COMMUN.STATIST.-THEORY METH.,26(8),2009-2019(1997)

A GENERALIZATION OF THE DONNELLY-TAVARÉ-GRIFFITHS FORMULA

Hajime Yamato

Dept of Math, Fac of Sci, Kagoshima University, Kagoshima 890, Japan

Key Words and Phrases: Random partition; Urn model; Waring distribution

ABSTRACT

We give a generalization of the one form of the Donnelly-Tavaré-Griffiths(DTG) formula. It contains not only this DTG formula but also the conditional distribution of the formula given the some first components. We can construct it using an simple urn model. For the generalization of the DTG formula, its probability distributions including marginal and conditional distributions, the related statistics and their asymptotic properties are discussed.

1. INTRODUCTION

Let $\mathcal{C}_n$ denote the set of all ordered partitions of a positive integer n, that is,

$$\mathcal{C}_n = \{(c_1, ..., c_k) : 1 \leq k \leq n, c_i > 0\, (i = 1, ..., k) \text{ and } c_1 + \cdots + c_k = n\}.$$

As a probability distribution on $\mathcal{C}_n$, the Donnelly-Tavaré-Griffiths formula is well-known (Ewens (1990)). The one form of this formula is a probability distribution of random ordered partition $C_n = (C_{n1}, ..., C_{nk})$ on $\mathcal{C}_n$ defined by

$$(1) \qquad P(C_n = (c_1, ..., c_k)) = \frac{\alpha^k}{\alpha^{[n]}} \cdot \frac{n!}{c_k(c_k + c_{k-1}) \cdots (c_k + c_{k-1} + \cdots + c_1)},$$

where α is a positive constant, $1 \leq k \leq n$, $(c_1, ..., c_k) \in \mathcal{C}_n$ and $\alpha^{[n]} = \alpha(\alpha + 1) \cdots (\alpha + n - 1)$. This distribution is obtained by the size-biased permutation of the Ewens sampling formula(Donnelly and Tavaré (1986)). Joyce and Tavaré (1987) uses the linear birth process with immigration to derive the distribution. The distribution can be characterized as the distribution of frequencies of order statistics from GEM distribution (Donnelly (1986), Donnelly and Tavaré (1991), Sibuya and Yamato (1995)). The distribution can be derived by using Pólya-like urn (Hoppe (1984), Sibuya and Yamato (1995)). It has equivalent models: random clustering process (Sibuya (1993)), urn with a continuum of colors and the sampling from Ferguson's Dirichlet process with a continuous parameter (Blackwell and MacQueen (1973), Yamato (1993)). The distribution can be also derived by using Pitman's Chinese restaurant process(see, for example, Donnelly and Tavaré (1990))

The other is a probability distribution of random ordered partition $D_n = (D_{n1}, ..., D_{nk})$ on $\mathcal{C}_n$ defined by

$$(2) \qquad P(D_n = (d_1, ..., d_k)) = \frac{\alpha^k}{\alpha^{[n]}} \cdot \frac{n!}{d_1(d_1 + d_2)\cdots(d_1 + d_2 + \cdots + d_k)},$$

where $(d_1, ..., d_k) \in \mathcal{C}_n$. This distribution is obtained from the n-coalescent with mutation (Donnelly and Tavaré (1986)). Ethier (1990) derives this distribution using diffusion model. Yamato (1996) gives the urn model yielding this distribution and its properties. Distinguishing between the distributions given by (1) and (2), we shall say the distribution given by (2) Donnelly-Tavaré-Griffiths II formula and abbreviate it DTGII(n,α).

