Studia Scientiarum Mathematicarum Hungarica 49 (3), 283–300 (2012)
DOI: 10.1556/SScMath.49.2012.3.1207

JOINT DISTRIBUTIONS OF CIRCULAR RUNS OF VARIOUS LENGTHS BASED ON MULTI-COLOUR PÓLYA URN MODEL

SONALI BHATTACHARYA

Symbiosis Centre for Management and Human Resource Development, Plot 15, Phase-I,
Hinjewadi, Pune-411057, Maharashtra, India
e-mail: sonali_bhattacharya@scmhrd.edu

Communicated by E. Csáki

(Received January 12, 2010; accepted October 25, 2011)

Abstract

In this paper, we have used Eryilmaz's (2008) multi-colour Pólya urn model to obtain joint distributions of runs of t-types of exact lengths $(k_1, k_2, \ldots, k_t)$, at least lengths $(k_1, k_2, \ldots, k_t)$, non-overlapping runs of lengths $(k_1, k_2, \ldots, k_t)$ and overlapping runs of lengths $(k_1, k_2, \ldots, k_t)$ when counting of runs is done in a circular set-up. We have also derived joint distributions of longest runs of various types under similar conditions. Distributions of runs have found applications in fields of reliability of consecutive-k-out-of n: F system, consecutive k-out-of-r-from n: F system, start-up demonstration test, molecular biology, radar detection, time sharing systems and quality control. The literature is profound in discussion of marginal distribution and joint distribution of runs of various types under linear and circular set up using techniques like urn model with balls of two or more colours, probability generating function and compounding discrete distribution with suitable beta functions. Through this paper for first time effort been made to discuss joint distributions of runs of various lengths and types using Multi-colour urn model.

1. Introduction

Circular Distributions of order k are based on runs of specific length when the sampling units are arranged circularly i.e., two ends of the linear sequence of size n are joined together and the counting is made at any point on the circle, so that, when sample is ordered circularly if we start counting from sampling unit which is supposed to be the second unit (in case of linear ordering) now will be counted as the nth unit. Makri and Philippou [5] intro-

2000 *Mathematics Subject Classification.* Primary 60C05; Secondary 62E15.

Key words and phrases. Multi-colour Pólya urn model, Pólya–Eggenberger sampling, circular Pólya distributions of order k, 'balls-into-cells' technique.

duced the circular binomial distribution based on the non-overlapping and overlapping success runs. Charalambides [1] gave the factorial moments of these distributions. Philippou and Tripsiannis [7] used urn model to obtain circular Pólya-Distributions of order k based on both urn model and also by compounding binomial distributions by beta distribution. Sen et al. [6] used Pólya–Eggenberger sampling scheme to obtain marginal as well as joint distributions of Type-I circular Pólya distribution of order k based on non-overlapping success runs and denoted by $N^c_{n,k}$, Type-II circular Pólya distribution of order k based on overlapping success runs and denoted by $M^c_{n,k}$, Type-III circular Pólya distribution of order k based on success runs of exact length k and denoted by $E^c_{n,k}$ and Type-IV circular Pólya distribution of order k based on success runs of at least length k and denoted by $G^c_{n,k}$. Sen et al. [6] also derived the distributions of various circular inverse Pólya distributions of order k and a special waiting time distribution of order k, $X^{(c,l,k)}_{y,z,u,v}$, which is waiting time for first success of run of length k succeeding the y non-overlapping success runs of length l, z overlapping success runs of length l, u success runs of exact length l and v success runs of at least length l.

Eryilmaz [2] used urn model with multicolour scheme to obtain joint distributions of runs. In this multicolor urn scheme, a ball is drawn from the urn initially containing m_j balls of color j, $j = 1, 2, \dots, t$, and its color is noted. If a ball of color j is drawn at a stage, s balls of color j, $j = 1, 2, \dots, t$, are added to the urn. Drawing a ball of color j is considered as a trial of type j, $j = 1, 2, \dots, t$. This scheme is repeated n times and a sequence consisting of trials, namely $\{1, 2, \dots, t\}$, is derived. This model was used to obtain joint distributions of runs of various lengths. The model was found to have applications in start-up demonstrations and reliability of consecutive systems. In this paper [2] multi-colour urn model is used to obtain the joint distributions of four types of circular distributions of order $(k_1, k_2, \dots, k_t)$ based on non-overlapping success runs, overlapping success runs, runs exactly of some specific lengths and runs of at least some specific lengths. We have also obtained joint distributions of longest runs of type i $(i = 1, 2, \dots, t)$.

Following lemmas have been used repeatedly in our work.

Lemma 1. *The number of distributions of ‘r’ similar balls in n different cells such that no cell has more than $(k-1)$ balls is given as:*

$$A_{(k)}(r,n) = \sum_{j=0}^{\left[\frac{r}{k}\right]} (-1)^j \binom{n}{j} \binom{n+r-jk-1}{r-jk} \qquad \text{(Riordan, [8])}. \tag{1.1}$$

LEMMA 2. *The number of ways of distributions of r similar balls in 'n' different cells such that no cell has exactly k ball is given as:*

(1.2)
$$C_{(k)}(r,n) = \sum_{u=0}^{[\frac{r}{k}]} (-1)^j \binom{n}{u} \binom{n+r-uk-u-1}{r-uk} \qquad \text{(Sen et al., [6]).}$$

LEMMA 3. *Total number of ways for getting r_1 runs of type 1, r_2 runs of type 2, and so on, without two adjacent runs being of the same type is given as:*

(1.3)
$$F_t(r_1, r_2, \ldots, r_t)$$
$$= (-1)^r \sum_{m_1=1}^{r_1} \sum_{m_2=1}^{r_2} \cdots \sum_{m_t}^{r_t} (-1)^m \binom{r_1-1}{m_1-1} \cdots \binom{r_t-1}{m_t-1} \binom{m}{m_1, m_2, \ldots, m_t}$$

where, $r = \sum_{r=1}^{t} r_i \leqq n$, $m = \sum_{r=1}^{t} m_i$ [4].

