldots, M_\ell$ containers, with $0 < M_\ell < \cdots < M_1 < N$, and the selection of M_{k+1} containers in subprocess

$k+1$ takes place among the M_k containers selected in subprocess k. This multi-stage retrieval and replenishment process is illustrated on Figure 1.1. There is a probability vector $f = \big(f(0), \ldots, f(\ell)\big)$ such that $f(k)$ is the probability that k is the number of times a container is selected. Specifically, with $M_0 := N$ and $M_{\ell+1} := 0$ for convenience, one can verify that $f(k) = (M_k - M_{k+1})/N$. We have $\ell = 2$ in the particular application that motivated this work. Its details are quite technical and better suited for a more specialized venue.

Theorem 2.1, the main result of this paper, is generally relevant to inventory management and other activities where one would be concerned with the steady-state pace of replenishing fixed-capacity containers. We discovered it independently, but a similar result from the point of view of polynomials appeared in the work of Major [3, Theorem 1]. Here, the result is more explicit about the asymptotic behavior and the proof is more self-contained.

2 The Result

Let $f = \big(f(0), \ldots, f(\ell)\big)$ be a probability vector; $f(0), \ldots, f(\ell) \in \mathbb{R}_{\geqslant 0}$ and $f(0) + \cdots + f(\ell) = 1$. For $n \in \mathbb{Z}_{\geqslant 0}$, the n-fold convolution f^{*n} is a probability vector of $n\ell + 1$ components over the numbers $0, \ldots, n\ell$. It represents the n-fold repetition of the process represented by f. Note the special case $n = 0$: we have the 0-fold convolution $f^{*0} = \big(f^{*0}(0)\big) = (1)$.

For $d \in \mathbb{Z}_{\geqslant 1}$ and $r = 0, \ldots, d-1$, let $\varphi(f, n, d, r)$ denote the probability that a component number in f^{*n} is congruent to r modulo d. By definition, we have

$$\varphi(f, n, d, r) \;=\; \sum_{\substack{0 \leqslant k \leqslant n\ell \\ k \equiv r \bmod d}} f^{*n}(k) \;=\; \sum_{q=0}^{\operatorname{floor}((n\ell - r)/d)} f^{*n}(r + d\,q)\ .$$

This paper is about the result that as the order n of convolution grows, this probability becomes independent of f and equidistributed with respect to r. We assemble these probabilities into a probability vector as follows.

$$\varphi(f, n, d) \;:=\; \big(\varphi(f, n, d, 0), \ldots, \varphi(f, n, d, d-1)\big)$$

In the case $d = 1$, this vector is $\varphi(f, n, 1) = \big(\varphi(f, n, 1, 0)\big) = (1)$, and the result is trivially true. So we suppose $d \geqslant 2$. We set

$$\omega_d \;:=\; \exp\left(\frac{2\pi\mathrm{i}}{d}\right) \qquad \text{and} \qquad P_{f,d}(X) \;:=\; \sum_{r=0}^{d-1} \varphi(f, 1, d, r) X^r\ ,$$

and then

$$\gamma(f, d) \;:=\; \max_{1 \leqslant r \leqslant d-1} \left|P_{f,d}\left(\omega_d^r\right)\right|\ .$$

Also, let

$$u_d \;:=\; \frac{1}{d}\,(\underbrace{1, \ldots, 1}_{d}) \quad \text{and} \quad e_d \;:=\; (1, \underbrace{0, \ldots, 0}_{d-1})\ .$$

Observe that $\varphi(f,0,d) = e_d$.