The conditional distribution of $C_n = (C_{n1}, ..., C_{nk})$ given $(C_{n1}, ..., C_{nr}) = (c_1, ..., c_r)$ is the same distribution as C_{n-c_0}, where a positive integer r is fixed and $c_0 = c_1 + \cdots + c_r$. For D_n having DTGII, the conditional distribution of $D_n = (D_{n1}, ..., D_{nk})$ given $(D_{n1}, ..., D_{nr}) = (d_1, ..., d_r)$ is not DTG II. We shall introduce a generalization of DTG II such that this conditional distribution and the distribution of D_n belong to the same class of distributions. The generalization of DTGII, which we introduce, is given by a probability distribution of random ordered partition $D_n = (D_{n1}, ..., D_{nk})$ on $\mathcal{C}_n$ defined by

$$(3) \quad P(D_n = (d_1, ..., d_k)) = \frac{\alpha^{k-1}}{(\alpha + \beta + 1)^{[n-1]}} \cdot \frac{(\beta + 1)^{[n]}}{(\beta + d_1)(\beta + d_1 + d_2)\cdots(\beta + d_1 + \cdots + d_k)},$$

where α is a positive constant, β is a non-negative constant, $1 \leq k \leq n$ and $(d_1, ..., d_k) \in \mathcal{C}_n$. We shall call this distribution generalized Donnelly-Tavaré-Griffiths II formula and abbreviate it GDTGII(n,α,β). DTGII(n,α) is equal to GDTGII(n,α,0). We shall show the properties of GDTG II.

In Section 2, we give a simple urn model and derive the GDTGII formula using this model.

In Section 3, we give the marginal distribution of D_n using a simple pure birth chain instead of the distribution (3) itself. Then we give the conditional distribution of $D_{n,r}$ given $D_{n1}, ..., D_{n,r-1}$ for $r = 1, ..., n-1$. These conditional distributions and the marginal distribution of D_{n1} are described using Waring distribution.

In Section 4, for the number k of distinct partions in D_n which is a random variable, its distribution is derived. For a positive integer r, the probability of $D_{n1} + \cdots + D_{nr}$ is derived. The asymptotic properties as $n \to \infty$ of these statistics are also given.

2. Generalized DTG II formula

We consider the following urn model (cf. Yamato (1990), Example 1.1 and Yamato (1997), Section 4). There are many red balls of mass one, a single red ball of mass $\beta \geq 0$ and a single black ball of mass $\alpha > 0$. An urn contains the red ball of mass β and the black ball at the beginning. A ball is randomly chosen from the urn in proportion to its mass and replaced along with a red ball of mass one. Let Y_1 be equal to 0 or 1, if the color of the ball chosen at the first trial is red or black, respectively. Let Y_{j+1} be equal to Y_j or $Y_j + 1$ if the color of the ball chosen at the $(j+1)$-th trial is red or black, respectively, for $j = 1, 2, ...$ Then we have a pure birth chain $\{Y_j; j = 1, 2, ...\}$ with states $0, 1, 2, ...$ Its initial state is $Y_1 = 0$ or 1 and the transition probabilities are

$$(4) \qquad P\{Y_{j+1} = y_j \mid Y_1 = y_1, ..., Y_j = y_j\} = \frac{\beta + j}{\alpha + \beta + j}$$

$$P\{Y_{j+1} = y_j + 1 \mid Y_1 = y_1, ..., Y_j = y_j\} = \frac{\alpha}{\alpha + \beta + j}$$

for $j = 1, 2, ...$ and all states $y_1, y_2, ..., y_j$. If we take a positive integer m as parameter β, the equivalent model is obtained from a Pólya-like urn after the first m trials and the sampling from Ferguson's Dirichlet process after the first m observations. It also obtained from Pitman's Chinese restaurant process after arriving the first m persons. For Hoppe's Pólya-like urn, we have the equivalent model by letting $Y_1 = 0$ or 1 if we have the previous color or a new color at the $(m+1)$-th trial, respectively and letting Y_{j+1} be equal to Y_j or $Y_j + 1$ $(j = 2, 3, ...)$ if we have the previous color or a new color at the $(m + j + 1)$-th trial, respectively, after the first m trials. For Chinese restaurant process, we let $Y_1 = 1$ or 0 if $(m+1)$-th person sits at a new empty table or not, respectively and let Y_{j+1} be equal to $Y_j + 1$ or Y_j $(j = 2, 3, ...)$ if the $(m + j + 1)$-th person sits at a new empty table or not, respectively, after the first m persons sat.

