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

Preprint revision B.4 5 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 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.

The generality of Theorem 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. Then 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 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 recently appeared in the work of Major (2011, Theorem 1). Here, thanks to Proposition 2, our approach is simpler, more explicit and 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 $n \in \mathbb{Z}_{\geqslant 1}$, 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.

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

$$\varphi(f, n, d, r) \quad = \quad \sum_{q=0}^{\mathrm{floor}((n\ell - r)/d)} f^{*n}(r + d\,q) \ . \tag{1}$$

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. The precise statement is as follows.

Theorem 1. *Given any positive probability vector $f = \big(f(0), \ldots, f(\ell)\big)$ with $\ell \geqslant 1$, and any $d, r \in \mathbb{Z}$ with $d \geqslant 1$ and $0 \leqslant r \leqslant d-1$, we have*

$$\lim_{n\to\infty} \varphi(f, n, d, r) \quad = \quad \frac{1}{d} \ . \tag{2}$$

The proof of Theorem 1 will follow from Proposition 2. The definition of φ in Equation (1) incorporates details relevant to its computation, an important practical consideration. However, these can be cumbersome in reasoning and manual algebraic calculations. For the purpose of formulating and proving Proposition 2, we find it more convenient to work with the vector $\bar{f} = \big(\bar{f}(k)\big)_{k\in\mathbb{Z}}$ that extends f with zeros. We have

$$\varphi(f,n,d,r) \;=\; \sum_{q\in\mathbb{Z}} \bar{f}^{*n}(r+d\,q) \;=\; \sum_{k\,\in\, r+d\mathbb{Z}} \bar{f}^{*n}(k)\;.$$

We define the d-vector $\varphi(f,n,d)$ by

$$\varphi(f,n,d) \;:=\; \big(\varphi(f,n,d,0),\ldots,\varphi(f,n,d,d-1)\big) \tag{3}$$

and the $d\times d$ matrix $\Phi(f,n,d) = \big(\Phi(f,n,d,u,v)\big)_{0\leqslant u,v\leqslant d-1}$ by

$$\Phi(f,n,d,u,v) \;:=\; \sum_{k\,\in\, v-u+d\mathbb{Z}} \bar{f}^{*n}(k)\;. \tag{4}$$

The vector $\varphi(f,n,d)$ is a probability vector. In the matrix $\Phi(f,n,d)$, the top row is the vector $\varphi(f,n,d)$ and each subsequent row is obtained from the preceding one by circularly shifting the entries rightward. Thus, $\Phi(f,n,d)$ is the circulant matrix associated with the vector $\varphi(f,n,d)$. Circulant matrices are pervasive in many areas of mathematics and engineering; see for instance the recent article of Kra and Simanca (2012) and the references therein.

Proposition 2. *Given any probability vector* $f = \big(f(0),\ldots,f(\ell)\big)$ *and any* $d,m,n\in\mathbb{Z}$ *with* $d\geqslant 1$ *and* $n>m\geqslant 1$*, we have*

$$\varphi(f,n,d) \;=\; \varphi(f,n-m,d)\cdot\Phi(f,m,d)\;. \tag{5}$$

Proving Proposition 2 amounts to performing the obvious algebraic manipulations:

$$\begin{aligned}
\varphi(f,n,d,v) &= \sum_{j\in v+d\mathbb{Z}} \bar{f}^{*n}(j)\\
&= \sum_{j\in v+d\mathbb{Z}} \Big(\bar{f}^{*(n-m)} * \bar{f}^{*m}\Big)(j)\\
&= \sum_{j\in v+d\mathbb{Z}}\sum_{i\in\mathbb{Z}} \bar{f}^{*(n-m)}(i)\,\bar{f}^{*m}(j-i)\\
&= \sum_{i\in\mathbb{Z}} \bar{f}^{*(n-m)}(i) \sum_{j\in v+d\mathbb{Z}} \bar{f}^{*m}(j-i)\\
&= \sum_{u=0}^{d-1}\sum_{i\in u+d\mathbb{Z}} \bar{f}^{*(n-m)}(i) \sum_{k\in v-i+d\mathbb{Z}} \bar{f}^{*m}(k)\\
&= \sum_{u=0}^{d-1}\sum_{i\in u+d\mathbb{Z}} \bar{f}^{*(n-m)}(i) \sum_{k\in v-u+d\mathbb{Z}} \bar{f}^{*m}(k)\\
&= \sum_{u=0}^{d-1} \varphi(f,n-m,d,u)\,\Phi(f,m,d,u,v)\;.
\end{aligned}$$

We now proceed to proving Theorem 1. We use classic notions of matrix algebra which may be found for instance in Horn and Johnson (1990).

A row-stochastic (respectively column-stochastic) matrix is a square nonnegative matrix whose row sums (respectively column sums) all equal one. A matrix is doubly stochastic if it is both row-stochastic and column-stochastic. Because it is the circulant matrix associated with a probability vector, the matrix $\Phi(f, m, d)$ is doubly stochastic.

Suppose that $m \geqslant (d-1)/\ell$. With Equation (1), we get that for every $r = 0, \ldots, d-1$, $\varphi(f, m, d, r) \;\geqslant\; f^{*m}(r) \;>\; 0$; the vector $\varphi(f, m, d)$ is positive. Therefore, the matrix $\Phi(f, m, d)$ is positive.

Let $U_0(d)$ and $U(d)$ be the d-vector and the $d \times d$ matrix whose entries all equal $1/d$. Because the matrix $\Phi(f, m, d)$ is positive and doubly stochastic, we have

$$\lim_{s\to\infty} \big(\Phi(f, m, d)\big)^s \;=\; U(d)\ .$$

This is an application of Perron-Frobenius theory, or alternatively of Major (2011, Lemma 2), where a very simple proof is presented.

It results from Proposition 2 that, for every $s, t \in \mathbb{Z}_{\geqslant 1}$,

$$\varphi(f, sm+t, d) \;=\; \varphi(f, t, d) \cdot \big(\Phi(f, m, d)\big)^s\ .$$

Therefore,

$$\lim_{s\to\infty} \varphi(f, sm+t, d) \;=\; \varphi(f, t, d) \cdot U(d)\ .$$

The product of any probability d-vector and the matrix $U(d)$ is equal to the vector $U_0(d)$, so

$$\lim_{s\to\infty} \varphi(f, sm+t, d) \;=\; U_0(d)\ .$$

Since this holds in particular for $t = 1, \ldots, m$, we have

$$\lim_{n\to\infty} \varphi(f, n, d) \;=\; U_0(d)\ .$$

Theorem 1 is thus proved. 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 (2011). But it is also easy to make the assertion of Theorem 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 3. Consider a probability vector $f = \big(f(0), \ldots, f(\ell)\big)$ with $\ell \geqslant 1$, and $d \in \mathbb{Z}$ with $d \geqslant 2$. Suppose that $f(k) = 0$ for every $k = 1, \ldots, \ell$ that is not divisible by d. Then we have $\varphi(f, n, d, 0) = 1$ and $\varphi(f, n, d, r) = 0$ for $r = 1, \ldots, d-1$.

Acknowledgments

László Major generously shared insight on his work leading up to and contained in his paper Major (2011), and helped contrast it with the approach employed here. The anonymous reviewer offered suggestions that helped improve the paper.

References

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

Kra, I. and Simanca, S. R., 2012. On Circulant Matrices. *Notices of the American Mathematical Society*, 59(3):368–377. http://www.ams.org/notices/201203/rtx120300368p.pdf.

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