Expectation Transposition for Grouped Bose-Einstein Allocation

Ken Suman
ksuman@winona.edu
Thu Apr 23 08:10:43 2020

Abstract

In this paper we derive formulas for transposing expected values of general statistics of certain random combinatorial objects (*e.g.* Bose-Einstein allocation of balls into urns, Pólya sampling and random compositions) which have a component structure of *dependent* random variables into equivalent expressions depending only on the joint distribution of *independent* random variables.

We use our transposition methods to derive results for unrestricted, urn restricted and group restricted Bose-Einstein allocation of identical balls into a single row of distinct urns where the urns have been partitioned into distinct groups. In each of these situations our variables are the number of balls allocated to each group of urns. We derive results for both finite and infinite allocations. We then extend these results to the case of multiple rows, with or without individual row total restrictions. We also extend these results to class size distributions where our variables are the number of groups with a given number of balls (as opposed to the number of balls in a given group). Finally, we derive transposition results for a discrete-time Pólya urn stochastic process, which we show is just an alternative description for grouped Bose-Einstein allocation.

Contents

1 Introduction 7
1.1 History 7
1.2 About This Paper 8
1.2.1 Binomial Coefficients 8
1.2.2 Appendices 9
1.3 Further Work 10
2 Grouped Bose-Einstein Allocation 10
3 Expectation Inversion 11
3.1 Identical Balls into Distinct Urns 12
3.2 Identical Balls into Distinct Urns with Urn Restrictions 14
3.3 Identical Balls into Distinct Urns with Group Restrictions 16
3.4 Asymptotics for Bose-Einstein Allocation 18
4 Class Size Distributions 21
4.1 Class Size Distribution for Bose-Einstein Allocation 21
4.2 Class Size Distribution for Bose-Einstein Allocation with No Empty Urns 24
4.3 Class Size Distribution for Bose-Einstein Allocation with No Empty Groups 26
5 Bose-Einstein Allocation with Two Rows of Urns 27
5.1 Allocation Without Individual Row Total Restrictions 28
5.2 Allocation With Individual Row Total Restrictions 31
6 Compositions 35
6.1 Random Compositions with a Fixed Number of Parts 36
6.2 Compositions with a Fixed Number of Parts, all Distinct 40
6.3 Random Compositions 41
7 Pólya Urn Model 47
7.1 Pólya Process 50
8 Proofs 59
8.1 Proof of Theorem 1 59
8.2 Proof of Example 2 62
8.3 Proof of Theorem 4 68
8.4 Proof of Example 5 71

8.5 Proof of Theorem 6 73
8.6 Proof of Theorem 7 76
8.7 Proof of Example 8 83
8.8 Proof of Theorem 9 85
8.9 Proof of Example 10 90
8.10 Proof of Theorem 11 92
8.11 Proof of Example 12 97
8.12 Proof of Theorem 13 98
8.13 Proof of Theorem 14 103
8.14 Proof of Theorem 15 110
8.15 Proof of Example 16 118
8.16 Proof of Example 17 122
8.17 Proof of Example 21 127
8.18 Proof of Example 22 131
8.19 Proof of Example 23 134
8.20 Proof of Theorem 24 137
8.21 Proof of Example 25 142
8.22 Proof of Theorems 26 and 27 144
8.23 Proof of Example 29 146
8.24 Proof of Example 30 149
9 Ready Reference 151
10 Probability Distributions: Definitions and Properties 159
11 References 162

1 Introduction

1.1 History

In this paper we refer to the general process of re-expressing expectations in terms of independent random variables as *expectation transposition*. Particular methods have gone by specific names in the literature. The method of determining the generating function

$$\mathrm{E}_\theta\big(\Psi(Z_1, Z_2, \dots, Z_n)\big) = \sum_r \mathrm{E}_r\big(\Psi(C_1, C_2, \dots, C_n)\big)\theta^r$$

and extracting the desired coefficient of θ is sometimes referred to as the method of *expectation inversion*. In those situations when the component structure distribution can be expressed as a compound distribution

$$f(C_1, \dots, C_n) = \int_\theta f_\theta(Z_1, \dots, Z_n) f(\theta) d\theta \tag{1}$$

then it will follow from the rule of iterated expectations, $\mathrm{E}(X) = \mathrm{E}\big(\mathrm{E}(X|Y)\big)$ (see Casella and Berger [11]), that

$$\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) = \int_\theta \mathrm{E}_\theta\big(\Psi(Z_1, \dots, Z_n)\big) f(\theta) d\theta.$$

This approach is referred to as the method of *compounding*. Finally, the method of determining functions g and h such that

$$\mathrm{E}_\theta\big(\Psi(Z_1, \dots, Z_n)\big) = \int_0^\infty \mathrm{E}\big(\Psi(rC_1, \dots, rC_n)\big) g(r) h(\theta) \mathrm{e}^{-\theta r} dr$$

and then extracting $\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$ through the inverse Laplace transform operation will be referred to as the method of *transforms*.

The term “expectation inversion” first appeared in Tweedie’s 1965 paper [46] but the method dates back at least to the 1952 paper of Domb [18]. Domb uses expectation inversion to solve a problem posed by Schrödinger [39] equivalent to the classical problem of finding the probability of m cells being occupied if r balls are randomly distributed among n cells. Dwass [20] uses expectation inversion to give a unified approach to determining the distributions of rank order statistics. Mohanty [34] devotes a section to the “Dwass Technique” and is a source for other references. Both Baclawski, Cerasoli, and Rota [5] and Baclawski, Rota, and Billey [6] elaborate on the application found in Domb [18] and without giving any references to explain why, label the approach as “Schrödinger’s Method”. Cerasoli [12] and Buoncristiani and Cerasoli [10] use expectation inversion in the case of multinomial allocation of balls into urns. Shepp and Lloyd [40] use expectation inversion in problems pertaining to cycle lengths in random permutations. Arratia and Tavaré [3, 4], Hansen [26, 27, 28], and Watterson [48, 49] use expectation inversion in a manner similar to Shepp and Lloyd [40] and greatly extend the potential of expectation inversion as an enumerative technique. The papers of Arratia and Tavaré [3, 4] refer to expectation inversion as the “Shepp and Lloyd Method”. We thought that “expectation inversion” was the most descriptive of the various names to be found in the literature and so we have followed Tweedie’s terminology in this paper.

The use of the compounding method dates back at least to Skellam [41]. Skellam gives some earlier references in his paper as well. Mosimann [36] and Gerstenkorn and Jarzebska [23] are some early references where the compounding method is used in a manner similiar to how it is used in this paper. We note that the terms compound, contagious and mixture are used interchangeably in the literature when referring to distributions formed as in (1).

The method of transforms as defined in this paper appeared in Dwass [19] and then was developed further by Steutel [43]. More recently it has been used by Huillet [33, 32].

1.2 About This Paper

1.2.1 Binomial Coefficients

A word of caution is in order before implementing results in this paper in your favorite computer algebra system. We have defined binomial coefficients by

$$\binom{x}{n} = \begin{cases} \dfrac{x_{[n]}}{n!} & \text{integer } n \geq 0, x \in \mathbb{R} \\ 0 & \text{else} \end{cases}$$

where for all $x \in \mathbb{R}$, $x_{[n]}$, the n^{th} falling factorial of x, is given by

$$x_{[n]} = \begin{cases} x(x-1)\cdots(x-n+1) & n = 1,2,\ldots \\ 1 & n = 0. \end{cases}$$

While this definition is widely used it is not universal. In particular, the computer algebra systems WolframAlpha [54] and Maxima [38] both return -2 instead of 0 to the query binomial$[-2,-3]$. The definition used by these computer programs does have the advantage of preserving the symmetry identity

$$\binom{x}{n} = \binom{x}{x-n}$$

so

$$\binom{-2}{-3} = \binom{-2}{-2-(-3)} = \binom{-2}{1} = -2.$$

But the definition for binomial coefficients used in this paper has several advantages of its own as enumerated in Graham, et al.'s *Concrete Mathematics* [25] and Wilf's *generatingfunctionology* [50]. Most importantly, using this definition can simplify indices of summations and can allow one to circumvent messy situations with an answer broken into a longish list of conditional cases. So the reader needs to be aware that they will have to redefine the query binomial$[a,b]$ before using these (and perhaps other) computer algebra systems to implement the results in this paper.

1.2.2 Appendices

This paper includes many theorems and examples. To aid the reader we only include "highlight" proofs of each in the main body of the paper and relegate the full proofs of each theorem and example to Appendix A.

The central theme in this paper of relating random combinatorial objects to probability distributions of independent random variables necessitates the use of several straightforward but lesser known results from probability theory. We have collected these results in Appendix B for easy reference for those cases where more than a simple citation is needed.

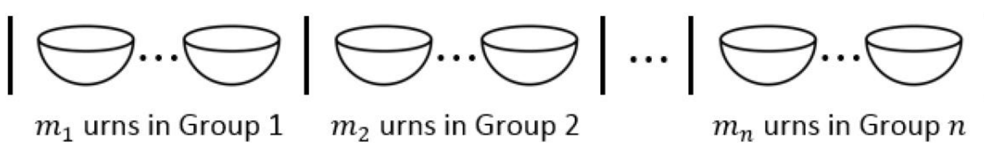
Finally, in Appendix C, for the readers convenience we have brought together the definitions, factorial moments, raw moments and distributional properties for each of the probability distributions used in this paper.

1.3 Further Work

This is the first of three planned related papers. In our second forthcoming paper we apply expectation transposition methods to a class of problems which includes sampling from a line or circle, coverage of a line or circle and success runs (linear, circular for both overlapping and non-overlapping cases) and we show how all these results relate to grouped Bose-Einstein allocation models. In our third planned paper we apply expectation transposition methods to the class of waiting time or quota fulfillment problems in the case of Pólya urn models. The particular waiting time problems considered in our third paper include inverse sampling, "sooner and later" waiting time problems and disaster-modified waiting time problems.

2 Grouped Bose-Einstein Allocation

Imagine that we distribute t identical balls into a row of M distinct urns in such a way that all $\binom{M+t-1}{t}$ possible allocations are equally likely to occur. Further suppose that each urn belongs to one of n distinct groups where the j^{th} group contains m_j urns with $m_1 + \cdots + m_n = M$.



(Note: An equivalent alternative formulation is to consider each group as a single urn where the j^{th} urn (group) has m_j ordered cells, where we allow any number of balls per cell.)

Now let C_j equal the number of balls that get distributed into the j^{th} group where $C_1 + C_2 + \cdots + C_n = t$. In this case the joint probability distribution of $(C_1, \dots, C_n)$ is

$$P(C_1 = c_1, \dots, C_n = c_n)$$

$$
= \begin{cases} \dfrac{\binom{c_1+m_1-1}{m_1-1}\cdots\binom{c_n+m_n-1}{m_n-1}}{\binom{M+t-1}{t}} & \begin{array}{c} c_1+\cdots+c_n=t \\ m_1+\cdots+m_n=M \\ c_j\in\{0,1,\dots\},\ j=1,\dots,n \\ m_j\in\{1,2,\dots\},\ j=1,\dots,n \end{array} \\ 0 & \text{else.} \end{cases}
$$

(2)

Distributing identical balls into distinguishable urns without restriction such that all possible allocations are equally likely to occur is referred to as Bose-Einstein allocation. So we can refer to this model where the distinguishable urns belong to distinguishable groups as a grouped Bose-Einstein allocation model.

A difficulty in using this joint distribution for involved combinatorial or probability problems is that the variables $C_1, \dots, C_n$ are dependent. Intuitively, problems would be easier to handle if we could remove this dependence. With that in mind, one might hope that (??) could be related to the similar looking joint distribution

$$
P(Z_1=c_1,\dots,Z_n=c_n)
$$

$$
= \begin{cases} \binom{c_1+m_1-1}{m_1-1}\cdots\binom{c_n+m_n-1}{m_n-1}(1-p)^M p^{c_1+\cdots+c_n} & \begin{array}{c} m_1+\cdots+m_n=M \\ c_j\in\{0,1,\dots\},\ j=1,\dots,n \\ m_j\in\{1,2,\dots\},\ j=1,\dots,n \end{array} \\ 0 & \text{else} \end{cases}
$$

(3)

where $Z_1, \cdots, Z_n$ are *independent* variables with $Z_j \sim$ negative binomial$(m_j,\ 1-p)$. (The symbol $\sim$ is used here as in statistical applications to mean "is distributed as".) We notice that apart from a constant factor there are only two differences in these joint distributions. The support of $(C_1, \dots, C_n)$ requires $c_1+\cdots+c_n=t$ and the distribution of $(Z_1, \dots, Z_n)$ contains the factor $p^{c_1+\cdots+c_n}$.

3 Expectation Inversion

In Theorem 1, which follows below, we show how we can exploit this close relationship between these two joint distributions. In particular, this theorem shows how to recast a problem involving the dependent variables $C_1, \dots, C_n$ into its equivalent problem involving the independent variables $Z_1, \dots, Z_n$. The insight on this theorem is to see that if we take the distribution of $(Z_1, \dots, Z_n)$ conditioned on $Z_1 + \cdots + Z_n = t$, the support of this conditional distribution will regain the requirement that $c_1 + \cdots + c_n = t$. Simultaneously, conditioning on $Z_1 + \cdots + Z_n = t$ will remove all traces of the parameter p, including the factor $p^{c_1+\cdots+c_n}$. This is not just a lucky coincidence. One can use standard methods in statistical inference to show that $Z_1 + \cdots + Z_n$ is a sufficient statistic for the parameter p and it is a result in statistical inference that conditioning on a sufficient statistic for a parameter necessarily removes that parameter from the resulting conditional distribution. Finally, the requirement that total probability equals 1 for $(Z_1, \dots, Z_n)$ conditioned on $Z_1 + \cdots + Z_n = t$ and for $(C_1, \dots, C_n)$ will necessitate that their constant factors match up.

3.1 Identical Balls into Distinct Urns

Theorem 1

For $P(C_1 = c_1, \dots, C_n = c_n)$ defined by (??) and for a general statistic $\Psi(\,)$,

$$\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) = \frac{1}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0} \tag{4}$$

where $Z_1, \dots, Z_n$ are independent, $Z_j \sim$ negative binomial$\big(m_j, 1-p\big)$ and $M = m_1 + \cdots + m_n$.

Proof. The proofs of all theorems are placed in Appendix A.

Example 2

Number of marked urns with a given number of balls

Suppose t identical balls are distributed among $n + 1$ distinguishable urns. The first n urns each contain m_1 (distinguishable) cells or compartments and the last urn contains m_2 (distinguishable) cells or compartments. There is no limit on the number of the number of balls that can go into any of the $m_1 n + m_2$ cells. Assume that all possible allocations of the t balls into the $m_1 n + m_2$ distinguishable cells are equally likely to have occurred.

Let D_i equal the number of urns among the first n which are occupied by exactly i balls. Then

$$P(D_i = d) = \frac{\binom{n}{d}}{\binom{m_1 n + m_2 + t - 1}{t}} \left(\sum_{j=d}^{n} (-1)^{j-d} \binom{n-d}{n-j} \right.$$

$$\left. \times \binom{m_1 + i - 1}{i}^j \binom{m_1(n-j) + m_2 + t - ij - 1}{t - ij} \right)$$

and $E\big((D_i)_{[r]}\big)$, the r^{th} falling factoring moment of D_i, is given by

$$E\big((D_i)_{[r]}\big) = \frac{\binom{i + m_1 - 1}{m_1 - 1}^r \binom{m_1(n-r) + m_2 + t - ir - 1}{t - ir}}{\binom{m_1 n + m_2 + t - 1}{t}} n_{[r]}.$$

(Note: With the definition for binomial coefficients we have adopted it is not necessary in the result for $P(D_i = d)$ above to make explicit in the summation those cases in

$$\binom{m_1(n-j) + m_2 + t - ij - 1}{t - ij}$$

where $m_1(n-j) + m_2 + t - ij - 1$ is positive but $t - ij$ is negative and we need this binomial coefficient to equal 0. With our definition, the binomial coefficient does the work for us.)

Charalambides and Koutras [13] consider this problem in the case $i = 0$ and Charalambides [14] considers this problem for general i but only for the case $m_1 = 1$.

Proof. (Only an outline of the proof is given here. A detailed proof of each example in this paper is placed in Appendix A.)

Let C_j equal the number of balls that go into the j^{th} urn. The first result follows directly from Theorem 1 and the generalized principle of inclusion-exclusion (Theorem 37 in Appendix B) with

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \mathbb{I}(\text{exactly } d \text{ of } (C_1, \dots, C_n) \text{ equal } i),$$

$Z_j \sim$ negative binomial$(m_1, 1-p), j = 1, \dots, n$ and $Z_{n+1} \sim$ negative binomial$(m_2, 1-p)$.

Throughout this paper we use the notation

$$\mathbb{I}(\mathcal{A}) = \begin{cases} 1 & \text{event } \mathcal{A} \text{ occurs} \\ 0 & \text{else.} \end{cases}$$

The second result also follows from Theorem 1 with

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \left(\mathbb{I}(C_1 = i) + \dots + \mathbb{I}(C_n = i)\right)_{[r]} = (D_i)_{[r]}$$

on applying the formula for falling factorial moments of sums of indicator functions (see Theorem 36 of Appendix B).

3.2 Identical Balls into Distinct Urns with Urn Restrictions

In practice the number of balls per urn in a Bose-Einstein (equally likely) allocation is often restricted. A common restriction is that no urn may be left empty. We consider the restriction of no empty urns in this section. Further restrictions, such as all urns must contain at least k balls, can be developed in a natural way from the case of no empty urns.

Imagine distributing t identical balls into a row of M distinct urns in such a way that all $\binom{t-1}{M-1}$ possible allocations with no urn left empty are equally likely to occur. We will again assume that each urn belongs to one of n distinct groups where the j^{th} group contains m_j urns with $m_1 + \dots + m_n = M$.

Let C_j equal the number of balls that get distributed into the j^{th} group where $C_1 + C_2 + \dots + C_n = t$. In this case the joint probability distribution of $(C_1, \dots, C_n)$ is

$$P(C_1 = c_1, \dots, C_n = c_n)$$
$$= \begin{cases} \dfrac{\binom{c_1 - 1}{m_1 - 1} \dots \binom{c_n - 1}{m_n - 1}}{\binom{t-1}{M-1}} & \begin{array}{l} c_1 + \dots + c_n = t \\ m_1 + \dots + m_n = M, \;\; M \geq t \\ c_j \in \{m_j, m_j + 1, \dots\}, j = 1, \dots, n \\ m_j \in \{1,2, \dots\}, \qquad j = 1, \dots, n \end{array} \\ 0 & \text{else.} \end{cases} \tag{5}$$

Now we are back to the question of how to choose independent random variables $(Z_1, \dots, Z_n)$ so that their joint distribution "looks similar" to the joint distribution of $(C_1, \dots, C_n)$ given in (??).

One way of distributing identical balls into distinct urns such that all possible allocations with no empty urns are equally likely is to initially put a single ball into each urn and then to go through a Bose-Einstein (equally likely) allocation of the remaining balls. This has the effect of shifting over the count of how many balls can land in any group. Specifically, the support for C_j starts at m_j instead of 0 as was the case in Section 3.1.

This suggests that we will need to shift over the support for Z_j in a comparable way. To do this we need the following definition.

Definition 3

For any nonnegative integer α, we will say that a random variable X has a α - *shifted negative binomial*$(m, 1-p)$ distribution if $X = Y + \alpha$ with $Y \sim$ negative binomial $(m, 1-p)$. In this case

$$P(X = x) = P(Y = x - \alpha) = \binom{m + (x-\alpha) - 1}{m-1}(1-p)^m p^{x-\alpha}. \tag{6}$$

Now let $Z_1, \dots, Z_n$ be independent random variables such that $Z_j \sim$ m_j - shifted negative binomial$(m_j, 1-p)$ for $j = 1, \dots, n$. Then the joint probability distribution of $(Z_1, Z_2, \dots, Z_n)$ is given by

$$P(Z_1 = c_1, \dots, Z_n = c_n)$$

$$= \begin{cases} \binom{c_1 - 1}{m_1 - 1} \cdots \binom{c_n - 1}{m_n - 1}(1-p)^M p^{(c_1 + \cdots + c_n) - M} & \begin{array}{l} m_1 + \cdots + m_n = M \\ c_j \in \{m_j, m_j + 1, \dots\} \\ m_j \in \{1,2,\dots\},\ j = 1,\dots,n \end{array} \\ 0 & \text{else} \end{cases} \tag{7}$$

which looks like (??) apart from constants and the missing restriction $c_1 + \cdots + c_n = t$ on the support.

Theorem 4

For $P(C_1 = c_1, \dots, C_n = c_n)$ defined by (??) and for a general statistic $\Psi(\,)$,

$$\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) = \frac{1}{\binom{t-1}{M-1} t!} \frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0} \tag{8}$$

where $Z_1, \dots, Z_n$ are independent random variables such that $Z_j \sim$ m_j- shifted negative binomial$\left(m_j, 1-p\right)$ for $j = 1, \dots, n$.

3.3 Identical Balls into Distinct Urns with Group Restrictions

The second restriction we will consider is the case when each *group* is required to have at least one ball. For this problem we will assume that $m_1 = \cdots = m_n = m$. As was the case in Section 3.2 further restrictions, such as general upper and lower bounds on the number of balls per group or general m_j can be developed in the same manner. In this section we will continue to assume that we are distributing t identical balls into a row of M distinct urns in such a way that all possible allocations with no group left empty are equally likely to occur. We will again assume that each urn belongs to one of n distinct groups where the j^{th} group contains m_j urns with $m_1 + \cdots + m_n = M$.

As a starting point towards a general result for finding $\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$ in this case where groups are required to have at least one ball, we need to use Theorem 1 to solve for $\mathfrak{Z}(t, m, n)$, the total number of ways to allocate the balls subject to the restriction that no group can be left empty.

We note that the numbers $(t!/n!)\mathfrak{Z}(t, m, n)$ are known as *associated Lah numbers* in the literature (Ahuja[2], Ahuja and Enneking[1], Nandi and Dutta[37]).

Example 5

Allocations with no empty groups

$$\mathfrak{Z}(t, m, n) = \sum_{j=1}^{n} (-1)^{n-j} \binom{n}{j} \binom{mj + t - 1}{t}.$$

Proof. (Only an outline of the proof is given here. A detailed proof of each example in this paper is placed in Appendix 8.)

Let C_j equal the number of balls that get distributed into the j^{th} group where $C_1 + C_2 + \cdots + C_n = t$. We can apply Theorem 1 with

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \mathbb{I}(C_1 \geq 1, C_2 \geq 1, \dots, C_n \geq 1)$$

to find

$$\mathrm{E}\big(\Psi(C_1,\dots,C_n)\big) = P(C_1 \geq 1, C_2 \geq 1, \dots, C_n \geq 1) = \frac{\mathfrak{Z}(t,m,n)}{\binom{mn+t-1}{t}}$$

and hence solve for $\mathfrak{Z}(t,m,n)$.

It follows from here that the joint probability distribution of $(C_1,\dots,C_n)$ is

$$P(C_1 = c_1,\dots,C_n = c_n)$$

$$= \begin{cases} \dfrac{\binom{c_1+m-1}{m-1}\cdots\binom{c_n+m-1}{m-1}}{\mathfrak{Z}(t,m,n)} & \begin{array}{l} c_1+\cdots+c_n = t \\ c_j \in \{1,2,\dots\},\ j=1,\dots,n \\ m \in \{1,2,\dots\} \end{array} \\ 0 & \text{else} \end{cases} \tag{9}$$

Theorem 6

For $P(C_1 = c_1,\dots,C_n = c_n)$ defined by (??)

$$\mathrm{E}\big(\Psi(C_1,\dots,C_n)\big) = \frac{1}{\mathfrak{Z}(t,m,n)\; t!}\,\frac{d^t}{dp^t}\left(\left(\frac{g(p,m)}{(1-p)^m}\right)^n \mathrm{E}\big(\Psi(Z_1,\dots,Z_n)\big)\right)\Bigg|_{p=0}$$

where $Z_1,\dots,Z_n$ are independent and identically distributed random variables such that

$$P\big(Z_j = z\big) = \frac{\binom{z+m-1}{m-1}(1-p)^m p^z}{g(p,m)}, \quad z \in \{1,\dots,\infty\}$$

with

$$g(p,m) = \sum_{z=1}^{\infty}\binom{z+m-1}{m-1}(1-p)^m p^z$$

and

$$\big(g(p,m)\big)^r = \sum_{t=r}^{\infty}\mathfrak{Z}(t,m,r)(1-p)^{mr}\,p^t.$$

3.4 Asymptotics for Bose-Einstein Allocation

Theorem 7

Suppose the joint distribution of $(C_1, C_2, \dots, C_n)$ is given by (??), (??) or (??). That is, where C_j equals the number of balls allocated to the j^{th} group in the case when (*i*) no restrictions are placed on the number of balls per urn, (*ii*) each urn must have at least one ball, or (*iii*) each group must have at least one ball. In all three cases we have that

$$\lim_{t\to\infty} \mathrm{E}\left(\Psi\left(\frac{C_1}{t}, \dots, \frac{C_n}{t}\right)\right) = (M-1)!\ \mathcal{L}^{-1}\left(\frac{\mathrm{E}\big(\Psi(Y_1, \dots, Y_n)\big)}{\lambda^M}\right)\Bigg|_{s=1}$$

for any statistic $\Psi(x_1, x_2, \dots, x_n)$ which is defined on all $\mathbb{R}^n_{\geq 0}$ and is bounded and continuous on the simplex $x_1 + x_2 + \cdots + x_n = 1$, $0 \leq x_j \leq 1$ for $j = 1,2, \dots, n$.

1. $\mathcal{L}^{-1}\big(G(\lambda)\big) = g(s)$ is the *inverse Laplace transform* of $G(\lambda)$. That is, for given function $G(\lambda)$, $g(s)$ is that function of s but not the parameter λ such that

$$G(\lambda) = \int_0^\infty g(s) e^{-\lambda s} ds;$$

2. $Y_1, Y_2, \dots, Y_n$ are independent random variables with $Y_j \sim \text{Gamma}\big(m_j, \lambda\big)$ for $j = 1,2, \dots, n$. That is, Y_j has the likelihood function

$$f_{Y_j}(y) = \frac{\lambda^{m_j}}{\big(m_j - 1\big)!} y^{m_j - 1} \mathrm{e}^{-\lambda y} \mathbb{I}(y > 0);$$

3. $M = m_1 + \cdots + m_n$.

Proof. The proof of this theorem is involved. It can be organized into four key steps which we will delineate here and will prove individually in Appendix 8.

Step 7a.

Let the joint distribution of

$(C_1, C_2, \dots, C_n)$ be given by (??) (i.e. unrestricted allocation),

$(C_1^*, C_2^*, \dots, C_n^*)$ be given by (??) (i.e. no empty urns), and

$(\hat{C}_1, \hat{C}_2, \dots, \hat{C}_n)$ be given by (??) (i.e. no empty groups).

Then

$$\lim_{t\to\infty} P\left(\frac{C_1}{t} \le v_1, \frac{C_2}{t} \le v_2, \dots, \frac{C_n}{t} \le v_n\right)$$

$$= \lim_{t\to\infty} P\left(\frac{C_1^*}{t} \le v_1, \frac{C_2^*}{t} \le v_2, \dots, \frac{C_n^*}{t} \le v_n\right)$$

$$= \lim_{t\to\infty} P\left(\frac{\hat{C}_1}{t} \le v_1, \frac{\hat{C}_2}{t} \le v_2, \dots, \frac{\hat{C}_n}{t} \le v_n\right).$$

Step 7b.

Let $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$ (see Appendix 10). Then

$$\lim_{t\to\infty} P\left(\frac{C_1}{t} \le v_1, \frac{C_2}{t} \le v_2, \dots, \frac{C_n}{t} \le v_n\right) = P(V_1 \le v_1, V_2 \le v_2, \dots, V_n \le v_n)$$

for all $(v_1, v_2, \dots, v_n)$ such that $v_1 + v_2 + \cdots + v_n = 1$ and $0 \le v_j \le 1$ for $j = 1,2, \dots, n$.

In the language of probability, we would say $\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right)$ converges in *distribution* to the Dirichlet$(m_1, m_2, \dots, m_n)$ distribution. We denote this as

$$\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right) \xrightarrow{d} (V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n).$$

Step 7c.

Suppose the joint distribution of $(C_1, C_2, \dots, C_n)$ is given by (??), (??), or (??) and let $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$. Then

$$\lim_{t\to\infty} \mathrm{E}\left(\Psi\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right)\right) = \mathrm{E}\big(\Psi(V_1, V_2, \dots, V_n)\big)$$

for any bounded and continuous function $\Psi(\,)$ on the common support, that is for all points $(v_1, v_2, \dots, v_n)$ such that $v_1 + v_2 + \cdots + v_n = 1$ and $0 \le v_j \le 1$ for $j = 1,2, \dots, n$.

Step 7d.

Let $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$ and let $Y_1, Y_2, \dots, Y_n$ be independent random variables with $Y_j \sim \text{Gamma}(m_j, \lambda)$ for $j = 1,2, \dots, n$. Then for any function $\Psi(\ )$,

$$\mathrm{E}\big(\Psi(V_1, V_2, \dots, V_n)\big) = (M-1)!\, \mathcal{L}^{-1}\left(\frac{\mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_n)\big)}{\lambda^M}\right)\Bigg|_{s=1}.$$

Example 8

For $P(C_1 = c_1, \dots, C_n = c_n)$ defined by (??) we can use Theorem 1 to derive the exact answer

$$\mathrm{E}\left(\left(C_1\right)^a\right) = \frac{1}{\binom{M+t-1}{t}} \sum_{k=1}^{a} S(a,k) m_1^{[k]} \binom{M+t-1}{t-k}$$

where $S(n,k)$ are the Stirling numbers of the second kind and $m_1^{[k]}$, the k^{th} rising factorial of m_1, is defined through

$$x^{[n]} = \begin{cases} x(x+1)\cdots(x+n-1) & n = 1,2, \dots \\ 1 & n = 0. \end{cases}$$

Then we can use Theorem 7 to derive the asymptotic result

$$\lim_{t\to\infty}\left(E\left(\left(\frac{C_1}{t}\right)^a\right)\right) = \frac{\binom{m_1+a-1}{a}}{\binom{M+a-1}{a}}.$$

It follows immediately from (10) that

$$\mathrm{E}\left(\left(C_1\right)^a\right) \approx t^a \frac{\binom{m_1+a-1}{a}}{\binom{M+a-1}{a}}$$

for large t. As one simple check we note that in the case $m_1 = 10$ and $M = 20$, the exact answer and the asymptotic approximation for $\mathrm{E}((C_1/t)^3)$ agree to 2 decimals for $t = 500$ and to 7 decimals for $t = 50$ million.

Proof. (Only an outline of the proof is given here. A detailed proof of each example in this paper is placed in Appendix 8.)

Factorial moments fit naturally with discrete random variables while raw moments (also known as moments about zero or just moments) fit naturally with continuous random variables.