LEMMA 4. *Let $A_1, A_2, \ldots, A_n$ be non-mutually exclusive and exhaustive events, then probability of union of $A_1, A_2, \ldots, A_n$ is given as:*

(1.4)
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_i P(A_i) - \sum_{\substack{(i,j)\\ i\neq j}} P(A_iA_j) + \sum_{\substack{(i,j,k)\\ i\neq j\neq k}} P(A_iA_jA_k)$$
$$+ \cdots + (-1)^{n-1} P(A_1A_2\ldots A_n) \qquad [3].$$

Also let $a^{(n,s)} = a(a+s)(a+2s)\ldots(a+\overline{n-1}s)$.

2. Circular Pólya distributions of order $(k_1, k_2, \ldots, k_t)$

2.1. Type-III circular Pólya distributions of order $(k_1, k_2, \ldots, k_t)$.

Let $E_{n,k_j}^{(j,C)}$ be the total number of runs of type j with length exactly equal to k_j $(j = 1, 2, \ldots, t)$.

Suppose we draw n balls from the urn and n_j $(j = 1, 2, \ldots, t)$ balls of colour j are drawn, such that $n_1 + n_2 + \cdots + n_t = n$ and there are r_j

$(j = 1, 2, \ldots, t)$ runs of j^{th} colour. Then we are interested in finding the probability

$$P\Big(E^{(1,C)}_{n,k_1} = x_1, E^{(2,C)}_{n,k_2} = x_2, \ldots, E^{(t,C)}_{n,k_t} = x_t\Big).$$

THEOREM 2.1. *Let* $(x_1 \geqq 0,\ x_2 \geqq 0,\ \ldots,\ x_t \geqq 0)$, $s \geqq -1$, $m_i > 0$ *and* $\sum_{i=1}^{t} m_i = m$.

(i) *For* $x_1 > 0$,

$$\text{(2.1)} \quad P\Big(E^{(1,C)}_{n,k_1} = x_1, E^{(2,C)}_{n,k_2} = x_2, \ldots, E^{(t,C)}_{n,k_t} = x_t\Big)$$
$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$
$$\left(F_t(r_2, r_3, \ldots r_t) \prod_{i=1}^{t} \left(\binom{r_i}{x_i} C_{(k_i - 1)}(n_i - k_i x_i - r_i + x_i, r_i - x_i) \right) \right.$$
$$\left. \times \left(A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1 \right) \left(\frac{x_1}{r_1} \right) \right) \left(\frac{\prod_{i=1}^{t} m_i^{(n_i, s)}}{m^{(n,s)}} \right) \right),$$

where, $\displaystyle\sum_{i=1}^{t} m_i = m$ *and* $\displaystyle\sum_{i=2}^{t} n_i = n - n_1.$

(ii) *For* $x_1 = 0$, $n_1 > 0$,

$$\text{(2.2)} \quad P\Big(E^{(1,C)}_{n,k_1} = x_1, E^{(2,C)}_{n,k_2} = x_2, \ldots, E^{(t,C)}_{n,k_t} = x_t\Big)$$
$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$
$$\left(F_t(r_2, r_3, \ldots, r_t) \prod_{i=2}^{t} \left(\binom{r_i}{x_i} C_{(k_i - 1)}(n_i - k_i x_i - r_i + x_i, r_i - x_i) \right) \right.$$
$$\times A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1 \right)$$

$$\times \sum_{\substack{u=1 \\ u\neq k_1}}^{n_1-r_1+1} \left(C_{(k_1-1)}(n_1-u-r_1+1, r_1-1)\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}} \right) \Bigg) ,$$

where, $\sum_{i=1}^{t} m_i = m$ *and* $\sum_{i=1}^{t} n_i = n.$

(iii)

$$P\Big(E_{n,k_1}^{(1,C)} = x_1, E_{n,k_2}^{(2,C)} = x_2, \ldots, E_{n,k_t}^{(t,C)} = x_t\Big) \tag{2.3}$$

$$= P\left(\bigcup_{\substack{j=1 \\ \sum_{j=1}^{t} n_j = n \\ x_j \geqq 0,\ n_j>0}}^{t} \Big(E_{n,k_1}^{(1,C)} = x_1, E_{n,k_2}^{(2,C)} = x_2, \ldots, E_{n,k_t}^{(t,C)} = x_t\Big) \right).$$