For the first n observations $Y_1, ..., Y_n$ of this chain $\{Y_j; j = 1, 2, ...\}$, we put

$$D_{n1} = l \text{ such that } Y_1 = \cdots = Y_l < Y_{l+1}, 1 \leq l \leq n,$$

$$D_{n2} = l \text{ such that } Y_{D_{n1}+1} = \cdots = Y_{D_{n1}+l} < Y_{D_{n1}+l+1}, D_{n1} + l \leq n$$

$$D_{ni} = l \text{ such that } Y_{D_{n1}+\cdots+D_{n,i-1}+1} = \cdots = Y_{D_{n1}+\cdots+D_{n,i-1}+l}$$

$$< Y_{D_{n1}+\cdots+D_{n,i-1}+l+1}, \quad D_{n1} + \cdots + D_{n,i-1} + l \leq n$$

for $i = 3, 4, ..., n$. That is, D_{n1} is the number of observations equal to Y_1, D_{n2} is the number of observations equal to the first one which exceeds Y_1, and so on.

Proposition 1 *For the pure birth chain given by* (4)*,* $D_n = (D_{n1}, ..., D_{nk})$ *has* GDTG II(n,α,β)*, where k is the number of distinct observations among $Y_1, Y_2, ..., Y_n$. That is, the probability distribution of D_n is given by* (3)*.*

Proof. For $(d_1, ..., d_k) \in \mathcal{C}_n$, we have

$$P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nk} = d_k, Y_1 = 0)$$

$$= P(Y_1 = \cdots = Y_{d_1} = 0, Y_{d_1+1} = \cdots = Y_{d_1+d_2} = 1, ..., Y_{d_1+\cdots+d_{k-1}+1} = \cdots = Y_n = k-1).$$

Writing the right-hand side as the products of the conditional probabilities and using the transition probabilities (4), this is equal to

$$\frac{\beta}{\alpha+\beta} \cdot \frac{\alpha^{k-1}}{(\beta+\alpha+1)^{[n-1]}} \cdot \frac{(\beta+1)^{[n]}}{(\beta+d_1)(\beta+d_1+d_2)\cdots(\beta+d_1+\cdots+d_k)}.$$

Similarly we have

$$P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nk} = d_k, Y_1 = 1)$$

$$= \frac{\alpha}{\alpha+\beta} \cdot \frac{\alpha^{k-1}}{(\beta+\alpha+1)^{[n-1]}} \cdot \frac{(\beta+1)^{[n]}}{(\beta+d_1)(\beta+d_1+d_2)\cdots(\beta+d_1+\cdots+d_k)}.$$

Taking the sum of these two probabilities we get (3). $\square$

$\{D_n; n = 1, 2, ...\}$ is a Markov chain by the construction itself. Its one-step transition probabilities are given by the following.

Proposition 2 $\{D_n; n = 1, 2, ...\}$ *is a Markov chain whose one-step transition probabilities are*

$$P(D_{n+1} = (d_1, ..., d_{k-1}, d_k + 1) \mid D_n = (d_1, ..., d_k)) = \frac{\beta+n}{\alpha+\beta+n},$$

$$P(D_{n+1} = (d_1, ..., d_k, 1) \mid D_n = (d_1, ..., d_k)) = \frac{\alpha}{\alpha+\beta+n}$$

for $(d_1, ..., d_k) \in \mathcal{C}_n$ *and* $n = 1, 2, ...$

Conversely, these transition probabilities determine the distribution of D_n, which is given by (3). $\{C_n; n = 1, 2, ...\}$ having the distribution (1) is consistent (Donnelly and Tavaré

(1991)). $\{D_n; n = 1, 2, ...\}$ is not consistent except for the case of $\beta = 0$. That is, it holds only for $\beta = 0$ that

$$P(D_{n-1} = (d_1, ..., d_k)) = \frac{1}{n}\{\sum_{j=1}^{n}(d_j + 1)P(D_n = (d_1, ..., d_j + 1, ..., d_k))$$

$$+P(D_n = (1, d_1, ..., d_k)) + \cdots + P(D_n = (d_1, d_2, ..., d_k, 1))\}, \ \ (d_1, ..., d_k) \in \mathcal{C}_n.$$

2. Marginal and conditional distributions

We shall consider the marginal and conditional distributions of $D_{n1}, ..., D_{nk}$ when $D_n = (D_{n1}, ..., D_{nk})$ has GDTG II (n,α,β).