To compare an exact answer derived from Theorem 1 based on the discrete negative binomial random variables with an asymptotic answer derived from Theorem 7 based on the continuous gamma random variables we will need to use the connection between raw moments and factorial moments via the definition of Stirling numbers of the second kind. Namely, for integer $n \geq 1$,

$$t^n = \sum_{k=1}^{n} S(n,k) t_{[n]}.$$

Let C_j equal the number of balls that go into the j^{th} urn. The first result follows directly from Theorem 1 and the generalized principle of inclusion-exclusion (Theorem 37 in Appendix 9) with

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \mathbb{I}(\text{exactly } d \text{ of } (C_1, \dots, C_n) \text{ equal } i),$$

$Z_j \sim$negative binomial$(m_1, 1-p)$, $j = 1, \dots, n$ and $Z_{n+1} \sim$negative binomial$(m_2, 1-p)$.

The proof of the second part of this example requires Theorem 7. The necessary condition that $\Psi(x_1, x_2, \dots, x_n) = x_1^a$ is bounded and continuous on the simplex $x_1 + x_2 + \cdots + x_n = 1$ and $0 \leq x_j \leq 1$ for $j = 1,2, \dots, n$ is clearly satisfied. Hence, for $Y_1 \sim$ gamma(m_1, λ), we have

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t} \right)^a \right) \right) = (M-1)!\, \mathcal{L}^{-1} \left(\frac{\mathrm{E}(Y_1^a)}{\lambda^M} \right) \Bigg|_{s=1}$$

and the final result follows on simplification.

4 Class Size Distributions

Let $(C_1 = c_1, \dots, C_n = c_n)$ be a vector of discrete random variables such that $C_j \in \{a, a+1, \dots, b\}$ for all $j = 1,2, \dots, n$. Define D_k as the number of C_j's which equal k, $k \in \{a, a+1, \dots, b\}$.

The distribution of $(D_a, D_{a+1}, \dots, D_b)$ is referred to as the *class size distribution* of $(C_1, C_2, \dots, C_n)$. Wilks [52, 53] discusses several applications of class size distributions. Tukey [45] (who refers to this distribution as a group size distribution) derives various moment results for some class size distributions arising in urn models.

4.1 Class Size Distribution for Bose-Einstein Allocation

For $P(C_1 = c_1, \ldots, C_n = c_n)$ defined by (??), let D_k equal the number of C_j's which equal $k,\ k \in \{0,1,\ \ldots\ ,\ t\}$.

We can easily derive the class size distribution for $(C_1, C_2, \ldots, C_n)$ defined by (??) in the case $m_1 = m_2 = \cdots = m_n = m$ and $M = mn$. For a given vector $(d_0, d_1, \ldots, d_t)$ define the set $\mathbb{B}_D$ by

$$\mathbb{B}_D = \left\{(c_1, \ldots, c_n) \middle| \sum_{j=1}^{n} \mathbb{I}(c_j = 0) = d_0, \ldots, \sum_{j=1}^{n} \mathbb{I}(c_j = t) = d_t \right\}.$$

Then

$$P(D_0 = d_0, D_1 = d_1, \ldots, D_t = d_t) = \sum \underset{\mathbb{B}_D}{\cdots} \sum P(C_1 = c_1, \ldots, C_n = c_n)$$

$$= \sum \underset{\mathbb{B}_D}{\cdots} \sum \frac{\binom{c_1 + m - 1}{m - 1} \cdots \binom{c_n + m - 1}{m - 1}}{\binom{t + mn - 1}{mn - 1}}.$$

$$= \sum \underset{\mathbb{B}_D}{\cdots} \sum \frac{\binom{0 + m - 1}{m - 1}^{d_0} \binom{1 + m - 1}{m - 1}^{d_1} \cdots \binom{t + m - 1}{m - 1}^{d_t}}{\binom{t + mn - 1}{mn - 1}}$$

$$= \frac{\binom{0 + m - 1}{m - 1}^{d_0} \binom{1 + m - 1}{m - 1}^{d_1} \cdots \binom{t + m - 1}{m - 1}^{d_t}}{\binom{t + mn - 1}{mn - 1}} \cdot \frac{n!}{d_0!\, d_1! \cdots d_t!} \tag{11}$$

where

$$d_0 + d_1 + \cdots + d_t = n$$

$$0d_0 + 1d_1 + \cdots + td_t = t$$

$$M = mn \text{ and } d_j = \{0,1,2, \ldots\}.$$

Following the approach used in Theorem 1 we would like to determine a set of *independent* random variables $Z_0, Z_1, \ldots, Z_t$ whose joint distribution is "similar looking" to (11) and would allow us to perform an expectation inversion as we did in Theorem 1.

We note that the part of (11) involving just d_j has the form $\frac{c^{d_j}}{d_j!}$ where $c = \binom{j + m - 1}{m - 1}$ is a constant.

However we recognize that this form is just the kernal of the distribution of a Poisson random variable with parameter c. That is, for $Z_j \sim \text{Poisson}(c)$,

$$P(Z_j = z_j) = \mathrm{e}^{-c} \cdot \frac{c^{z_j}}{z_j!}.$$

But as in Theorem 1, our independent random variables needs to include parameter(s) that allow us to introduce the necessary restrictions on the support when we condition on the sufficient statistic(s). In this case we need to impose two restrictions on the support, namely, $z_0 + z_1 + \cdots + z_t = n$ and $0z_0 + 1z_1 + \cdots + tz_t = t$. Therefore, the distribution of Z_j should include two parameters chosen to make $z_0 + z_1 + \cdots + z_t$ and $0z_0 + 1z_1 + \cdots + tz_t$ jointly sufficient statistics so that when we condition on these statistics the parameters will fall out but will leave us the appropriate restrictions on the support.

We can achieve this by taking $Z_j \sim \text{Poisson}\left(\binom{j+m-1}{m-1}\theta\lambda^j\right)$. The details of this process are given in the next theorem.

Theorem 9

For the class size distribution given in (11) we have

$$\mathrm{E}\big(\Psi(D_0, D_1, \dots, D_t)\big)$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\mathrm{e}^{\frac{\theta}{(1-\lambda)^m}} \mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big) \right)\Bigg|_{\substack{\lambda=0\\ \theta=0}} \tag{12}$$

where $Z_0, Z_1, \dots, Z_t$ are independent and $Z_j \sim \text{Poisson}\left(\binom{j+m-1}{m-1}\theta\lambda^j\right)$.

Example 10 (Joint Factorial Moments of the Count Size Distribution for Bose-Einstein Allocation)

Suppose t identical balls are allocated to a row of mn distinct urns according to the Bose-Einstein scheme (all allocations equally likely) where each urn belongs to one of n distinct groups with m urns per group.

Let D_j equal the number of groups containing exactly j balls. Then

$$\mathrm{E}\big((D_0)_{[r_0]}\cdots(D_t)_{[r_t]}\big) = \frac{\binom{m(n-R)+(t-S)-1}{t-S}\prod_{j=0}^{t}\binom{j+m-1}{m-1}^{r_j}}{\binom{mn+t-1}{t}} n_{[R]}$$

where $R = r_0 + \cdots + r_t$ and $S = 0r_0 + 1r_1 + \cdots tr_t$. Recall that $a_{[0]} = 1$ by definition so this result can be used to find the joint factorial moments of any subset of the D_j's. The special case when $r_k = 1, r_j = 0$ for all $j \neq k$, and $m = 1$ is solved in Hardy [29].

Proof. The result follows on direct application of Theorem 9 if we take

$$\Psi(D_0, D_1, \dots, D_t) = (D_0)_{[r_0]}\cdots(D_t)_{[r_t]}$$

and take $Z_0, Z_1, \dots, Z_t$ to be independent random variables with

$$Z_j \sim \text{Poisson}\left(\theta\binom{j+m-1}{m-1}\lambda^j\right),\ \ j = 0,1,2,\dots.$$

4.2 Class Size Distribution for Bose-Einstein Allocation with No Empty Urns

For $P(C_1 = c_1, \dots, C_n = c_n)$ defined by (??) in the case $m_1 = \cdots = m_n = m$ and $M = mn$, let D_k equal the number of C_j's which equal k, $k \in \{m, m+1, \dots, t - m(n-1)\}$.

For a given vector $\big(d_m, d_{m+1}, \dots, d_{t-m(n-1)}\big)$ define the set $\mathbb{B}_D$ by

$$\mathbb{B}_D = \left\{(c_1, \dots, c_n) \middle| \sum_{j=1}^{n} \mathbb{I}\big(c_j = m\big) = d_m, \dots, \sum_{j=1}^{n} \mathbb{I}\left(c_j = t - m(n-1)\right) = d_{t-m(n-1)}\right\}.$$

Then

$$P\big(D_m = d_m, D_{m+1} = d_{m+1}, \dots, D_{t-m(n-1)} = d_{t-m(n-1)}\big)$$

$$= \sum_{\mathbb{B}_D}\cdots\sum P(C_1 = c_1, \dots, C_n = c_n)$$

$$= \sum_{\mathbb{B}_D} \cdots \sum \frac{\binom{c_1 - 1}{m-1} \cdots \binom{c_n - 1}{m-1}}{\binom{t-1}{mn-1}}$$

$$= \sum_{\mathbb{B}_D} \cdots \sum \frac{\binom{m-1}{m-1}^{d_m} \binom{(m+1)-1}{m-1}^{d_{m+1}} \cdots \binom{\left(t-m(n-1)\right)-1}{m-1}^{d_{t-m(n-1)}}}{\binom{t-1}{mn-1}}$$

$$= \frac{\binom{m-1}{m-1}^{d_m} \binom{(m+1)-1}{m-1}^{d_{m+1}} \cdots \binom{\left(t-m(n-1)\right)-1}{m-1}^{d_{t-m(n-1)}}}{\binom{t-1}{mn-1}} \times \frac{n!}{d_m!\, d_{m+1}! \cdots d_{t-m(n-1)}!}$$

(13)

where

$$d_m + d_{m+1} + \cdots + d_{t-m(n-1)} = n$$

$$m d_m + (m+1) d_{m+1} + \cdots + \left(t - m(n-1)\right) d_{t-m(n-1)} = t$$

$$M = mn \text{ and } d_j = \{0,1,2,\ldots\}.$$

Theorem 11

For the class size distribution given in (13) we have

$$\mathrm{E}\left(\Psi\left(D_m, D_{m+1}, \ldots, D_{t-m(n-1)}\right)\right)$$

$$= \frac{1}{\binom{t-1}{mn-1} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\mathrm{e}^{\frac{\theta\lambda^m}{(1-\lambda)^m}} \mathrm{E}\left(\Psi\left(Z_m, Z_{m+1}, \ldots, Z_{t-m(n-1)}\right)\right) \right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

(14)

where $Z_m, Z_{m+1}, \ldots, Z_{t-m(n-1)}$ are independent and $Z_j \sim \text{Poisson}\left(\binom{j-1}{m-1}\theta\lambda^j\right)$.

4.3 Class Size Distribution for Bose-Einstein Allocation with No Empty Groups

For $P(C_1 = c_1, \dots, C_n = c_n)$ defined by (??), let $m_1 = \cdots = m_n = m$, let $M = mn$, and let D_k equal the number of C_j's which equal k, $k \in \{1, \dots, t-n+1\}$. To obtain a general result for $\mathrm{E}\big(\Psi(D_1, \dots, D_{t-n+1})\big)$ in this situation we need to first use Theorem 6 to find the joint distribution of the D_j's.

Example 12

$$P(D_1 = d_1, \dots, D_{t-n+1} = d_{t-n+1})$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)} \frac{n!}{d_1! \cdots d_{t-n+1}!} \binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}}$$

(15)

where $d_1 + \cdots + d_{t-n+1} = n$ and $1d_1 + \cdots + (t-n+1)d_{t-n+1} = t$.

Proof. This result follows on direct application of Theorem 6 with

$$\Psi(C_1, \dots, C_n) = \mathbb{I}(D_1 = d_1, \dots, D_{t-n+1} = d_{t-n+1}).$$

We note that $D_k = 0$ for all $k > t-n+1$ when we require that no group can be left empty.

Theorem 13

For the class size distribution given in (15) we have

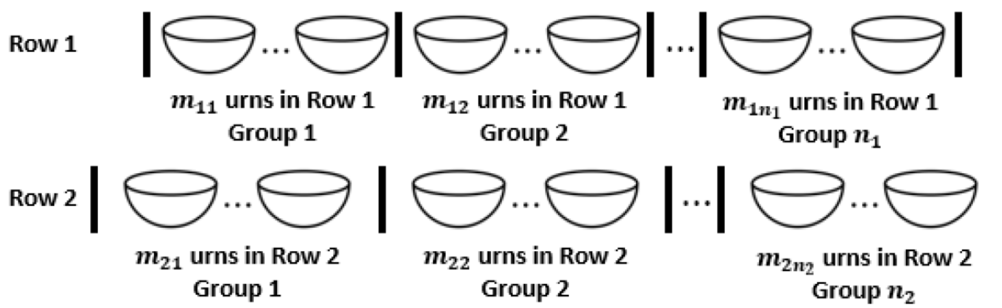
$$\mathrm{E}\big(\Psi(D_1, \dots, D_{t-n+1})\big)$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \mathrm{E}\big(\Psi(Z_1, \dots, Z_{t-n+1})\big) \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

where $Z_1, \ldots, Z_{t-n+1}$ are independent and $Z_j \sim \text{Poisson}\left(\binom{j+m-1}{m-1}\theta\lambda^j\right)$, $j = 1, \ldots, t-n+1$.

5 Bose-Einstein Allocation with Two Rows of Urns

Now imagine that we distribute t identical balls into two rows of urns. We will assume that the top row contains a total of M_1 urns and the bottom row contains a total of M_2 urns, that all $M_1 + M_2$ urns are distinct. Further suppose that each urn in the top row belongs to one of n_1 distinct groups where the j^{th} group contains m_{1j} urns with $m_{11} + \cdots + m_{1n_1} = M_1$. In the same way, we suppose that each urn in the bottom row belongs to one of n_2 groups where the j^{th} group contains m_{2j} urns with $m_{21} + \cdots + m_{2n_2} = M_2$.



Define the variables

C_{ij} : the number of balls in group j of row i

and

D_{ij} : the number of groups in row i containing j balls.

(By adopting this notation it becomes straightforward to generalize all the results in this section to cases of more than two rows.)

We consider two different methods for the distribution of the t identical balls into these $M_1 + M_2$ distinct urns. The first method is when all $\binom{M_1 + M_2 + t - 1}{t}$ possible allocations of these t balls are equally likely to occur.

The second method occurs when we impose the restriction that the first row is allocated a total of t_1 balls and the second row is allocated a total of t_2 balls where $t_1 + t_2 = t$. In this case we will assume that all $\binom{M_1 + t_1 - 1}{t_1}\binom{M_2 + t_2 - 1}{t_2}$ possible allocations of these $t = t_1 + t_2$ balls, subject to these row total restrictions, are equally likely to occur.

5.1 Allocation Without Individual Row Total Restrictions

In this section we suppose that t identical balls are distributed into the $M_1 + M_2$ distinct urns with no restrictions on the number of balls per row and where all $\binom{M_1 + M_2 + t - 1}{t}$ possible allocations of these t balls are equally likely to occur. In this case we have the following results.

Theorem 14

If $m_{11} = \cdots = m_{1n_1} = m_1$ and $m_{21} = \cdots = m_{2n_2} = m_2$, then

$$\mathrm{E}\Big(\Psi\big((D_{10}, D_{11}, \dots, D_{1t}), (D_{20}, D_{21}, \dots, D_{2t})\big)\Big)$$
$$= \frac{1}{\binom{m_1 n_1 + m_2 n_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda)^{m_1}} + \frac{\theta_2}{(1-\lambda)^{m_2}}} \times \mathrm{E}\Big(\Psi\big((Z_{10}, Z_{11}, \dots, Z_{1t}), (Z_{20}, Z_{21}, \dots, Z_{2t})\big)\Big) \right)\Bigg|_{\substack{\lambda=0 \\ \theta_1=0 \\ \theta_2=0}}$$

where $Z_{ij} \sim \text{Poisson}\left(\binom{j + m_i - 1}{m_i - 1} \theta_i \lambda^j\right)$ and where all random variables are independent.

Theorem 15

If $m_{11} = \cdots = m_{1n_1} = m_1$ and if we define $M_2 = m_{21} + \cdots + m_{2n_2}$, then,

$$\mathrm{E}\Big(\Psi\Big((D_{10}, D_{11}, \dots, D_{1t}), \big(C_{21}, \dots, C_{2n_2}\big)\Big)\Big) = \frac{1}{\binom{m_1 n_1 + M_2 + t - 1}{t} t!}$$
$$\times \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\left(\frac{1}{1-\lambda}\right)^{M_2} \mathrm{e}^{\frac{\theta_1}{(1-\lambda)^{m_1}}} \mathrm{E}\Big(\Psi\Big((Z_{10}, Z_{11}, \dots, Z_{1t}), \big(Z_{21}, \dots, Z_{2n_2}\big)\Big)\Big) \right)\Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

where $Z_{1j} \sim \text{Poisson}\left(\binom{j + m_1 - 1}{m_1 - 1} \theta_1 \lambda^j\right)$, $Z_{2j} \sim$ negative binomial$\big(m_{2j}, 1 - \lambda\big)$ and where all random variables are independent.

Note: An additional theorem for $\mathrm{E}\left(\Psi\left(\left(C_{11}, C_{12}, \dots, C_{1n_1}\right), \left(C_{21}, \dots, C_{2n_2}\right)\right)\right)$ is not necessary because this would be equivalent to Theorem 1 when there are not separate row total restrictions.

Example 16

Suppose we have two rows of empty urns. Assume that Row 1 contains n_1 groups of urns with m_1 urns in each group. Assume that Row 2 contains m_2 urns (*i.e.* $n_2 = 1$ group with m_2 urns in that group). We will assume that all $m_1 n_1 + m_2$ urns are distinct (*e.g.* the urns could be numbered). A total of t identical balls are distributed into these $m_1 n_1 + m_2$ distinct urns in such a way that all possible allocations are equally likely to occur.

What is the probability that Row 1 contains exactly j balls and exactly k of the n_1 groups in Row 1 are empty? In terms of the notation set up for Theorem 15, this question asks for $P(D_{10} = k, C_{21} = t - j)$.

Solution. We can use Theorem 15 with

$$\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), \left(C_{21}, \dots, C_{2n_2}\right)\right) = \mathbb{I}(D_{10} = k, C_{21} = t - j)$$

to show

$$P(D_{10} = k, C_{21} = t - j)$$

$$= \frac{\binom{n_1}{k}\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t}} \sum_{i=0}^{n_1 - k} (-1)^{n_1 - k - i} \binom{n_1 - k}{i} \binom{m_1 i + j - 1}{j}.$$

We can use Theorem 15 with

$$\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), \left(C_{21}, \dots, C_{2n_2}\right)\right) = (D_{10})_{[\nu]} (C_{21})_{[\delta]}$$

to show that

$$\mathrm{E}\big((D_{10})_{[\nu]}(C_{21})_{[\delta]}\big) = \frac{\binom{n_1}{\nu}\binom{m_2+\delta-1}{\delta}\binom{m_1(n_1-\nu)+m_2+t-1}{t-\delta}}{\binom{m_1n_1+m_2+t-1}{t}}\ \nu!\,\delta!.$$

Example 17

Suppose we randomly distribute t identical balls into a row of mn identical urns partitioned into n groups with m urns per group. Assume that all distinct allocations are equally likely to occur. If we are given the information that j of the first n_1 groups are empty then we can expect that

$$\frac{\displaystyle\sum_{k=0}^{n_2}\sum_{i=0}^{n-j-k} k\cdot(-1)^{n-i-j-k}\binom{n-j-k}{i}\binom{im+t-1}{t}\binom{n_2}{k}}{\displaystyle\sum_{i=0}^{n_1-j}(-1)^{n_1-j-i}\binom{n_1-j}{i}\binom{(n_2+i)m+t-1}{t}}$$

of the remaining $n_2 = n - n_1$ groups will be empty.

Proof. In the language of Theorem 14 we can think of the first n_1 groups of urns as our “top row” and the remaining $n_2 = n - n_1$ groups of urns as our “bottom row”. In the notation of Theorem 14 the problems asks for $\mathrm{E}(D_{20}|D_{10}=j)$.

We have that

$$\mathrm{E}(D_{20}|D_{10}=j) = \sum_{k=0}^{n_2} k\,\frac{P(D_{10}=j, D_{20}=k)}{P(D_{10}=j)}$$

and we can use Theorem 14 to find each of these probabilities. For the numerator probability we can apply Theorem 14 with

$$\Psi\big((D_{10},D_{11},\dots,D_{1t}),(D_{20},D_{21},\dots,D_{2t})\big) = \mathbb{I}(D_{10}=j, D_{20}=k).$$

For the denominator probability we can apply Theorem 14 with

$$\Psi\big((D_{10},D_{11},\dots,D_{1t}),(D_{20},D_{21},\dots,D_{2t})\big) = \mathbb{I}(D_{10}=j).$$

5.2 Allocation With Individual Row Total Restrictions

In this section we suppose that t_1 identical balls are distributed into the M_1 distinct urns in the first row and that t_2 identical balls are distributed into the M_2 distinct urns in the second row in a manner such that all $\binom{M_1+t_1-1}{t_1}\binom{M_2+t_2-1}{t_2}$ possible allocations of these $t_1+t_2=t$ balls are equally likely to occur. In this case we have the following results.

Theorem 18

Recall that we have defined C_{ij} to equal the number of balls allocated to the j^{th} group in row i. In this case,

$$
\begin{aligned}
&\mathrm{E}\left(\Psi\left(\left(C_{11},\ldots,C_{1n_1}\right),\left(C_{21},\ldots,C_{2n_2}\right)\right)\right)\\
&=\frac{1}{\binom{M_1+t_1-1}{t_1}\binom{M_2+t_2-1}{t_2}t_1!\,t_2!}\;\frac{d_1^t}{dp_1^{t_1}}\frac{d_2^t}{dp_2^{t_2}}\left(\left(\frac{1}{1-p_1}\right)^{M_1}\left(\frac{1}{1-p_2}\right)^{M_2}\right.\\
&\left.\times\,\mathrm{E}\left(\Psi\left(\left(Z_{11},\ldots,Z_{1n_1}\right),\left(Z_{21},\ldots,Z_{2n_2}\right)\right)\right)\right)\Big|_{\substack{p_1=0\\p_2=0}}
\end{aligned}
$$

where

$$Z_{1j}\sim \text{negative binomial}\left(m_{1j},1-p_1\right),\;\; j=1,2,\ldots,n_1$$

$$Z_{2j}\sim \text{negative binomial}\left(m_{2j},1-p_2\right),\;\; j=1,2,\ldots,n_2$$

$$Z_{11},\ldots,Z_{1n_1},Z_{21},\ldots,Z_{2n_2} \text{ are all independent}$$

and

$$M_1=m_{11}+\cdots+m_{1n_1},\; M_2=m_{21}+\cdots+m_{2n_2}.$$

Theorem 19

Recall that we have defined D_{ij} to equal the number of groups in row i containing j balls. In this case,

$$\mathrm{E}\left(\Psi\left(\left(D_{10},\dots,D_{1t_1}\right),\left(D_{20},\dots,D_{2t_2}\right)\right)\right)$$

$$= \frac{1}{\binom{m_1 n_1 + t_1 - 1}{t_1}\binom{m_2 n_2 + t_2 - 1}{t_2} t_1!\, t_2!} \frac{d^{t_1}}{d\lambda_1^{t_1}} \frac{d^{t_2}}{d\lambda_2^{t_2}} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda_1)^{m_1}} + \frac{\theta_2}{(1-\lambda_2)^{m_2}}} \right.$$

$$\left. \mathrm{E}\left(\Psi\left(\left(Z_{10},\dots,Z_{1t_1}\right),\left(Z_{20},\dots,Z_{2t_2}\right)\right)\right) \right)\Bigg|_{\substack{\lambda_1=0\\ \lambda_2=0\\ \theta_1=0\\ \theta_2=0}}$$

where

$$Z_{1j} \sim \text{Poisson}\left(\binom{j+m_1-1}{m_1-1}\theta_1\lambda_1^j\right),\ j = 0,1,\dots,t_1$$

$$Z_{2j} \sim \text{Poisson}\left(\binom{j+m_2-1}{m_2-1}\theta_2\lambda_2^j\right),\ j = 0,1,\dots,t_2$$

and

$Z_{10},\dots,Z_{1t_1},Z_{20},\dots,Z_{2t_2}$ are all independent.

Theorem 20

$$\mathrm{E}\left(\Psi\left(\left(D_{10},\dots,D_{1t_1}\right),\left(C_{21},\dots,C_{2n_2}\right)\right)\right)$$

$$= \frac{1}{\binom{m_1 n_1 + t_1 - 1}{t_1}\binom{M_2 + t_2 - 1}{t_2} t_1!\, t_2!} \frac{d^{t_1}}{d\lambda_1^{t_1}} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{dp_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda_1)^{m_1}}} \right.$$

$$\times \left(\frac{1}{1-p_2}\right)^{M_2} \mathrm{E}\left(\Psi\left(\left(Z_{10}, \dots, Z_{1t_1}\right), \left(Z_{21}, \dots, Z_{2n_2}\right)\right)\right)\Bigg|_{\substack{\lambda_1=0 \\ \theta_1=0 \\ p_2=0}} \tag{17}$$

where

$$Z_{1j} \sim \text{Poisson}\left(\binom{j+m_1-1}{m_1-1}\theta_1\lambda_1^j\right),\ j = 0,1,\dots,t_1$$

$$Z_{2j} \sim \text{negative binomial}\left(m_{2j}, 1-p_2\right),\ \ j = 1,2,\dots,n_2$$

$$Z_{10}, \dots, Z_{1t_1}, Z_{21}, \dots, Z_{2n_2} \text{ are all independent}$$

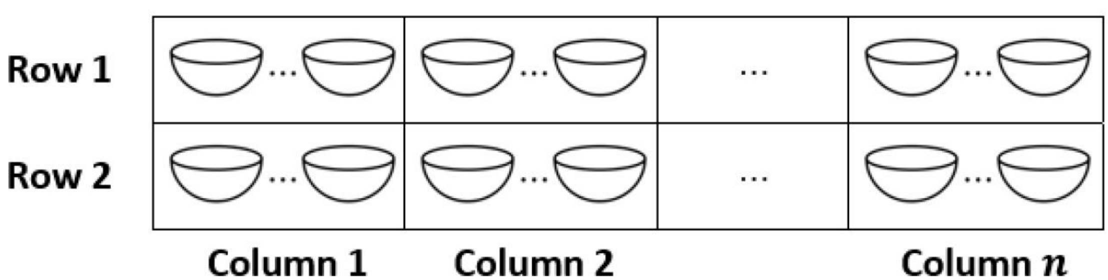
and

$$m_{11} = \cdots = m_{1n_1} = m_1 \text{ and } M_2 = m_{21} + \cdots + m_{2n_2}.$$

We will not provide proofs of these last three theorems in as much as Theorem 18 follows directly from Theorem 1, Theorem 19 follows directly from Theorem 9 and Theorem 20 follows from Theorem 15 with only obvious modifications.

Example 21

Consider a $2 \times n$ matrix of urns with m_1 urns per cell in the first row and m_2 urns per cell in the second row. Assume that all $m_1 n + m_2 n$ urns are distinguishable, that urns are arranged linearly within a cell (as shown in the diagram) and that the position of an urn within the matrix, including its position within a cell, is fixed.



A. Suppose that t identical balls are randomly distributed among the $m_1n + m_2n$ urns in such a way that all distinguishable allocations are equally likely to occur. Let W_A equal the number of empty columns. (We will say that a column is empty if all $m_1 + m_2$ urns in that column are empty.) Then

$$\mathrm{E}\big((W_A)_{[r_0]}\big) = \frac{\binom{(m_1+m_2)(n-r_0)+t-1}{t}}{\binom{(m_1+m_2)n+t-1}{t}} n_{[r_0]}.$$

B. Suppose that t_1 identical balls are randomly distributed among the urns in the first row and that t_2 identical balls are randomly distributed among the urns in the second row that in such a way that all distinguishable allocations are equally likely to occur. Let W_B equal the number of empty columns in this case. Then

$$\mathrm{E}\big((W_B)_{[r_0]}\big) = \frac{\binom{m_1(n-r_0)+t_1-1}{t_1}\binom{m_2(n-r_0)+t_2-1}{t_2}}{\binom{m_1n+t_1-1}{t_1}\binom{m_2n+t_2-1}{t_2}} n_{[r_0]}.$$

For any fixed value of $t = t_1 + t_2$ in part B, choosing t_1 and t_2 as

$$t_1 = \left\lfloor\frac{m_1t+m_2}{m_1+m_2}\right\rfloor \quad \text{and} \quad t_2 = t - \left\lfloor\frac{m_1t+m_2}{m_1+m_2}\right\rfloor$$

will minimize the expected number of empty columns.

Proof. If we identify each column of urns as a group, then the allocation model in Part A is equivalent to an unrestricted grouped Bose-Einstein allocation of t identical balls into a single row of urns divided into n groups (columns) with $m_1 + m_2$ urns per group.

Thus, W_A is equivalent to D_0 in Section 4.1 and

$$(W_A)_{[r_0]} = (D_0)_{[r_0]} = (D_0)_{[r_0]}(D_1)_{[0]}(D_2)_{[0]}\cdots(D_t)_{[0]}.$$

So, the result for $\mathrm{E}\big((W_A)_{[r_0]}\big)$ follows immediately from our work in Example 10 in Section 4.1.

We can express W_B as

$$W_B = \sum_{j=1}^{n} \mathbb{I}\big(C_{1j} + C_{2j} = 0\big)$$

and so the result for $\mathrm{E}\big((W_B)_{[r_0]}\big)$ in Part B follows on application of Theorem 18 with

$$\Psi((C_{11}, \dots, C_{1n}), (C_{21}, \dots, C_{2n})) = \left(\sum_{j=1}^{n} \mathbb{I}\big(C_{1j} + C_{2j} = 0\big)\right)_{[r_0]}.$$

Finally, to verify that for any fixed value of $t = t_1 + t_2$

$$h(t_1) = \mathrm{E}\big((W_B)_{[1]}\big) = \frac{\binom{m_1(n-1)+t_1-1}{t_1}\binom{m_2(n-1)+t-t_1-1}{t-t_1}}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t - t_1 - 1}{t - t_1}} n$$

is minimized at $t_1 = \left\lfloor \frac{m_1 t + m_2}{m_1 + m_2} \right\rfloor$ it suffices to algebraically verify that

$$\frac{h(t_1)}{h(t_1 - 1)} < 1 \Leftrightarrow t_1 < \frac{m_1 t + m_2}{m_1 + m_2}.$$

6 Compositions

A composition of the positive integer t is defined as any solution $(x_1, x_2, \dots)$ of $x_1 + x_2 + \cdots = t$ where $x_j \in \{1,2,\dots\}$. The numbers $x_1, x_2, \dots$ are called the parts of the composition. A composition of t can consist of anywhere from 1 to t parts. In a composition, the order of the parts is considered important. That is, the composition $1 + 3$ of $t = 4$ is considered to be distinct from the composition $3 + 1$. (A solution $x_1 + x_2 + \cdots = t$ where the order of parts is not considered important is called a partition of t.) The set of (distinct) elements in the list of parts of a composition are called the part sizes of that composition. The multiplicity of a part size j in a composition $(x_1, x_2, \dots)$ is defined as the number of times that part size j occurs in the ordered list of its parts $(x_1, x_2, \dots)$. The number of part sizes with a given multiplicity (i.e. the multiplicity of a part size multiplicity) is also a concept defined in the study of compositions and partitions. See for example [16, 22, 30].