Proof. (i) For $x_1 > 0$, let us suppose we start the computation from a run of exact length k_1 of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \ldots, t)$. Total number of ways for getting r_2 runs of type 2, r_3 runs of type 3, and so on and r_k runs of type k, without two adjacent runs being of the same type is given as:

$$F_t(r_2, r_3, \ldots, r_t). \tag{2.4}$$

Number of selection of x_i runs of type i from r_i runs which has exactly k_i balls if type i $(i = 2, 3, \ldots, t)$ is given as

$$\binom{r_i}{x_i}. \tag{2.5}$$

The remaining $n_i - k_i x_i$balls of type i $(i = 2, 3, \ldots, t)$ are to be arranged in $r_i - x_i$ cells such that each cell has at least 1 ball and not exactly k_i balls of type i $(i = 2, 3, \ldots, t)$ in

$$C_{(k_i-1)}(n_i - k_i x_i - r_i + x_i, r_i - x_i) \quad \text{ways.} \tag{2.6}$$

$(r_1 - 1)$ runs type 1 are to be distributed in the $\big(\sum_{i=2}^{t} r_i - 2\big)$ cells created by the $\sum_{i=2}^{t} r_i$ runs of types i $(i = 2, 3, \ldots, t)$, such that no cell contain more than 1 run in $A_{(2)}\big(r_1 - 1, \sum_{i=2}^{t} r_i - 1\big)$ ways.

Number of selections of $(x_1 - 1)$ runs of type 1 from $(r_1 - 1)$ runs which has exactly k_1 balls of type 1 is given as $\binom{r_1-1}{x_1-1}$ ways.

The remaining $n_1 - k_1x_1$ balls of type 1 are to be arranged in $r_1 - x_1$ cells such that each cell has at least 1 ball and not exactly k_1 balls of type 1 in $C_{(k_1-1)}(n_1 - k_1x_1 - r_1 + x_1, r_1 - x_1)$ ways.

Number of drawing n_i, $(i = 2, 3, \dots, t)$ balls of type i from an urn containing m_i balls of type i $(i = 2, 3, \dots, t)$ by Pólya–Eggenberger sampling scheme is given as:

$$\frac{\prod_{i=1}^t m_i^{(n_i,s)}}{m^{(n,s)}}, \qquad \sum_{i=1}^t m_i = m \text{ and } \sum_{i=1}^t n_i = n. \tag{2.7}$$

Hence, by combining (2.4), (2.5), (2.6), and (2.7) we get the probability given in (2.1)

(ii) For $x_1 = 0$, let us suppose we start the computation from a run of exact length $u(\neq k_1) = 1, 2, \dots, n_1 - r_1 + 1$ of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \dots, t)$.

The remaining $n_1 - u$ balls of type 1 are to be arranged in $r_1 - 1$ cells, such that each cell has at least 1 ball and not exactly k_1 balls of type 1 in $C_{(k_1-1)}(n_1 - u - r_1 + 1, r_1 - 1)$ ways.

Hence, we get the probability given in (2.2). □

2.2. Type-IV circular Pólya distributions of order $(k_1, k_2, \dots, k_t)$.

Let $G_{n,k_j}^{(j,C)}$ be the total number of runs of type j with length at least equal to k_j $(j = 1, 2, \dots, t)$.

Suppose we draw n balls from the urn and n_j $(j = 1, 2, \dots, t)$ balls of colour j are drawn, such that $n_1 + n_2 + \dots + n_t = n$ and there are r_j $(j = 1, 2, \dots, t)$ runs of j^{th} colour. Then, we are interested in finding the probability

$$P\Big(G_{n,k_1}^{(1,C)} = x_1, G_{n,k_2}^{(2,C)} = x_2, \dots, G_{n,k_t}^{(t,C)} = x_t\Big).$$

Theorem 2.2. *Let* $(x_1 \geqq 0,\ x_2 \geqq 0,\ \dots,\ x_t \geqq 0)$, $s \geqq -1$, $m_i > 0$ *and* $\sum_{i=1}^t m_i = m$.

(i) *For* $x_1 > 0$,

$$P\Big(G_{n,k_1}^{(1,C)} = x_1, G_{n,k_2}^{(2,C)} = x_2, \dots, G_{n,k_t}^{(t,C)} = x_t\Big) \tag{2.8}$$
$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \dots, r_t) \prod_{i=1}^{t} \binom{r_i}{x_i} \Bigg(\sum_{v_i=0}^{n_i - k_i x_i - r_i - x_i} \binom{x_i + v_i - 1}{v_i}$$

$$\times A_{(k_i-1)}(n_i - k_i x - r_i + x_i - v_i, r_i - x_i) \Bigg)$$

$$\times A_{(2)} \Bigg(r_1 - 1, \sum_{i=2}^{t} r_i - 1 \Bigg) \left(\frac{x_1}{r_1} \right) \frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}},$$

where, $\sum_{i=1}^{t} m_i = m$ *and* $\sum_{i=2}^{t} n_i = n - n_1.$

(ii) *For* $x_1 = 0$, $n_1 > 0$

(2.9) $$P\Big(G_{n,k_1}^{(1,C)} = x_1, G_{n,k_2}^{(2,C)} = x_2, \dots, G_{n,k_t}^{(t,C)} = x_t \Big)$$

$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \dots, r_t) \prod_{i=2}^{t} \Bigg(\binom{r_i}{x_i} \sum_{v_i=0}^{n_i - k_i x_i - r_i - x_i} \binom{x_i + v_i - 1}{v_i}$$

$$A_{(k_i-1)}(n_i - k_i x - r_i + x_i - v_i, r_i - x_i) \Bigg)$$

$$\times \Bigg(A_{(2)} \Bigg(r_1 - 1, \sum_{i=2}^{t} r_i - 1 \Bigg)$$

$$\times \sum_{\substack{u=1 \\ u \neq k_1}}^{n_1 - r_1 + 1} A_{(k_1-1)}(n_1 - u - r_1 + 1, r_1 - 1) \frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}} \Bigg),$$

where, $\sum_{i=1}^{t} m_i = m$ *and* $\sum_{i=1}^{t} n_i = n.$

(iii)