Proposition 3 *Suppose that D_n have* GDTG II (n,α,β)*. Let r be a positive integer such that $1 \le r \le n$. Then, for positive integers $d_1, d_2, ..., d_r$ satisfying $d(r) = d_1 + \cdots + d_r < n$, $D_{n1}, D_{n2}, ..., D_{nr}$ has the probability given by*

(5)
$$P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nr} = d_r)$$

$$= \frac{\alpha^r}{(\alpha + \beta + 1)^{[d(r)]}} \cdot \frac{(\beta + 1)^{[d(r)]}}{(\beta + d_1)(\beta + d_1 + d_2) \cdots (\beta + d_1 + \cdots + d_r)}$$

For positive integers $d_1, d_2, ..., d_r$ satisfying $d_1 + \cdots + d_r = n$, the probability $P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nr} = d_r)$ is given by (3) *with r instead of k.*

Proof. In order to derive the marginal distributions of D_n, we use the pure birth chain defined by (4). For positive integers $d_1, ..., d_r$ satisfying $d(r) < n$, we have $r < k(\le n)$ and

$$P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nr} = d_r, Y_1 = 0) = P(Y_1 = \cdots = Y_{d_1} = 0, Y_{d_1+1} = \cdots$$

$$= Y_{d_1+d_2} = 1, ..., Y_{d_1+\cdots+d_{r-1}+1} = \cdots = Y_{d_1+\cdots+d_r} = r - 1, Y_{d_1+\cdots+d_r+1} = r).$$

We can similarly write the probability $P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nr} = d_r, Y_1 = 1)$ by the random variables $Y_1, ..., Y_{d_1+\cdots+d_r+1}$. Thus we get relation (5) by the similar method to the proof of Proposition 1. In case of $d_1 + \cdots + d_r = n$, we have $r = k$ and the probability $P(D_{n1} = d_1, D_{n2} = d_2, ..., D_{nr} = d_r)$ is given by (3). $\Box$

Before giving the corollary, we state Waring and bounded Waring distributions. The Waring distribution is the probability distribution of the random variable W taking the values $0, 1, 2, ...$ such that

$$P(W = x) = (c - a)\frac{a^{[x]}}{c^{[x+1]}}, \ \ x = 0, 1, 2, ...,$$

where c, a are positive constants such that $c > a$ (see, for example, Johnson et al. (1992), 6.10.4.). We shall denote this Waring distribution by $\mathrm{Wa}(c, a)$. By grouping the events $\{W = n\}$,$\{W = n+1\}$,$\{W = n+2\}$,... with respect to W having $\mathrm{Wa}(c, a)$ for a non-negative integer n, we have the probability distribution given by

$$P(W = x) = (c - a)\frac{a^{[x]}}{c^{[x+1]}}, \quad x = 0, 1, 2, ..., n - 1,$$

$$\frac{a^{[n]}}{c^{[n]}}, \quad x = n.$$

We shall call this distribution bounded Waring distribution and denote it by $\mathrm{BWa}(n; c, a)$ (Yamato(1997)).

Corollary 1 *Suppose that* $D_n = (D_{n1}, ..., D_{nk})$ *have* GDTG II(n, α, β). *Then, we have*

$$P(D_{n1} - 1 = x) = \frac{\alpha(\beta + 1)^{[x]}}{(\alpha + \beta + 1)^{[x+1]}}, \quad x = 0, 1, ..., n - 2,$$

$$\frac{(\beta + 1)^{[n-1]}}{(\alpha + \beta + 1)^{[n-1]}}, \quad x = n - 1.$$

That is, $D_{n1} - 1$ *has the bounded Waring distribution* $\mathrm{BWa}(n - 1, \alpha + \beta + 1, \beta + 1)$.