By way of example, the composition $2 + 1 + 1 + 3 + 2 + 1 + 1 + 4$ of 15 has 8 parts. The set of distinct numbers (part sizes) in this composition are $1,2,3$ and 4. Part size 1 occurs four times so we say the multiplicity of 1 is 4. Similarly, the multiplicity of part size 2 is 2 and the multiplicity of part sizes 3 and 4 are both 1. The list of part size multiplicities is $4,2,1$ and 1. So two part sizes have multiplicity 1, one part size has multiplicity 2 and one part size has multiplicity 4.

The number of compositions of t into n parts is the same as the number of ways to distribute t identical balls into n distinct urns with no empty urns. So there are $\binom{t-1}{n-1}$ distinct compositions of t into n parts. Thus there are a total of $\sum_{n=1}^{t} \binom{t-1}{n-1} = 2^{t-1}$ compositions of t. A random composition of t is defined as a composition picked at random from the set of all 2^{t-1} compositions of t.

6.1 Random Compositions with a Fixed Number of Parts

We can view a random composition of t with n parts as a Bose-Einstein allocation of t identical balls into n distinct urns with the restriction that no urn can be left empty. That is, Theorems 4 and 11, with $m_1 = \cdots = m_n = 1$ can be applied to random compositions with a fixed number of parts.

Example 22 Ordered values in a composition with a fixed number of parts

Let $(X_1, X_2, \dots, X_n)$ be a random composition of t with n parts. Let $X_{(1:n)} \le X_{(2:n)} \le \cdots \le X_{(n:n)}$ be the order statistics (ordered values) of our data $X_1, X_2, \dots, X_n$. Then for nonnegative integers m

$$\mathrm{E}\left(\left(X_{(j:n)}\right)^m\right) = \frac{1}{\binom{t-1}{n-1}} \sum_{k=n-j+1}^{n} \sum_{z=0}^{\lfloor \frac{t-n}{k} \rfloor} (-1)^{k-n+j-1} \binom{t-zk-1}{t-zk-n} \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^m - z^m\right).$$

As a check, one can verify that $\sum_{j=1}^{n} E\left(X_{(j:n)}\right) = \sum_{j=1}^{n} E\left(X_j\right) = t.$

Proof. The result follows from Theorem 4 if we identify the C_i in Theorem 4 with the X_i in this example and take $\Psi(C_1, \dots, C_n) = \Psi(X_1, \dots, X_n) = (X_{(j:n)})^m$. When applying Theorem 4 we will have

$$\mathrm{E}\left(\left(Z_{(j:n)}\right)^m\right) = \sum_{z=0}^{\infty} P\left(Z_{(j:n)} > z\right)\left((z+1)^m - z^m\right)$$

by Theorem 39 in Appendix 9 (a general result for expressing raw moments in terms of the cumulative distribution function for a discrete random variable defined on the nonnegative integers) and we can solve for

$$P\left(Z_{(j:n)} > z\right) = P(\text{at least } n-j+1 \text{ of } Z_1, \dots, Z_n > z)$$

using the generalized inclusion-exclusion principle.

In the next example we will use Theorem 7 to derive the asymptotic value of $\mathrm{E}\left(\left(X_{(j:n)}/t\right)^m\right)$ for m a nonnegative integer. We can then track the large t approximation for $E\left(\left(X_{(j:n)}\right)^m\right)$ against the exact value found in Example 22.

Example 23

Let $(X_1, X_2, \dots, X_n)$ be a random composition of t with n parts. Then for large t and nonnegative integer m,

$$\mathrm{E}\left(\left(X_{(j:n)}\right)^m\right) \approx \frac{t^m}{\binom{n+m-1}{m}} \sum_{r=n-j+1}^{n} (-1)^{r-n+j-1} \binom{r-1}{n-j} \binom{n}{r} \frac{1}{r^m}.$$

Proof. The joint distribution of a random composition $(X_1, X_2, \dots, X_n)$ is given by (??) with $m_1 = \cdots = m_n = 1$ where our X_j are the C_j in (??). Therefore, as we explain in Section 3.4, Theorem 7 is applicable to the problem of random compositions with a fixed number of parts whenever our chosen statistic $\Psi(x_1, x_2, \dots, x_n)$ is defined on all $\mathbb{R}^n_{\geq 0}$ and is bounded and continuous on the simplex $x_1 + x_2 + \cdots + x_n = 1$, $0 \leq x_j \leq 1$ for $j = 1,2, \dots, n$.

We have $0 \leq X_j/t \leq 1$ for each $j = 1,2, \dots, n$ in this problem, so

$$\Psi\left(\frac{X_1}{t}, \dots, \frac{X_n}{t}\right) = \left(\frac{X_{(j:n)}}{t}\right)^m$$

is clearly a bounded function in the variables $(X_1/t, X_2/t, \dots, X_n/t)$. To show that $\Psi(X_1/t, \dots, X_n/t) = (X_{(j:n)}/t)^m$ is a continuous function in the variables $(X_1/t, X_2/t, \dots, X_n/t)$ requires a few extra steps. First, the well-known identity

$$\min(a, b) = \frac{(a+b) - |a-b|}{2}$$

makes it clear that

$$\min\big(f(x_1, \dots, x_n), g(x_1, \dots, x_n)\big) \tag{18}$$

is continuous provided $f(\,)$ and $g(\,)$ are continuous. Second, the identity

$$\min(x_1, x_2, x_3) = \min(\min(x_1, x_2), x_3)$$

along with (18) and an induction argument shows that $\min(x_1, x_2, \ldots, x_n)$ is a continuous function in the variables $(x_1, x_2, \ldots, x_n)$.

Finally, letting $g(x) = x^m$ in the pointwise identity (i.e. valid for any given set of numbers $\{x_1, x_2, \ldots, x_n\}$)

$$g(x_{(j:n)}) = \sum_{r=n-j+1}^{n} \sum_{(k_1,\ldots,k_r)\in\mathbb{C}_r} (-1)^{r-n+j-1} \binom{r-1}{n-j} g\left(\min(x_{k_1}, \ldots, x_{k_r})\right)$$

(19)

given in Suman [44], where $\mathbb{C}_r$ is defined to be the set of all subsets of $\{1,2,\ldots,n\}$ with r elements shows that $\Psi(X_1/t, \ldots, X_n/t) = (X_{(j:n)}/t)^m$ is a continuous function in the variables $(X_1/t, X_2/t, \ldots, X_n/t)$. So we can conclude that Theorem 7 is applicable in this example.

By Theorem 7, we have that

$$\lim_{t\to\infty} \mathrm{E}\left(\left(\frac{X_{(j:n)}}{t}\right)^m\right) = (n-1)!\,\mathcal{L}^{-1}\left(\frac{\mathrm{E}((Y_{(j:n)})^m)}{\lambda^n}\right)\Bigg|_{s=1}$$

(20)

where $Y_{(j:n)}$ is the j^{th} order statistic of $Y_1, Y_2, \ldots, Y_n$ *iid* $\text{Gamma}(1, \lambda) \equiv \text{Exponential}(\lambda)$. The final result follows on simplification of (20).

The following table shows one example ($n = 8$ and $j = 5$) of how the asymptotic approximation developed in Example 23 and the exact answer in Example 22 track for $\mathrm{E}(X_{(5:8)})$ at increasing values of t.

t	**Exact**	**Approximation**
8	1	0.88

50	5.57	5.53
100	11.11	11.06
500	55.34	55.28
5000	552.89	552.83
50,000	5528.33	5528.27

$E(X_{(5:8)})$

6.2 Compositions with a Fixed Number of Parts, all Distinct

To start with, we recall that $\prod_{j=1}^{\infty}\left(1+\theta^{j}\right)=\sum_{t=0}^{\infty} d(t)\theta^{t}$ where $d(t)$ is the number of partitions of the integer t with distinct parts.

e.g. $d(5)=3$ because there are 3 partitions of the integer 5 where each part is distinct. These three partitions with distinct parts are $\{1,4\}$, $\{2,3\}$ and $\{5\}$.

Similarly, it is well known that $\prod_{j=1}^{\infty}\left(1+\lambda\,\theta^{j}\right)=\sum_{t=0}^{\infty}\sum_{n=1}^{t} d(t,n)\lambda^{n}\theta^{t}$ where $d(t,n)$ is the number of partitions of the integer t into exactly n distinct parts.

e.g. $d(5,1)=1$ and $d(5,2)=2$. In particular $\{5\}$ is the only partition of 5 with one part and $\{1,4\}$ and $\{2,3\}$ are the two partitions of 5 with exactly 2 distinct parts.

Define $c_d(t,n)$ as the number of compositions (ordered partitions) of the integer t into n distinct parts. Clearly partitions with n distinct parts and compositions with n distinct parts are related by $c_d(t,n)=n!\ d(t,n)$.

It is also clear that the set of all $d(t,n)$ partitions of the integer t into n distinct parts corresponds to the set of all solutions $(y_1,y_2,\dots,y_t)$ to the pair of equations $1y_1+2y_2+\cdots+ty_t=t$ and $y_1+y_2+\cdots+y_t=n$ with $y_j\in\{0,1\}, j=1,\dots,t$.

Thus, the probability distribution for a partition of the integer t into n distinct parts randomly picked from the set of all $d(t,n)$ such partitions has the form

$$P(Y_1 = y_1, \dots, Y_t = y_t) = \begin{cases} \dfrac{1}{d(t,n)} & \begin{matrix} 1y_1 + \cdots + ty_t = t \\ y_1 + \cdots + y_t = n \\ y_j \in \{0,1\},\ j = 1, \dots, t \end{matrix} \\ 0 & \text{else.} \end{cases}$$

Theorem 24

For $(Y_1, \dots, Y_t)$ as defined above we have

$$\mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_t)\big)$$
$$= \frac{1}{d(t,n)n!\,t!} \frac{d^n}{d\lambda^n} \frac{d^t}{d\theta^t} \left(\left(\prod_{j=1}^{\infty} \left(1 + \lambda\theta^j\right) \right) \mathrm{E}\big(\Psi(Z_1, Z_2, \dots, Z_t)\big) \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

where $Z_1, Z_2, \dots$ is an infinite sequence of independent random variables with $Z_j \sim$ Binomial $\left(1, \frac{\lambda\theta^j}{1+\lambda\theta^j}\right)$.

Example 25

$$P(k \text{ occurs in a random composition of } t \text{ into } n \text{ parts, all distinct})$$
$$= \frac{1}{d(t,n)} \sum_{v=0}^{\min\left(\left\lfloor \frac{t-k}{k} \right\rfloor, n-2\right)} (-1)^v d(t - k - kv, n - v - 1).$$

Proof. The result follows directly from Theorem 24 with $\Psi(Y_1, Y_2, \dots, Y_t) = \mathbb{I}(Y_k = 1)$.

6.3 Random Compositions

We can use the rule of iterated expectations, $E(X) = E\big(E(X|Y)\big)$, in conjunction with Theorems 4 and 11 (with $m_1 = \cdots = m_n = 1$), to develop companion results in the case of compositions without a fixed number of parts.

Theorem 26

If C_j represents the size of the j^{th} part in a random composition of the positive integer t, then for a general statistic $\Psi(\ )$,

$$\mathrm{E}\big(\Psi(C_1, C_2, \dots, C_t)\big)$$
$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t}\frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^n \mathrm{E}\big(\Psi^*(Z_1, Z_2, \dots, Z_n)\big)\right)\Bigg|_{p=0} \tag{21}$$

where $Z_1, Z_2, \dots$ are independent random variables such that $Z_j \sim 1\text{-shifted geometric}(1-p)$ and

$$\Psi^*(a_1, a_2, \dots, a_n) = \begin{cases}\Psi(a_0, a_1, \dots, a_n, 0, \dots, 0) & n < t\\ \Psi(a_0, a_1, \dots, a_t) & n = t\end{cases}$$

with $a_{n+1}, \dots, a_t$ all being replaced with 0's in the case $n < t$.

Theorem 27

If D_j represents the multiplicity of j in a random composition of the positive integer t, then for a general statistic $\Psi(\ )$,

$$\mathrm{E}\big(\Psi(D_1, D_2, \dots, D_t)\big)$$
$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t}\frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\frac{\theta\lambda}{1-\lambda}}\mathrm{E}\big(\Psi(Z_1, Z_2, \dots, Z_t)\big)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

(22)

where $Z_1, Z_2, \dots, Z_t$ are independent and $Z_j \sim \text{Poisson}\left(\theta\lambda^j\right)$.

Example 28

Suppose that we select a part from a random composition of t in a manner such that any particular part is selected with probability proportional to its size. To put this in some context, again think of a random composition of t as a random distribution of t identical balls into ordered urns (i.e. the ordered parts). Then selecting a ball uniformly at random from the set of t balls as a method of identifying an urn would mean that each urn (part) is selected with a probability proportional to its size.

Let Y represent the size of this randomly selected part. Then

$$P(Y = y) = \begin{cases} \dfrac{y(t-y+3)}{t\,2^{y+1}} & y \in \{1,2,\dots,t-1\} \\ \dfrac{1}{2^{t-1}} & y = t. \end{cases}$$

As a corollary to the proof of this result we have that

$$\mathrm{E}\left(D_y\right) = \begin{cases} \dfrac{t-y+3}{2^{y+1}} & y \in \{1,2,\dots,t-1\} \\ \dfrac{1}{2^{t-1}} & y = t \end{cases}$$

where D_y represents the multiplicity of y in a random composition of t. As a check on our work we can verify that $\sum_{y=1}^{t} P(Y = y) = 1$ and $\mathrm{E}(1D_1 + 2D_2 + \cdots + tD_t) = t$.

Proof. Continuing with the balls in urns analogy mentioned in the statement of the problem, if d urns contained exactly y balls, then the probability of selecting a ball from an urn with y balls would just be $(yd)/t$. Interpreted back into the language of compositions this means that

$$P\big(Y = y|D_y = d\big) = \frac{yd}{t}$$

and

$$P(Y = y) = \sum_{d=1}^{t} P\big(Y = y|D_y = d\big)P\big(D_y = d\big) = \frac{y}{t}\mathrm{E}\big(D_y\big).$$

Then we can find $\mathrm{E}\big(D_y\big)$ and hence $P(Y = y)$ by direct application of Theorem 27 with $\Psi(D_1, D_2, \ldots, D_t) = D_y$.

Example 29 (Number of part sizes with a given multiplicity in a random composition)

1. Let V_w^t represent the number of part sizes with multiplicity w in a random composition of t.

Then

$$\mathrm{E}(V_w^t) = \frac{1}{2^{t-1}} \sum_{n=\max\{1,w\}}^{t} \sum_{r=1}^{t} \sum_{j=0}^{n-w} (-1)^{n-w-j} \binom{n}{w}\binom{n-w}{j}\binom{t-r(n-j)-1}{t-r(n-j)-j}.$$

As a check on this result, one can verify that $\sum_{w=0}^{t} \mathrm{E}(V_w^t) = t$ as required.

2. The number of distinct part sizes in a random composition of t equals $t - V_0^t$ and

$$\mathrm{E}(t - V_0^t) = t - \frac{1}{2^{t-1}} \sum_{n=1}^{t} \sum_{r=1}^{t} \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j}\binom{t-r(n-j)-1}{t-r(n-j)-j}.$$

Hitczenko and Savage [30] derive an asymptotic approximation for $E(V_w^t)$. They show

$$\mathrm{E}(V_w^t) \approx \binom{\lfloor t/2 \rfloor}{w} \sum_{j=1}^{\infty} \left(2^{-j}\right)^w \left(1-2^{-j}\right)^{\lfloor t/2 \rfloor - w} \approx \frac{1}{w\ \ln(2)}$$

for large t. Hitczenko and Stengle [31] show that for large t

$$E(t - V_0^t) \approx \log_2(t) - \frac{\gamma}{\ln(2)} - \frac{3}{2} + g\left(\log_2(t)\right) + o(1)$$

where $\gamma \approx 0.5772156649$ is Euler's constant and g is a mean-zero function of period 1 satisfying $|g| \leq 0.0000016$. Goh and Schmutz [24] and Wilf [51] derive companion asymptotic results for random partitions and we will compare their results with exact expressions in a separate paper.

The following tables show how Hitczenko and Savage's [30] and Hitczenko and Stengle's [31] asymptotic approximations track with the exact value found in Example 29.

t	Exact value of $\mathrm{E}(V_3^t)$	t	Exact value of $\mathrm{E}(V_3^t)$
50	0.480824	140	0.481384
60	0.481359	150	0.481355
70	0.481297	160	0.481199
80	0.480844	170	0.480979
90	0.480460	175	0.480865
100	0.480395	200	0.480476
110	0.480618	250	0.480876
120	0.480959	300	0.481299
130	0.481246	350	0.481008
Asymptotic Approximation $\lim_{t\to\infty} \mathrm{E}(V_3^t) \approx \frac{1}{3\ln(2)} = 0.480898$			

w	Exact Value of $\mathrm{E}(V_w^{100})$	Asymptotic Approximation
5	0.290129	0.288539
10	0.138741	0.144270
15	0.100250	0.096180
20	0.058805	0.072135
25	0.071275	0.057708
30	0.048332	0.048090
35	0.016881	0.041220

t	Exact Value of $\mathrm{E}(t - V_0^t)$	Asymptotic Approximation
25	3.980203	3.976602
50	4.978260	4.976602
75	5.562663	5.561565
100	5.977430	5.976602
200	6.977014	6.976602

Proof

Part A.

V_w^t, the number of part sizes with multiplicity w, can be expressed as

$$V_w^t = \sum_{r=1}^{t} \mathbb{I}(D_r = w).$$

Therefore $\mathrm{E}(V_w^t)$ follows from Theorem 27 with

$$\Psi(D_1, D_2, \dots, D_t) = \sum_{r=1}^{t} \mathbb{I}(D_r = w).$$

Part B.

V_0^t equals the number of part sizes with multiplicity 0 in a random composition of t. That is, V_0^t equals the number of part sizes that do not occur in a random composition of t. Thus $t - V_0^t$ equals the number of part sizes which do occur in a random composition of t. Therefore,

$E(t - V_0^t) = t - E(V_0^t)$ equals the expected number of part sizes which occur in a random composition of t.

7 Pólya Urn Model

Suppose an urn initially contains m_j balls of color $j, j = 1, \dots, n$. Balls are drawn at random and then returned along with another ball of the same color. Let C_j equal the number of times a ball of color j was selected in the first t draws where $C_1 + C_2 + \cdots + C_n = t$. In this case the joint probability distribution of $(C_1, \dots, C_n)$ is the same as the joint distribution given by (??) for the grouped Bose-Einstein allocation model. Namely,

$$P(C_1 = c_1, \dots, C_n = c_n)$$

$$= \begin{cases} \dfrac{\binom{c_1 + m_1 - 1}{m_1 - 1} \cdots \binom{c_n + m_n - 1}{m_n - 1}}{\binom{M + t - 1}{t}} & \begin{array}{c} c_1 + \cdots + c_n = t \\ m_1 + \cdots + m_n = M \\ c_j \in \{0,1,\dots\},\ j = 1, \dots, n \\ m_j \in \{1,2,\dots\},\ j = 1, \dots, n \end{array} \\ 0 & \text{else.} \end{cases}$$

It is a straightforward proof by induction on t to see how this comes about. We note that

$$P_t(C_1 = c_1, \dots, C_n = c_n)$$
$$= P_{t-1}(C_1 = c_1 - 1, C_2 = c_2, \dots, C_n = c_n) \cdot \left(\frac{m_1 + c_1 - 1}{M + t - 1}\right)$$
$$+ \cdots + \ P_{t-1}(C_1 = c_1, \dots, C_{n-1} = c_{n-1}, C_n = c_n - 1) \cdot \left(\frac{m_n + c_n - 1}{M + t - 1}\right).$$

The given result is trivial for the case $t = 1$ and assuming the result is true for the case $t - 1$, we get

$$P_t(C_1 = c_1, \ldots, C_n = c_n)$$

$$= \frac{\binom{m_1 + c_1 - 2}{c_1 - 1}\binom{m_2 + c_2 - 1}{c_2}\cdots\binom{c_n + m_n - 1}{c_n}}{\binom{M + t - 2}{t - 1}} \cdot \left(\frac{m_1 + c_1 - 1}{M + t - 1}\right)$$

$$+ \cdots + \frac{\binom{m_1 + c_1 - 1}{c_1}\cdots\binom{m_{n-1} + c_{n-1} - 1}{c_{n-1}}\binom{m_n + c_n - 2}{c_n - 1}}{\binom{M + t - 2}{t - 1}} \cdot \left(\frac{m_n + c_n - 1}{M + t - 1}\right)$$

$$= \binom{m_1 + c_1 - 1}{c_1}\cdots\binom{m_n + c_n - 1}{c_n} \cdot \frac{c_1 + \cdots + c_n}{\binom{M + t - 2}{t - 1}(M + t - 1)}$$

$$= \frac{\binom{m_1 + c_1 - 1}{c_1}\cdots\binom{m_n + c_n}{c_n}}{\binom{M + t - 1}{t}}.$$

It follows that all the results developed in Sections 3, 4, 5, and 6, properly translated, apply to the Pólya urn model.

Example 30 Ordered Values in the Pólya Urn Model

Let the joint probability distribution of $(C_1, C_2, \ldots, C_n)$ be given by (??) with $m_1 = m_2 = \cdots = m_n = 1$. Let $C_{(1:n)} \leq C_{(2:n)} \leq \cdots \leq C_{(n:n)}$ be the order statistics of $C_1, C_2, \ldots, C_n$.

Then

$$\mathrm{E}\left(\left(C_{(j:n)}\right)^r\right) = \frac{1}{\binom{n+t-1}{t}} \sum_{k=n-j+1}^{n} \sum_{z=0}^{\lfloor\frac{t-k}{k}\rfloor} (-1)^{k-n+j-1} \binom{n+t-k(z+1)-1}{t-k(z+1)} \times \binom{k-1}{n-j}\binom{n}{k} \left((z+1)^r - z^r\right).$$

(23)

As a check, one can verify that $\sum_{j=1}^{n} E\left(C_{(j:n)}\right) = \sum_{j=1}^{n} E\left(C_j\right) = t$. Recall that we proved in Example 23 that for a random composition $(X_1, X_2, \dots, X_n)$ of t with n parts that

$$\lim_{t\to\infty} \mathrm{E}\left(\left(\frac{X_{(j:n)}}{t}\right)^r\right) = \frac{1}{\binom{n+r-1}{r}} \sum_{i=n-j+1}^{n} (-1)^{i-n+j-1} \binom{i-1}{n-j}\binom{n}{i}\frac{1}{i^r}.$$

(24)

Compositions with a fixed number of parts belong to the case of urn restricted (no empty urns) Bose-Einstein allocation while this example belongs to the case of unrestricted Bose-Einstein allocation. However, we proved as part of the formula given in Theorem 7 that unrestricted, urn restricted and group restricted Bose-Einstein allocation all have the same asymptotic expectations. Hence (24) applies to this example as well.

Therefore,

$$\lim_{t\to\infty} \mathrm{E}\left(\left(\frac{C_{(j:n)}}{t}\right)^r\right) = \frac{1}{\binom{n+r-1}{r}} \sum_{i=n-j+1}^{n} (-1)^{i-n+j-1} \binom{i-1}{n-j}\binom{n}{i}\frac{1}{i^r}.$$

Chen [15] derives recurrence relationships for calculating

$$\lim_{t\to\infty} \mathrm{E}\left(\left(\frac{C_{(n:n)}}{t}\right)^r\right) = \frac{1}{\binom{n+r-1}{r}} \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i}\frac{1}{i^r}.$$

Chen also includes numerical calculations of

$$\mathrm{E}\left(C_{(n:n)}\right)=\frac{1}{\binom{n+t-1}{t}}\sum_{k=1}^{n}\sum_{z=0}^{\lfloor\frac{t-k}{k}\rfloor}(-1)^{k-1}\binom{n+t-k(z+1)-1}{t-k(z+1)}\binom{n}{k}$$

based on computer simulations for ranges of t from 100 to 20000 and with n ranging from 1 to 25.

Proof. The proof follows the same line developed in Example 22 (ordered values in a composition with a fixed number of parts) but starting with Theorem 1 (unrestricted Bose-Einstein allocation) instead of Theorem 4 (urn restricted Bose-Einstein allocation).

7.1 Pólya Process

Suppose we select a ball at random from an urn initially contains m_j balls of color j, $j = 1,2,\dots,n$ at times $i = 1,2,\dots$. After each draw we note the color of the ball and then replace the ball along with an additional ball of the same color (i.e. Pólya sampling).

Let $X_i = (X_{i1},\dots,X_{in})$ where

$$X_{ij}=\begin{cases}1 & \text{if } i^{th} \text{ ball drawn has color } j\\ 0 & \text{else.}\end{cases}$$

Then $\{X_i, i = 1,2,3,\dots\}$ forms a discrete-time Pólya stochastic process.

Let $Y_i = (Y_{i1},\dots,Y_{in})$ where $Y_i|p_1,p_2,\dots,p_n \sim$ Multivariate Bernoulli $(p_1,p_2,\dots,p_n)$. That is, for any given vector $(p_1,p_2,\dots,p_n)$ such that $0 \le p_j \le 1$, and $p_1+p_2+\cdots+p_n = 1$,

$$P\left(Y_{i1}=y_{i1},Y_{i2}=y_{i2},\dots,Y_{in}=y_{in}|p_1,p_2,\dots,p_n\right)=p_1^{y_{i1}}p_2^{y_{i2}}\cdots p_n^{y_{in}}$$

where $y_{ij} \in \{0,1\}$, $y_{i1} + y_{i2} + \cdots + y_{in} = 1$. Then $\{Y_i | p_1, p_2, \dots, p_n, i = 1,2,3, \dots\}$ forms a discrete-time multivariate Bernoulli stochastic process.

Our next theorem shows how these two processes are related.

Theorem 31

$$\mathrm{E}\big(\Psi(X_1, X_2, \dots, X_t)\big) = \int_{\text{simplex}} \cdots \int \mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_t) | p_1, p_2, \dots, p_n\big)$$
$$\times \frac{\Gamma(m_1 + m_2 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)} p_1^{m_1 - 1} \cdots p_n^{m_n - 1} dp_1 dp_2 \cdots dp_n.$$

Proof.

It is well known (see for example Feller [21]) and straightforward to establish that

$$P\big(X_1 = (x_{11}, \dots, x_{1n}), X_2 = (x_{21}, \dots, x_{2n}), \dots, X_t = (x_{t1}, \dots, x_{tn})\big)$$
$$= \frac{m_1^{[x_{11} + \cdots + x_{t1}]} m_2^{[x_{12} + \cdots + x_{t2}]} \cdots m_n^{[x_{1n} + \cdots + x_{tn}]}}{(m_1 + \cdots + m_n)^{[(x_{11} + \cdots + x_{t1}) + \cdots + (x_{1n} + \cdots + x_{tn})]}}$$

where $x_{1j} + \cdots + x_{tj}$ represents the total number of times a ball of color j is selected in these $(x_{11} + \cdots + x_{t1}) + \cdots + (x_{1n} + \cdots + x_{tn}) = t$ draws. Recall that $a^{[b]}$ is our notation for a rising factorial.

But suppose we let $(p_1, \dots, p_n) \sim \text{Dirichlet}(m_1, \dots, m_n)$. That is, let the joint density function of $(p_1, p_2, \dots, p_n)$ be given by

$$f(p_1, p_2, \dots, p_n) = \frac{\Gamma(m_1 + m_2 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)} p_1^{m_1 - 1} \cdots p_n^{m_n - 1}$$

for all points $(p_1, p_2, \dots, p_n)$ such that $p_1 + p_2 + \cdots + p_n = 1$ and $0 \le p_j \le 1$ for $j = 1,2,\dots,n$. Then it follows that

$$P(Y_1 = (x_{11}, \dots, x_{1n}), Y_2 = (x_{21}, \dots, x_{2n}), \dots, Y_t = (x_{t1}, \dots, x_{tn}))$$

$$= \underset{\text{simplex}}{\int \cdots \int} P\big(Y_1 = (x_{11}, \dots, x_{1n}), \dots, Y_t = (x_{t1}, \dots, x_{tn}) | p_1, \dots, p_n\big)$$

$$\times f(p_1, p_2, \dots, p_n) dp_1 dp_2 \cdots dp_n$$

$$= \underset{\text{simplex}}{\int \cdots \int} p_1^{x_{11}+\cdots+x_{t1}} \cdots p_n^{x_{1n}+\cdots+x_{tn}} \frac{\Gamma(m_1 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)} p_1^{m_1-1} \cdots p_n^{m_n-1} dp_1 \cdots dp_n$$

$$= \frac{\Gamma(m_1 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)}$$

$$\times \frac{\Gamma(m_1 + (x_{11} + \cdots + x_{t1})) \cdots \Gamma(m_n + (x_{1n} + \cdots + x_{tn}))}{\Gamma(m_1 + \cdots + m_n + (x_{11} + \cdots + x_{t1}) + \cdots + (x_{1n} + \cdots + x_{tn}))}$$

$$= \frac{m_1^{[x_{11}+\cdots+x_{t1}]} m_2^{[x_{12}+\cdots+x_{t2}]} \cdots m_n^{[x_{1n}+\cdots+x_{tn}]}}{(m_1 + \cdots + m_n)^{[(x_{11}+\cdots+x_{t1})+\cdots+(x_{1n}+\cdots+x_{tn})]}}.$$

Thus,

$$\mathrm{E}\big(\Psi(X_1, X_2, \dots, X_t)\big) = \mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_t)\big)$$

$$= \mathrm{E}\Big(\mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_t)\big) \big| p_1, \dots, p_n\Big)$$

$$= \underset{\text{simplex}}{\int \cdots \int} \mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_t) | p_1, p_2, \dots, p_n\big)$$

$$\times \frac{\Gamma(m_1 + m_2 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)} p_1^{m_1-1} \cdots p_n^{m_n-1} dp_1 dp_2 \cdots dp_n.$$

As part of this proof we can see that the random variables $X_1, X_2, \ldots, X_t$ are not independent because $(x_{11} + \cdots + x_{t1}) + \cdots + (x_{1n} + \cdots + x_{tn})$ must equal t so the support cannot be rectangular. But $X_1, X_2, \ldots, X_t$ are exchangeable random variables. That is, the probability of drawing a certain sequence of colored balls depends on the number of each color but not on the order in which they occur. So, the probability of drawing (Black, Yellow, Green, Yellow) in that order is the same as the probability of drawing (Green, Black, Yellow, Yellow), etc.