Theorem 2.1. *Let $f = \big(f(0),\dots,f(\ell)\big)$ be a positive probability vector with $\ell \geqslant 1$, and let $d \in \mathbb{Z}_{\geqslant 2}$. We have*

$$\gamma(f,d) \quad < \quad 1 \tag{2.1}$$

and

$$\forall\, n \in \mathbb{Z}_{\geqslant 0}\,,\ \big\|\varphi(f,n,d) - u_d\big\|_2 \ \leqslant\ \big(\gamma(f,d)\big)^n \sqrt{\frac{d-1}{d}}\ . \tag{2.2}$$

Consequently,

$$\lim_{n\to\infty} \varphi(f,n,d) \quad = \quad u_d\ , \tag{2.3}$$

i.e.

$$\forall\, r = 0,\dots,d-1,\ \lim_{n\to\infty} \varphi(f,n,d,r) \ = \ \frac{1}{d}\ . \tag{2.4}$$

Theorem 2.1 is proved in the next section. It is sufficient for the application problem that motivated this work. Yet one may wonder whether it is necessary to require that the probability vector f be positive. It is not, as noted in Major [3]. But it is also easy to make the assertions of Theorem 2.1 fail when f is just required to be nonnegative. Suppose for instance that $f(k) = 0$ whenever k is odd. Then, an obvious inductive argument shows that, for all $n \in \mathbb{Z}_{\geqslant 1}$, $f^{*n}(k) = 0$ when k is odd, whence $\varphi(f,n,2,0) = 1$ and $\varphi(f,n,2,1) = 0$. More generally, we have the following result, the proof of which presents no difficulty, and is omitted.

Remark 2.2. Consider a probability vector $f = \big(f(0),\dots,f(\ell)\big)$ with $\ell \geqslant 1$, and $d \in \mathbb{Z}_{\geqslant 2}$. Suppose that $f(k) = 0$ for every $k = 1,\dots,\ell$ that is not divisible by d. Then we have $\varphi(f,n,d,0) = 1$ and $\varphi(f,n,d,r) = 0$ for $r = 1,\dots,d-1$, i.e. $\varphi(f,n,d) = e_d$, for all $n \in \mathbb{Z}_{\geqslant 1}$.

3 The Proof

In this section, we state and prove some intermediate results, which we then use to prove Theorem 2.1 at the end. We use classic notions of matrix algebra which may be found for instance in Horn and Johnson [1], and known facts about circulant matrices which may be found in Kra and Simanca [2] and references therein.

We extend the definition of φ for convenience as follows.

$$\varphi(g,n,d,r) \ := \sum_{k\,\in\, r+d\mathbb{Z}} g^{*n}(k) \ = \ \sum_{q\in\mathbb{Z}} g^{*n}(r + d\,q)$$

for any finitely supported $\mathbb{Z}$-indexed vector $g = \big(g(k)\big)_{k\in\mathbb{Z}}$ and $n \in \mathbb{Z}_{\geqslant 0}$. We then define the d-vector $\varphi(g,n,d)$ by

$$\varphi(g,n,d) \quad := \quad \big(\varphi(g,n,d,0),\dots,\varphi(g,n,d,d-1)\big)\ .$$

Note the special case $n = 0$: the 0-fold convolution g^{*0} is the $\mathbb{Z}$-indexed vector with $g^{*0}(0) = 1$ and $g^{*0}(k) = 0$ for $k \neq 0$; and $\varphi(g, 0, d) = e_d$.

Let $\Phi(g, d)$ be the circulant matrix associated with the vector $\varphi(g, 1, d)$. By definition, $\Phi(g, d)$ is a $d\times d$ matrix, its top row is the vector $\varphi(g, 1, d)$, and each subsequent row is obtained from the preceding one by circularly shifting the entries rightward.

Lemma 3.1. *Let $g = \big(g(k)\big)_{k\in\mathbb{Z}}$ be finitely supported and let $d \in \mathbb{Z}_{\geqslant 1}$ and $n \in \mathbb{Z}_{\geqslant 0}$. The matrix $\big(\Phi(g, d)\big)^n$ is the circulant matrix associated with the vector $\varphi(g, n, d)$. In particular, the top row of the matrix $\big(\Phi(g, d)\big)^n$ is the vector $\varphi(g, n, d)$, i.e.*

$$\varphi(g, n, d) \quad = \quad e_d \cdot \big(\Phi(g, d)\big)^n \; .$$

Proof. Let $\Psi(g, n, d) = \big(\Psi(g, n, d, r, s)\big)_{0\leqslant r,s\leqslant d-1}$ be the circulant matrix associated with the vector $\varphi(g, n, d)$. The entries are given by

$$\Psi(g, n, d, r, s) \quad = \quad \sum_{k \,\in\, s-r+d\mathbb{Z}} g^{*n}(k) \; .$$