Proposition 4 *Suppose that* D_n *have* GDTG II (n, α, β). *Then given* $D_{n1} = d_1, ..., D_{nr} = d_r$, $(D_{n,r+1}, ..., D_{nk})$ *has* GDTG II $(n - d(r), \alpha, \beta + d(r))$, *where* $r = 1, 2, ..., n-1$, $d_1, ..., d_r = 1, 2, ..., n - 1$ *and* $d(r) = d_1 + \cdots + d_r < n$. *Especially, if* D_n *have* DTG II (n, α), *then given* $D_{n1} = d_1, ..., D_{nr} = d_r$, $(D_{n,r+1}, ..., D_{nk})$ *has* GDTG II $(n - d(r), \alpha, d(r))$.

Proof. Dividing the probability (3) by (5), we get the conditional probability,

$$P(D_{n,r+1} = d_{r+1}, ..., D_{nk} = d_k \mid D_{n1} = d_1, ..., D_{nr} = d_r)$$

$$= \frac{\alpha^{k-r-1}}{(\alpha + \beta + d(r) + 1)^{[n-d(r)-1]}} \cdot \frac{(\beta + d(r) + 1)^{[n-d(r)]}}{(\beta + d(r) + d_{r+1}) \cdots (\beta + d(r) + d_{r+1} + \cdots + d_k)}. \quad \square$$

We let $D(r) = D_{n1} + \cdots + D_{nr}$ for a positive integer $r \leq k$. Since the conditional probability of Proposition 4 depends on $d_1, ..., d_r$ only through the sum $d(r) = d_1 + \cdots + d_r$, we have the following.

Corollary 2 *Suppose that* D_n *have* GDTG II (n, α, β). *Then, given* $D(r) = d(r)$, $(D_{n,r+1}, ..., D_{nk})$ *has* GDTG II $(n - d(r), \alpha, \beta + d(r))$, *where* $r = 2, ..., n-1$ *and* $d(r) < n$. *Especially, if* D_n *have* DTG II (n, α), *then, given* $D(r) = d(r)$, $(D_{n,r+1}, ..., D_{nk})$ *has* GDTGII $(n - d(r), \alpha, d(r))$

By applying Corollary 1 to Proposition 4 and Corollary 2, we have the followings.

Corollary 3 *Suppose that* D_n *have* GDTG II (n, α, β). *Then given* $D_{n1} = d_1, ..., D_{nr} = d_r$, $D_{n,r+1} - 1$ *has the bounded Waring distribution* $\mathrm{BWa}(n - d(r) - 1; \alpha + \beta + d(r) + 1, \beta + d(r) + 1)$, *where* $r = 1, ..., n-1$, $d_1, ..., d_r = 1, 2, ..., n-1$ *and* $d(r) = d_1 + \cdots + d_r < n$.

Corollary 4 *Suppose that* D_n *have* GDTG II(n, α, β). *Then given* $D(r) = d(r)$, $D_{n,r+1} - 1$ *has the bounded Waring distribution* $\mathrm{BWa}(n - d(r) - 1; \alpha + \beta + d(r) + 1,\ \beta + d(r) + 1)$, *where* $r = 1, ..., n-1$, $d_1, ..., d_r = 1, 2, ..., n-1$ *and* $d(r) < n$.

3. Related statistics

For D_n having GDTG II(n, α, β), in this section we denote k by K_n to express explicitly that k is a random variable. K_n is equal to $Y_n + 1$ if $Y_1 = 0$ and Y_n if $Y_1 = 1$, where $\{Y_j; j = 1, 2, ...\}$ is the pure birth chain stated in the first paragraph of Section 2. We shall consider the properties of K_n. Using relation derived by the poperties of K_n, we give the probability of $D(r)$. Before we derive the distribution of K_n, we shall note the relation $(\lambda + y)^{[n]} = \sum_{i=0}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) \lambda^{[n-i]} y^{[i]}$, where λ, y are arbitary numbers and n is a positive integer. This relation is shown using that the sum of the probability of the hypergeometric distribution is equal to one (see, for example, Johnson et al. (1992), p.205, (5.16)). Using the unsigned Stirling number of the first kind [ ], we have $y^{[i]} = \sum_{j=0}^{i} \left[\begin{array}{c} i \\ j \end{array} \right] y^j$. Thus we get

$$(6) \qquad (\lambda + y)^{[n]} = \sum_{j=0}^{n} R_1(n, j, \lambda) y^j$$

where $R_1(n, j, \lambda) = \sum_{i=j}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) \left[\begin{array}{c} i \\ j \end{array} \right] \lambda^{[n-i]}$.