However, the random variables $Y_1, Y_2, \ldots, Y_t$ are independent. So, Theorem 31 is another example of expectation transposition.

Example 32 Equalization of Colors for the First Time

Suppose an urn initially contains b black and w white balls, with $b > w$. Balls are drawn at random and then returned along with another ball of the same color (Pólya sampling). Then the probability that the number of black balls and the number of white in the urn become equal for the first time on the t draw equals

$$\frac{b-w}{t}\binom{t}{\frac{1}{2}\left(t-(b-w)\right)}\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}\cdot\frac{\Gamma\left(\frac{1}{2}(t+b+w)\right)\Gamma\left(\frac{1}{2}(t+b+w)\right)}{\Gamma(t+b+w)}.$$

Proof. Consider a sequence of independent Bernoulli trials (success probability p, failure probability $1-p$). A random walker starts at position $z \geq 0$ on a number line and at each step either goes up one (success) or down one (failure) according to the outcome of that Bernoulli trial. Feller [21] shows that the probability $d(z,t)$ that this random walker will reach the origin on this number line for the first time on the t^{th} trial is given by

$$d(z,t) = \frac{z}{t}\binom{t}{\frac{1}{2}(t-z)}p^{\frac{1}{2}(t-z)}(1-p)^{\frac{1}{2}(t+z)}.$$

Therefore, it follows from Theorem 31 that

P(the number of black balls and the number of white balls in the urn

are the same for the first time on the t^{th} draw)

$$= \int_0^1 P\Big(\text{random walker starting at } b-w \text{ will reach the origin for the first time on the } t^{th} \text{ trial}\Big)\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^1 d(b-w,t)\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^1 \frac{b-w}{t}\binom{t}{\frac{1}{2}\big(t-(b-w)\big)}p^{\frac{1}{2}(t-(b-w))}(1-p)^{\frac{1}{2}(t+(b-w))} \times \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \frac{b-w}{t}\binom{t}{\frac{1}{2}\big(t-(b-w)\big)}\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \times \int_0^1 p^{\frac{1}{2}(t-(b-w))}(1-p)^{\frac{1}{2}(t+(b-w))}p^{b-1}(1-p)^{w-1}dp$$

$$= \frac{b-w}{t}\binom{t}{\frac{1}{2}\big(t-(b-w)\big)}\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \times \int_0^1 p^{\frac{1}{2}(t-(b-w))+b-1}(1-p)^{\frac{1}{2}(t+(b-w))+w-1}dp$$

$$= \frac{b-w}{t}\binom{t}{\frac{1}{2}\big(t-(b-w)\big)}\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \times \frac{\Gamma\left(\frac{1}{2}\big(t-(b-w)\big)+b\right)\Gamma\left(\frac{1}{2}\big(t+(b-w)\big)+w\right)}{\Gamma\left(\frac{1}{2}\big(t-(b-w)\big)+b+\frac{1}{2}\big(t+(b-w)\big)+w\right)}$$

$$= \frac{b-w}{t}\binom{t}{\frac{1}{2}\left(t-(b-w)\right)}\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \cdot \frac{\Gamma\left(\frac{1}{2}(t+b+w)\right)\Gamma\left(\frac{1}{2}(t+b+w)\right)}{\Gamma(t+b+w)}.$$

Example 33 Equalization of Colors at any Time

Suppose an urn initially contains b black and w white balls, with $b > w$. Balls are drawn at random and then returned along with another ball of the same color (Pólya sampling). Then the probability that the number of black balls and the number of white in the urn will ever be the same equals

$$2 \cdot \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \int_0^{1/2} p^{b-1}(1-p)^{w-1}dp.$$

This problem was comes from Wallstrom [47] who solves it by a different method.

Proof.

Consider a sequence of independent Bernoulli trials (success probability p, failure probability $1-p$). A random walker starts at position $z \geq 0$ on a number line and at each step either goes up one (success) or down one (failure) according to the outcome of that Bernoulli trial. Feller [21] has shown that the probability $\tilde{q}(z)$ that this random walker will at any point in the future reach the origin on this number line is given by

$$\tilde{q}(z) = \begin{cases} 1 & \text{if } p \leq \frac{1}{2} \\ \left(\frac{1-p}{p}\right)^z & \text{if } p > \frac{1}{2}. \end{cases}$$

Therefore, it follows from Theorem 31 that

$$P\Big(\text{at any point in the future the number of black ball } s$$
$$\text{and the number of white balls in the urn are the same}\Big)$$

$$= \lim_{t\to\infty} P\Big(\text{at some point in the first } t \text{ trials the number}$$
$$\text{of black balls and the number of white balls in the urn are the same}\Big)$$

$$= \lim_{t\to\infty}\int_0^1 P\Big(\text{random walker starting at } b-w \text{ will reach the origin}$$
$$\text{at some point in the first } t \text{ trials}\Big)\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^1 \lim_{t\to\infty} P\Big(\text{random walker starting at } b-w \text{ will reach the origin}$$
$$\text{at some point in the first } t \text{ trials}\Big)\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

(pulling the limit inside the integral is justified by the bounded convergence theorem because every probability is uniformly bounded)

$$= \int_0^1 P\Big(\text{random walker starting at } b-w \text{ will ever reach the origin}\Big)$$
$$\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^1 \tilde{q}(b-w)\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^{1/2} 1 \cdot \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} p^{b-1}(1-p)^{w-1} dp$$

$$+ \int_{1/2}^{1} \left(\frac{1-p}{p}\right)^{b-w} \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} p^{b-1}(1-p)^{w-1} dp$$

$$= \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \int_0^{1/2} p^{b-1}(1-p)^{w-1} dp + \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \int_{1/2}^{1} p^{w-1}(1-p)^{b-1} dp.$$

But by making the change of variable $p^* = 1 - p$ we see that

$$\int_{1/2}^{1} p^{w-1}(1-p)^{b-1} dp = \int_0^{1/2} (p^*)^{b-1}(1-p^*)^{w-1} dp^*.$$

So

$$P\Big(\text{at any point the number of black balls}$$
$$\text{and the number of white balls in the urn are the same}\Big)$$

$$= 2 \cdot \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} \int_0^{1/2} p^{b-1}(1-p)^{w-1} dp.$$

Example 34 Patterns in a Pólya Process

Suppose an urn initially contains b black and w white balls. Balls are drawn at random and then returned along with another ball of the same color (Pólya sampling). The probability that we will get a run of r consecutive black balls drawn before we get a run of s consecutive white balls drawn equals

$$\frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}\int_0^1 \left(\frac{p^{r-1}(1-(1-p)^s)}{p^{r-1}+(1-p)^{s-1}-p^{r-1}(1-p)^{s-1}}\right)p^{b-1}(1-p)^{w-1}dp.$$

Proof.

$$P(\text{a run of } r \text{ consecutive black balls drawn occurs before } s \text{ consecutive white balls})$$

$$= \lim_{t\to\infty} P(\text{within the next } t \text{ draws a run of } r \text{ consecutive black balls drawn occurs before } s \text{ consecutive white balls})$$

$$= \lim_{t\to\infty}\int_0^1 P(\text{within the next } t \text{ draws a run of } r \text{ successes will occur before a run of } s \text{ failures}) \cdot \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} p^{b-1}(1-p)^{w-1}dp$$

$$= \int_0^1 \lim_{t\to\infty} P(\text{within the next } t \text{ draws a run of } r \text{ successes will occur before a run of } s \text{ failures}) \cdot \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)} p^{b-1}(1-p)^{w-1}dp$$

$$= \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}\int_0^1 P(\text{a run of } r \text{ successes will occur before a run of } s \text{ failures}) \cdot p^{b-1}(1-p)^{w-1}dp$$

$$= \frac{\Gamma(b+w)}{\Gamma(b)\Gamma(w)}\int_0^1 \left(\frac{p^{r-1}(1-(1-p)^s)}{p^{r-1}+(1-p)^{s-1}-p^{r-1}(1-p)^{s-1}}\right)p^{b-1}(1-p)^{w-1}dp.$$

The result that

$$P(\text{a run of } r \text{ successes will occur before a run of } s \text{ failures})$$

$$= \frac{p^{r-1}(1-(1-p)^s)}{p^{r-1}+(1-p)^{s-1}-p^{r-1}(1-p)^{s-1}}$$

for a sequence of independent Bernoulli trials with success probability p and failure probability $1-p$ is from Bizley [8], page 150.

8 Proofs

8.1 Proof of Theorem 1

Proof of Theorem. It is a standard result that $Z_1 + \cdots + Z_n \sim \text{Negative Binomial}(M, 1-p)$ for the Z_j as defined above (see Appendix 10). Let $\mathbb{S}^n$ be the product space $\{0,1,\dots\} \times \cdots \times \{0,1,\dots\}$ and let $\mathbb{S}_t^n$ be the set of all vectors $(s_1, \dots, s_n)$ in $\mathbb{S}^n$ such that $s_1 + \dots + s_n = t$. Now let $(c_1, \dots, c_n)$ be a vector in $\mathbb{S}_t^n$. We can easily show formally what we deduced in the discussion above. Namely, the distribution of $(C_1, \dots, C_n)$ is equivalent to the distribution of $(Z_1, \dots, Z_n)$ conditioned on $Z_1 + \cdots + Z_n = t$.

$$P(Z_1 = c_1, \dots, Z_n = c_n | Z_1 + \cdots + Z_n = t)$$

$$= \frac{P(Z_1 = c_1, \dots, Z_n = c_n, Z_1 + \cdots + Z_n = t)}{P(Z_1 + \cdots + Z_n = t)}$$

$$= \frac{P(Z_1 = c_1, \dots, Z_n = c_n)}{P(Z_1 + \cdots + Z_n = t)}$$

$$= \frac{\binom{c_1+m_1-1}{m_1-1} \cdots \binom{c_n+m_n-1}{m_n-1} (1-p)^{m_1+\cdots+m_n} p^{c_1+\cdots+c_n}}{(1-p)^M p^t \binom{t+M-1}{M-1}}$$

$$= \frac{\binom{c_1 + m_1 - 1}{m_1 - 1} \cdots \binom{c_n + m_n - 1}{m_n - 1} (1-p)^M \, p^t}{(1-p)^M \, p^t \binom{t + M - 1}{M - 1}}$$

$$= \frac{\binom{c_1 + m_1 - 1}{m_1 - 1} \cdots \binom{c_n + m_n - 1}{m_n - 1}}{\binom{t + M - 1}{M - 1}} = P(C_1 = c_1, \dots, C_n = c_n).$$

Now we can use the general rule of iterated expectations, $\mathrm{E}(X) = \mathrm{E}\big(\mathrm{E}(X|Y)\big)$, to find $\mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)$ in terms of $\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$.

$$\mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) = \mathrm{E}\left(\mathrm{E}\left(\Psi(Z_1, \dots, Z_n) \middle| \sum_{i=1}^{n} Z_i \right)\right)$$

$$= \sum_{r=0}^{\infty} \mathrm{E}\left(\Psi(Z_1, \dots, Z_n) \middle| \sum_{i=1}^{n} Z_i = r \right) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=0}^{\infty} \left(\sum \cdots \sum_{(c_1, \dots, c_n) \in \mathbb{S}_r^n} \Psi(c_1, \dots, c_n) P\left(Z_1 = c_1, \dots, Z_n = c_n \middle| \sum_{i=1}^{n} Z_i = r \right)\right) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=0}^{\infty} \left(\sum \cdots \sum_{(c_1, \dots, c_n) \in \mathbb{S}_r^n} \Psi(c_1, \dots, c_n) P(C_1 = c_1, \dots, C_n = c_n) \right) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=0}^{\infty} E\big(\Psi(C_1, \dots, C_n)\big) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=0}^{\infty} E\big(\Psi(C_1, \dots, C_n)\big)(1-p)^M p^r \binom{r + M - 1}{M - 1}.$$

But this is backwards from what we what, namely, $\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$ in terms of $\mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)$. We need to invert the above relationship between these two expectations. Hence the name *expectation inversion* for this general method.

First, we can move constants to the other side.

$$\left(\frac{1}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) = \sum_{r=0}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r+M-1}{M-1} p^r.$$

From here it follows that

$$\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0}$$

$$= \frac{d^t}{dp^t}\left(\sum_{r=0}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r+M-1}{M-1} p^r\right)\Bigg|_{p=0}$$

$$= \sum_{r=0}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r+M-1}{M-1} \left(\frac{d^t}{dp^t} p^r\Bigg|_{p=0}\right)$$

$$= \sum_{r=0}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r+M-1}{M-1} \big(t!\ \mathbb{I}(r=t)\big)$$

$$= \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{t+M-1}{M-1} t!.$$

Thus,

$$\mathrm{E}\big(\Psi(C_1,\dots,C_n)\big) = \frac{1}{\binom{t+M-1}{M-1}t!}\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1,\dots,Z_n)\big)\right)\Bigg|_{p=0}.$$

8.2 Proof of Example 2

Proof.

Let C_j equal the number of balls that go into the j^{th} urn. The first result follows directly from Theorem 1 and the principle of inclusion-exclusion with

$$\Psi(C_1,\dots,C_n,C_{n+1}) = \mathbb{I}(\text{exactly } d \text{ of } (C_1,\dots,C_n) \text{ equal } i),$$

$Z_j \sim$ negative binomial$(m_1, 1-p)$, $j = 1,\dots,n$ and $Z_{n+1} \sim$ negative binomial$(m_2, 1-p)$.

Throughout this paper we use the notation

$$\mathbb{I}(\mathcal{A}) = \begin{cases} 1 & \text{event } \mathcal{A} \text{ occurs} \\ 0 & \text{else.} \end{cases}$$

In this case,

$$\mathrm{E}\big(\Psi(Z_1,\dots,Z_{n+1})\big)$$
$$= 1 \cdot P(\text{exactly } d \text{ of } (Z_1,\dots,Z_n) \text{ equal } i)$$
$$+ 0 \cdot \big(1 - P(\text{exactly } d \text{ of } (Z_1,\dots,Z_n) \text{ equal } i)\big)$$

$$= P(\text{exactly } d \text{ of } (Z_1,\dots,Z_n) \text{ equal } i)$$

$$=\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} P\left(Z_1 = i, \dots, Z_j = i\right)$$

(By the generalized principle of inclusion-exclusion, Theorem 37, in Appendix B.)

$$=\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \left(P(Z_1 = i)\right)^j$$

$$=\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \left(\binom{m_1 + i - 1}{i} (1-p)^{m_1} p^i\right)^j$$

$$=\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \binom{m_1 + i - 1}{i}^j (1-p)^{m_1 j} p^{ij}.$$

Thus,

$$\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^M \mathrm{E}\left(\Psi(Z_1, \dots, Z_{n+1})\right)\right)\Bigg|_{p=0}$$

$$=\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1 n + m_2} \sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \binom{m_1 + i - 1}{i}^j (1-p)^{m_1 j} p^{ij}\right)\Bigg|_{p=0}$$

$$=\frac{d^t}{dp^t}\left(\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \binom{m_1 + i - 1}{i}^j \left(\frac{1}{1-p}\right)^{m_1(n-j)+m_2} p^{ij}\right)\Bigg|_{p=0}$$

$$=\frac{d^t}{dp^t}\left(\sum_{j=d}^{n} (-1)^{j-d} \binom{j}{d}\binom{n}{j} \binom{m_1 + i - 1}{i}^j\right.$$

$$\times\left(\sum_{k=0}^{\infty}\binom{m_1(n-j)+m_2+k-1}{k}p^k\right)p^{ij}\Bigg)\Bigg|_{p=0}$$

$$=\frac{d^t}{dp^t}\left(\sum_{j=d}^{n}\sum_{k=0}^{\infty}(-1)^{j-d}\binom{j}{d}\binom{n}{j}\binom{m_1+i-1}{i}^j\right.$$

$$\left.\times\binom{m_1(n-j)+m_2+k-1}{k}p^{ij+k}\right)\Bigg|_{p=0}$$

$$=\sum_{j=d}^{n}\sum_{k=0}^{\infty}(-1)^{j-d}\binom{j}{d}\binom{n}{j}\binom{m_1+i-1}{i}^j$$

$$\times\binom{m_1(n-j)+m_2+k-1}{k}\left(\frac{d^t}{dp^t}p^{ij+k}\Big|_{p=0}\right)$$

$$=\sum_{j=d}^{n}\sum_{k=0}^{\infty}(-1)^{j-d}\binom{j}{d}\binom{n}{j}\binom{m_1+i-1}{i}^j$$

$$\times\binom{m_1(n-j)+m_2+k-1}{k}\big(t!\ \mathbb{I}(t=ij+k)\big)$$

$$=\sum_{j=d}^{n}\sum_{k=0}^{\infty}(-1)^{j-d}\binom{j}{d}\binom{n}{j}\binom{m_1+i-1}{i}^j$$

$$\times\binom{m_1(n-j)+m_2+k-1}{k}\big(t!\ \mathbb{I}(k=t-ij)\big)$$

$$=\sum_{j=d}^{n}(-1)^{j-d}\binom{j}{d}\binom{n}{j}\binom{m_1+i-1}{i}^j\binom{m_1(n-j)+m_2+t-ij-1}{t-ij}t!\,.$$

Finally, we note that

$$\binom{j}{d}\binom{n}{j} = \binom{n}{d}\binom{n-d}{n-j}$$

which brings us to

$$\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_{n+1})\big)\right)\Bigg|_{p=0}$$

$$= \binom{n}{d}\sum_{j=d}^{n} (-1)^{j-d}\binom{n-d}{n-j}\binom{m_1+i-1}{i}^j\binom{m_1(n-j)+m_2+t-ij-1}{t-ij}t!\,.$$

Thus,

$$\mathrm{E}\big(\Psi(C_1, \dots, C_{n+1})\big)$$

$$= \frac{1}{\binom{m_1 n+m_2+t-1}{t}t!}\frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1 n+m_2} \mathrm{E}\big(\Psi(Z_1, \dots, Z_{n+1})\big)\right)\Bigg|_{p=0}$$

$$= \frac{\binom{n}{d}}{\binom{m_1 n+m_2+t-1}{t}}\sum_{j=d}^{n} (-1)^{j-d}\binom{n-d}{n-j}\binom{m_1+i-1}{i}^j$$
$$\times \binom{m_1(n-j)+m_2+t-ij-1}{t-ij}.$$

The second result also follows from Theorem 1 with

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \left(\mathbb{I}(C_1 = i) + \cdots + \mathbb{I}(C_n = i)\right)_{[r]} = (D_i)_{[r]}$$

on applying the formula for falling factorial moments of sums of indicator functions (see Theorem 36 of Appendix 9).

Proof. On taking

$$\Psi(C_1, \dots, C_n, C_{n+1}) = \left(\mathbb{I}(C_1 = i) + \cdots + \mathbb{I}(C_n = i)\right)_{[r]}$$

in Theorem 1 we have

$$\mathrm{E}\left(\Psi(C_1, \dots, C_n, C_{n+1})\right)$$

$$= \frac{1}{\binom{t + m_1 n + m_2 - 1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1 n + m_2} \times \mathrm{E}\left(\left(\mathbb{I}(Z_1 = i) + \cdots + \mathbb{I}(Z_n = i)\right)_{[r]}\right)\right)\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t + m_1 n + m_2 - 1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1 n + m_2} r! \times \sum_{\{j_1, \dots, j_r\}} \cdots \sum P\left(Z_{j_1} = i, \dots, Z_{j_r} = i\right)\right)\Bigg|_{p=0}$$

(where the above sum is taken over all r subsets $\{j_1, \dots, j_r\}$ of $\{1, \dots, n\}$)

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1n+m_2} r! \binom{n}{r} (P(Z_1 = i))^r\right)\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \frac{d^t}{dp^t}\Bigg(\left(\frac{1}{1-p}\right)^{m_1n+m_2} r! \binom{n}{r}$$

$$\times \binom{i+m_1-1}{m_1-1}^r (1-p)^{m_1 r} p^{ir}\Bigg)\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^{m_1(n-r)+m_2} r! \binom{n}{r} \binom{i+m_1-1}{m_1-1}^r p^{ir}\right)\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \frac{d^t}{dp^t}\Bigg(\sum_{j=0}^{\infty} \binom{m_1(n-r)+m_2+j-1}{j} r!$$

$$\times \binom{n}{r} \binom{i+m_1-1}{m_1-1}^r p^{ir+j}\Bigg)\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \sum_{j=0}^{\infty} \binom{m_1(n-r)+m_2+j-1}{j} r!$$

$$\times \binom{n}{r} \binom{i+m_1-1}{m_1-1}^r \frac{d^t}{dp^t}(p^{ir+j})\Bigg|_{p=0}$$

$$= \frac{1}{\binom{t+m_1n+m_2-1}{t} t!} \sum_{j=0}^{\infty} \binom{m_1(n-r)+m_2+j-1}{j} r!$$

$$\times \binom{n}{r} \binom{i+m_1-1}{m_1-1}^r t!\, \mathbb{I}(t = ir+j)$$

$$= \frac{1}{\binom{t + m_1 n + m_2 - 1}{t}} \binom{m_1(n-r) + m_2 + t - ir - 1}{t - ir} r! \binom{n}{r} \binom{i + m_1 - 1}{m_1 - 1}^r$$

$$= \frac{\binom{i + m_1 - 1}{m_1 - 1}^r \binom{m_1(n-r) + m_2 + t - ir - 1}{t - ir}}{\binom{m_1 n + m_2 + t - 1}{t}} n_{[r]}.$$

8.3 Proof of Theorem 4

Proof. We need the following lemma to begin the proof.

Lemma 35 If $Z_1, \dots, Z_n$ are independent and if $Z_j \sim \alpha_j$-shifted negative binomial(m_j, p), then $S = Z_1 + \cdots + Z_n \sim \alpha^*$-shifted negative binomial(M, p), where $\alpha^* = \alpha_1 + \cdots + \alpha_n$ and $M = m_1 + \cdots + m_n$.

Proof of Lemma Let $Y_1, \dots, Y_n$ be independent random variables with $Y_j \sim$ negative binomial(m_j, p) for $j = 1,2, \dots, n$. Then by definition

$$Y_j + \alpha_j \sim \alpha_j\text{-shifted negative binomial}(m_j, 1-p).$$

$$P(S = s) = P\big((Y_1 + \alpha_1) + \cdots + (Y_n + \alpha_n) = s\big)$$
$$= P(Y_1 + \cdots + Y_n = s - \alpha^*).$$

However, we know that $Y_1 + \cdots + Y_n \sim$ negative binomial$(M, 1-p)$. Therefore

$$P(S = s) = \binom{M + (s - \alpha^*) - 1}{M - 1} (1-p)^M p^{s-\alpha^*}. \tag{25}$$

But we recognize from (6) that this is just the probability distribution for the α^* - shifted negative binomial$(M, 1-p)$ random variable.

Thus it follows from Equations (??) and (25) that for independent random variables $Z_1, \ldots, Z_n$ with $Z_j \sim \alpha_j$-shifted negative binomial$\left(m_j, p\right)$ and for all $(c_1, \ldots, c_n)$ such that $c_1 + \cdots + c_n = t$

$$P(Z_1 = c_1, \ldots, Z_n = c_n | Z_1 + \cdots + Z_n = t)$$

$$= \frac{P(Z_1 = c_1, \ldots, Z_n = c_n)}{P(Z_1 + \cdots + Z_n = t)}$$

$$= \frac{\binom{c_1 - 1}{m_1 - 1} \cdots \binom{c_n - 1}{m_n - 1} (1-p)^M p^{(c_1 + \cdots + c_n) - M}}{\binom{M + (t - M) - 1}{M - 1} (1-p)^M p^{t-M}}$$

$$= \frac{\binom{c_1 - 1}{m_1 - 1} \cdots \binom{c_n - 1}{m_n - 1}}{\binom{t - 1}{M - 1}} = P(C_1 = c_1, \ldots, C_n = c_n).$$

Now we are ready to find $\mathrm{E}\left(\Psi(Z_1, \ldots, Z_n)\right)$ in terms of $\mathrm{E}\left(\Psi(C_1, \ldots, C_n)\right)$. By the general rule of iterated expectations, $\mathrm{E}(X) = \mathrm{E}\left(\mathrm{E}(X|Y)\right)$,

$$\mathrm{E}\left(\Psi(Z_1, \ldots, Z_n)\right) = \mathrm{E}\left(\mathrm{E}\left(\Psi(Z_1, \ldots, Z_n) \middle| \sum_{i=1}^{n} Z_i\right)\right)$$

$$= \sum_{r=M}^{\infty} \mathrm{E}\left(\Psi(Z_1, \ldots, Z_n) \middle| \sum_{i=1}^{n} Z_i = r\right) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=M}^{\infty} \left(\sum \cdots \sum_{(c_1, \ldots, c_n)} \Psi(c_1, \ldots, c_n) P\left(Z_1 = c_1, \ldots, Z_n = c_n \middle| \sum_{i=1}^{n} Z_i = r\right)\right) P\left(\sum_{i=1}^{n} Z_i = r\right)$$

$$= \sum_{r=M}^{\infty} \left(\sum_{(c_1,\dots,c_n)} \cdots \sum \Psi(c_1, \dots, c_n) P(C_1 = c_1, \dots, C_n = c_n) \right) P\left(\sum_{i=1}^{n} Z_i = r \right)$$

$$= \sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) P\left(\sum_{i=1}^{n} Z_i = r \right)$$

$$= \sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r-1}{M-1} (1-p)^M p^{r-M}.$$

Now we need to invert these expectations to make $\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$ a function of $\mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)$. Moving constants to the other side we have

$$\left(\frac{p}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) = \sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r-1}{M-1} \, p^r.$$

It follows that

$$\frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^M \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) \right)\Bigg|_{p=0}$$

$$= \frac{d^t}{dp^t}\left(\sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r-1}{M-1} \, p^r \right)\Bigg|_{p=0}$$

$$= \sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r-1}{M-1} \left(\frac{d^t}{dp^t} \, p^r \Big|_{p=0} \right)$$

$$= \sum_{r=M}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{r-1}{M-1} \big(t!\, \mathbb{I}(r=t)\big)$$

$$= \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \binom{t-1}{M-1} t!.$$

Thus,

$$\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) = \frac{1}{\binom{t-1}{M-1} t!} \frac{d^t}{dp^t} \left(\left(\frac{p}{1-p}\right)^M E\big(\Psi(Z_1, \dots, Z_n)\big) \right) \Bigg|_{p=0}.$$

8.4 Proof of Example 5

Proof. Let C_j equal the number of balls that get distributed into the j^{th} group where $C_1 + C_2 + \cdots + C_n = t$. Then

$$P(C_1 \geq 1, C_2 \geq 1, \dots, C_n \geq 1) = \frac{\mathfrak{Z}(t, m, n)}{\binom{mn + t - 1}{t}}.$$

By Theorem 1, for $Z_1, \dots, Z_n$ are independent, $Z_j \sim$ negative binomial$(m,\ 1-p)$ and $M = mn$, we have

$$P(C_1 \geq 1, C_2 \geq 1, \dots, C_n \geq 1)$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^{mn} P(Z_1 \geq 1, Z_2 \geq 1, \dots, Z_n \geq 1) \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^{mn} \left(P(Z_1 \geq 1) \right)^n \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^{mn} (1-(1-p)^m)^n \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} \frac{1}{(1-p)^{mj}} \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{dp^t} \left(\sum_{j=0}^{n} \sum_{i=0}^{\infty} (-1)^{n-j} \binom{n}{j} \binom{mj+i-1}{i} p^i \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \sum_{j=0}^{n} \sum_{i=0}^{\infty} (-1)^{n-j} \binom{n}{j} \binom{mj+i-1}{i} \left(\frac{d^t}{dp^t} (p^i) \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \sum_{j=0}^{n} \sum_{i=0}^{\infty} (-1)^{n-j} \binom{n}{j} \binom{mj+i-1}{i} t! \, \mathbb{I}(t=i)$$

$$= \frac{1}{\binom{mn+t-1}{t}} \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} \binom{mj+t-1}{t}$$

$$= \frac{1}{\binom{mn+t-1}{t}} \sum_{j=1}^{n} (-1)^{n-j} \binom{n}{j} \binom{mj+t-1}{t}.$$

Therefore,

$$\mathfrak{Z}(t,m,n) = \sum_{j=1}^{n} (-1)^{n-j} \binom{n}{j} \binom{mj+t-1}{t}.$$

8.5 Proof of Theorem 6

Proof. Let $S = Z_1 + \cdots + Z_n$ and let $g(p,m) = \sum_{z=1}^{\infty} \binom{z+m-1}{m-1} (1-p)^m p^z$. In this case

$$P(S=s) = \sum_{\substack{(z_1,\ldots,z_n)\ni \\ z_1+\cdots+z_n=s \\ 1\le z_j\le\infty\, j=1,\ldots,n}} \cdots \sum P(Z_1 = z_1, \ldots, Z_n = z_n)$$

$$= \left(\frac{1}{g(p,m)}\right)^n \sum_{\substack{(z_1,\ldots,z_n)\ni \\ z_1+\cdots+z_n=s \\ 1\le z_j\le\infty\, j=1,\ldots,n}} \cdots \sum \binom{z_1+m-1}{m-1} \cdots \binom{z_n+m-1}{m-1} (1-p)^{mn} p^s$$

$$= \left(\frac{1}{g(p,m)}\right)^n (1-p)^{mn} p^s \ \mathfrak{Z}(s,m,n).$$

Also,

$$P(Z_1 = z_1, \ldots, Z_n = z_n \mid S = s)$$
$$= \frac{P(Z_1 = z_1, \ldots, Z_n = z_n)}{P(S=s)} \mathbb{I}(z_1 + \cdots + z_n = s)$$

$$= \frac{\left(\frac{1}{g(p,m)}\right)^n \binom{z_1 + m - 1}{m-1} \cdots \binom{z_n + m - 1}{m-1} (1-p)^{mn} p^s}{\left(\frac{1}{g(p,m)}\right)^n (1-p)^{mn} p^s \ \mathfrak{Z}(s,m,n)}$$

$$= \frac{\binom{z_1 + m - 1}{m-1} \cdots \binom{z_n + m - 1}{m-1}}{\mathfrak{Z}(s,m,n)}$$

$$= P(C_1 = c_1, \dots, C_n = c_n).$$

Therefore,

$$\mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) = \mathrm{E}\big(\mathrm{E}(\Psi(Z_1, \dots, Z_n)|S)\big)$$

$$= \sum_{s=n}^{\infty} \mathrm{E}(\Psi(Z_1, \dots, Z_n)|S = s) P(S = s)$$

$$= \sum_{s=n}^{\infty} \left(\sum \cdots \sum_{(c_1, \dots, c_n)} \Psi(c_1, \dots, c_n) P(Z_1 = c_1, \dots, Z_n = c_n | S = s) \right) P(S = s)$$

$$= \sum_{s=n}^{\infty} \left(\sum \cdots \sum_{(c_1, \dots, c_n)} \Psi(c_1, \dots, c_n) P(C_1 = c_1, \dots, C_n = c_n) \right) P(S = s)$$

$$= \sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) P(S = s)$$

$$= \sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) \left(\frac{1}{g(p,m)}\right)^n (1-p)^{mn} p^s \ \mathfrak{Z}(s,m,n).$$

Now we need to invert these expectations. Moving constants to the other side we have

$$\left(\frac{g(p,m)}{(1-p)^m}\right)^n \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big) = \sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)\mathfrak{Z}(s,m,n)p^s.$$

It follows that

$$\frac{d^t}{dp^t}\left(\left(\frac{g(p,m)}{(1-p)^m}\right)^n \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0}$$

$$= \frac{d^t}{dp^t}\left(\sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)\mathfrak{Z}(s,m,n)p^s\right)\Bigg|_{p=0}$$

$$= \sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)\mathfrak{Z}(s,m,n)\left(\frac{d^t}{dp^t}\, p^s\Big|_{p=0}\right)$$

$$= \sum_{s=n}^{\infty} \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)\mathfrak{Z}(s,m,n)\big(t!\ \mathbb{I}(s=t)\big)$$

$$= \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)\mathfrak{Z}(t,m,n)t!.$$

Thus,

$$\mathrm{E}\big(\Psi(C_1, \dots, C_n)\big) = \frac{1}{\mathfrak{Z}(t,m,n)t!}\frac{d^t}{dp^t}\left(\left(\frac{g(p,m)}{(1-p)^m}\right)^n \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0}.$$

Finally, we note that

$$\left(g(p,m)\right)^r = \left(\sum_{z=1}^{\infty} \binom{z+m-1}{m-1} (1-p)^m p^z\right)^r$$

$$= \sum_{z_1=1}^{\infty} \cdots \sum_{z_r=1}^{\infty} \binom{z_1+m-1}{m-1} \cdots \binom{z_r+m-1}{m-1} (1-p)^{mr} p^{z_1+\cdots+z_r}$$

$$= \sum_{t=r}^{\infty} (1-p)^{mr} p^t \left(\sum_{\substack{(z_1,\dots,z_r)\ni \\ z_1+\cdots+z_r=t \\ 1\le z_j \le \infty\, j=1,\dots,r}} \cdots \sum \binom{z_1+m-1}{m-1} \cdots \binom{z_r+m-1}{m-1} \right)$$

$$= \sum_{t=r}^{\infty} \mathfrak{Z}(t,m,r)(1-p)^{mr}\, p^t.$$

8.6 Proof of Theorem 7

8.6.1 Step 7a.

Proof. Let the joint distribution of

$(C_1, C_2, \dots, C_n)$ be given by (??) (i.e. unrestricted allocation),

$(C_1^*, C_2^*, \dots, C_n^*)$ be given by (??) (i.e. no empty urns), and

$(\hat{C}_1, \hat{C}_2, \dots, \hat{C}_n)$ be given by (??) (i.e. no empty groups).