The matrices Ψ are related as follows.

$$\forall\, m, n \in \mathbb{Z}_{\geqslant 0}, \Psi(g, m, d) \cdot \Psi(g, n, d) = \Psi(g, m + n, d) \, .$$

This results from the following sequence of algebraic transformations.

$$\begin{aligned}
\Psi(g, m + n, d, r, t) \quad &= \quad \sum_{j\in t-r+d\mathbb{Z}} g^{*(m+n)}(j) \\
&= \quad \sum_{j\in t-r+d\mathbb{Z}} \big(g^{*m} * g^{*n}\big)(j) \\
&= \quad \sum_{j\in t-r+d\mathbb{Z}} \sum_{i\in\mathbb{Z}} g^{*m}(i) g^{*n}(j - i) \\
&= \quad \sum_{i\in\mathbb{Z}} g^{*m}(i) \sum_{j\in t-r+d\mathbb{Z}} g^{*n}(j - i) \\
&= \quad \sum_{s=0}^{d-1} \sum_{i\in s-r+d\mathbb{Z}} g^{*m}(i) \sum_{j\in t-r+d\mathbb{Z}} g^{*n}(j - i) \\
&= \quad \sum_{s=0}^{d-1} \sum_{i\in s-r+d\mathbb{Z}} g^{*m}(i) \sum_{j\in t-r-i+d\mathbb{Z}} g^{*n}(j) \\
&= \quad \sum_{s=0}^{d-1} \sum_{i\in s-r+d\mathbb{Z}} g^{*m}(i) \sum_{j\in t-s+d\mathbb{Z}} g^{*n}(j) \\
&= \quad \sum_{s=0}^{d-1} \Psi(g, m, d, r, s) \Psi(g, n, d, s, t) \; .
\end{aligned}$$

The relation implies (and in fact is equivalent to) that for all $n \in \mathbb{Z}_{\geqslant 0}$, $\Psi(g, n, d) = \big(\Psi(g, 1, d)\big)^n$. Since $\Psi(g, 1, d) = \Phi(g, d)$, the proof of Lemma 3.1 is complete. □

Lemma 3.2. *Let $g = \big(g(k)\big)_{k\in\mathbb{Z}}$ be finitely supported and nonnegative. Suppose that $g(0),\ldots,g(\ell) > 0$ for some $\ell \geqslant 1$. Then the matrix $\Phi(g,d)$ is primitive.*

Proof. The matrix $\Phi(g,d)$ is nonnegative, so the assertion that it is primitive is equivalent to the existence of $n \in \mathbb{Z}_{\geqslant 1}$ such that the matrix $\big(\Phi(g,d)\big)^n$ is positive. This in turn is equivalent to the existence of $n \in \mathbb{Z}_{\geqslant 1}$ such that the vector $\varphi(g,n,d)$ is positive, because by Lemma 3.1, $\big(\Phi(g,d)\big)^n$ is the circulant matrix associated with the $\varphi(g,n,d)$. With $g(k) > 0$ for $0 \leqslant k \leqslant \ell$, we have $g^{*n}(k) > 0$ for $0 \leqslant k \leqslant n\ell$. Suppose that $n \geqslant (d-1)/\ell$. Let $r = 0,\ldots,d-1$. We have $0 \leqslant r \leqslant n\ell$, so $g^{*n}(r) > 0$. But $\varphi(g,n,d,r) \geqslant g^{*n}(r)$ because g is nonnegative. So $\varphi(g,n,d,r) > 0$. Hence, the vector $\varphi(g,n,d)$ is positive. □

Proof of Theorem 2.1.