R_1 is the function introduced by Carlitz (1980a). For $\lambda = 0, 1$, R_1 is equal to the Stirling number of the first kind,

$$R_1(n, j, 0) = \left[\begin{array}{c} n \\ j \end{array} \right], \quad R_1(n, j, 1) = \left[\begin{array}{c} n+1 \\ j+1 \end{array} \right].$$

(See Carlitz (1980a,b).)

Proposition 5 *Suppose that* D_n *have* GDTG II (n, α, β). *Then for* $k = 1, 2, ..., n$

$$(7) \qquad P(K_n = k) = R_1(n-1, k-1, \beta+1) \frac{\alpha^{k-1}}{(\alpha + \beta + 1)^{[n-1]}}.$$

For $\beta = 0$, this probability is given by Ewens (1972).

Proof. From the distribution given by (3), we have

$$P(K_n = k) = \sum_{(d_1, ..., d_k) \in \mathcal{C}_n} P(D_n = (d_1, ..., d_k)) = f(n, k, \beta) \frac{\alpha^{k-1}}{(\alpha + \beta + 1)^{[n-1]}},$$

where the summation Σ is taken over all distinct ordered partitions $(d_1, ..., d_k)$ of n with k fixed and

$$f(n,k,\beta) = \sum_{(d_1,\ldots,d_k)\in\mathcal{C}_n} \frac{(\beta+1)^{[n]}}{(\beta+d_1)(\beta+d_1+d_2)\cdots(\beta+d_1+\cdots+d_k)}.$$

Since $\sum_{k=1}^{n} P(K_n = k) = 1$, we have $(\alpha+\beta+1)^{[n-1]} = \sum_{k=1}^{n} \alpha^{k-1} f(n,k,\beta)$. Therefore by (6), we get $f(n,k,\beta) = R_1(n-1,k-1,\beta+1)$. $\Box$

Let T_i be the time of appearance of the i-th state among the first n trials, where $i = 2,3,...,n$. T_i has the following probabilities.

Corollary 5 *For $i = 2,3,...,n$ and $l = i,i+1,...,n$ we have*

$$P(T_i = l) = R_1(l-2,i-2,\beta+1)\frac{\alpha^{i-1}}{(\alpha+\beta+1)^{[l-1]}},$$

The probability of the event that the i-th state does not occur among the first n trials is

$$\sum_{j=1}^{i-1} R_1(n-1,j-1,\beta+1)\frac{\alpha^{j-1}}{(\alpha+\beta+1)^{[n-1]}}.$$

Proof. By (4) and (7), for $i = 2,3,...$ and $n = i,i+1,...$ we have $P(T_i = n) = P(K_{n-1} = i-1, Y_n = Y_{n-1}+1) = P(K_{n-1} = i-1)E[P(Y_n = Y_{n-1}+1 \mid Y_1,...,Y_{n-1}) \mid K_{n-1} = i-1] = R_1(n-2,i-2,\beta+1)\ \alpha^{i-1}/(\alpha+\beta+1)^{[n-1]}$. Since the event that the i-th state does not occur among the first n trials is written as $\{K_n \le i-1\}$, by (7) its probability is $P(K_n \le i-1) = \sum_{j=1}^{i-1} R_1(n-1,j-1,\beta+1)\alpha^{j-1}/(\alpha+\beta+1)^{[n-1]}$. $\Box$

From the proof of Proposition 5, we have the following algebraic relation.

Corollary 6

$$(8) \quad R_1(n-1,l-1,\beta+1) = \sum_{(d_1,\ldots,d_l)\in\mathcal{C}_n} \frac{(\beta+1)^{[n]}}{(\beta+d_1)(\beta+d_1+d_2)\cdots(\beta+d_1+\cdots+d_l)}.$$

Using this relation to Proposition 3, we have the probability for $D(r)$.