(2.10)

$$P\Big(G^{(1,C)}_{n,k_1} = x_1, G^{(2,C)}_{n,k_2} = x_2, \ldots, G^{(t,C)}_{n,k_t} = x_t\Big)$$
$$= P\left(\bigcup_{\substack{j=1 \\ \sum_{j=1}^{t} n_j = n,\ x_j \geqq 0,\ n_j > 0}}^{t} \Big(G^{(1,C)}_{n,k_1} = x_1, G^{(2,C)}_{n,k_2} = x_2, \ldots, G^{(t,C)}_{n,k_t} = x_t\Big)\right).$$

PROOF. For $x_1 > 0$, let us suppose we start the computation of a run of at least length k_1 of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \ldots, t)$. Total number of ways for getting r_2 runs of type 2, r_3 runs of type 3, and so on r_k runs of type k, without two adjacent runs being of the same type is given as:

$$F_t(r_2, r_3, \ldots, r_t). \tag{2.11}$$

Number of selection of x_i runs of type i from r_i runs which has at least k_i balls if type i $(i = 2, 3, \ldots, t)$ is given as $\binom{r_i}{x_i}$.

Since each of the r_i cells having balls of type i $(i = 2, 3, \ldots, t)$ is required to have at least 1 ball, let us assume that there are v_i $(= 0, 1, \ldots, n_i - k_i x_i - r_i + x_i)$ balls which are to be distributed in x_i cells in $\binom{x_i + v_i - 1}{v_i}$ ways.

The remaining $n_i - k_i x - r_i + x_i - v_i$ balls of type i $(i = 2, 3, \ldots, t)$ are to be arranged in $r_i - x_i$ cells such that each cell has at the most $(k_i - 2)$ balls of type i $(i = 2, 3, \ldots, t)$ in $A_{(k_i-1)}(n_i - k_i x - r_i + x_i - v_i, r_i - x_i)$ ways.

$(r_1 - 1)$ runs type1 are to be distributed in the $\Big(\sum_{i=2}^{t} r_i - 2\Big)$ cells created by the $\sum_{i=2}^{t} r_i$ runs of types i $(i = 2, 3, \ldots, t)$, such that no cell contain more than one run in $A_{(2)}\Big(r_1 - 1, \sum_{i=2}^{t} r_i - 1\Big)$ ways.

Number of selections of $(x_1 - 1)$ runs of type 1 from $(r_1 - 1)$ runs, which has at least k_1 balls of type 1 is given as $\binom{r_1 - 1}{x_1 - 1}$ ways.

Since each of the r_1 cells having balls of type 1 is required to have at least 1 ball, let us assume there are v_1 $(= 0, 1, \ldots, n_1 - k_1 x_1 - r_1 + x_1)$ balls that are to be distributed in x_1 cells in $\binom{x_1 + v_1 - 1}{v_1}$ ways.

The remaining $n_1 - k_1 x_1 - r_1 + x_1 - v_1$ balls of type 1 are to be arranged in $r_1 - x_1$ cells such that each cell has at the most $(k_1 - 2)$ balls of type 1 in $A_{(k_1-1)}(n_1 - k_1 x_1 - r_1 + x_1 - v_1, r_1 - x_1)$ ways.

Number of drawing n_i, $(i = 2, 3, \ldots, t)$ balls of type i from an urn containing m_i balls of type i $(i = 2, 3, \ldots, t)$ by Pólya–Eggenberger sampling scheme is given as

$$\text{(2.12)}\qquad \frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}}, \qquad \sum_{i=1}^{t} m_i = m \text{ and } \sum_{i=1}^{t} n_i = n.$$

Hence, we get the probability given in (2.8).

(ii) For $x_1 = 0$, let us suppose we start the computation of from a run of length $u(\neq k_1) = 1, 2, \ldots, n_1 - r_1 + 1)$ of type 1 flanked on both sides balls of any one of the types i $(i = 2, 3, \ldots, t)$.

The remaining $n_1 - u$ balls of type 1 are to be arranged in $r_1 - 1$ cells, such that each cell has at least 1 ball and at most $k_1 - 1$ balls of type 1 in $A_{(k_1-1)}(n_1 - u - r_1 + 1, r_1 - 1)$ ways.

Hence, we get the probability given in (2.9). □

2.3. Type-I circular Pólya distributions of order $(k_1, k_2, \ldots, k_t)$.

Let $N_{n,k_j}^{(j,C)}$ be the number of non-overlapping runs of type j of length k_j $(j = 1, 2, \ldots, t)$.

Suppose we draw n balls from the urn and n_j $(j = 1, 2, \ldots, t)$ balls of colour j are drawn, such that $n_1 + n_2 + \cdots + n_t = n$ and there are r_j $(j = 1, 2, \ldots, t)$ runs of j^{th} colour. Then, we are interested in finding the probability

$$P\Big(N_{n,k_1}^{(1,C)} = x_1, N_{n,k_2}^{(2,C)} = x_2, \ldots, N_{n,k_t}^{(t,C)} = x_t\Big).$$

THEOREM 2.3. *Let* $\big(0 \leqq x_i \leqq \big[\frac{n}{k}\big],\ i = 1, 2, \ldots, t\big)$, $s \geqq -1$, $m_i > 0$ *and* $\sum_{i=1}^{t} m_i = m$.