Now suppose that we distribute t identical balls into a row of M distinct urns in such a way that all possible (unrestricted) allocations are equally likely to occur. Assume that each urn belongs to one of n distinct groups where the j^{th} group contains m_j urns with $m_1 + \cdots + m_n = M$.

Let X_{ij} be the number of balls allocated to the j^{th} urn in the i^{th} group, $i = 1,2, \dots, n$, $j = 1,2, \dots, m_j$.

We note that by definition, for every vector $\mathbf{c}$,

$$P(\mathbf{C}^* \leq \mathbf{c}) = P\big(\mathbf{C} \leq \mathbf{c} | X_{ij} > 0 \text{ for all } i \text{ and } j\big)$$

and

$$P\big(\hat{\mathbf{C}} \leq \mathbf{c}\big) = P(\mathbf{C} \leq \mathbf{c} | C_i > 0 \text{ for all } i).$$

Now let $\mathbf{C}/\mathbf{t}$ represent the random vector $\left(\frac{C_1}{t}, \dots, \frac{C_n}{t}\right)$. It follows from the law of total probability for each $t > 0$ and every vector $\mathbf{v}$ that

$$P(\mathbf{C}/\mathbf{t} \leq \mathbf{v})$$

$$= P\big(\mathbf{C}/\mathbf{t} \leq \mathbf{v} \,| X_{ij} > 0 \text{ for all } i \text{ and } j\big) P\big(X_{ij} > 0 \text{ for all } i \text{ and } j\big)$$

$$+ P\big(\mathbf{C}/\mathbf{t} \leq \mathbf{v} \,| X_{ij} = 0 \text{ for some } ij \text{ pair}\big) P\big(X_{ij} = 0 \text{ for some } ij \text{ pair}\big).$$

Thus,

$$\lim_{t\to\infty}\big(P(\mathbf{C}/\mathbf{t} \leq \mathbf{v})\big)$$
$$= \lim_{t\to\infty}\Big(P\big(\mathbf{C}/\mathbf{t} \leq \mathbf{v} \,| X_{ij} > 0 \text{ for all } i \text{ and } j\big)\Big) \cdot \lim_{t\to\infty}\Big(P\big(X_{ij} > 0 \text{ for all } i \text{ and } j\big)\Big)$$
$$+ \lim_{t\to\infty}\Big(P\big(\mathbf{C}/\mathbf{t} \leq \mathbf{v} \,| X_{ij} = 0 \text{ for some } ij \text{ pair}\big)\Big) \cdot \lim_{t\to\infty}\Big(P\big(X_{ij} = 0 \text{ for some } ij \text{ pair}\big)\Big).$$

We can also see that

$$\lim_{t\to\infty}\Big(P\big(X_{ij} = 0 \text{ for some } ij \text{ pair}\big)\Big) \leq \lim_{t\to\infty}\big(M \cdot P(X_{11} = 0)\big)$$

$$= \lim_{t\to\infty}\left(M\cdot\frac{\binom{(M-1)+t-1}{t}}{\binom{M+t-1}{t}}\right) = \lim_{t\to\infty}\left(M\cdot\frac{M-1}{M+t-1}\right) = 0.$$

Hence,

$$\lim_{t\to\infty}\big(P(\mathbf{C}/\mathbf{t}\le\mathbf{v})\big)$$

$$= \lim_{t\to\infty}\Big(P\big(\mathbf{C}/\mathbf{t}\le\mathbf{v}|\ X_{ij}>0 \text{ for all } i \text{ and } j\big)\Big)\cdot 1$$

$$+ \lim_{t\to\infty}\Big(P\big(\mathbf{C}/\mathbf{t}\le\mathbf{v}|\ X_{ij}=0 \text{ for some } ij \text{ pair}\big)\Big)\cdot 0$$

$$= \lim_{t\to\infty}\Big(P\big(\mathbf{C}/\mathbf{t}\le\mathbf{v}|\ X_{ij}>0 \text{ for all } i \text{ and } j\big)\Big) = \lim_{t\to\infty}\big(P(\mathbf{C}^*/\mathbf{t}\le\mathbf{v})\big).$$

That is,

$$\lim_{t\to\infty}P\left(\frac{C_1}{t}\le v_1,\frac{C_2}{t}\le v_2,\dots,\frac{C_n}{t}\le v_n\right) = \lim_{t\to\infty}P\left(\frac{C_1^*}{t}\le v_1,\frac{C_2^*}{t}\le v_2,\dots,\frac{C_n^*}{t}\le v_n\right).$$

We can make the same kind of argument to show that

$$\lim_{t\to\infty}P\left(\frac{C_1}{t}\le v_1,\frac{C_2}{t}\le v_2,\dots,\frac{C_n}{t}\le v_n\right) = \lim_{t\to\infty}P\left(\frac{\hat{C}_1}{t}\le v_1,\frac{\hat{C}_2}{t}\le v_2,\dots,\frac{\hat{C}_n}{t}\le v_n\right).$$

8.6.2 Step 7b.

Proof. Suppose we allocate t identical balls into $M = m_1 + m_2 + \cdots + m_n$ distinct (e.g. numbered) urns where each urn belongs to one of n distinct (e.g. different colored) groups where the j^{th} group contains m_j urns. In our (unrestricted) grouped Bose-Einstein allocation scheme all possible allocations of these t balls into the M distinct urns are equally likely events.

Let X_{ij} be the number of balls allocated to the j^{th} urn in the i^{th} group, $i = 1,2, \dots, n$, $j = 1,2, \dots, m_j$. Let $C_j = X_{i1} + \cdots + X_{im_i}$ represent the total number of balls allocated to the i^{th} group.

Let $(V_1, V_2, \dots, V_n)$ be a random vector following the Dirichlet$(m_1, m_2, \dots, m_n)$ distribution. That is, let the joint density function of $(V_1, V_2, \dots, V_n)$ be given by

$$f(v_1, v_2, \dots, v_n) = \frac{\Gamma(m_1 + m_2 + \cdots + m_n)}{\Gamma(m_1) \cdots \Gamma(m_n)} v_1^{m_1 - 1} \cdots v_n^{m_n - 1}$$

for all points $(v_1, v_2, \dots, v_n)$ such that $v_1 + v_2 + \cdots + v_n = 1$ and $0 \le v_j \le 1$ for $j = 1,2, \dots, n$.

Blackwell and MacQueen [9] showed that in this case

$$\left(\left(\frac{X_{11}}{t}, \dots, \frac{X_{1m_1}}{t}\right), \left(\frac{X_{21}}{t}, \dots, \frac{X_{2m_2}}{t}\right), \dots, \left(\frac{X_{n1}}{t}, \dots, \frac{X_{nm_m}}{t}\right)\right)$$

converges in distribution to

$$\left((Y_{11}, \dots, Y_{1m_1}), (Y_{21}, \dots, Y_{2m_2}), \dots, (Y_{n1}, \dots, Y_{nm_n})\right)$$
$$\sim \text{Dirichlet}((1, \dots, 1), (1, \dots, 1), \dots, (1, \dots, 1))$$

Thus, it follows the Cramér-Wold Theorem [7] that

$$\left(\left(\frac{X_{11}}{t} + \cdots + \frac{X_{1m_1}}{t}\right), \left(\frac{X_{21}}{t} + \cdots + \frac{X_{2m_2}}{t}\right), \dots, \left(\frac{X_{n1}}{t} + \cdots + \frac{X_{nm_n}}{t}\right)\right) = \left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right)$$

converges in distribution to

$$\left((Y_{11} + \cdots + Y_{1m_1}), (Y_{21} + \cdots + Y_{2m_2}), \dots, (Y_{n1} + \cdots + Y_{nm_n})\right).$$

But by the aggregation property of the Dirichlet distribution, (Wilks [53], Theorem 7.7.5) if

$$\left((Y_{11}, \dots, Y_{1m_1}), (Y_{21}, \dots, Y_{2m_2}), \dots, (Y_{n1}, \dots, Y_{nm_n})\right)$$
$$\sim \text{Dirichlet}((1, \dots, 1), (1, \dots, 1), \dots, (1, \dots, 1))$$

then

$$\left((Y_{11} + \cdots + Y_{1m_1}), (Y_{21} + \cdots + Y_{2m_2}), \dots, (Y_{n1} + \cdots + Y_{nm_n})\right)$$
$$\sim \text{Dirichlet}((1 + \cdots + 1), (1 + \cdots + 1), \dots, (1 + \cdots + 1))$$
$$= \text{Dirichlet}(m_1, m_2, \dots, m_n).$$

So, for any positive integers $m_1, m_2, \dots, m_n$,

$$\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right) \xrightarrow{d} (V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$$

as $t \to \infty$.

(26)

8.6.3 Step 7c.

Proof. Suppose the joint distribution of $(C_1, C_2, \dots, C_n)$ is given by (??), (??), or (??) and let $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$. Then by the Portmanteau Theorem [7],

$$\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right) \xrightarrow{d} (V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$$

as $t \to \infty$ implies

(27)

$$\lim_{t \to \infty} \mathrm{E}\left(\Psi\left(\frac{C_1}{t}, \frac{C_2}{t}, \dots, \frac{C_n}{t}\right)\right) = \mathrm{E}\left(\Psi(V_1, V_2, \dots, V_n)\right)$$

for any bounded and continuous function $\Psi(\ )$ on the common support, that is for all points $(v_1, v_2, \dots, v_n)$ such that $v_1 + v_2 + \cdots + v_n = 1$ and $0 \le v_j \le 1$ for $j = 1,2, \dots, n$.

8.6.4 Step 7d.

Proof. Let $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$ and let $Y_1, Y_2, \dots, Y_n$ be independent random variables with $Y_j \sim \text{Gamma}(m_j, \lambda)$ for $j = 1,2, \dots, n$. It is a standard exercise to show that

$$\sum_{i=1}^{n} Y_i \sim \text{Gamma}(M, \lambda)$$

and

$$\left(Y_1, Y_2, \dots, Y_n \middle| \sum Y_i = s\right) \stackrel{d}{=} (sV_1, sV_2, \dots, sV_n)$$

where $(V_1, V_2, \dots, V_n) \sim \text{Dirichlet}(m_1, m_2, \dots, m_n)$. Therefore,

$$\mathrm{E}\big(\Psi(Y_1, \dots, Y_n)\big) = \mathrm{E}\left(\mathrm{E}\left(\Psi(Y_1, \dots, Y_n) \middle| \sum_{i=1}^{n} Y_i\right)\right)$$

$$= \int_0^\infty \mathrm{E}\left(\Psi(Y_1, \dots, Y_n) \middle| \sum_{i=1}^{n} Y_i = s\right) \frac{1}{(M-1)!} s^{M-1} e^{-\lambda s} \lambda^M ds$$

$$= \int_0^\infty E\left(\Psi\big((sV_1, sV_2, \dots, sV_n)\big)\right) \frac{1}{(M-1)!} s^{M-1} e^{-\lambda s} \lambda^M ds.$$

Thus,

$$\frac{(M-1)!}{\lambda^M} E\big(\Psi(Y_1, Y_2, \dots, Y_n)\big) = \int_0^\infty E\left(\Psi\big((sV_1, sV_2, \dots, sV_n)\big)\right) s^{M-1} e^{-\lambda s} ds.$$

(28)

One can also establish that $\sum Y_i$ is a sufficient statistic for $Y_1, Y_2, \dots, Y_n$. Therefore, $\mathrm{E}\big(\Psi(sV_1, sV_2, \dots, sV_n)\big)$ is free of the parameter λ.

This tells us that (28) is of the form

$$G(\lambda) = \int_0^\infty g(s) e^{-\lambda s} ds$$

with

$$G(\lambda) = \frac{(M-1)!}{\lambda^M} \mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_n)\big)$$

and

$$g(s) = \mathrm{E}\Big(\Psi\big((sV_1, sV_2, \dots, sV_n)\big)\Big) s^{M-1}.$$

That is,

$$g(s) = \mathrm{E}\Big(\Psi\big((sV_1, sV_2, \dots, sV_n)\big)\Big) s^{M-1} = \mathcal{L}^{-1}\big(G(\lambda)\big).$$

Thus,

$$\mathrm{E}\big(\Psi(V_1, V_2, \dots, V_n)\big) = \frac{1}{s^{M-1}} \mathcal{L}^{-1}\big(G(\lambda)\big)\Big|_{s=1}$$

$$= \mathcal{L}^{-1}\left((M-1)!\ \frac{\mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_n)\big)}{\lambda^M}\right)\Bigg|_{s=1}$$

$$= (M-1)!\ \ \mathcal{L}^{-1}\left(\frac{\mathrm{E}\big(\Psi(Y_1, Y_2, \dots, Y_n)\big)}{\lambda^M}\right)\Bigg|_{s=1}.$$

8.7 Proof of Example 8

Proof. We begin by recalling the connection between raw moments and factorial moments via the definition of Stirling numbers of the second kind, namely, for integer $n \geq 1$,

$$t^n = \sum_{k=1}^{n} S(n,k) t_{[n]}.$$

By Theorem 1, for $Z_1 \sim$negative binomial$(m_1, 1-p)$ and $M = m_1 + \cdots + m_n$, we have

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^M \mathrm{E}((Z_1)^a) \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^M \sum_{k=1}^{a} S(a,k) \mathrm{E}\big(Z_{[k]}\big) \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^M \sum_{k=1}^{a} S(a,k) \left(\frac{p}{1-p} \right)^k m_1^{[k]} \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \sum_{k=1}^{a} S(a,k) m_1^{[k]} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^{M+k} p^k \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \sum_{k=1}^{a} S(a,k) m_1^{[k]} \frac{d^t}{dp^t} \left(\sum_{j=0}^{\infty} \binom{M+k+j-1}{j} p^{k+j} \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t} t!} \sum_{k=1}^{a} S(a,k) m_1^{[k]}$$

$$\times \left(\sum_{j=0}^{\infty} \binom{M+k+j-1}{j} p^{k+j} t!\, \mathbb{I}(t = k+j) \right) \Bigg|_{p=0}$$

$$\mathrm{E}\big((C_1)^a\big) = \frac{1}{\binom{M+t-1}{t}} \sum_{k=1}^{a} S(a,k) m_1^{[k]} \binom{M+t-1}{t-k}.$$

Now we can verify the second part of this example using Theorem 7. First note that the condition that $\Psi(x_1, x_2, \dots, x_n) = (x_1)^a$ is bounded and continuous on the simplex $x_1 + x_2 + \cdots + x_n = 1$ and $0 \le x_j \le 1$ for $j = 1,2, \dots, n$ is clearly satisfied. Hence, for $Y_1 \sim \text{Gamma}(m_1, \lambda)$, we have

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t}\right)^a \right) \right) = (M-1)!\ \mathcal{L}^{-1} \left(\frac{\mathrm{E}((Y_1)^a)}{\lambda^M} \right) \Bigg|_{s=1}$$

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t}\right)^a \right) \right) = (M-1)!\, \mathcal{L}^{-1} \left(\frac{\Gamma(m_1+a)}{\Gamma(m_1)\lambda^a \lambda^M} \right) \Bigg|_{s=1}$$

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t}\right)^a \right) \right) = (M-1)!\ \frac{\Gamma(m_1+a)}{\Gamma(m_1)}\ \mathcal{L}^{-1} \left(\frac{1}{\lambda^{a+M}} \right) \Bigg|_{s=1}$$

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t}\right)^a \right) \right) = (M-1)!\ \frac{\Gamma(m_1+a)}{\Gamma(m_1)}\ \frac{s^m}{(a+M-1)!} \Bigg|_{s=1}$$

$$\lim_{t\to\infty} \left(\mathrm{E}\left(\left(\frac{C_1}{t}\right)^a \right) \right) = \frac{\binom{m_1+a-1}{a}}{\binom{M+a-1}{a}}.$$

8.8 Proof of Theorem 9

Proof. We begin by deriving the joint probability distribution of $Z_0, Z_1, \ldots$. We have

$$P(Z_0 = z_0, Z_1 = z_1, \ldots) = e^{-\theta\left(\binom{0+m-1}{m-1}\lambda^0 + \binom{1+m-1}{m-1}\lambda^1 + \cdots\right)}$$

$$\times \frac{\theta^{z_0+z_1+\cdots}\left(\binom{0+m-1}{m-1}\lambda^0\right)^{z_0}\left(\binom{1+m-1}{m-1}\lambda^1\right)^{z_1}\cdots}{z_0!\,z_1!\cdots}$$

$$= e^{-\frac{\theta}{(1-\lambda)^m}} \frac{\theta^{z_0+z_1+\cdots}\left(\binom{0+m-1}{m-1}\lambda^0\right)^{z_0}\left(\binom{1+m-1}{m-1}\lambda^1\right)^{z_1}\cdots}{z_0!\,z_1!\cdots}.$$

Therefore,

$$P(Z_0 + Z_1 + \cdots = n, 0Z_0 + 1Z_1 + \cdots = t)$$

$$= \sum_{\substack{(z_0,z_1,\ldots)\ni \\ z_j\in\{0,1,\ldots\}\ j=0,1,\ldots \\ z_0+z_1+\cdots=n \\ 0z_0+1z_1+\cdots=t}} \cdots \sum P(Z_0 = z_0, Z_1 = z_1, \ldots)$$

$$= e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^t\, \theta^n \sum_{\substack{(z_0,z_1,\ldots)\ni \\ z_j\in\{0,1,\ldots\}\ j=0,1,\ldots \\ z_0+z_1+\cdots=n \\ 0z_0+1z_1+\cdots=t}} \cdots \sum \frac{\binom{0+m-1}{m-1}^{z_0}\binom{1+m-1}{m-1}^{z_1}\cdots}{z_0!\,z_1!\cdots}$$

$$= e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^t \theta^n \sum_{\substack{(z_0, z_1, \dots, z_t) \ni \\ z_j \in \{0,1,\dots\}\ j=0,1,\dots,t \\ z_0+z_1+\cdots+z_t=n \\ 0z_0+1z_1+\cdots+tz_t=t}} \cdots \sum \frac{\binom{0+m-1}{m-1}^{z_0} \cdots \binom{t+m-1}{m-1}^{z_t}}{z_0! \cdots z_t!}$$

[as $z_j \equiv 0$ for all $j \geq t+1$ by the restriction $0z_0 + 1z_1 + \cdots = t$]

$$= e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^t \theta^n \cdot \frac{\binom{t+mn-1}{mn-1}}{n!}.$$

This last simplification follows on noting that

$$1 = \sum_{\substack{(d_0, d_1, \dots, d_t) \ni \\ d_j \in \{0,1,\dots\}\ j=0,1,\dots,t \\ d_0+d_1+\cdots+d_t=n \\ 0d_0+1d_1+\cdots+td_t=t}} \cdots \sum P(D_0 = d_0, D_1 = d_1, \dots, D_t = d_t)$$

$$= \frac{n!}{\binom{t+mn-1}{mn-1}} \sum_{\substack{(d_0, d_1, \dots, d_t) \ni \\ d_j \in \{0,1,\dots\}\ j=0,1,\dots,t \\ d_0+d_1+\cdots+d_t=n \\ 0d_0+1d_1+\cdots+td_t=t}} \cdots \sum \frac{\binom{0+m-1}{m-1}^{d_0} \cdots \binom{t+m-1}{m-1}^{d_t}}{d_0! \cdots d_t!}.$$

Hence, for all $(z_0, z_1, \dots) \ni z_j \in \{0,1,\dots\}\ j = 0,1,\dots,\ \ z_0 + z_1 + \cdots = n$ and $0z_0 + 1z_1 + \cdots = t$ we have

$$P(Z_0 = z_0, Z_1 = z_1, \dots | Z_0 + Z_1 + \cdots = n, 0Z_0 + 1Z_1 + \cdots = t)$$

$$= \frac{e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \frac{\binom{0+m-1}{m-1}^{z_0} \cdots \binom{t+m-1}{m-1}^{z_t}}{z_0! \cdots z_t!}}{e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \frac{\binom{t+mn-1}{mn-1}}{n!}}$$

$$= \frac{\binom{0+m-1}{m-1}^{z_0} \cdots \binom{t+m-1}{m-1}^{z_t}}{\binom{t+mn-1}{mn-1}} \cdot \frac{n!}{z_0! \cdots z_t!}$$

$$= P(D_0 = z_0, D_1 = z_1, \dots, D_t = z_t).$$

With these preliminary results we are in position to finish the proof.

$$\mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big)$$

$$= \mathrm{E}\left(\mathrm{E}\left(\Psi(Z_0, Z_1, \dots, Z_t) \middle| \sum_{j=0}^{\infty} Z_j, \sum_{j=0}^{\infty} jZ_j\right)\right)$$

$$= \sum_s \sum_r \mathrm{E}\left(\Psi(Z_0, Z_1, \dots, Z_t) \middle| \sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right)$$

$$\times P\left(\sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_0, z_1, \dots) \ni \\ z_j \in \{0,1,\dots\} \;\; j=0,1,\dots \\ z_0 + z_1 + \cdots = r \\ 0z_0 + 1z_1 + \cdots = s}} \cdots \sum \Psi(z_0, z_1, \dots, z_t) \right.$$

$$\times\ P\left(Z_0 = z_0, Z_1 = z_1, \dots \middle| \sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right)$$

$$\left. \times P\left(\sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right) \right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_0, z_1, \dots, z_s) \ni \\ z_j \in \{0,1,\dots\} \ j=0,1,\dots,s \\ z_0+z_1+\dots+z_s=r \\ 0z_0+1z_1+\dots+sz_s=s}} \cdots \sum \Psi^*(\mathbf{z}) \right.$$

$$\times\ P(D_0 = z_0, D_1 = z_1, \dots, D_s = z_s)$$

$$\left. \times\ P\left(\sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right) \right)$$

where

$$\Psi^*(\mathbf{z}) = \begin{cases} \Psi(z_0, z_1, \dots, z_t) & t \le s \\ \Psi(z_0, z_1, \dots, z_s, 0, \dots, 0) & t > s \end{cases}$$

with $z_{s+1}, \dots, z_t$ all being replaced with 0's in the case $t > s$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) P\left(\sum_{j=0}^{\infty} Z_j = r, \sum_{j=0}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \cdot e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^s \, \theta^r \frac{\binom{s+mr-1}{mr-1}}{r!}.$$

So, we have now established that

$$\mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big) = \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \cdot e^{-\frac{\theta}{(1-\lambda)^m}} \lambda^s \, \theta^r \frac{\binom{s+mr-1}{mr-1}}{r!}$$

or

$$e^{\frac{\theta}{(1-\lambda)^m}} \mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big) = \sum_s \sum_r \left(\mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s+mr-1}{mr-1}}{r!} \right) \cdot \lambda^s \, \theta^r.$$

It follows that

$$\frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta}{(1-\lambda)^m}} \mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big) \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s+mr-1}{mr-1}}{r!} \cdot \lambda^s \, \theta^r \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s+mr-1}{mr-1}}{r!} \cdot \left(\frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \lambda^s \, \theta^r \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s+mr-1}{mr-1}}{r!} \cdot t! \; n! \; \mathbb{I}(s=t) \mathbb{I}(r=n)$$

$$= \mathrm{E}\big(\Psi(D_0, D_1, \dots, D_t)\big) \frac{\binom{t+mn-1}{mn-1}}{n!} \cdot t!\, n!$$

or

$$\mathrm{E}\big(\Psi(D_0, D_1, \dots, D_t)\big) = \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta}{(1-\lambda)^m}} \mathrm{E}\big(\Psi(Z_0, Z_1, \dots, Z_t)\big) \right) \Bigg|_{\substack{\lambda=0\\ \theta=0}} .$$

8.9 Proof of Example 10

Proof.