Proposition 6 *Suppose that D_n have* GDTG II (n,α,β). *Let r be a positive integer.*

$$(9) \quad P(D(r) = j, r < k) = R_1(j-1,r-1,\beta+1)\frac{\alpha^r}{(\alpha+\beta+1)^{[j]}}, \quad j = r,r+1,...,n-1$$

$$P(D(r) = n) = R_1(n-1,r-1,\beta+1)\frac{\alpha^{r-1}}{(\alpha+\beta+1)^{[n-1]}}.$$

Proof. From Proposition 3,

$$P(D(r) = j, r < k) = \frac{\alpha^r}{(\alpha+\beta+1)^{[j]}} \sum_{(d_1,\ldots,d_r)\in\mathcal{C}_j} \frac{(\beta+1)^{[j]}}{\prod_{l=1}^{r}(\beta+\sum_{i=1}^{l} d_i)}.$$

By (8), we have $\sum_{(d_1,...,d_r)\in\mathcal{C}_j}(\beta+1)^{[j]}/\prod_{l=1}^{r}(\beta+\sum_{i=1}^{l}d_i) = R_1(j-1,r-1,\beta+1)$. For $D(r)=n$, we have $k=r$. From the distribution of D_n given by (3) and the relation (8), we have $P(D(r)=n) = R_1(n-1,r-1,\beta+1)\alpha^{r-1}/(\alpha+\beta+1)^{[n-1]}$. □

For the urn model stated at the first paragraph of section 2, we let Z_j be 0 or 1 if the color of the ball chosen at the j-th trial is red or black, respectively, for $j=1,2,...$ Immediately after the j-th trial ($j=1,2,...$), the urn contains the black ball of mass α, the red ball of mass β and j red ball of mass one, no matter what the results of the previous j trials are. Thus Z_{j+1} are independent of $Z_1,...,Z_j$ ($j=1,2,...$) and

$$P(Z_{j+1}=1)=\frac{\alpha}{\alpha+\beta+j},\ P(Z_{j+1}=0)=\frac{\beta+j}{\alpha+\beta+j},\quad j=1,2,...$$

We put $Z(n)=Z_1+\cdots+Z_n$ for $n=1,2,...$ Then by the second Borel-Cantelli lemma, $Z(n)$ diverges to $+\infty$ with probability one(cf. Korwar and Hollander (1973), Corol. 2.2, Donnelly and Tavaré (1986), (6.4)). We can prove the strong law of large numbers for independent random variables $Z_1,Z_2,...$ and $E(Z(n)/\mathrm{log}n)$ converges to α as $n\to\infty$, by the similar method to the proof of Theorem 2.3 of Korwar and Hollander (1973). Thus $Z(n)/\mathrm{log}n$ coverges to α with probability one. Since $K_n=Z(n)+1$ if $Z_1=0$ and $K_n=Z(n)$ if $Z_1=1$, we have the following.

Proposition 7 *Suppose that D_n have* GDTG II (n,α,β). *Then K_n diverges to $+\infty$ with probability one and $K_n/\mathrm{log}n$ coverges to α with probability one.*

For the asymptotic distributions as $n\to\infty$, by Propositions 3, 6, 7 and Corollaries 1, 3, 5, we have the following.

Proposition 8 *Suppose that D_n have* DTG II(n,α,β). *Let r be a positive integer. Then*
(i) *$D_{n1}-1$ has the Waring distribution* $\mathrm{Wa}(\alpha+\beta+1,\beta+1)$ *asymptotically as $n\to\infty$.*
(ii) *$(D_{n1},...,D_{nr})$ has the asymptotic distribution given by*

$$P(D_{n1}=d_1,...,D_{nr}=d_r)$$
$$=\frac{\alpha^r}{(\alpha+\beta+1)^{[d_1+\cdots+d_r]}}\cdot\frac{(d_1+\cdots+d_r)!}{(\beta+d_1)(\beta+d_1+d_2)\cdots(\beta+d_1+\cdots+d_r)},\quad d_1,...,d_r=1,2,...$$