(i) *For* $x_1 > 0$,

(2.13)

$$P\Big(N_{n,k_1}^{(1,C)} = x_1, N_{n,k_2}^{(2,C)} = x_2, \ldots, N_{n,k_t}^{(t,C)} = x_t\Big)$$

$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \ldots, r_t) \prod_{i=1}^{t} \Bigg(\sum_{u_i=1}^{\min(r_i, x_i)} \binom{r_i}{u_i} \binom{x_i - 1}{u_i - 1} (\delta_i)$$

$$\times \sum_{v_i=0}^{n_i-k_ix_i-r_i+u_i} A_{(k_i)}(v_i,u_i)A_{(k_i-1)}(n_i-k_ix_i-r_i+u_i-v_i,r_i-u_i)\Bigg)$$

$$\times A_{(2)}\left(r_1-1,\sum_{i=2}^{t} r_i-1\right)\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}},$$

$$\textit{where,}\qquad \sum_{i=1}^{t} m_i=m,\quad \sum_{i=2}^{t} n_i=n-n_1 \quad \textit{and}\quad \delta_i=\begin{cases}\dfrac{u_1}{r_1}, & \textit{if}\ \ i=1\\ 1, & \textit{otherwise.}\end{cases}$$

(ii) *For* $x_1=0$, $n_1>0$,

(2.14)

$$P\Big(N_{n,k_1}^{(1,C)}=x_1,N_{n,k_2}^{(2,C)}=x_2,\dots,N_{n,k_t}^{(t,C)}=x_t\Big)$$

$$=\sum_{n_1\geqq r_1}\ \sum_{n_2\geqq r_2}\cdots\sum_{n_t\geqq r_t}\ \sum_{r_1\geqq x_1}\ \sum_{r_2\geqq x_2}\cdots\sum_{r_t\geqq x_t}$$

$$F_t(r_2,r_3,\dots,r_t)\left(\prod_{i=2}^{t}\sum_{u_i=1}^{\min(r_i,x_i)}\binom{r_i}{u_i}\binom{x_i-1}{u_i-1}\sum_{v_i=0}^{n_i-k_ix_i-r_i+u_i} A_{(k_i)}(v_i,u_i)\right.$$

$$\left.\times A_{(k_i-1)}(n_i-k_ix_i-r_i+u_i-v_i,r_i-u_i)\right)$$

$$\times A_{(2)}\left(r_1-1,\sum_{i=2}^{t} r_i-1\right)\sum_{u=1}^{k_1-\max(r_1,1)} A_{(k_1-2)}(n_1-u-r_1+1,r_1-1)$$

$$\times\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}},\qquad \textit{where,}\qquad \sum_{i=1}^{t} m_i=m\quad\textit{and}\quad\sum_{i=2}^{t} n_i=n-n_1.$$

(iii)

(2.15)

$$P\Big(N_{n,k_1}^{(1,C)}=x_1,N_{n,k_2}^{(2,C)}=x_2,\dots,N_{n,k_t}^{(t,C)}=x_t\Big)$$

$$= P\left(\bigcup_{\substack{j=1 \\ \sum_{j=1}^{t} n_j = n,\ x_j \geqq 0,\ n_j > 0}}^{t} \left(N_{n,k_1}^{(1,C)} = x_1, N_{n,k_2}^{(2,C)} = x_2, \ldots, N_{n,k_t}^{(t,C)} = x_t \right) \right).$$

PROOF. For $x_1 > 0$, let us suppose we start the computation from a run of at least length k_1 of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \ldots, t)$. Total number of ways for getting r_2 runs of type 2, r_3 runs of type 3, and so on r_k runs of type k, without two adjacent runs being of the same type is given as:

$$F_t(r_2, r_3, \ldots, r_t). \tag{2.16}$$

Suppose there are u_i runs of type i which contributes to x_i non-overlapping runs of length k_i and of type i $(i = 2, 3, \ldots, t)$. Number of ways of selection of these u_i runs of type i from r_i runs and then distributing the x_i non-overlapping runs of type i in these u_i cells (by Balls-into-cells technique) can be given as:

$$\binom{r_i}{u_i}\binom{x_i - 1}{u_i - 1} \quad \text{where,} \quad 1 \leqq u_i \leqq \min(r_i, x_i) \quad (i = 2, 3, \ldots, t). \tag{2.17}$$

Since each of the remaining $r_i - u_i$ runs having balls of type i $(i = 2, 3, \ldots, t)$ is required to have at least 1 ball of type i, let us assume there are v_i balls that are to be distributed in u_i cells such that each cell has at the most $(k_i - 1)$ balls by 'balls-into-cells' technique can be given as

$$A_{(k_i)}(v_i, u_i). \tag{2.18}$$

The remaining $n_i - k_i x_i - r_i + u_i - v_i$ balls can be distributed in $r_i - u_i$ runs such that no runs have more than $(k_i - 2)$ balls in $A_{(k_i-1)}(n_i - k_i x_i - r_i + u_i - v_i, r_i - u_i)$, $(i = 2, 3, \ldots, t)$ ways.