Let

$$\Psi(D_0, D_1, \dots, D_t) = (D_0)_{[r_0]} \cdots (D_t)_{[r_t]}$$

and let $Z_0, Z_1, \dots, Z_t$ be independent random variables with

$$Z_j \sim \text{Poisson}\left(\theta \binom{j+m-1}{m-1} \lambda^j \right), \qquad j = 0,1,2, \dots.$$

Then by Theorem 9 we have

$$\mathrm{E}\big(\Psi(D_0, D_1, \dots, D_t)\big)$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta}{(1-\lambda)^m}} \mathrm{E}\big((Z_0)_{[r_0]} (Z_1)_{[r_1]} \cdots (Z_t)_{[r_t]}\big) \right) \Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\mathrm{e}^{\frac{\theta}{(1-\lambda)^m}} \prod_{j=0}^{t} \mathrm{E}\left(\left(Z_j\right)_{[r_j]} \right) \right) \Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta}{(1-\lambda)^m}} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j} \theta^{r_j} \lambda^{j r_j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

(applying the formula for factorial moments of a Poisson random variable given in Appendix 10)

$$= \frac{1}{\binom{mn+t-1}{t} t!} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta}{(1-\lambda)^m}} \theta^R \lambda^S \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\sum_{i=0}^{\infty} \sum_{v=0}^{\infty} \frac{1}{i!} \binom{mi+v-1}{v} \theta^{i+R} \lambda^{v+S} \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{1}{\binom{mn+t-1}{t} t!} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j} \sum_{i=0}^{\infty} \sum_{v=0}^{\infty} \Bigg(\frac{1}{i!} \binom{mi+v-1}{v}$$

$$\times \; n! \; \mathbb{I}(n = i + R) \; t! \; \mathbb{I}(t = v + S) \Bigg)$$

$$= \frac{1}{\binom{mn+t-1}{t}} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j} \binom{m(n-R)+(t-S)-1}{t-S} \frac{n!}{(n-R)!}$$

$$= \frac{\binom{m(n-R)+(t-S)-1}{t-S} \prod_{j=0}^{t} \binom{j+m-1}{m-1}^{r_j}}{\binom{mn+t-1}{t}} n_{[R]}.$$

8.10 Proof of Theorem 11

Proof. We begin by deriving the joint probability distribution of $Z_m, Z_{m+1}, \dots$. We have

$$P(Z_m = z_m, Z_{m+1} = z_{m+1}, \dots) = e^{-\theta\left(\binom{m-1}{m-1}\lambda^m + \binom{(m+1)-1}{m-1}\lambda^{m+1} + \cdots\right)}$$

$$\times \frac{\theta^{z_m + z_{m+1} + \cdots}\left(\binom{m-1}{m-1}\lambda^m\right)^{z_m}\left(\binom{(m+1)-1}{m-1}\lambda^{m+1}\right)^{z_{m+1}}\cdots}{z_m!\, z_{m+1}!\cdots}$$

$$= e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \cdot \frac{\theta^{z_m + z_{m+1} + \cdots}\left(\binom{m-1}{m-1}\lambda^m\right)^{z_m}\left(\binom{(m+1)-1}{m-1}\lambda^{m+1}\right)^{z_{m+1}}\cdots}{z_m!\, z_{m+1}!\cdots}.$$

Therefore,

$$P(Z_m + Z_{m+1} + \cdots = n, mZ_m + (m+1)Z_{m+1} + \cdots = t)$$

$$= \sum_{\substack{(z_m, z_{m+1}, \dots) \ni \\ z_j \in \{0,1,\dots\}\ j = m, m+1, \dots \\ z_m + z_{m+1} + \cdots = n \\ mz_m + (m+1)z_{m+1} + \cdots = t}} \cdots \sum P(Z_m = z_m, Z_{m+1} = z_{m+1}, \dots)$$

$$= e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \lambda^t\, \theta^n \cdot \sum_{\substack{(z_m, z_{m+1}, \dots) \ni \\ z_j \in \{0,1,\dots\}\ j = m, m+1, \dots \\ z_m + z_{m+1} + \cdots = n \\ mz_m + (m+1)z_{m+1} + \cdots = t}} \cdots \sum \frac{\binom{m-1}{m-1}^{z_m}\binom{m}{m-1}^{z_{m+1}}\cdots}{z_m!\, z_{m+1}!\cdots}$$

$$= e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \sum_{\substack{(z_m,\dots,z_{t-m(n-1)})\ni \\ z_j\in\{0,1,\dots\}\ j=m,\dots,t-m(n-1) \\ z_m+\dots+z_{t-m(n-1)}=n \\ mz_m+\dots+(t-m(n-1))z_{t-m(n-1)}=t}} \cdots \sum \frac{\binom{m-1}{m-1}^{z_m} \cdots \binom{\left(t-m(n-1)\right)-1}{m-1}^{z_{t-m(n-1)}}}{z_m! \cdots z_{t-m(n-1)}!}$$

[as $z_j \equiv 0$ for all $j \geq t-m(n-1)$ by the restrictions $z_m + \cdots + z_{t-m(n-1)} = n$ and $mz_m + \cdots + \left(t-m(n-1)\right)z_{t-m(n-1)} = t$ taken together]

$$= e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \frac{\binom{t-1}{mn-1}}{n!}.$$

This last simplification follows on noting that

$$1 = \sum_{\substack{(d_m,\dots,d_{t-m(n-1)})\ni \\ d_j\in\{0,1,\dots\}\ j=m,\dots,t-m(n-1) \\ d_m+\dots+d_{t-m(n-1)}=n \\ md_m+\dots+(t-m(n-1))d_{t-m(n-1)}=t}} \cdots \sum P\left(D_m = d_m, \dots, D_{t-m(n-1)} = d_{t-m(n-1)}\right)$$

$$= \frac{n!}{\binom{t-1}{mn-1}} \sum_{\substack{(d_m,\dots,d_{t-m(n-1)})\ni \\ d_j\in\{0,1,\dots\}\ j=m,\dots,t-m(n-1) \\ d_m+\dots+d_{t-m(n-1)}=n \\ md_m+\dots+(t-m(n-1))d_{t-m(n-1)}=t}} \cdots \sum \frac{\binom{m-1}{m-1}^{d_m} \cdots \binom{t-m(n-1)-1}{m-1}^{d_{t-m(n-1)}}}{d_m! \cdots d_{t-m(n-1)}!}.$$

Hence, for all $(z_m, z_{m+1}, \dots) \ni z_j \in \{0,1,\dots\}\ j = m, m+1, \dots, z_m + z_{m+1} + \cdots = n$ and $mz_m + (m+1)z_{m+1} + \cdots = t$ we have

$$P\left(Z_m = z_m, Z_{m+1} = z_{m+1}, \dots \middle| Z_m + Z_{m+1} + \cdots = n, mZ_m + (m+1)Z_{m+1} + \cdots = t\right)$$

$$= \frac{e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \frac{\binom{m-1}{m-1}^{z_m} \cdots \binom{t-m(n-1)-1}{m-1}^{z_{t-m(n-1)}}}{z_m! \cdots z_{t-m(n-1)}!}}{e^{-\frac{\theta\lambda^m}{(1-\lambda)^m}} \lambda^t \, \theta^n \cdot \frac{\binom{t-1}{mn-1}}{n!}}$$

$$= \frac{\binom{m-1}{m-1}^{z_m} \cdots \binom{t-m(n-1)-1}{m-1}^{z_{t-m(n-1)}}}{\binom{t+mn-1}{mn-1}} \cdot \frac{n!}{z_m! \cdots z_{t-m(n-1)}!}$$

$$= P\left(D_m = z_m, D_{m+1} = z_{m+1}, \dots, D_{t-m(n-1)} = z_{t-m(n-1)}\right).$$

With these preliminary results we are in position to finish the proof.

$$\mathrm{E}\left(\Psi\left(Z_m, \dots, Z_{t-m(n-1)}\right)\right)$$

$$= \mathrm{E}\left(\mathrm{E}\left(\Psi\left(Z_m, \dots, Z_{t-m(n-1)}\right) \middle| \sum_{j=m}^{\infty} Z_j, \sum_{j=m}^{\infty} jZ_j\right)\right)$$

$$= \sum_s \sum_r \mathrm{E}\left(\Psi\left(Z_m, \dots, Z_{t-m(n-1)}\right) \middle| \sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right)$$

$$\times P\left(\sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_m, z_{m+1}, \dots) \ni \\ z_j \in \{0,1,\dots\} \;\; j=m,m+1,\dots \\ z_m + z_{m+1} + \cdots = r \\ m z_m + (m+1) z_{m+1} + \cdots = s}} \!\!\!\cdots \sum \quad \Psi\left(z_m, \dots, z_{t-m(n-1)}\right) \right.$$

$$\times\ P\left(Z_m = z_m, Z_{m+1} = z_{m+1}, \ldots \middle| \sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right)$$

$$\left. \times\ P\left(\sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right) \right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_m,\ldots,z_{s-m(r-1)})\ni \\ z_j\in\{0,1,\ldots\}\ j=m,\ldots,s-m(r-1) \\ z_m+\cdots+z_{s-m(r-1)}=r \\ mz_m+\cdots+(s-m(r-1))z_{s-m(r-1)}=s}} \cdots \sum \Psi^*(\mathbf{z}) \right.$$

$$\times\ P\left(D_m = z_m, \ldots, D_{s-m(r-1)} = z_{s-m(r-1)}\right)$$

$$\left. \times\ P\left(\sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right) \right)$$

where

$$\Psi^*(\mathbf{z}) = \begin{cases} \Psi\left(z_m, z_{m+1}, \ldots, z_{t-m(n-1)}\right) & t - m(n-1) \le s - m(r-1) \\ \Psi\left(z_m, z_{m+1}, \ldots, z_{s-m(r-1)}, 0, \ldots, 0\right) & t - m(n-1) > s - m(r-1) \end{cases}$$

with $z_{s-m(r-1)+1}, \dots, z_{t-m(n-1)}$ all being replaced with 0's in the case $t-m(n-1) > s-m(r-1)$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) P\left(\sum_{j=m}^{\infty} Z_j = r, \sum_{j=m}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) e^{\frac{-\theta\lambda^m}{(1-\lambda)^m}} \lambda^s \theta^r \frac{\binom{s-1}{mr-1}}{r!}.$$

So, we have now established that

$$\mathrm{E}\left(\Psi\big(Z_m, \dots, Z_{t-m(n-1)}\big)\right) = \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) e^{\frac{-\theta\lambda^m}{(1-\lambda)^m}} \lambda^s \theta^r \frac{\binom{s-1}{mr-1}}{r!}$$

or

$$e^{\frac{\theta\lambda^m}{(1-\lambda)^m}} \mathrm{E}\left(\Psi\big(Z_m, \dots, Z_{t-m(n-1)}\big)\right) = \sum_s \sum_r \left(\mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s-1}{mr-1}}{r!}\right) \cdot \lambda^s \theta^r.$$

It follows that

$$\frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n}\left(e^{\frac{\theta\lambda^m}{(1-\lambda)^m}} \mathrm{E}\left(\Psi\big(Z_m, \dots, Z_{t-m(n-1)}\big)\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n}\left(\sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s-1}{mr-1}}{r!} \cdot \lambda^s \theta^r\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s-1}{mr-1}}{r!} \left(\frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \lambda^s\, \theta^r\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\binom{s-1}{mr-1}}{r!} \cdot t!\, n!\, \mathbb{I}(s=t)\mathbb{I}(r=n)$$

$$= \mathrm{E}\left(\Psi\big(D_m, \dots, D_{t-m(n-1)}\big)\right) \frac{\binom{t-1}{mn-1}}{n!} \cdot t!\, n!$$

or

$$\mathrm{E}\left(\Psi\big(D_m, \dots, D_{t-m(n-1)}\big)\right)$$

$$= \frac{1}{\binom{t-1}{mn-1} t!} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(e^{\frac{\theta\lambda^m}{(1-\lambda)^m}} \mathrm{E}\left(\Psi\big(Z_m, \dots, Z_{t-m(n-1)}\big)\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}} .$$

8.11 Proof of Example 12

Proof. Using Theorem 6 we have

$$P(D_1 = d_1, \dots, D_{t-n+1} = d_{t-n+1})$$

$$= \mathrm{E}\big(\Psi(C_1, \dots, C_n)\big)$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)t!} \frac{d^t}{dp^t} \left(\left(\frac{g(p,m)}{(1-p)^m}\right)^n \mathrm{E}\big(\Psi(Z_1, \dots, Z_n)\big)\right)\Bigg|_{p=0}$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)t!} \frac{d^t}{dp^t} \left(\left(\frac{g(p,m)}{(1-p)^m}\right)^n \frac{n!}{d_1! \cdots d_{t-n+1}!}\right.$$

$$\times \binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}} \frac{(1-p)^{mn} p^{(1d_1+\cdots+(t-n+1)d_{t-n+1})}}{\left(g(p,m)\right)^n} \Bigg)\Bigg|_{p=0}$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)t!} \frac{n!}{d_1! \cdots d_{t-n+1}!} \binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}} \left(\frac{d^t}{dp^t}(p^t)\Big|_{p=0} \right)$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)\ t!} \frac{n!}{d_1! \cdots d_{t-n+1}!} \binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}} t!$$

$$= \frac{1}{\mathfrak{Z}(t,m,n)} \frac{n!}{d_1! \cdots d_{t-n+1}!} \binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}}$$

where $d_1 + \cdots + d_{t-n+1} = n$ and $1d_1 + \cdots + (t-n+1)d_{t-n+1} = t$.

(29)

8.12 Proof of Theorem 13

Proof. The joint probability distribution of $Z_1, Z_2, \ldots$ can be determined as

$$P(Z_1 = z_1, Z_2 = z_2, \ldots) = e^{-\theta\left(\binom{1+m-1}{m-1}\lambda^1 + \binom{2+m-1}{m-1}\lambda^2 + \cdots\right)}$$

$$\times \frac{\theta^{z_1+z_2+\cdots} \left(\binom{1+m-1}{m-1}\lambda^1\right)^{z_1} \left(\binom{2+m-1}{m-1}\lambda^1\right)^{z_2} \cdots}{z_1!\, z_2! \cdots}$$

$$= e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1}\lambda^i} \cdot \frac{\theta^{z_1+z_2+\cdots} \left(\binom{1+m-1}{m-1}\lambda^1\right)^{z_1} \left(\binom{2+m-1}{m-1}\lambda^2\right)^{z_2} \cdots}{z_1!\, z_2! \cdots}.$$

Therefore,

$P(Z_1 + Z_2 + \cdots = n, 1Z_1 + 2Z_2 + \cdots = t)$

$$= \sum_{\substack{(z_1, z_2, \dots) \ni \\ z_j \in \{0,1,\dots\} \;\; j=1,2,\dots \\ z_1 + z_2 + \cdots = n \\ 1z_1 + 2z_2 + \cdots = t}} \cdots \sum P(Z_1 = z_1, Z_2 = z_2, \dots)$$

$$= e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \lambda^t \, \theta^n \cdot \sum_{\substack{(z_1, z_2, \dots) \ni \\ z_j \in \{0,1,\dots\} \;\; j=1,2,\dots \\ z_1 + z_2 + \cdots = n \\ 1z_1 + 2z_2 + \cdots = t}} \cdots \sum \frac{\binom{1+m-1}{m-1}^{z_1} \binom{2+m-1}{m-1}^{z_2} \cdots}{z_1! \, z_2! \cdots}$$

$$= e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \lambda^t \, \theta^n \sum_{\substack{(z_1, z_2, \dots, z_{t-n+1}) \ni \\ z_j \in \{0,1,\dots\} \;\; j=1,2,\dots,t-n+1 \\ z_1 + z_2 + \cdots + z_{t-n+1} = n \\ 1z_1 + 2z_2 + \cdots + (t-n+1) z_{t-n+1} = t}} \cdots \sum \frac{\binom{1+m-1}{m-1}^{z_1} \cdots \binom{t-n+m}{m-1}^{z_{t-n+1}}}{z_1! \cdots z_{t-n+1}!}$$

[as $z_j \equiv 0$ for all $j \geq t - n + 2$ by the restrictions $z_1 + z_2 + \cdots + z_{t-n+1} = n$ and $1z_1 + 2z_2 + \cdots + (t - n + 1) z_{t-n+1} = t$ taken together]

$$= e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \lambda^t \, \theta^n \cdot \frac{\mathfrak{Z}(t, m, n)}{n!}.$$

This last simplification follows on noting that

$$1 = \sum_{\substack{(d_1,d_2,\dots,d_{t-n+1})\ni \\ d_j\in\{0,1,\dots\}\ j=1,2,\dots,t-n+1 \\ d_1+d_2+\cdots+d_{t-n+1}=n \\ 1d_1+2d_2+\cdots+(t-n+1)d_{t-n+1}=t}} \cdots \sum P(D_1 = d_1, D_2 = d_2, \dots, D_{t-n+1} = d_{t-n+1})$$

$$= \frac{n!}{\mathfrak{Z}(t,m,n)} \sum_{\substack{(d_1,d_2,\dots,d_{t-n+1})\ni \\ d_j\in\{0,1,\dots\}\ j=1,2,\dots,t-n+1 \\ d_1+d_2+\cdots+d_{t-n+1}=n \\ 1d_1+2d_2+\cdots+(t-n+1)d_{t-n+1}=t}} \cdots \sum \frac{\binom{1+m-1}{m-1}^{d_1} \cdots \binom{t-n+m}{m-1}^{d_{t-n+1}}}{d_1! \cdots d_{t-n+1}!}.$$

Hence, for all $(z_1, z_2, \dots) \ni z_j \in \{0,1,\dots\}\ j = 1,2,\dots,\ \ z_1 + z_2 + \cdots = n$ and $1z_1 + 2z_2 + \cdots = t$ we have

$$P\left(Z_1 = z_1, Z_2 = z_2, \dots \,\middle|\, Z_1 + Z_2 + \cdots = n, 1Z_1 + 2Z_2 + \cdots = t\right)$$

$$= \frac{e^{-\theta\sum_{i=1}^{\infty}\binom{i+m-1}{m-1}\lambda^i}\lambda^t\,\theta^n \cdot \dfrac{\binom{1+m-1}{m-1}^{z_1} \cdots \binom{t-n+m}{m-1}^{z_{t-n+1}}}{z_1! \cdots z_{t-n+1}!}}{e^{-\theta\sum_{i=1}^{\infty}\binom{i+m-1}{m-1}\lambda^i}\lambda^t\,\theta^n \cdot \dfrac{\mathfrak{Z}(t,m,n)}{n!}}$$

$$= \frac{\binom{1+m-1}{m-1}^{z_1} \cdots \binom{t-n+m}{m-1}^{z_{t-n+1}}}{\mathfrak{Z}(t,m,n)} \cdot \frac{n!}{z_1! \cdots z_{t-n+1}!}$$

$$= P(D_1 = z_1, D_2 = z_2, \dots, D_{t-n+1} = z_{t-n+1}).$$

With these preliminary results we are in position to finish the proof.

$$\mathrm{E}\big(\Psi(Z_1, \dots, Z_{t-n+1})\big)$$

$$= \mathrm{E}\left(\mathrm{E}\left(\Psi(Z_1, \dots, Z_{t-n+1}) \middle| \sum_{j=0}^{\infty} Z_j, \sum_{j=0}^{\infty} jZ_j\right)\right)$$

$$= \sum_s \sum_r \mathrm{E}\left(\Psi(Z_1, \dots, Z_{t-n+1}) \middle| \sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$\times P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_1, z_2, \dots) \ni \\ z_j \in \{0,1,\dots\} \;\; j=1,2,\dots \\ z_1 + z_2 + \cdots = r \\ 1z_1 + 2z_2 + \cdots = s}} \cdots \sum \Psi(z_1, \dots, z_{t-n+1}) \right.$$

$$\times \; P\left(Z_1 = z_1, Z_2 = z_2, \dots \middle| \sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$\left. \times P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right) \right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_1, \dots, z_{s-r+1}) \ni \\ z_j \in \{0,1,\dots\} \;\; j=1,\dots,s-r+1 \\ z_1 + \cdots + z_{s-r+1} = r \\ 1z_1 + \cdots + (s-r+1) z_{s-r+1} = s}} \cdots \sum \Psi^*(\mathbf{z}) \, P(D_1 = z_1, \dots, D_{s-r+1} = z_{s-r+1}) \right)$$

$$\times\ P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

where

$$\Psi^*(\mathbf{z}) = \begin{cases} \Psi(z_1, z_2, \dots, z_{t-n+1}) & t-n+1 \le s-r+1 \\ \Psi(z_1, z_2, \dots, z_{s-r+1}, 0, \dots, 0) & t-n+1 > s-r+1 \end{cases}$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \cdot e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \lambda^s\, \theta^r \cdot \frac{\mathfrak{Z}(s,m,r)}{r!}.$$

So, we have now established that

$$\mathrm{E}\big(\Psi(Z_1, \dots, Z_{t-n+1})\big) = \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big) \cdot e^{-\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i} \lambda^s\, \theta^r \cdot \frac{\mathfrak{Z}(s,m,r)}{r!}$$

or

$$e^{\theta \sum_{i=1}^{\infty} \binom{i+m-1}{m-1} \lambda^i}\, \mathrm{E}\big(\Psi(Z_1, \dots, Z_{t-n+1})\big) = \sum_s \sum_r \left(\mathrm{E}\big(\Psi^*(\mathbf{D})\big) \frac{\mathfrak{Z}(s,m,r)}{r!}\right) \cdot \lambda^s\, \theta^r.$$

It follows that

$$\frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\theta\sum_{i=1}^{\infty}\binom{i+m-1}{m-1}\lambda^i}\mathrm{E}\big(\Psi(Z_1,\dots,Z_{t-n+1})\big)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$=\frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(\sum_s\sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big)\frac{\mathfrak{Z}(s,m,r)}{r!}\cdot\lambda^s\,\theta^r\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$=\sum_s\sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big)\frac{\mathfrak{Z}(s,m,r)}{r!}\cdot\left(\frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\lambda^s\,\theta^r\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$=\sum_s\sum_r \mathrm{E}\big(\Psi^*(\mathbf{D})\big)\frac{\mathfrak{Z}(s,m,r)}{r!}\cdot t!\;n!\;\mathbb{I}(s=t)\mathbb{I}(r=n)$$

$$=\mathrm{E}\big(\Psi(D_1,\dots,D_{t-n+1})\big)\frac{\mathfrak{Z}(t,m,n)}{n!}\cdot t!\,n!$$

or

$$\mathrm{E}\big(\Psi(D_1,\dots,D_{t-n+1})\big)=\frac{1}{\mathfrak{Z}(t,m,n)t!}\frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\theta\sum_{i=1}^{\infty}\binom{i+m-1}{m-1}\lambda^i}\mathrm{E}\big(\Psi(Z_1,\dots,Z_{t-n+1})\big)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}.$$

8.13 Proof of Theorem 14

Proof.

Following the ideas developed in the proofs of Theorems 9, 11 and 13, we can see that

$$P\Big(\big((D_{10},\dots,D_{1t}),(D_{20},\dots,D_{2t})\big)=\big((d_{10},\dots,d_{1t}),(d_{20},\dots,d_{2t})\big)\Big)$$

$$=P\begin{pmatrix}\text{exactly } d_{1j} \text{ of the } n_1 \text{ groups in row 1 contain } j \text{ balls, } j=0,1,\dots,t\\ \text{and}\\ \text{exactly } d_{2j} \text{ of the } n_2 \text{ groups in row 2 contain } j \text{ balls, } j=0,1,\dots,t\end{pmatrix}$$

$$=\left(\frac{\dfrac{n_1!}{(d_{10})!\cdots(d_{1t})!}\dbinom{0+m_1-1}{m_1-1}^{d_{10}}\cdots\dbinom{t+m_1-1}{m_1-1}^{d_{1t}}}{\dbinom{m_1n_1+m_2n_2+t-1}{t}}\right)$$

$$\times\ \frac{n_2!}{(d_{20})!\cdots(d_{2t})!}\binom{0+m_2-1}{m_2-1}^{d_{10}}\cdots\binom{t+m_2-1}{m_2-1}^{d_{2t}}.$$

Also,

$$P\Big(\big((Z_{10},Z_{11},\dots),(Z_{20},Z_{21},\dots)\big)=\big((z_{10},z_{11},\dots),(z_{20},z_{21},\dots)\big)\Big)$$

$$=e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}-\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\left(\prod_{j=0}^{\infty}\binom{j+m_2-1}{m_2-1}^{z_{2j}}\right)$$

$$\times\ \frac{\theta_1^{z_{10}+z_{11}+\cdots}\ \ \theta_2^{z_{20}+z_{21}+\cdots}\ \ \lambda^{(0z_{10}+1z_{11}+\cdots)+(0z_{20}+1z_{21}+\cdots)}}{(z_{10})!\,(z_{11})!\cdots(z_{20})!\,(z_{21})!\cdots}.$$

Therefore,

$$P\begin{pmatrix} (0Z_{10}+1Z_{11}+\cdots)+(0Z_{20}+1Z_{21}+\cdots)=t, \\ Z_{10}+Z_{11}+\cdots=n_1, \\ Z_{20}+Z_{21}+\cdots=n_2 \end{pmatrix}$$

$$= \sum_{\substack{((z_{10},z_{11},\dots),(z_{20},z_{21},\dots))\ni \\ (0z_{10}+1z_{11}+\cdots)+(0z_{20}+1z_{21}+\cdots)=t \\ z_{10}+z_{11}+\cdots=n_1 \\ z_{20}+z_{21}+\cdots=n_2}} \cdots \sum P\Big(\big((Z_{10},Z_{11},\dots),(Z_{20},Z_{21},\dots)\big) = \big((z_{10},z_{11},\dots),(z_{20},z_{21},\dots)\big)\Big)$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}-\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}} \left(\sum_{\substack{((z_{10},z_{11},\dots),(z_{20},z_{21},\dots))\ni \\ (0z_{10}+1z_{11}+\cdots)+(0z_{20}+1z_{21}+\cdots)=t \\ z_{10}+z_{11}+\cdots=n_1 \\ z_{20}+z_{21}+\cdots=n_2}} \cdots \sum \theta_1^{z_{10}+z_{11}+\cdots}\ \theta_2^{z_{20}+z_{21}+\cdots} \right.$$

$$\left. \times\ \lambda^{(0z_{01}+1z_{11}+\cdots)+(0z_{20}+1z_{21}+\cdots)}\ \frac{\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\left(\prod_{j=0}^{\infty}\binom{j+m_2-1}{m_2-1}^{z_{2j}}\right)}{(z_{10})!\,(z_{11})!\cdots(z_{20})!\,(z_{21})!\cdots} \right)$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}-\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\theta_1^{n_1}\theta_2^{n_2}\lambda^t$$

$$\times \sum_{\substack{((z_{10},z_{11},\dots),(z_{20},z_{21},\dots))\ni \\ (0z_{10}+1z_{11}+\cdots)+(0z_{20}+1z_{21}+\cdots)=t \\ z_{10}+z_{11}+\cdots=n_1 \\ z_{20}+z_{21}+\cdots=n_2}} \cdots \sum \frac{\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\left(\prod_{j=0}^{\infty}\binom{j+m_2-1}{m_2-1}^{z_{2j}}\right)}{(z_{10})!\,(z_{11})!\cdots(z_{20})!\,(z_{21})!\cdots}$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}-\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\theta_1^{n_1}\theta_2^{n_2}\lambda^t$$

$$\times \sum_{\substack{((z_{10},z_{11},\dots,z_{1t}),(z_{20},z_{21},\dots,z_{2t}))\ni \\ (0z_{10}+1z_{11}+\cdots+tz_{1t})+(0z_{20}+1z_{21}+\cdots+tz_{2t})=t \\ z_{10}+z_{11}+\cdots+z_{1t}=n_1 \\ z_{20}+z_{21}+\cdots+z_{2t}=n_2}} \cdots \sum \frac{\left(\prod_{j=0}^{t}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\left(\prod_{j=0}^{t}\binom{j+m_2-1}{m_2-1}^{z_{2j}}\right)}{(z_{10})!\cdots(z_{1t})! \ \ (z_{20})!\cdots(z_{2t})!}$$

[as $z_{1j} = z_{2j} = 0$ for all $j \geq t+1$ by the restriction $(0z_{10} + 1z_{11} + \cdots) + (0z_{20} + 1z_{21} + \cdots) = t$]

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}-\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\theta_1^{n_1}\theta_2^{n_2}\lambda^t \frac{\binom{m_1n_1+m_2n_2+t-1}{t}}{n_1!\,n_2!}.$$

Hence, for all $((z_{10}, z_{11}, \dots), (z_{20}, z_{21}, \dots)) \ni z_{1j} \in \{0,1,\dots\}$ and $z_{2j} \in \{0,1,\dots\}$ $j = 0,1,\dots,$ $z_{10} + z_{11} + \cdots = n_1,$ $z_{20} + z_{21} + \cdots = n_2$ and $(0z_{10} + 1z_{11} + \cdots) + (0z_{20} + 1z_{21} + \cdots) = t$ we have

$$P\left(((Z_{10}, Z_{11}, \dots), (Z_{20}, Z_{21}, \dots)) = ((z_{10}, z_{11}, \dots), (z_{20}, z_{21}, \dots)) \ \middle| \ \begin{matrix} (0Z_{10} + 1Z_{11} + \cdots) + (0Z_{20} + 1Z_{21} + \cdots) = t \\ Z_{10} + Z_{11} + \cdots = n_1 \\ Z_{20} + Z_{21} + \cdots = n_2 \end{matrix}\right)$$

$$= \frac{\binom{0+m_1-1}{m_1-1}^{z_{10}}\cdots\binom{t+m_1-1}{m_1-1}^{z_{1t}}\binom{0+m_2-1}{m_2-1}^{z_{20}}\cdots\binom{t+m_2-1}{m_2-1}^{z_{2t}} \ n_1!\,n_2!}{\binom{m_1n_1+m_2n_2+t-1}{t}(z_{10})!\cdots(z_{1t})!\,(z_{20})!\cdots(z_{2t})!}$$

$$= P\left(\left((D_{10}, \dots, D_{1t}), (D_{20}, \dots, D_{2t})\right) = \left((z_{10}, \dots, z_{1t}), (z_{20}, \dots, z_{2t})\right)\right).$$

With these preliminary results we are in position to finish the proof.

$$\mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots, Z_{1t}), (Z_{20}, Z_{21}, \dots, Z_{2t})\right)\right)$$

$$= \mathrm{E}\left(\mathrm{E}\left(\Psi\left((Z_{10}, \dots, Z_{1t}), (Z_{20}, \dots, Z_{2t})\right) \middle| \sum_{j=0}^{\infty} Z_{1j}, \sum_{j=0}^{\infty} Z_{2j}, \sum_{j=0}^{\infty} j\left(Z_{1j} + Z_{2j}\right)\right)\right)$$

$$= \sum_{k} \sum_{q} \sum_{v} \left(\mathrm{E}\left(\Psi\left((Z_{10}, \dots, Z_{1t}), (Z_{20}, \dots, Z_{2t})\right) \middle| \sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} Z_{2j} = v, \sum_{j=0}^{\infty} j\left(Z_{1j} + Z_{2j}\right) = k \right) \times P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} Z_{2j} = v, \sum_{j=0}^{\infty} j\left(Z_{1j} + Z_{2j}\right) = k\right)\right)$$

$$= \sum_{k} \sum_{q} \sum_{v} \left(\sum_{\substack{((z_{10}, z_{11}, \dots), (z_{20}, z_{21}, \dots)) \ni \\ (0z_{10} + 1z_{11} + \cdots) + (0z_{20} + 1z_{21} + \cdots) = k \\ z_{10} + z_{11} + \cdots = q \\ z_{20} + z_{21} + \cdots = v}} \cdots \sum \Psi\left((z_{10}, \dots, z_{1t}), (z_{20}, \dots, z_{2t})\right) \right.$$

$$\times\ P\left(\left((Z_{10}, \dots, Z_{1t}), (Z_{20}, \dots, Z_{2t})\right) = \left((z_{10}, \dots, z_{1t}), (z_{20}, \dots, z_{2t})\right)\right.$$

$$\left| \sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} Z_{2j} = v, \sum_{j=0}^{\infty} j\big(Z_{1j} + Z_{2j}\big) = k \right)$$

$$\left. \times \; P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} Z_{2j} = v, \sum_{j=0}^{\infty} j\big(Z_{1j} + Z_{2j}\big) = k\right) \right)$$

$$= \sum_{k} \sum_{q} \sum_{v} \left(\sum_{\substack{((z_{10},z_{11},\dots,z_{1k}),(z_{20},z_{21},\dots,z_{2k}))\ni \\ (0z_{10}+1z_{11}+\dots+kz_{1k})+(0z_{20}+1z_{21}+\dots+kz_{2k})=k \\ z_{10}+z_{11}+\dots+z_{1k}=q \\ z_{20}+z_{21}+\dots+z_{2k}=v}} \cdots \sum \Psi^{*}(\mathbf{z}_1, \mathbf{z}_2) \right.$$

$$\times \; P\Bigg(\big((Z_{10}, \dots, Z_{1t}), (Z_{20}, \dots, Z_{2t})\big) = \big((z_{10}, \dots, z_{1t}), (z_{20}, \dots, z_{2t})\big)$$

$$\left| \sum_{j=0}^{t} Z_{1j} = q, \sum_{j=0}^{t} Z_{2j} = v, \sum_{j=0}^{t} j\big(Z_{1j} + Z_{2j}\big) = k \right)$$

$$\left. \times \; P\left(\sum_{j=0}^{t} Z_{1j} = q, \sum_{j=0}^{t} Z_{2j} = v, \sum_{j=0}^{t} j\big(Z_{1j} + Z_{2j}\big) = k\right) \right)$$

where

$$\Psi^{*}(\mathbf{z}_1, \mathbf{z}_2) = \begin{cases} \Psi((z_{10}, \dots, z_{1t}), (z_{20}, \dots, z_{2t})) & t \le k \\ \Psi((z_{10}, \dots, z_{1k}, 0, \dots, 0), (z_{20}, \dots, z_{2k}, 0, \dots, 0)) & t > k \end{cases}$$

with $z_{1(k+1)}, \dots, z_{1t}, z_{2(k+1)}, \dots, z_{2t}$ all being replaced with 0's in the case $t > k$

$$= \sum_k \sum_q \sum_v \mathrm{E}\big(\Psi^*(\mathbf{D}_1, \mathbf{D}_2)\big) P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} Z_{2j} = v, \sum_{j=0}^{\infty} j\big(Z_{1j} + Z_{2j}\big) = k\right)$$

$$= \sum_k \sum_q \sum_v \mathrm{E}\big(\Psi^*(\mathbf{D}_1, \mathbf{D}_2)\big) e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1} - \theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}} \theta_1^q \theta_2^v \lambda^k \frac{\binom{m_1 q + m_2 v + k - 1}{k}}{q!\, v!}.$$