(iii) *Given $D_{n1}=d_1,...,D_{n,r}=d_r$, $D_{n,r+1}-1$ has the Waring distribution* $\mathrm{Wa}(\alpha+\beta+d(r)+1,\beta+d(r)+1)$ *asymptotically, where $d(r)=d_1+\cdots+d_r$.*
(iv) *$D(r)=D_{n1}+\cdots+D_{nr}$ has the asymptotic distribution given by*

$$P(D(r)=j)=R_1(j-1,r-1,\beta+1)\frac{\alpha^r}{(\alpha+\beta+1)^{[j]}},\quad j=r,r+1,...$$

(v) T_i *has the asymptotic distribution given by*

$$P(T_i = l) = R_1(l-2, i-2, \beta+1)\frac{\alpha^{i-1}}{(\alpha+\beta+1)^{[l-1]}} \quad i = 2, 3, ..., \quad l = i, i+1, ...$$

For $\beta = 0$, the last probability is given by Hoppe (1987), p.141.

BIBLIOGRAPHY

Blackwell, D. and MacQueen, J. (1973). "Ferguson distribution via Pólya urn schemes," *Ann. Statist.*, **1**, 353–355.

Carlitz, L. (1980a). "Weighted Stirling numbers of the first and second kind–I," *Fibonacci Quart.*, **18**, 147–162.

Carlitz, L. (1980b). "Weighted Stirling numbers of the first and second kind–II," *Fibonacci Quart.*, **18**, 242–257.

Donnelly, P. (1986). "Partition structures, Polya urns, the Ewens sampling formula, and the ages of alles," *Theor. Popul. Biol.*, **30**, 271–288.

Donnelly, P. and Joyce, P. (1991). "Consistent ordered sampling distribution: Characterization and convergence," *Adv. Appl. Prob.*, **23**, 229–258.

Donnelly, P. and Tavaré, S. (1986). "The ages of alleles and a coalescent," *Adv. Appl. Prob.*, **18**, 1–19.

Donnelly, P. and Tavaré, S. (1990). *Chinese restaurant process*, Chap. 2 of unpublished lecture notes.

Ethier, S.N. (1990). "The infinitely-many-neutral-alleles-diffusion model with ages," *Adv. Appl.Prob.*, **22**, 1–24

Ewens, W.J. (1972). "The sampling theory of selectively neutral alleles," *Theor. Popul. Biol.*, **3**, 87–112.

Ewens, W.J. (1990). "Population genetics theory – the past and the future," In *Mathematical and Statistical Developments of Evolutionary Theory*, ed. S. Lessard, Kluwer Academic Publishers, Dordrecht, 177– 227.

Hoppe, F.M. (1984). "Pólya-like urns and the Ewens sampling formula,*J. Math. Biol.*, **20**, 91–99.

Hoppe, F.M. (1987). "The sampling theory of neutral alleles and an urn model in population genetics," *J. Math. Biol.* **25**, 123–159.

Johnson, N.L., Kotz, S. and Kemp, A.W. (1992). *Univariate discrete distributions*, Wiley, New York.

Joyce, P. and Tavaré, S. (1987). "Cycles, permutation and the structure of the Yule process with immigration," *Stoch. Processes Appl.*, **25**, 309–314.

Korwar R.M. and Hollander, M. (1973). "Contribution to the theory of Dirichlet processes," Ann. Probab., **1**, 705–711.

Sibuya, M. (1993). "A random clustering process," *Ann. Inst. Statist. Math.*, **45**, 459–465.

Sibuya, M. and Yamato, H.(1995). "Ordered and unordered random partitions of an integer and the GEM distribution," *Statist. Prob. Letters*, **25**, 177-183.

Yamato, H.(1990). "A probabilistic approach to Stirling numbers of the first kind," *Commun. Statist.-Theory Meth*, **19**, 3915–3923.

Yamato, H.(1993). "A Pólya urn model with a continuum of colors," *Ann. Inst. Statist. Math.*,**45**, 453–458.

Yamato, H. (1996). "On the Donnelly-Tavaré-Griffiths formula associated with the coalescent," to appear in *Commun. Statist.-Theory Meth.*