$(r_1 - 1)$ runs type 1 are to be distributed in the $\left(\sum_{i=2}^{t} r_i - 2\right)$ cells created by $\sum_{i=2}^{t} r_i$ runs of types i $(i = 2, 3, \ldots, t)$ such that no cell contain more than 1 run in $A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1\right)$ ways.

Suppose there are u_1 runs of type 1 which contributes to x_1 non-overlapping runs of length k_1 and of type 1. Number of ways of selection of $u_1 - 1$ runs of type 1 from $r_1 - 1$ runs and then distributing the $u_1 - x_1$ balls of type 1 in these u_1 cells (by Balls-into-cells technique) can be given as:

$$\binom{r_1 - 1}{u_1 - 1}\binom{x_1 - 1}{u_1 - 1}. \tag{2.19}$$

Assuming distribution of v_1 balls in u_1 cells such that each cell has at the most (k_1-1) balls by balls-into cells technique can be given as

$$A_{(k_1)}(v_1, u_1). \tag{2.20}$$

Since each of the remaining r_1-u_1 runs having balls of type 1 is required to have at least 1 ball of type 1, the remaining $n_1-k_1x_1-r_1+u_1-v_1$ balls can be distributed in r_1-u_1 runs such that no runs have more than (k_1-2) balls in $A_{(k_1-1)}(n_1-k_1x_1-r_1+u_1-v_1, r_1-u_1)$ ways.

Number of drawing n_i, i $(i=1,2,3,\dots,t)$ balls of type i from an urn containing m_i balls of type i $(i=2,3,\dots,t)$ by Pólya–Eggenberger sampling scheme is given as

$$\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}}, \qquad \sum_{i=1}^{t} m_i = m \;\text{ and }\; \sum_{i=1}^{t} n_i = n. \tag{2.21}$$

Hence, we get the probability given in (2.13).

(ii) For $x_1 = 0$, let us suppose we start the computation from a run of length u $\big(u = 1,2,\dots,\big(k_1 - \max(r_1,1)\big)\big)$ of type 1 flanked on both sides by balls of any one of the types i $(i=2,3,\dots,t)$.

The remaining n_1-u balls of type 1 are to be arranged in r_1-1 cells such that each cell has at least 1 ball and at most k_1-1 balls of type 1 in $A_{(k_1-2)}(n_1-u-r_1+1, r_1-1)$ ways.

Hence, we get the probability given in (2.14). □

2.4. Type-IV circular Pólya distributions of order $(k_1, k_2, \dots, k_t)$.

Let $M_{n,k_j}^{(j,C)}$ be the number of overlapping runs of type j of length k_j $(j=1,2,\dots,t)$.

Suppose we draw n balls from the urn and n_j $(j=1,2,\dots,t)$ balls of colour j are drawn, such that $n_1+n_2+\dots+n_t = n$ and there are r_j $(j=1,2,\dots,t)$ runs of j^{th} colour. Then, we are interested in finding the probability

$$P\Big(M_{n,k_1}^{(1,C)} = x_1, M_{n,k_2}^{(2,C)} = x_2, \dots, M_{n,k_t}^{(t,C)} = x_t\Big).$$

THEOREM 2.4. *Let* $(0 \leqq x_i \leqq n-k+1,\ i=1,2,\dots,t)$*,* $s \geqq -1$*,* $m_i > 0$ *and* $\sum_{i=1}^{t} m_i = m$*.*

(i) *For* $x_1 > 0$,

(2.22)

$$P\Big(M^{(1,C)}_{n,k_1} = x_1, M^{(2,C)}_{n,k_2} = x_2, \ldots, M^{(t,C)}_{n,k_t} = x_t\Big)$$

$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \ldots, r_t) \prod_{i=1}^{t} \left(\sum_{u_i=1}^{\min(r_i, x_i)} \binom{r_i}{u_i} \binom{x_i - 1}{u_i - 1} (\delta_i) \right.$$

$$\left. \times A_{(k_1-1)}\big(n_1 - (k_1 - 2)u_1 - x_1 - r_1, r_1 - u_1\big) \right)$$

$$A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1\right) \left(\frac{\prod_{i=1}^{t} m_i^{(n_i, s)}}{m^{(n,s)}} \right),$$

$$\textit{where,} \quad \sum_{i=1}^{t} m_i = m, \quad \sum_{i=2}^{t} n_i = n - n_1 \quad \textit{and} \quad \delta_i = \begin{cases} \dfrac{u_1}{r_1}, & \textit{if } \ i = 1 \\ 1, & \textit{otherwise.} \end{cases}$$

(ii) *For* $x_1 = 0$, $n_1 > 0$,

(2.23)
$$P\Big(M^{(1,C)}_{n,k_1} = x_1, M^{(2,C)}_{n,k_2} = x_2, \ldots, M^{(t,C)}_{n,k_t} = x_t\Big)$$

$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \ldots, r_t) \prod_{i=2}^{t} \left(\sum_{u_i=1}^{\min(r_i, x_i)} \binom{r_i}{u_i} \binom{x_i - 1}{u_i - 1} \right.$$

$$\left. \times A_{(k_i-1)}(n_i - (k_i - 2)u_i - x_i - r_i, r_i - u_i) \right)$$

$$\times A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1\right)$$

$$\times \left(\sum_{u=1}^{k_1-\max(r_1,1)} A_{(k_1-2)}(n_1-u-r_1+1, r_1-1) \right) \left(\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}} \right),$$

$$\textit{where,} \qquad \sum_{i=1}^{t} m_i = m \quad \textit{and} \quad \sum_{i=2}^{t} n_i = n - n_1.$$