So, we have now established that

$$\mathrm{E}\left(\Psi\big((Z_{10}, Z_{11}, \dots, Z_{1t}), (Z_{20}, Z_{21}, \dots, Z_{2t})\big)\right)$$

$$= \sum_k \sum_q \sum_v \mathrm{E}(\Psi^*(\mathbf{D}_1, \mathbf{D}_2)) e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1} - \theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}} \theta_1^q \theta_2^v \lambda^k \frac{\binom{m_1 q + m_2 v + k - 1}{k}}{q!\, v!}$$

or

$$e^{\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1} + \theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}} \mathrm{E}\left(\Psi\big((Z_{10}, Z_{11}, \dots, Z_{1t}), (Z_{20}, Z_{21}, \dots, Z_{2t})\big)\right)$$

$$= \sum_k \sum_q \sum_v \left(\mathrm{E}(\Psi^*(\mathbf{D}_1, \mathbf{D}_2)) \frac{\binom{m_1 q + m_2 v + k - 1}{k}}{q!\, v!}\right) \theta_1^q \theta_2^v \lambda^k.$$

It follows that

$$\frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\frac{d^{n_2}}{d\theta_2^{n_2}}\left(e^{\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}+\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\right.$$

$$\left.\times\ \mathrm{E}\left(\Psi\big((Z_{10},Z_{11},\dots,Z_{1t}),(Z_{20},Z_{21},\dots,Z_{2t})\big)\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}$$

$$= t!\ n_1!\ n_2!\left(\mathrm{E}\left(\Psi\big((D_{10},D_{11},\dots,D_{1t}),(D_{20},D_{21},\dots,D_{2t})\big)\right)\frac{\binom{m_1n_1+m_2n_2+t-1}{t}}{n_1!\,n_2!}\right)$$

or

$$\mathrm{E}\left(\Psi\big((D_{10},D_{11},\dots,D_{1t}),(D_{20},D_{21},\dots,D_{2t})\big)\right)=\frac{1}{\binom{m_1n_1+m_2n_2+t-1}{t}\ t!}$$

$$\times\ \frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\frac{d^{n_2}}{d\theta_2^{n_2}}\left(e^{\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}+\theta_2\left(\frac{1}{1-\lambda}\right)^{m_2}}\right.$$

$$\left.\times\ \mathrm{E}\left(\Psi\big((Z_{10},Z_{11},\dots,Z_{1t}),(Z_{20},Z_{21},\dots,Z_{2t})\big)\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}.$$

8.14 Proof of Theorem 15

Proof.

$$P\Big(\big((D_{10},\dots,D_{1t}),(C_{21},\dots,C_{2n_2})\big)=\big((d_{10},\dots,d_{1t}),(c_{21},\dots,c_{2n_2})\big)\Big)$$

$$= P\begin{pmatrix} \text{exactly } d_{1j} \text{ of the } n_1 \text{ groups in row 1 contain } i \text{ balls, } j = 0,1,\dots,t \\ \text{and} \\ \text{there are } c_{2j} \text{ balls in the } j^{th} \text{ group in row 2, } j = 0,1,\dots,t \end{pmatrix}$$

$$= \frac{\frac{n_1!}{(d_{10})!\cdots(d_{1t})!}\binom{0+m_1-1}{m_1-1}^{d_{10}}\cdots\binom{t+m_1-1}{m_1-1}^{d_{1t}} \times \binom{c_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{c_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}}{\binom{m_1n_1+M_2+t-1}{t}}.$$

Also,

$$P\Big(\big((Z_{10},Z_{12},\dots),(Z_{21},\dots,Z_{2n_2})\big) = \big((z_{10},z_{11},\dots),(z_{21},\dots,z_{2n_2})\big)\Big)$$

$$= e^{-\theta_1 \sum_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}\lambda^j}\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\frac{\theta_1^{z_{10}+z_{11}+\cdots}\ \lambda^{0z_{10}+1z_{11}+\cdots}}{(z_{10})!\,(z_{11})!\cdots}$$

$$\times \binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}(1-\lambda)^{M_2}\lambda^{z_{21}+\cdots+z_{2n_2}}$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}}\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\frac{\theta_1^{z_{10}+z_{11}+\cdots}\ \lambda^{0z_{10}+1z_{11}+\cdots}}{(z_{10})!\,(z_{11})!\cdots}$$

$$\times \binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}(1-\lambda)^{M_2}\lambda^{z_{21}+\cdots+z_{2n_2}}.$$

Therefore,

$$P\begin{pmatrix}(0Z_{10}+1Z_{11}+\cdots)+\left(Z_{21}+\cdots+Z_{2n_2}\right)=t,\\ Z_{10}+Z_{11}+\cdots=n_1\end{pmatrix}$$

$$=\sum_{\substack{\left((z_{10},z_{11},\dots),\left(z_{21},\dots,z_{2n_2}\right)\right)\ni\\ (0z_{10}+1z_{11}+\cdots)+\left(z_{21}+\cdots+z_{2n_2}\right)=t\\ z_{10}+z_{11}+\cdots=n_1}}\cdots\sum P\left(\left((Z_{10},Z_{11},\dots),\left(Z_{21},\dots,Z_{2n_2}\right)\right)=\left((z_{10},z_{11},\dots),\left(z_{21},\dots,z_{2n_2}\right)\right)\right)$$

$$=e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}}(1-\lambda)^{M_2}\left(\sum_{\substack{\left((z_{10},z_{11},\dots),\left(z_{21},\dots,z_{2n_2}\right)\right)\ni\\ (0z_{10}+1z_{11}+\cdots)+\left(z_{21}+\cdots+z_{2n_2}\right)=t\\ z_{10}+z_{11}+\cdots=n_1}}\cdots\sum\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\right.$$

$$\left.\times\frac{\binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}}{(z_{10})!\,(z_{11})!\cdots}\theta_1^{z_{10}+z_{11}+\cdots}\,\lambda^{(0z_{10}+1z_{11}+\cdots)+\left(z_{21}+\cdots+z_{2n_2}\right)}\right)$$

$$=e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}}(1-\lambda)^{M_2}\left(\sum_{\substack{\left((z_{10},z_{11},\dots),\left(z_{21},\dots,z_{2n_2}\right)\right)\ni\\ (0z_{10}+1z_{11}+\cdots)+\left(z_{21}+\cdots+z_{2n_2}\right)=t\\ z_{10}+z_{11}+\cdots=n_1}}\cdots\sum\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\right.$$

$$\times \frac{\binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}}{(z_{10})!\,(z_{11})!\cdots}\theta_1^{n_1}\lambda^t \Bigg)$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}}(1-\lambda)^{M_2}\theta_1^{n_1}\lambda^t\left(\sum_{\substack{\left((z_{10},z_{11},\dots),(z_{21},\dots,z_{2n_2})\right)\ni\\ (0z_{10}+1z_{11}+\cdots)+(z_{21}+\cdots+z_{2n_2})=t\\ z_{10}+z_{11}+\cdots=n_1}}\cdots\sum\left(\prod_{j=0}^{\infty}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\right.$$

$$\left.\times \frac{\binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}}{(z_{10})!\,(z_{11})!\cdots}\right)$$

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}}(1-\lambda)^{M_2}\theta_1^{n_1}\lambda^t$$

$$\times\left(\sum_{\substack{\left((z_{10},z_{11},\dots,z_{1t}),(z_{21},\dots,z_{2n_2})\right)\ni\\ (0z_{10}+1z_{11}+\cdots+tz_{1t})+(z_{21}+\cdots+z_{2n_2})=t\\ z_{10}+z_{11}+\cdots+z_{1t}=n_1}}\cdots\sum\left(\prod_{j=0}^{t}\binom{j+m_1-1}{m_1-1}^{z_{1j}}\right)\right.$$

$$\left.\times \frac{\binom{z_{21}+m_{21}-1}{m_{21}-1}\cdots\binom{z_{2n_2}+m_{2n_2}-1}{m_{2n_2}-1}}{(z_{10})!\cdots(z_{1t})!}\right)$$

[as $z_{1j} = 0$ for all $j \geq t+1$ by the restriction $0z_{10} + 1z_{11} + \cdots = t$]

$$= e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{M_2}\theta_1^{n_1}\lambda^t \cdot \binom{m_1 n_1 + M_2 + t - 1}{t}\frac{1}{n_1!}.$$

Hence, for all $\left((z_{10}, z_{11}, \dots), (z_{21}, \dots, z_{2n_2})\right) \ni z_{1j} \in \{0,1,\dots\}$ $j = 0,1,\dots$ and $z_j \in \{0,1,\dots\}$, $j = 1,\dots,r,$ $z_{10} + z_{11} + \cdots = n_1,$ and $(0z_{10} + 1z_{11} + \cdots) + \left(z_{21} + \cdots + z_{2n_2}\right) = t$ we have

$$P\left(\left((Z_{10}, Z_{11}, \dots), (Z_{21}, \dots, Z_{2n_2})\right) = \left((z_{10}, z_{11}, \dots), (z_{21}, \dots, z_{2n_2})\right) \middle| \begin{matrix} (0Z_{10} + 1Z_{11} + \cdots) + \left(Z_{21} + \cdots + Z_{2n_2}\right) = t, \\ Z_{10} + Z_{11} + \cdots = n_1 \end{matrix}\right)$$

$$= \frac{\binom{0 + m_1 - 1}{m_1 - 1}^{z_{10}} \cdots \binom{t + m_1 - 1}{m_1 - 1}^{z_{1t}} \binom{z_{21} + m_{21} - 1}{m_{21} - 1} \cdots \binom{z_{2n_2} + m_{2n_2} - 1}{m_{2n_2} - 1} n_1!}{\binom{m_1 n_1 + M_2 + t - 1}{t} (z_{10})! \cdots (z_{1t})!}$$

$$= P\left(\left((D_{10}, \dots, D_{1t}), (C_{21}, \dots, C_{2n_2})\right) = \left((z_{10}, \dots, z_{1t}), (z_{21}, \dots, z_{2n_2})\right)\right).$$

Therefore,

$$\mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots\ Z_{1t}), (Z_{21}, \dots, Z_{2n_2})\right)\right)$$

$$= \mathrm{E}\left(\mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots\ Z_{1t}), (Z_{21}, \dots, Z_{2n_2})\right) \middle| \sum_{j=0}^{\infty} Z_{1j}, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j}\right)\right)$$

$$= \sum_k \sum_q \mathrm{E}\left(\Psi\Big((Z_{10}, Z_{11}, \dots\ Z_{1t}), \big(Z_{21}, \dots, Z_{2n_2}\big)\Big) \right.$$

$$\left| \sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j} = k \right)$$

$$\times\ P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j} = k \right)$$

$$= \sum_k \sum_q \left[\left(\sum_{\substack{\left((z_{10}, z_{11}, \dots), (z_{21}, \dots, z_{2n_2})\right) \ni \\ (0z_{10} + 1z_{11} + \cdots) + (z_{21} + \cdots + z_{2n_2}) = k \\ z_{10} + z_{11} + \cdots = q}} \cdots \sum \Psi\Big((z_{10}, z_{11}, \dots\ z_{1t}), \big(z_{21}, \dots, z_{2n_2}\big)\Big) \right. \right.$$

$$\cdot\ P\left(\Big((Z_{10}, \dots, Z_{1t}), \big(Z_{21}, \dots, Z_{2n_2}\big)\Big) = \Big((z_{10}, \dots, z_{1t}), \big(z_{21}, \dots, z_{2n_2}\big)\Big) \right.$$

$$\left. \left| \begin{matrix} (0Z_{10} + 1Z_{11} + \cdots) + \big(Z_{21} + \cdots + Z_{2n_2}\big) = k \\ Z_{10} + Z_{11} + \cdots = q \end{matrix} \right) \right)$$

$$\left. \cdot\ P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j} = k \right) \right]$$

$$= \sum_k \sum_q \left[\left(\sum_{\substack{\left((z_{10}, z_{11}, \dots, z_{1k}), (z_{21}, \dots, z_{2n_2})\right) \ni \\ (0z_{10} + 1z_{11} + \cdots + kz_{1k}) + (z_{21} + \cdots + z_{2n_2}) = k \\ z_{10} + z_{11} + \cdots + z_{1k} = q}} \cdots \sum \Psi^*(\mathbf{z}_1, \mathbf{z}_2) \right. \right.$$

$$\left. \cdot\ P\left(\left((D_{10}, \dots, D_{1t}), (C_{21}, \dots, C_{2n_2})\right) = \left((z_{10}, \dots, z_{1t}), (z_{21}, \dots, z_{2n_2})\right)\right) \right)$$

$$\left. \cdot\ P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j} = k\right) \right]$$

where

$$\Psi^*(\mathbf{z}_1, \mathbf{z}_2) = \begin{cases} \Psi\left((z_{10}, z_{11}, \dots\ z_{1t}), (z_{21}, \dots, z_{2n_2})\right) & t \le k \\ \Psi\left((z_{10}, z_{11}, \dots\ z_{1k}, 0, \dots, 0), (z_{21}, \dots, z_{2n_2})\right) & t > k \end{cases}$$

with $z_{1(k+1)}, z_{1(k+2)}, \dots, z_{1t}$ all being replaced with 0’s in the case $t > k$

$$= \sum_k \sum_q \mathrm{E}\left(\Psi^*(\mathbf{D}_1, \mathbf{C}_2)\right) P\left(\sum_{j=0}^{\infty} Z_{1j} = q, \sum_{j=0}^{\infty} jZ_{1j} + \sum_{j=1}^{n_2} Z_{2j} = k\right)$$

$$= \sum_k \sum_q \mathrm{E}\big(\Psi^*(\mathbf{D}_1, \mathbf{C}_2)\big) e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{M_2} \theta_1^q \lambda^k \binom{m_1 q + M_2 + k - 1}{k} \frac{1}{q!}.$$

So, we have now established that

$$\mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots\ Z_{1t}), \left(Z_{21}, \dots, Z_{2n_2}\right)\right)\right)$$

$$= \sum_k \sum_q \mathrm{E}\big(\Psi^*(\mathbf{D}_1, \mathbf{C}_2)\big) e^{-\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{M_2} \theta_1^q \lambda^k \binom{m_1 q + M_2 + k - 1}{k} \frac{1}{q!}$$

or

$$e^{\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{-M_2} \mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots\ Z_{1t}), \left(Z_{21}, \dots, Z_{2n_2}\right)\right)\right)$$

$$= \sum_k \sum_q \mathrm{E}\big(\Psi^*(\mathbf{D}_1, \mathbf{C}_2)\big) \binom{m_1 q + M_2 + k - 1}{k} \frac{1}{q!} \theta_1^q \lambda^k.$$

It follows that

$$\frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(e^{\theta_1\left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{-M_2} \mathrm{E}(\Psi((Z_{10}, Z_{11}, \dots\ Z_{1t}), \left(Z_{21}, \dots, Z_{2n_2}\right))) \right)\Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \mathrm{E}\left(\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), \left(C_{21}, \dots, C_{2n_2}\right)\right)\right) \binom{m_1 n_1 + M_2 + t - 1}{t} \frac{n_1!\, t!}{n_1!}$$

or

$$\mathrm{E}\left(\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), (C_{21}, \dots, C_{2n_2})\right)\right)$$

$$= \frac{1}{\binom{m_1 n_1 + M_2 + t - 1}{t}\, t!} \left(\frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(e^{\theta_1 \left(\frac{1}{1-\lambda}\right)^{m_1}} (1-\lambda)^{-M_2} \right.\right.$$

$$\left.\left. \mathrm{E}\left(\Psi\left((Z_{10}, Z_{11}, \dots Z_{1t}), (Z_{21}, \dots, Z_{2n_2})\right)\right)\right)\Big|_{\substack{\lambda=0\\ \theta_1=0}} \right).$$

8.15 Proof of Example 16

Part 1. If we take

$$\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), (C_{21}, \dots, C_{2n_2})\right) = \mathbb{I}(D_{10} = k, C_{21} = t - j)$$

in Theorem 15 then

$$P(D_{10} = k, C_{21} = t - j) = \mathrm{E}\left(\Psi\left((D_{10}, D_{11}, \dots, D_{1t}), (C_{21}, \dots, C_{2n_2})\right)\right)$$

$$= \frac{1}{\binom{m_1 n_1 + m_2 + t - 1}{t}\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\left(\frac{1}{1-\lambda}\right)^{m_2} e^{\frac{\theta_1}{(1-\lambda)^{m_1}}} P(Z_{10} = k, Z_{21} = t - j) \right)\Big|_{\substack{\lambda=0\\ \theta_1=0}}$$

where $Z_{10} \sim$ Poisson(θ_1), $Z_{21} \sim$ negative binomial$(m_{21}, 1-\lambda)$ and where Z_{10} and Z_{21} are independent random variables

$$= \frac{1}{\binom{m_1 n_1 + m_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\left(\frac{1}{1-\lambda} \right)^{m_2} e^{\frac{\theta_1}{(1-\lambda)^{m_1}}} \frac{e^{-\theta_1} (\theta_1)^k}{k!} \right.$$

$$\left. \times \binom{t - j + m_2 - 1}{m_2 - 1} (1-\lambda)^{m_2} \lambda^{t-j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(e^{\frac{\theta_1}{(1-\lambda)^{m_1}}} e^{-\theta_1} (\theta_1)^k \lambda^{t-j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(e^{\theta_1 \left(\frac{1}{(1-\lambda)^{m_1}} - 1 \right)} (\theta_1)^k \lambda^{t-j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\sum_{v=0}^{\infty} \frac{(\theta_1)^v \left(\frac{1}{(1-\lambda)^{m_1}} - 1 \right)^v}{v!} (\theta_1)^k \lambda^{t-j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\sum_{v=0}^{\infty} \sum_{i=0}^{v} (-1)^{v-i} \binom{v}{i} \right.$$

$$\left. \times \frac{1}{v!\, (1-\lambda)^{m_1 i}} (\theta_1)^{k+v} \lambda^{t-j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{\binom{t - j + m_2 - 1}{m_2 - 1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\sum_{v=0}^{\infty} \sum_{i=0}^{v} \sum_{w=0}^{\infty} (-1)^{v-i} \binom{v}{i} \right.$$

$$\times \frac{1}{v!}\binom{m_1 i + w - 1}{w}(\theta_1)^{k+v}\lambda^{t-j+w}\Bigg)\Bigg|_{\substack{\lambda=0\\ \theta_1=0}}$$

$$= \frac{\binom{t-j+m_2-1}{m_2-1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!}\left(\sum_{v=0}^{\infty}\sum_{i=0}^{v}\sum_{w=0}^{\infty}(-1)^{v-i}\binom{v}{i}\frac{1}{v!}\binom{m_1 i + w - 1}{w}\right.$$

$$\left.\times \frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\left((\theta_1)^{k+v}\lambda^{t-j+w}\right)\Big|_{\substack{\lambda=0\\ \theta_1=0}}\right)$$

$$= \frac{\binom{t-j+m_2-1}{m_2-1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!}\left(\sum_{v=0}^{\infty}\sum_{i=0}^{v}\sum_{w=0}^{\infty}(-1)^{v-i}\binom{v}{i}\frac{1}{v!}\binom{m_1 i + w - 1}{w}\right.$$

$$\left.\times\ t!\, n_1!\, \mathbb{I}(n_1 = k+v)\mathbb{I}(t = t-j+w)\right)$$

$$= \frac{\binom{t-j+m_2-1}{m_2-1}}{\binom{m_1 n_1 + m_2 + t - 1}{t} k!\, t!}\sum_{i=0}^{n_1-k}(-1)^{n_1-k-i}\binom{n_1-k}{i}\frac{1}{(n_1-k)!}\binom{m_1 i + j - 1}{j} t!\, n_1!$$

$$= \frac{\binom{t-j+m_2-1}{m_2-1}\binom{n_1}{k}}{\binom{m_1 n_1 + m_2 + t - 1}{t}}\sum_{i=0}^{n_1-k}(-1)^{n_1-k-i}\binom{n_1-k}{i}\binom{m_1 i + j - 1}{j}.$$

Proof of Part 2.

If we take

$$\Psi\left((D_{10},D_{11},\dots,D_{1t}),(C_{21},\dots,C_{2n_2})\right)=(D_{10})_{[\nu]}(C_{21})_{[\delta]}$$

in Theorem 15 then

$$\mathrm{E}\left((D_{10})_{[\nu]}(C_{21})_{[\delta]}\right)$$

$$=\frac{1}{\binom{m_1n_1+m_2+t-1}{t}t!}\frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\left(\left(\frac{1}{1-\lambda}\right)^{m_2}e^{\frac{\theta_1}{(1-\lambda)^{m_1}}}\mathrm{E}\left((Z_{10})_{[\nu]}(Z_{21})_{[\delta]}\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta_1=0}}$$

where $Z_{10}\sim$ Poisson(θ_1), $Z_{21}\sim$ negative binomial$(m_2,1-\lambda)$ and where Z_{10} and Z_{21} are independent random variables

$$=\frac{1}{\binom{m_1n_1+m_2+t-1}{t}t!}\frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\left(\left(\frac{1}{1-\lambda}\right)^{m_2}e^{\frac{\theta_1}{(1-\lambda)^{m_1}}}\mathrm{E}\left((Z_{10})_{[\nu]}\right)\mathrm{E}\left((Z_{21})_{[\delta]}\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta_1=0}}.$$

But we know (see Appendix 10) that

$$\mathrm{E}\left((Z_{10})_{[\nu]}\right)=(\theta_1)^{\nu}$$

and

$$\mathrm{E}\left((Z_{21})_{[\delta]}\right)=\left(\frac{\lambda}{1-\lambda}\right)^{\delta}m_2^{[\delta]}.$$

Therefore,

$$\mathrm{E}\left((D_{10})_{[\nu]}(C_{21})_{[\delta]}\right)$$

$$= \frac{m_2^{[\delta]}}{\binom{m_1 n_1 + m_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left(\left(\frac{1}{1-\lambda} \right)^{m_2} e^{\frac{\theta_1}{(1-\lambda)^{m_1}}} (\theta_1)^{\nu} \left(\frac{\lambda}{1-\lambda} \right)^{\delta} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}}$$

$$= \frac{m_2^{[\delta]}}{\binom{m_1 n_1 + m_2 + t - 1}{t} t!} \left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{i!} \binom{m_1 i + m_2 + \delta + j - 1}{j} \right.$$

$$\left. \times \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \left((\theta_1)^{\nu+i} (\lambda)^{\delta+j} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0}} \right)$$

$$= \frac{m_2^{[\delta]}}{\binom{m_1 n_1 + m_2 + t - 1}{t} t!} \frac{t!\, n_1!}{(n_1 - \nu)!} \binom{m_1(n_1 - \nu) + m_2 + \delta + (t - \delta) - 1}{t - \delta}$$

$$= \frac{m_2^{[\delta]}}{\binom{m_1 n_1 + m_2 + t - 1}{t}} \binom{n_1}{\nu} \nu! \binom{m_1(n_1 - \nu) + m_2 + t - 1}{t - \delta}$$

$$= \frac{1}{\binom{m_1 n_1 + m_2 + t - 1}{t}} \binom{n_1}{\nu} \binom{m_2 + \delta - 1}{\delta} \binom{m_1(n_1 - \nu) + m_2 + t - 1}{t - \delta} \nu!\, \delta!.$$

8.16 Proof of Example 17

Proof. In the language of Theorem 14 we can think of the first n_1 groups of urns as our "top row" and the remaining $n_2 = n - n_1$ groups of urns as our "bottom row". In the notation of Theorem 14 the problems asks for $\mathrm{E}(D_{20}|D_{10} = j)$.

We have that

$$\mathrm{E}(D_{20}|D_{10} = j)$$
$$= \sum_{k=0}^{n_2} kP(D_{20} = k|D_{10} = j)$$
$$= \sum_{k=0}^{n_2} k \cdot \frac{P(D_{10} = j, D_{20} = k)}{P(D_{10} = j)}$$

and we can use Theorem 14 to find each of these probabilities. For the numerator probability we take

$$\Psi\big((D_{10}, D_{11}, \ldots, D_{1t}), (D_{20}, D_{21}, \ldots, D_{2t})\big) = \mathbb{I}(D_{10} = j, D_{20} = k).$$

In this case

$$P(D_{10} = j, D_{20} = k)$$
$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda)^m} + \frac{\theta_2}{(1-\lambda)^m}} P(Z_{10} = j, Z_{20} = k) \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0 \\ \theta_2=0}}$$

where $Z_{10} \sim$ Poisson$\left(\binom{0 + m - 1}{m - 1} \theta_1 \lambda^0\right)$ and $Z_{20} \sim$ Poisson$\left(\binom{0 + m - 1}{m - 1} \theta_2 \lambda^0\right)$ and where Z_{10} and Z_{20} are independent

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda)^m} + \frac{\theta_2}{(1-\lambda)^m}} \frac{e^{-\theta_1} \theta_1^j}{j!} \frac{e^{-\theta_2} \theta_2^k}{k!} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0 \\ \theta_2=0}}$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \sum_{i=0}^{r+s} (-1)^{r+s-i} \binom{r+s}{i} \frac{1}{(1-\lambda)^{im}} \theta_1^{r+j} \theta_2^{s+k} \frac{1}{j!\,k!\,r!\,s!} \Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \sum_{i=0}^{r+s} \sum_{b=0}^{\infty} \Bigg((-1)^{r+s-i} \frac{1}{j!\,k!\,r!\,s!} \binom{r+s}{i}$$

$$\times \binom{im + b - 1}{b} \theta_1^{r+j} \theta_2^{s+k} \lambda^b \Bigg) \Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \sum_{i=0}^{r+s} \sum_{b=0}^{\infty} \Bigg((-1)^{r+s-i} \frac{1}{j!\,k!\,r!\,s!} \binom{r+s}{i} \binom{im + b - 1}{b}$$

$$\times \; t!\,n_1!\,n_2! \;\; \mathbb{I}(t = b) \mathbb{I}(n_1 = r + j) \mathbb{I}(n_2 = s + k) \Bigg)$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \sum_{i=0}^{(n_1-j)+(n_2-k)} \Bigg((-1)^{(n_1-j)+(n_2-k)-i} \frac{1}{j!\,k!\,(n_1-j)!\,(n_2-k)!}$$

$$\times \binom{(n_1 - j) + (n_2 - k)}{i} \binom{im + t - 1}{t} \; t!\,n_1!\,n_2! \Bigg)$$

$$= \frac{\binom{n_1}{j}\binom{n_2}{k}}{\binom{mn_1 + mn_2 + t - 1}{t}}$$

$$\times \sum_{i=0}^{n_1+n_2-j-k} (-1)^{n_1+n_2-j-k-i} \binom{n_1 + n_2 - j - k}{i} \binom{im + t - 1}{t}.$$

For the denominator probability we take

$$\Psi\big((D_{10}, D_{11}, \dots, D_{1t}), (D_{20}, D_{21}, \dots, D_{2t})\big) = \mathbb{I}(D_{10} = j).$$

In this case

$$P(D_{10} = j)$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!} \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda)^m} + \frac{\theta_2}{(1-\lambda)^m}} P(Z_{10} = j) \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0 \\ \theta_2=0}}$$

where $Z_{10} \sim \text{Poisson}\left(\binom{0 + m - 1}{m - 1} \theta_1 \lambda^0 \right)$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \frac{d^t}{d\lambda^t} \frac{d^{n_1}}{d\theta_1^{n_1}} \frac{d^{n_2}}{d\theta_2^{n_2}} \left(e^{\frac{\theta_1}{(1-\lambda)^m} + \frac{\theta_2}{(1-\lambda)^m}} \frac{e^{-\theta_1} \theta_1^j}{j!} \right) \Bigg|_{\substack{\lambda=0 \\ \theta_1=0 \\ \theta_2=0}}$$

$$= \frac{1}{\binom{mn_1 + mn_2 + t - 1}{t} t!}$$

$$\times \frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\frac{d^{n_2}}{d\theta_2^{n_2}}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{i=0}^{r}(-1)^{r-i}\binom{r}{i}\frac{1}{(1-\lambda)^{m(s+i)}}\theta_1^{r+j}\theta_2^{s}\frac{1}{j!\,r!\,s!}\Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}$$

$$=\frac{1}{\binom{mn_1+mn_2+t-1}{t}t!}$$

$$\times \frac{d^t}{d\lambda^t}\frac{d^{n_1}}{d\theta_1^{n_1}}\frac{d^{n_2}}{d\theta_2^{n_2}}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{i=0}^{r}\sum_{b=0}^{\infty}\Bigg((-1)^{r-i}\frac{1}{j!\,r!\,s!}\binom{r}{i}$$

$$\times \binom{m(s+i)+b-1}{b}\theta_1^{r+j}\theta_2^{s}\lambda^{b}\Bigg)\Bigg|_{\substack{\lambda=0\\ \theta_1=0\\ \theta_2=0}}$$

$$=\frac{1}{\binom{mn_1+mn_2+t-1}{t}t!}$$

$$\times \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{i=0}^{r}\sum_{b=0}^{\infty}\Bigg((-1)^{r-i}\frac{1}{j!\,r!\,s!}\binom{r}{i}\binom{m(s+i)+b-1}{b}$$

$$\times\ t!\,n_1!\,n_2!\ \ \mathbb{I}(t=b)\mathbb{I}(n_1=r+j)\mathbb{I}(n_2=s)\Bigg)$$

$$=\frac{1}{\binom{mn_1+mn_2+t-1}{t}t!}$$

$$\times \sum_{i=0}^{n_1-j}\Bigg((-1)^{n_1-j-i}\frac{1}{j!\,(n_1-j)!\,n_2!}$$

$$\times \binom{n_1-j}{i}\binom{m(n_2+i)+t-1}{t}\ t!\,n_1!\,n_2!\Bigg)$$

$$=\frac{\binom{n_1}{j}}{\binom{mn_1+mn_2+t-1}{t}}\sum_{i=0}^{n_1-j}(-1)^{n_1-j-i}\binom{n_1-j}{i}\binom{m(n_2+i)+t-1}{t}.$$

Therefore,

$$\mathrm{E}(D_{20}|D_{10}=j)=\sum_{k=0}^{n_2} k\cdot\frac{P(D_{10}=j,D_{20}=k)}{P(D_{10}=j)}$$

$$=\frac{\displaystyle\sum_{k=0}^{n_2}\sum_{i=0}^{n-j-k} k\cdot(-1)^{n-i-j-k}\binom{n-j-k}{i}\binom{im+t-1}{t}\binom{n_2}{k}}{\displaystyle\sum_{i=0}^{n_1-j}(-1)^{n_1-j-i}\binom{n_1-j}{i}\binom{(n_2+i)m+t-1}{t}}.$$

8.17 Proof of Example 21

Proof.

If we identify each column of urns as a group, then the allocation model in Part A is equivalent to an unrestricted grouped Bose-Einstein allocation of t identical balls into a single row of urns divided into n groups (columns) with m_1+m_2 urns per group.