Where a $\mathbb{Z}$-indexed vector is expected and we put f, one should read $\bar{f} = \big(\bar{f}(k)\big)_{k\in\mathbb{Z}}$, the vector extending f with zeros over the whole $\mathbb{Z}$. The matrix $\Phi(f,d)$ is the circulant matrix associated with the vector $\varphi(f,1,d)$, so as defined, $P_{f,d}$ is the representer polynomial of $\Phi(f,d)$. For $r = 0,\ldots,d-1$, let

$$\lambda_{f,d,r} := P_{f,d}\left(\omega_d^{-r}\right) \qquad \text{and} \qquad v_{d,r} := \frac{1}{\sqrt{d}}\left(1,\omega_d^{r},\omega_d^{2r},\ldots,\omega_d^{(d-1)r}\right) .$$

The following is well known: the eigenvalues of $\Phi(f,d)$ are $\lambda_{f,d,0},\lambda_{f,d,1},\ldots,\lambda_{f,d,d-1}$; $v_{d,r}$ is a left $\lambda_{f,d,r}$-eigenvector of $\Phi(f,d)$; and the vectors $v_{d,0},v_{d,1},\ldots,v_{d,d-1}$ form an orthonormal basis of $\mathbb{C}^n$. Note that $\lambda_{f,d,0} = 1$ and $v_{d,0} = \big(1/\sqrt{d}\big)u_d$.

The vector $\varphi(f,d)$ is a probability vector, so $\Phi(f,d)$ is nonnegative and all its row sums and column sums equal one; $\Phi(f,d)$ is doubly stochastic. It is also primitive by Lemma 3.2. By application of Perron-Frobenius theory, we obtain that the eigenvalue $\lambda_{f,d,0} = 1$ is simple and that for $r = 1,\ldots,d-1$, $|\lambda_{f,d,r}| < 1$. Equation (2.1) in Theorem 2.1 is thus proved.

Let

$$\mathcal{H}_d := \left\{z = (z_1,\ldots,z_d) \in \mathbb{C}^d : z_1 + \cdots + z_d = 0\right\} .$$

This is a hyperplane in $\mathbb{C}^d$ containing the eigenvectors $v_{d,1},\ldots,v_{d,d-1}$. Therefore, these $d-1$ vectors form an orthonormal basis of $\mathcal{H}_d$, $\mathcal{H}_d$ is stable under $\Phi(f,d)$, and

$$\forall\, z \in \mathcal{H}_d\,,\ \left\|z \cdot \Phi(f,d)\right\|_2 \leqslant \gamma(f,d)\,\|z\|_2 .$$

It follows that

$$\forall\, n \in \mathbb{Z}_{\geqslant 0}\,,\ \forall\, z \in \mathcal{H}_d\,,\ \left\|z \cdot \big(\Phi(f,d)\big)^n\right\|_2 \leqslant \big(\gamma(f,d)\big)^n\,\|z\|_2 .$$

Using Lemma 3.1, we obtain

$$\varphi(f,n,d) - u_d = e_d \cdot \big(\Phi(f,d)\big)^n - u_d \cdot \big(\Phi(f,d)\big)^n = \big(e_d - u_d\big) \cdot \big(\Phi(f,d)\big)^n .$$

We have $e_d - u_d \in \mathcal{H}_d$ and $\left\|e_d - u_d\right\|_2 = \sqrt{(d-1)/d}$, so

$$\forall\, n \in \mathbb{Z}_{\geqslant 0}\,,\ \left\|\varphi(f,n,d) - u_d\right\|_2 \leqslant \big(\gamma(f,d)\big)^n \left\|e_d - u_d\right\|_2 = \big(\gamma(f,d)\big)^n \sqrt{(d-1)/d} .$$

The proof of Theorem 2.1 is complete. □

References

[1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990, ISBN 0521386322.

[2] I. Kra and S. R. Simanca, *On Circulant Matrices*, Notices of the American Mathematical Society **59** (2012), no. 3, 368–377, http://www.ams.org/notices/201203/rtx120300368p.pdf.

[3] L. Major, *On the Distribution of Coefficients of Powers of Positive Polynomials*, Australasian Journal of Combinatorics **49** (2011), 239–243, http://ajc.maths.uq.edu.au/?page=get_volumes&volume=49.