(iii)

(2.24)

$$P\Big(M_{n,k_1}^{(1,C)} = x_1, M_{n,k_2}^{(2,C)} = x_2, \ldots, M_{n,k_t}^{(t,C)} = x_t\Big)$$

$$= P\left(\bigcup_{\substack{j=1 \\ \sum_{j=1}^{t} n_j = n,\ x_j \geqq 0,\ n_j > 0}}^{t} \Big(M_{n,k_1}^{(1,C)} = x_1, M_{n,k_2}^{(2,C)} = x_2, \ldots, M_{n,k_t}^{(t,C)} = x_t\Big) \right).$$

Proof. For $x_1 > 0$, let us suppose we start the computation from a run of at least length k_1 of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \ldots, t)$. Total number of ways for getting r_2 runs of type 2, r_3 runs of type 3, and so on r_k runs of type k, without two adjacent runs being of the same type is given as:

$$F_t(r_2, r_3, \ldots r_t). \tag{2.25}$$

Suppose there are u_i runs of type i which contributes to x_i non-overlapping runs of length k_i and of type i $(i = 2, 3, \ldots, t)$. Each of these u_i runs of type i will contribute to at least u_i runs of length k_i. Hence, number of ways of selecting u_i runs of type i and then distributing the $x_i - u_i$ balls of type 'i' in these u_i cells (by Balls-into-cells technique) can be given as:

$$\binom{r_i}{u_i}\binom{x_i-1}{u_i-1} \qquad \text{where,} \qquad 1 \leqq u_i \leqq \min(r_i, x_i) \quad (i = 2, 3, \ldots, t). \tag{2.26}$$

The remaining $n_i - u_i k_i - (x_i - u_i) - r_i + u_i$ balls are to be distributed in $r_i - u_i$ cells such that there are at most $(k_i - 2)$ balls in a cell in $A_{(k_i-1)}\big(n_i - (k_i-2)u_i - x_i - r_i, r_i - u_i\big)$ ways.

$(r_1 - 1)$ runs of type 1 are to be distributed in $\left(\sum_{i=2}^{t} r_i - 2\right)$ cells created by $\sum_{i=2}^{t} r_i$ runs of types i $(i = 2, 3, \ldots, t)$ such that no cell contain more than 1 run in $A_{(2)}\big(r_1 - 1, \sum_{i=2}^{t} r_i - 1\big)$ ways.

Suppose there are u_1 runs of type 1 which contributes to $x_1$overlapping runs of length k_1 and of type 1. Number of ways of selection of $u_1 - 1$ runs of type 1 from $r_1 - 1$ runs and then distributing the $x_1 - u_1$ balls of type 1 in these u_1 cells (by Balls-into-cells technique) can be given as:

$$\binom{r_1 - 1}{u_1 - 1}\binom{x_1 - 1}{u_1 - 1}. \tag{2.27}$$

The remaining $n_1 - u_1k_1 - (x_1 - u_1) - r_1 + u_1$ balls are to be distributed in $r_1 - u_1$ cells such that there are at most $(k_i - 2)$ balls in a cell in $A_{(k_1-1)}\big(n_1 - (k_1 - 2)u_1 - x_1 - r_1, r_1 - u_1\big)$ ways.

Number of ways of drawing n_i, i $(i = 1, 2, 3, \dots, t)$ balls of type i from an urn containing m_i balls of type i $(i = 2, 3, \dots, t)$ by Pólya–Eggenberger sampling scheme is given as:

$$\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}}, \qquad \sum_{i=1}^{t} m_i = m \text{ and } \sum_{i=1}^{t} n_i = n. \tag{2.28}$$

Hence, we get the probability given in (2.22).

(ii) For $x_1 = 0$, let us suppose we start the computation from a run of length u $\Big(u = 1, 2, \dots, \big(k_1 - \max(r_1, 1)\big)\Big)$ of type 1 flanked on both sides by balls of any one of the types i $(i = 2, 3, \dots, t)$.

The remaining $n_1 - u$ balls of type 1 are to be arranged in $r_1 - 1$ cells such that each cell has at least 1 ball and at most $k_1 - 1$ balls of type 1 in $A_{(k_1-2)}(n_1 - u - r_1 + 1, r_1 - 1)$ ways.

Hence, we get the probability given in (2.23). □

2.5. Joint Distribution of Longest Runs of type i $(i = 1, 2, \dots, t)$

Let $L_j^{(C)}$be the longest run of type j $(j = 1, 2, \dots, t)$.

Suppose we draw n balls from the urn and n_j $(j = 1, 2, \dots, t)$ balls of colour j are drawn, such that $n_1 + n_2 + \cdots + n_t = n$ and there are r_j $(j = 1, 2, \dots, t)$ runs of j^{th} colour. Then we are interested in finding the probability

$$P\Big(L_1^{(C)} = l_1, L_2^{(C)} = l_2, \dots, L_t^{(C)} = l_t\Big).$$

THEOREM 2.5. *Let* $(0 \leqq l_i \leqq n,\ i = 1, 2, \dots t)$, $\sum_{i=1}^{t} l_i \leqq n$, $s \geqq -1$, $m_i > 0$ *and* $\sum_{i=1}^{t} m_i = m$.

