Statistics and Probability Letters 118 (2016) 37–44

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Deconvolution of a discrete uniform distribution

Anatoly Zhigljavsky [a,b,*], Nina Golyandina [c], Svyatoslav Gryaznov [c]

[a] School of Mathematics, Cardiff University, CF24 4AG, UK
[b] Lobachevskii Nizhnii Novgorod State University, Russia
[c] St.Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg, 199034, Russia

ARTICLE INFO

Article history:
Received 28 March 2016
Received in revised form 9 June 2016
Accepted 9 June 2016
Available online 16 June 2016

Keywords:
Deconvolution
Discrete uniform
Cyclotomic polynomial
Roots of unity

ABSTRACT

Let ξ be a discrete random variable (r.v.) with uniform distribution on the support set $\{0, 1, \ldots, N\}$. We study the problem of construction of non-degenerate independent r.v.'s ξ_1 and ξ_2 such that $\xi = \xi_1 + \xi_2$, if these r.v.'s exist. We describe a general form for the solutions to this problem, offer some analytic constructions and develop algorithms for computing the distributions of ξ_1 and ξ_2.

1. Introduction

Let ξ be a discrete r.v. with uniform distribution on the support set $\{0, 1, \ldots, N\}$, where $N > 1$ is given. We investigate the problem of existence of non-degenerate independent r.v.'s ξ_1 and ξ_2 such that

$$\xi = \xi_1 + \xi_2 \tag{1}$$

and develop general schemes for construction of these r.v.'s. In Section 2, we reformulate the problem in terms of convolutions of vectors, introduce generating functions and roots of unity. In Section 3, we prove our main theorem which says that the r.v.'s ξ_1 and ξ_2 must have uniform distributions on specific sets of integers. In Section 4, we develop several schemes for analytic and numerical construction of the distributions of the r.v.'s ξ_1 and ξ_2, which are expressed in terms of these specific sets of integers. In particular, we shall establish a connection between our main problem and the problem of ordered factorization of integers into primes. Our motivation for studying the problem is given in Section 2.2.

For simplicity of notation, all vectors in this paper are rows (rather than columns).

2. Reformulation of the main problem

2.1. Vectors and their convolutions

If F, F_1 and F_2 denote the distribution functions of ξ, ξ_1 and ξ_2, respectively, then (1) is equivalent to $F = F_1 \star F_2$, where $\star$ denotes the convolution of distribution functions.

* Corresponding author at: School of Mathematics, Cardiff University, CF24 4AG, UK.
E-mail addresses: ZhigljavskyAA@cardiff.ac.uk (A. Zhigljavsky), n.golyandina@spbu.ru (N. Golyandina), Svyatoslav.Gryaznov@statmod.ru (S. Gryaznov).

http://dx.doi.org/10.1016/j.spl.2016.06.006

If ξ_1 and ξ_2 satisfying (1) exist, then they have to be supported on the sets of integers $\{0, 1, \ldots, L\}$ and $\{0, 1, \ldots, K\}$ for some integral L and $K = N - L$; otherwise, we get an easy contradiction.

Denote

$$q_l = \mathrm{P}\{\xi_1 = l\}, \qquad r_k = \mathrm{P}\{\xi_2 = k\}; \quad l = 0, 1, \ldots, L; \; k = 0, 1, \ldots, K. \tag{2}$$

These numbers obviously satisfy

$$q_l \geq 0 \; (l = 0, 1, \ldots, L), \quad \sum_{l=0}^{L} q_l = 1; \qquad r_k \geq 0 \; (k = 0, 1, \ldots, K), \quad \sum_{k=0}^{K} r_k = 1. \tag{3}$$

In terms of numbers (2), the relation $F = F_1 \star F_2$ can be written as

$$\frac{1}{N+1} = \sum_{l=\max\{0,n-K\}}^{\min\{n,L\}} q_l r_{n-l} \quad \text{for all } n = 0, 1, \ldots, N. \tag{4}$$

Lemma 1. *The numbers (2) satisfying (3) and (4) exist if and only if there exist nonnegative numbers $a_0, \ldots, a_L$ and $b_0, \ldots, b_K$ satisfying*

$$1 = \sum_{l=\max\{0,n-K\}}^{\min\{n,L\}} a_l b_{n-l} \quad \textit{for all } n = 0, 1, \ldots, N. \tag{5}$$

Proof. (i) If the numbers (3) satisfying (4) exist, then we can simply set $a_l = (N+1)q_l$ $(l = 0, 1, \ldots, L)$ and $b_k = r_k$ $(k = 0, 1, \ldots, K)$.

(ii) Assume that there exist nonnegative numbers $a_0, \ldots, a_L$ and $b_0, \ldots, b_K$ satisfying (5). Set

$$n_1 = \sum_{l=0}^{L} a_l, \qquad n_2 = \sum_{k=0}^{K} b_k. \tag{6}$$

Since

$$\sum_{n=0}^{N} \sum_{l=\max\{0,n-K\}}^{\min\{n,L\}} a_l b_{n-l} = \sum_{l=0}^{L} a_l \cdot \sum_{k=0}^{K} b_k,$$

it follows from (5) that

$$n_1 n_2 = N + 1. \tag{7}$$

Set $q_l = a_l/n_1$ for all $l = 0, 1, \ldots, L$ and $r_k = b_k/n_2$ for all $k = 0, 1, \ldots, K$. Then q_l and r_k satisfy (3) and (4). The proof is complete. □

Consider the sets of numbers $a_0, \ldots, a_L$ and $b_0, \ldots, b_K$ from Lemma 1. Set

$$\mathbb{A} = (a_0, \ldots, a_L) \in \mathrm{R}_+^{L+1}, \qquad \mathbb{B} = (b_0, \ldots, b_K) \in \mathrm{R}_+^{K+1} \quad \text{and} \quad \mathbb{C}_N = (1, \ldots, 1) \in \mathrm{R}^{N+1}, \tag{8}$$

where for any M we denote by R_+^M the set of row-vectors of size M with nonnegative components. Then the relation (5) is simply

$$\mathbb{A} \star \mathbb{B} = \mathbb{C}_N. \tag{9}$$

We now formalize the main problem as follows.

The problem. *For given $N > 1$ and $0 \leq L \leq N$, establish the existence of vectors $\mathbb{A} \in \mathrm{R}_+^{L+1}$ and $\mathbb{B} \in \mathrm{R}_+^{N-L+1}$ so that the relation (9) holds and, if such vectors exist, construct them.*

Note that this problem is more difficult than the problem stated in the abstract as it requires deconvoluting the uniform distribution for any given N and L, whereas in the abstract only N was assumed to be fixed.

It follows from (9) that $a_0 b_0 = 1$. The validity of (9) for some vectors $\mathbb{A}$ and $\mathbb{B}$ is equivalent to the validity of $\widetilde{\mathbb{A}} \star \widetilde{\mathbb{B}} = \mathbb{C}_N$ with $\widetilde{\mathbb{A}} = c\mathbb{A}$ and $\widetilde{\mathbb{B}} = \mathbb{B}/c$, for any $c > 0$. We choose $\mathbb{A}$ so that $a_0 = 1$ (the equation $a_0 b_0 = 1$ does not allow a_0 to be zero). Then $a_0 b_0 = 1$ yields $b_0 = 1$. Summarizing this paragraph, without loss of generality we assume

$$a_0 = b_0 = 1. \tag{10}$$