(i) *For* $n_1 > 0$,

(2.29)

$$P\Big(L_1^{(C)} = l_1, L_2^{(C)} = l_2, \ldots, L_t^{(C)} = l_t\Big)$$

$$= \sum_{n_1 \geqq r_1} \sum_{n_2 \geqq r_2} \cdots \sum_{n_t \geqq r_t} \sum_{r_1 \geqq x_1} \sum_{r_2 \geqq x_2} \cdots \sum_{r_t \geqq x_t}$$

$$F_t(r_2, r_3, \ldots, r_t) \sum_{u_i=1}^{\min\left(r_i, [\frac{n_i}{l_i}]\right)} \left(\binom{r_i}{u_i} A_{(l_i-2)}(n_i - l_i u_i - r_i + u_i, r_i - u_i) \right.$$

$$\left. \times A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1\right)\right)$$

$$\times A_{(l_1)}(n_1 - l_1 - r_1 + 1, r_1 - 1) \frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}},$$

where, $\displaystyle\sum_{i=1}^{t} m_i = m$ *and* $\displaystyle\sum_{i=2}^{t} n_i = n - n_1$.

(ii)

$$P\Big(L_1^{(C)} = l_1, L_2^{(C)} = l_2, \ldots, L_t^{(C)} = l_t\Big) \tag{2.30}$$

$$= P\left(\bigcup_{\substack{j=1 \\ \sum_{j=1}^{t} n_j = n,\ n_j > 0}}^{t} \Big(L_1^{(C)} = l_1, L_2^{(C)} = l_2, \ldots, L_t^{(C)} = l_t\Big) \right).$$

Proof. For $n_1 > 0$, let us suppose we start the computation from a run of longest run of type 1 flanked on both sides by balls of any one of the types i ($i = 2, 3, \ldots, t$). Total number of ways for getting r_2 runs of type 2, r_3 runs of type 3, and so on r_k runs of type k, without two adjacent runs being of the same type is given as:

$$F_t(r_2, r_3, \ldots, r_t). \tag{2.31}$$

Number of ways of distribution of n_i balls of type i in r_i cells such that at least one cell has exactly l_i balls and rest of the cells should have at least 1 and at most $(l_i - 1)$ balls is given as

$$\sum_{u_i=1}^{(r_i,[\frac{n_i}{l_i}])} \binom{r_i}{u_i} A_{(l_i-2)}(n_i - l_i u_i - r_i + u_i, r_i - u_i) \quad \text{ways.} \tag{2.32}$$

$(r_1 - 1)$ runs of type 1 are to be distributed in the $\left(\sum_{i=2}^{t} r_i - 2\right)$ cells created by $\sum_{i=2}^{t} r_i$ runs of types i $(i = 2, 3, \ldots, t)$ such that no cell contain more than 1 run in $A_{(2)}\left(r_1 - 1, \sum_{i=2}^{t} r_i - 1\right)$ ways.

Number of ways of distributing $n_1 - l_1 - r_1 + 1$ balls in $r_1 - 1$ cells such that each cell has at most $(l_1 - 1)$ cells can be obtained as:

$$A_{(l_1)}(n_1 - l_1 - r_1 + 1, r_1 - 1). \tag{2.33}$$

Number of ways of drawing n_i, i $(i = 1, 2, 3, \ldots, t)$ balls of type i from an urn containing m_i balls of type i $(i = 2, 3, \ldots, t)$ by Pólya–Eggenberger sampling scheme is given as

$$\frac{\prod_{i=1}^{t} m_i^{(n_i,s)}}{m^{(n,s)}}, \qquad \sum_{i=1}^{t} m_i = m \ \text{ and } \ \sum_{i=1}^{t} n_i = n. \tag{2.34}$$

Hence, we get probability given in (2.29). □

Remark. For $t = 1$, Theorems 2.1, 2.2, 2.3, 2.4 and 2.5 will lead to circular Pólya distributions of order k discussed in [6].

REFERENCES

[1] Charalambides, Ch. A., *Combinatorial methods in discrete distributions,* John Wiley & Sons (New York, 2005).

[2] Eryilmaz, S., Run statistics defined on multicolor urn model, *J. Appl. Prob.,* **45** (2008), 1007–1023.

[3] Feller, W., *An introduction to probability theory and applications,* Vol. 1 (3^{rd} edition), Wiley (New York, 1968).

[4] Kong, Y., Distribution of runs and longest runs: a new generating function approach, *J. Amer. Statist. Assoc.,* **101** (2006), no. 1, 253–1263.

[5] Makri, F. S. and Philippou, A. N., Binomial distributions of order k on the circle, in: *Runs and Patterns in Probability: Selected Papers* (eds.: Godbole, A. P. and Papastravridis, S. G.), Kluwer (1994), 65–81.

[6] Sen, K., Agarwal, M. and Bhattacharya, S., On circular distributions of order k based on Pólya–Eggenberger sampling scheme, *J. Math. Sci.,* **2** (2003), 34–54.

[7] Philippou, A. N. and Tripsiannis, G. A., Circular Polya distributions of order k, *International Journal of Mathematics and Mathematical Sciences,* No. 25 (2003), 1563–1575.

[8] Riordan, J., *An Introduction to Combinatorial Analysis,* John Wiley and Sons, Inc., New York, Chapman and Hall (1958): *MR* **20**#3077. Reprinted: Princeton University Press, Princeton, NJ, 1980, **m81e**:05002.