Thus, W_A is equivalent to D_0 in Section 4.1. So, it follows immediately from our work in Example 10 in Section 4.1 along with the identity $a_{[0]}\equiv 1$, that

$$\mathrm{E}\big((W_A)_{[r_0]}\big)=\mathrm{E}\big((D_0)_{[r_0]}\big)=\mathrm{E}\big((D_0)_{[r_0]}(D_1)_{[0]}(D_2)_{[0]}\cdots(D_t)_{[0]}\big)$$

$$=\frac{\binom{(m_1+m_2)(n-r_0)+(t-0)-1}{t-0}\binom{0+(m_1+m_2)-1}{(m_1+m_2)-1}^{r_0}}{\binom{(m_1+m_2)n+t-1}{t}}n_{[r_0]}$$

$$= \frac{\binom{(m_1+m_2)(n-r_0)+t-1}{t}}{\binom{(m_1+m_2)n+t-1}{t}} n_{[r_0]}.$$

(Part B)

To find $\mathrm{E}\big((W_B)_{[r_0]}\big)$ we need to use the results developed in Section 5.2. For C_{ij} as defined in Theorem 18 we can express W_B as

$$W_B = \sum_{j=1}^{n} \mathbb{I}\big(C_{1j} + C_{2j} = 0\big).$$

Taking $\Psi((C_{11},\dots,C_{1n}),(C_{21},\dots,C_{2n})) = \left(\sum_{j=1}^{n} \mathbb{I}\big(C_{1j}+C_{2j}=0\big)\right)_{[r_0]}$ in Theorem 18, we have

$$\mathrm{E}\big((W_B)_{[r_0]}\big)$$

$$= \mathrm{E}\left(\Psi\left(\left(C_{11},\dots,C_{1n_1}\right),\left(C_{21},\dots,C_{2n_2}\right)\right)\right)$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \frac{d_1^t}{dp_1^{t_1}} \frac{d_2^t}{dp_2^{t_2}} \left(\left(\frac{1}{1-p_1}\right)^{m_1 n} \left(\frac{1}{1-p_2}\right)^{m_2 n} \right.$$

$$\left.\left. \times\ \mathrm{E}\left(\left(\sum_{j=1}^{n} \mathbb{I}\big(Z_{1j} + Z_{2j} = 0\big)\right)_{[r_0]} \right)\right)\right|_{\substack{p_1=0\\ p_2=0}}$$

where

$$Z_{1j} \sim \text{negativebinomial}(m_1, 1-p_1),\ \ j = 1,2,\dots,n$$

$$Z_{2j} \sim \text{negativebinomial}(m_2, 1-p_2), \;\; j = 1,2,\dots,n$$

$$Z_{11},\dots,Z_{1n},Z_{21},\dots,Z_{2n} \text{ are all independent}$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \; \frac{d_1^t}{dp_1^{t_1}} \frac{d_2^t}{dp_2^{t_2}} \Bigg(\left(\frac{1}{1-p_1}\right)^{m_1 n} \left(\frac{1}{1-p_2}\right)^{m_2 n}$$

$$\times \; r_0! \binom{n}{r_0} \big(P(Z_{11} + Z_{21} = 0)\big)^{r_0} \Bigg)\Bigg|_{\substack{p_1=0\\ p_2=0}}$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \; \frac{d_1^t}{dp_1^{t_1}} \frac{d_2^t}{dp_2^{t_2}} \Bigg(\left(\frac{1}{1-p_1}\right)^{m_1 n} \left(\frac{1}{1-p_2}\right)^{m_2 n}$$

$$\times r_0! \binom{n}{r_0} \big(P(Z_{11} = 0) P(Z_{21} = 0)\big)^{r_0} \Bigg)\Bigg|_{\substack{p_1=0\\ p_2=0}}$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \; \frac{d_1^t}{dp_1^{t_1}} \frac{d_2^t}{dp_2^{t_2}} \Bigg(\left(\frac{1}{1-p_1}\right)^{m_1 n} \left(\frac{1}{1-p_2}\right)^{m_2 n}$$

$$\times r_0! \binom{n}{r_0} (1-p_1)^{m_1 r_0} (1-p_2)^{m_2 r_0} \Bigg)\Bigg|_{\substack{p_1=0\\ p_2=0}}$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \;\; r_0! \binom{n}{r_0}$$

$$\times \; \frac{d_1^t}{dp_1^{t_1}} \frac{d_2^t}{dp_2^{t_2}} \Bigg(\left(\frac{1}{1-p_1}\right)^{m_1 (n - r_0)} \left(\frac{1}{1-p_2}\right)^{m_2 (n - r_0)} \Bigg)\Bigg|_{\substack{p_1=0\\ p_2=0}}$$

$$= \frac{1}{\binom{m_1 n + t_1 - 1}{t_1}\binom{m_2 n + t_2 - 1}{t_2} t_1!\, t_2!} \;\; r_0! \binom{n}{r_0} \; \times$$

$$\frac{d_1^t}{dp_1^{t_1}}\frac{d_2^t}{dp_2^{t_2}}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{m_1(n-r_0)+i-1}{i}\binom{m_2(n-r_0)+j-1}{j}p_1^i p_2^j\right)\Bigg|_{\substack{p_1=0\\p_2=0}}$$

$$=\frac{\binom{m_1(n-r_0)+t_1-1}{t_1}\binom{m_2(n-r_0)+t_2-1}{t_2}}{\binom{m_1n+t_1-1}{t_1}\binom{m_2n+t_2-1}{t_2}}n_{[r_0]}.$$

Finally, to verify that

$$h(t_1)=\frac{\binom{m_1(n-1)+t_1-1}{t_1}\binom{m_2(n-1)+t-t_1-1}{t-t_1}}{\binom{m_1n+t_1-1}{t_1}\binom{m_2n+t-t_1-1}{t-t_1}}n$$

is minimized at $t_1=\left\lfloor\frac{m_1t+m_2}{m_1+m_2}\right\rfloor$ it suffices to show that

$$\frac{h(t_1)}{h(t_1-1)}<1\Leftrightarrow t_1<\frac{m_1t+m_2}{m_1+m_2}.$$

After simplification we have

$$\frac{h(t_1)}{h(t_1-1)}=\frac{(m_1(n-1)+t_1-1)(m_2n+t-t_1)}{(m_1n+t_1-1)(m_2(n-1)+t-t_1)}$$

and

$$\frac{h(t_1)}{h(t_1-1)}<1$$

$$\Leftrightarrow \frac{(m_1(n-1)+t_1-1)(m_2n+t-t_1)}{(m_1n+t_1-1)(m_2(n-1)+t-t_1)} < 1$$

$$\Leftrightarrow (m_1(n-1)+t_1-1)(m_2n+t-t_1)-(m_1n+t_1-1)(m_2(n-1)+t-t_1) < 0$$

$$\Leftrightarrow (m_1+m_2)t_1 - m_1t - m_2 < 0$$

$$\Leftrightarrow t_1 < \frac{m_1t+m_2}{m_1+m_2}.$$

8.18 Proof of Example 22

Proof. By Theorem 4

$$\mathrm{E}\left(\left(X_{(j:n)}\right)^m\right) = \frac{1}{\binom{t-1}{n-1}t!}\frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^n \mathrm{E}\left(\left(Z_{(j:n)}\right)^m\right)\right)\Bigg|_{p=0}$$

where $Z_1, \dots, Z_n$ are independent random variables such that $Z_j \sim 1$-shifted geometric$(1-p)$ for $j = 1, \dots, n$.

And by Theorem 39 in Appendix 9 we have that

$$\mathrm{E}\left(\left(Z_{(j:n)}\right)^m\right) = \sum_{z=0}^{\infty} P\left(Z_{(j:n)} > z\right)\left((z+1)^m - z^m\right).$$

Now

$$P\left(Z_{(j:n)} > z\right) = P(\text{at least } n-j+1 \text{ of } Z_1, \dots, Z_n > z)$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)}\binom{k-1}{(n-j+1)-1}\binom{n}{k}P(Z_1 > z, \dots, Z_k > z)$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(P(Z_1 > z)\right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(\sum_{i=z+1}^{\infty} P(Z_1 = i) \right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(\sum_{i=z+1}^{\infty} (1-p)p^{i-1} \right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} p^{zk}.$$

Therefore,

$$\mathrm{E}\left(\left(Z_{(j:n)}\right)^m\right) = \sum_{z=0}^{\infty} P\left(Z_{(j:n)} > z\right)\left((z+1)^m - z^m\right)$$

$$= \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} p^{zk}\left((z+1)^m - z^m\right).$$

Thus,

$$\mathrm{E}\left(\left(X_{(j:n)}\right)^m\right)$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \frac{d^t}{dp^t} \left(\left(\frac{p}{1-p} \right)^n \mathrm{E}\left(\left(Z_{(j:n)}\right)^m\right) \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \frac{d^t}{dp^t} \Bigg(\left(\frac{p}{1-p} \right)^n \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)}$$

$$\times \binom{k-1}{(n-j+1)-1} \binom{n}{k} p^{zk}\left((z+1)^m - z^m\right) \Bigg) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \frac{d^t}{dp^t} \left(\sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} (-1)^{k-(n-j+1)} \binom{n+i-1}{i} \right.$$

$$\left. \left. \times \binom{k-1}{n-j} \binom{n}{k} p^{i+n+zk} \left((z+1)^m - z^m \right) \right) \right|_{p=0}$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} (-1)^{k-(n-j+1)} \binom{n+i-1}{i}$$

$$\times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^m - z^m \right) \left(\left. \frac{d^t}{dp^t} p^{i+n+zk} \right|_{p=0} \right)$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} (-1)^{k-(n-j+1)} \binom{n+i-1}{i}$$

$$\times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^m - z^m \right) t! \; \mathbb{I}(t = i + n + zk)$$

$$= \frac{1}{\binom{t-1}{n-1} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{n + (t-n-zk) - 1}{t-n-zk}$$

$$\times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^m - z^m \right) t! \, \mathbb{I}(t - n - zk \geq 0)$$

$$= \frac{1}{\binom{t-1}{n-1}} \sum_{k=n-j+1}^{n} \sum_{z=0}^{\lfloor \frac{t-n}{k} \rfloor} (-1)^{k-n+j-1} \binom{t-zk-1}{t-zk-n}$$

$$\times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^m - z^m \right).$$

8.19 Proof of Example 23

Proof. The joint distribution of a random composition $(X_1, X_2, \dots, X_n)$ is given by (??) with $m_1 = \cdots = m_n = 1$ where our X_j are the C_j in (??). Therefore, as we explain in Section 3.4, Theorem 7 is applicable to the problem of random compositions with a fixed number of parts whenever our chosen statistic $\Psi(x_1, x_2, \dots, x_n)$ is defined on all $\mathbb{R}^n_{\geq 0}$ and is bounded and continuous on the simplex $x_1 + x_2 + \cdots + x_n = 1$, $0 \leq x_j \leq 1$ for $j = 1,2, \dots, n$.

We have $0 \leq X_j/t \leq 1$ for each $j = 1,2, \dots, n$ in this problem, so

$$\Psi\left(\frac{X_1}{t}, \dots, \frac{X_n}{t}\right) = \left(\frac{X_{(j:n)}}{t}\right)^m$$

is clearly a bounded function in the variables $(X_1/t, X_2/t, \dots, X_n/t)$.

To show that $\Psi(X_1/t, \dots, X_n/t) = (X_{(j:n)}/t)^m$ is a continuous function in the variables $(X_1/t, X_2/t, \dots, X_n/t)$ requires a few extra steps. First, the well-known identity

$$\min(a, b) = \frac{(a+b) - |a-b|}{2}$$

makes it clear that

$$\min(f(x_1, \dots, x_n), g(x_1, \dots, x_n)) \tag{30}$$

is continuous provided $f(\ )$ and $g(\ )$ are continuous. Second, the identity

$$\min(x_1, x_2, x_3) = \min(\min(x_1, x_2), x_3)$$

along with (30) and an induction argument shows that $\min(x_1, x_2, \dots, x_n)$ is a continuous function in the variables $(x_1, x_2, \dots, x_n)$.

Finally, letting $g(x) = x^m$ in the pointwise identity (i.e. valid for any given set of numbers $\{x_1, x_2, \dots, x_n\}$)

$$g(x_{(j:n)}) = \sum_{r=n-j+1}^{n} \sum_{(k_1,\dots,k_r)\in\mathbb{C}_r} (-1)^{r-n+j-1} \binom{r-1}{n-j} g\left(\min(x_{k_1}, \dots, x_{k_r})\right) \tag{31}$$

given in Suman [44], where $\mathbb{C}_r$ is defined to be the set of all subsets of $\{1,2, \dots, n\}$ with r elements shows that $\Psi(X_1/t, \dots, X_n/t) = (X_{(j:n)}/t)^m$ is a continuous function in the variables $(X_1/t, X_2/t, \dots, X_n/t)$. So, we can conclude that Theorem 7 is applicable in this example.

By Theorem 7, we have that

$$\lim_{t\to\infty} E\left(\left(\frac{X_{(j:n)}}{t}\right)^m\right) = (n-1)!\,\mathcal{L}^{-1}\left(\frac{E\left(\left(Y_{(j:n)}\right)^m\right)}{\lambda^n}\right)\Bigg|_{s=1} \tag{32}$$

where $Y_{(j:n)}$ is the j^{th} order statistic of $Y_1, Y_2, \dots, Y_n$ iid Gamma$(1, \lambda) \equiv$ Exponential(λ).

But it is well known (see for example [17]) that in this case $Y_{(1:n)} \sim$ Gamma$(1, n\lambda) \equiv$ Exponential$(n\lambda)$. Furthermore, it is a standard result (see Mood, Graybill and Boes [35]) that

$$\mathrm{E}(X^m) = \frac{\Gamma(r+m)}{\beta^m \Gamma(r)}$$

for $X \sim \text{Gamma}(r, \beta)$. It follows from these two well known results that

$$\mathrm{E}\left(\left(Y_{(1:n)}\right)^m\right) = \frac{m!}{\lambda^m n^m}.$$

Additionally, Srikantan [42] has proven that for general independent and identically distributed random variables $X_1, \dots, X_n$,

$$\mathrm{E}\left(g\left(X_{(j:n)}\right)\right) = \sum_{r=n-j+1}^{n} (-1)^{r-n+j-1} \binom{r-1}{n-j} \binom{n}{r} \mathrm{E}\left(g\left(X_{(1:r)}\right)\right). \tag{33}$$

We might note that (33) is also an immediate consequence of (31) which also gives some intuition into this result.

So, we have

$$\mathrm{E}\left(\left(Y_{(j:n)}\right)^m\right)$$

$$= \sum_{r=n-j+1}^{n} (-1)^{r-n+j-1} \binom{r-1}{n-j} \binom{n}{r} \mathrm{E}\left(\left(Y_{(1:r)}\right)^m\right)$$

$$= \sum_{r=n-j+1}^{n} (-1)^{r-n+j-1} \binom{r-1}{n-j} \binom{n}{r} \frac{m!}{\lambda^m r^m}.$$

Therefore,

$$\lim_{t\to\infty} \mathrm{E}\left(\left(\frac{X_{(j:n)}}{t}\right)^m\right) = (n-1)!\,\mathcal{L}^{-1}\left(\frac{\mathrm{E}(\left(Y_{(j:n)}\right)^m)}{\lambda^n}\right)\Bigg|_{s=1}$$

$$= (n-1)!\,\mathcal{L}^{-1}\left(\frac{\sum_{r=n-j+1}^{n} (-1)^{r-n+j-1}\binom{r-1}{n-j}\binom{n}{r}\frac{m!}{\lambda^m r^m}}{\lambda^n}\right)\Bigg|_{s=1}$$

$$= (n-1)!\left(\sum_{r=n-j+1}^{n} (-1)^{r-n+j-1}\binom{r-1}{n-j}\binom{n}{r}\frac{m!}{r^m}\right)\mathcal{L}^{-1}\left(\frac{1}{\lambda^{n+m}}\right)\Bigg|_{s=1}$$

$$= (n-1)!\left(\sum_{r=n-j+1}^{n} (-1)^{r-n+j-1}\binom{r-1}{n-j}\binom{n}{r}\frac{m!}{r^m}\right)\left(\frac{s^{n+m-1}}{(n+m-1)!}\right)\Bigg|_{s=1}$$

$$= \frac{\sum_{r=n-j+1}^{n} (-1)^{r-n+j-1}\binom{r-1}{n-j}\binom{n}{r}\frac{1}{r^m}}{\binom{n+m-1}{m}}.$$

Hence, for large t,

$$\mathrm{E}\left(\left(X_{(j:n)}\right)^m\right) \approx \frac{t^m}{\binom{n+m-1}{m}} \sum_{r=n-j+1}^{n} (-1)^{r-n+j-1}\binom{r-1}{n-j}\binom{n}{r}\frac{1}{r^m}.$$

8.20 Proof of Theorem 24

Proof.

Let $Z_1, Z_2, \ldots$ be an infinite sequence of independent random variables with $Z_j \sim$ Binomial $\left(1, \frac{\lambda\theta^j}{1+\lambda\theta^j}\right)$. In this case we see that

$$P(Z_1 = y_1, Z_2 = y_2, \ldots) = \prod_{j=1}^{\infty} P(Z_j = y_j)$$

$$= \prod_{j=1}^{\infty} \binom{1}{y_j} \left(\frac{\lambda\theta^j}{1+\lambda\theta^j}\right)^{y_j} \left(1 - \frac{\lambda\theta^j}{1+\lambda\theta^j}\right)^{1-y_j}$$

$$= \prod_{j=1}^{\infty} \binom{1}{y_j} \left(\frac{1}{1+\lambda\theta^j}\right) \lambda^{y_j} \theta^{\,j y_j}$$

$$= \frac{1}{\prod_{j=1}^{\infty} \left(1+\lambda\theta^j\right)} \; \lambda^{(y_1+y_2+\cdots)} \theta^{(1y_1+2y_2+\cdots)}.$$

It follows from this joint distribution that

$$P\left(Z_1 + Z_2 + \cdots = n \text{ and } 1Z_1 + 2Z_2 + \cdots = t\right)$$

$$= \sum_{\substack{(y_1,y_2,\ldots)\ni \\ y_j\in\{0,1\} \;\; j=1,2,\ldots \\ y_1+y_2+\cdots=n \\ y_1+2y_2+\cdots=t}} \cdots \sum P(Z_1 = y_1, Z_2 = y_2, \ldots)$$

$$= \sum_{\substack{(y_1,y_2,\ldots)\ni \\ y_j\in\{0,1\} \;\; j=1,2,\ldots \\ y_1+y_2+\cdots=n \\ y_1+2y_2+\cdots=t}} \cdots \sum \frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \; \lambda^{(y_1+y_2+\cdots)} \; \theta^{(1y_1+2y_2+\cdots)}$$

$$= \frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \ \lambda^n \, \theta^t \cdot \left(\sum \cdots \sum_{\substack{(y_1,y_2,\ldots)\ni \\ y_j\in\{0,1\} \ \ j=1,2,\ldots \\ y_1+y_2+\cdots=n \\ y_1+2y_2+\cdots=t}} 1 \right)$$

$$= \frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \ \lambda^n \, \theta^t \cdot \left(\sum \cdots \sum_{\substack{(y_1,\ldots,y_t)\ni \\ y_j\in\{0,1\} \ \ j=1,\ldots,t \\ y_1+\cdots+y_t=n \\ y_1+2y_2+\cdots+ty_t=t}} 1 \right)$$

$$= \frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \ \lambda^n \, \theta^t \cdot d(t,n).$$

Thus, for all $(y_1, y_2, \ldots)$ such that $y_j \in \{0,1\}$ $j = 1,2,\ldots$, $y_1 + y_2 + \cdots = n$ and $1y_1 + 2y_2 + \cdots = t$ we have that

$$P\left(Z_1 = y_1, Z_2 = y_2, \ldots \middle| Z_1 + Z_2 + \cdots = n \text{ and } 1Z_1 + 2Z_2 + \cdots = t\right)$$

$$= \frac{P(Z_1 = y_1, Z_2 = y_2, \ldots \text{ and } Z_1 + Z_2 + \cdots = n \text{ and } 1Z_1 + 2Z_2 + \cdots = t)}{P(Z_1 + Z_2 + \cdots = n \text{ and } 1Z_1 + 2Z_2 + \cdots = t)}$$

$$= \frac{P(Z_1 = y_1, Z_2 = y_2, \ldots)}{P(Z_1 + Z_2 + \cdots = n \text{ and } 1Z_1 + 2Z_2 + \cdots = t)}$$

$$= \frac{\frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \lambda^n \theta^t}{\frac{1}{\prod_{j=1}^{\infty} (1+\lambda\theta^j)} \ \lambda^n \, \theta^t \cdot d(t,n)}$$

$$= \frac{1}{d(t,n)} = P(Y_1 = y_1, \dots, Y_t = y_t).$$

With these preliminary results we are in position to finish the proof.

$$\mathrm{E}\big(\Psi(Z_1, Z_2, \dots, Z_t)\big)$$

$$\mathrm{E}\left(\mathrm{E}\left(\Psi(Z_1, Z_2, \dots, Z_t)\middle|\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)\right)$$

$$= \sum_s \sum_r \mathrm{E}\left(\Psi(Z_1, Z_2, \dots, Z_t)\middle|\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$\times P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \left(\sum_{\substack{(z_1, z_2 \dots, z_s) \ni \\ z_j \in \{0,1\} \;\; j=1,2,\dots,s \\ z_1 + z_2 + \cdots + z_s = r \\ 1z_1 + 2z_2 + \cdots + sz_s = s}} \cdots \sum \Psi^*(\mathbf{z}) \right.$$

$$\times \;\; P\left(Z_1 = z_1, Z_2 = z_2, \dots\middle|\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)\Bigg)$$

$$\times \;\; P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

where

$$\Psi^*(\mathbf{z}) = \begin{cases} \Psi(z_1, \dots, z_t) & t \le s \\ \Psi(z_1, \dots, z_s, 0, \dots, 0) & t > s \end{cases}$$

with $z_{s+1}, \dots, z_t$ all being replaced with 0's in the case $t > k$

$$\sum_s \sum_r \left(\sum_{\substack{(z_1, z_2 \dots, z_s) \ni \\ z_j \in \{0,1\} \;\; j=1,2,\dots,s \\ z_1+z_2+\dots+z_s=r \\ 1z_1+2z_2+\dots+sz_s=s}} \cdots \sum \Psi^*(\mathbf{z}) \times P(Y_1 = z_1, Y_2 = z_2, \dots, Y_s = z_s) \right) P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$= \sum_r \sum_s \mathrm{E}\big(\Psi^*(\mathbf{Y})\big) P\left(\sum_{j=1}^{\infty} Z_j = r, \sum_{j=1}^{\infty} jZ_j = s\right)$$

$$= \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{Y})\big) \cdot \frac{1}{\prod_{j=1}^{\infty} (1 + \lambda\theta^j)} d(s,r) \lambda^r \theta^s.$$

So, we have now established that

$$\mathrm{E}\big(\Psi(Z_1, Z_2, \dots, Z_t)\big) = \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{Y})\big) \cdot \frac{1}{\prod_{j=1}^{\infty} (1 + \lambda\theta^j)} d(s,r) \lambda^r \theta^s$$

or

$$\left(\prod_{j=1}^{\infty} \left(1 + \lambda\theta^j\right)\right) \mathrm{E}\big(\Psi(Z_1, Z_2, \dots, Z_t)\big) = \sum_s \sum_r \mathrm{E}\big(\Psi^*(\mathbf{Y})\big) d(s,r) \lambda^r \theta^s.$$

It follows that

$$\frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\left(\prod_{j=1}^{\infty}\left(1+\lambda\theta^j\right)\right)\mathrm{E}\big(\Psi(Z_1,Z_2,\dots,Z_t)\big)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \mathrm{E}\big(\Psi(Y_1,Y_2,\dots,Y_t)\big)d(t,n)\cdot n!\ \ t!$$

or

$$\mathrm{E}\big(\Psi(Y_1,Y_2,\dots,Y_t)\big) = \frac{1}{d(t,n)n!\ \ t!}\ \frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\left(\prod_{j=1}^{\infty}\left(1+\lambda\theta^j\right)\right)\cdot\mathrm{E}\big(\Psi(Z_1,Z_2,\dots,Z_t)\big)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}.$$

8.21 Proof of Example 25

Proof. By Theorem 24,

$$P\big(k \text{ occurs in a random composition of } t \text{ into } n \text{ parts, all distinct}\big)$$

$$= P(Y_k = 1)$$

$$= \frac{1}{d(t,n)\,n!\,t!}\ \frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\left(\prod_{j=1}^{\infty}\left(1+\lambda\theta^j\right)\right)\cdot P(Z_k=1)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{d(t,n)\, n!\, t!} \frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\left(\prod_{j=1}^{\infty}\left(1+\lambda\theta^j\right)\right)\cdot\left(\frac{\lambda\theta^k}{1+\lambda\theta^k}\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{d(t,n)\, n!\, t!} \frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\sum_{s=0}^{\infty}\sum_{r=1}^{s} d(s,r)\lambda^r\theta^s\left(\frac{\lambda\theta^k}{1+\lambda\theta^k}\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{d(t,n)n!\ t!} \frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}\left(\sum_{s=0}^{\infty}\sum_{r=1}^{s}\sum_{v=0}^{\infty} (-1)^v d(s,r)\lambda^{r+v+1}\theta^{s+k+kv}\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{d(t,n)n!\ t!} \sum_{s=0}^{\infty}\sum_{r=1}^{s}\sum_{v=0}^{\infty} (-1)^v d(s,r)\frac{d^n}{d\lambda^n}\frac{d^t}{d\theta^t}(\lambda^{r+v+1}\ \theta^{s+k+kv})\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{d(t,n)n!\ t!} \sum_{s=0}^{\infty}\sum_{r=1}^{s}\sum_{v=0}^{\infty} (-1)^v d(s,r)n!\,t!\,\mathbb{I}(n=r+v+1)\mathbb{I}(t=s+k+kv)$$

$$= \frac{1}{d(t,n)n!\ t!} \sum_{s=0}^{\infty}\sum_{r=1}^{s}\sum_{v=0}^{\infty} (-1)^v d(s,r)n!\,t!\,\mathbb{I}(r=n-v-1)\mathbb{I}(s=t-k-kv)$$

$$= \frac{1}{d(t,n)}\sum_{v=0}^{\infty} (-1)^v d(t-k-kv, n-v-1)\mathbb{I}(n-v-1\geq 1)\mathbb{I}(t-k-kv\geq 0)$$

$$= \frac{1}{d(t,n)} \sum_{v=0}^{\min\left(\left\lfloor\frac{t-k}{k}\right\rfloor, n-2\right)} (-1)^v d(t-k-kv, n-v-1).$$

8.22 Proof of Theorems 26 and 27

Proof. We will take the proofs of these two theorems together because they are so closely related. Let the random variable N represent the number of parts in a random composition of the integer t. Then

$$\begin{aligned}
&\mathrm{E}\big(\Psi(C_1, C_2, \dots, C_t)\big) \\
&= \mathrm{E}\big(\mathrm{E}(\Psi(C_1, C_2, \dots, C_t)|N)\big) \\
&= \sum_{n=1}^{t} \mathrm{E}(\Psi^*(C_1, C_2, \dots, C_n)|N = n) \cdot P(N = n) \\
&= \sum_{n=1}^{t} \frac{1}{\binom{t-1}{n-1} t!} \frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^n \mathrm{E}\big(\Psi^*(Z_1, Z_2, \dots, Z_n)\big)\right)\Bigg|_{p=0} \cdot \frac{\binom{t-1}{n-1}}{2^{t-1}} \\
&= \frac{1}{2^{t-1} t!} \sum_{n=1}^{t} \frac{d^t}{dp^t}\left(\left(\frac{p}{1-p}\right)^n \mathrm{E}\big(\Psi^*(Z_1, Z_2, \dots, Z_n)\big)\right)\Bigg|_{p=0}.
\end{aligned}$$

The proof of Theorem 27 follows in the same way.

8.23 Proof of Example 28

Proof.

$$P(Y = y) = \sum_{d=1}^{t} P\big(Y = y \big| D_y = d\big) P\big(D_y = d\big)$$

$$= \sum_{d=1}^{t} \frac{yd}{t} P\big(D_y = d\big) = \frac{y}{t} \sum_{d=1}^{t} d \cdot P\big(D_y = d\big) = \frac{y}{t} E\big(D_y\big).$$

But from Theorem 27

$$\mathrm{E}(D_y) = \frac{1}{2^{t-1}t!}\sum_{n=1}^{t} \frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\frac{\theta\lambda}{1-\lambda}}\mathrm{E}(Z_y)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

where $Z_y \sim$ Poisson$(\theta\lambda^y)$. However,

$$e^{\frac{\theta\lambda}{1-\lambda}} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{i+j-1}{j}\frac{1}{i!}\theta^i\lambda^{i+j} \quad \text{and} \quad \mathrm{E}(Z_y) = \theta\lambda^y.$$

Therefore,

$$\mathrm{E}(D_y) = \frac{1}{2^{t-1}t!}\sum_{n=1}^{t} \frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{i+j-1}{j}\frac{1}{i!}\theta^{i+1}\lambda^{i+j+y}\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{i+j-1}{j}\frac{t!\,n!}{i!}\ \mathbb{I}(n=i+1)\mathbb{I}(t=i+j+y)$$

$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\binom{i+j-1}{j}\frac{t!\,n!}{i!}\ \mathbb{I}(i=n-1)\mathbb{I}\big(j=t-y-(n-1)\big)$$

$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t}\left(\binom{(n-1)+\big(t-y-(n-1)\big)-1}{t-y-(n-1)}\frac{t!\,n!}{(n-1)!}\right.$$
$$\left.\times\ \mathbb{I}(n-1\geq 0)\mathbb{I}(t-y-n+1\geq 0)\right)$$

$$= \frac{1}{2^{t-1} t!} \sum_{n=1}^{t} \binom{t-y-1}{t-y-n+1} \frac{t!\, n!}{(n-1)!} \mathbb{I}(1 \le n \le t-y+1)$$

$$= \frac{1}{2^{t-1}} \sum_{n=1}^{t-y+1} n \binom{t-y-1}{t-y-(n-1)} \mathbb{I}(y \le t).$$

So, we have shown that

$$\mathrm{E}(D_y) = \frac{1}{2^{t-1}} \sum_{n=1}^{t-y+1} n \binom{t-y-1}{t-y-(n-1)}$$

for $y \in \{1,2,\dots,t\}$. However, after some simplification, we find

$$\sum_{n=1}^{t-y+1} n \binom{t-y-1}{t-y-(n-1)} = \mathbb{I}(t=y) + 2^{t-y-2}(t-y+3).$$

Therefore,

$$\mathrm{E}(D_y) = \begin{cases} \frac{t-y+3}{2^{y+1}} & y \in \{1,2,\dots,t-1\} \\ \frac{1}{2^{t-1}} & y = t \end{cases}$$

and

$$P(Y=y) = \begin{cases} \frac{y(t-y+3)}{t\, 2^{y+1}} & y \in \{1,2,\dots,t-1\} \\ \frac{1}{2^{t-1}} & y = t. \end{cases}$$

8.23 Proof of Example 29

Proof of Part A. Recall that D_j in Theorem 27 is defined as the multiplicity of the part size j in a random composition of t. It follows that, V_w^t, the number of part sizes with multiplicity w, can be expressed as

$$V_w^t = \sum_{r=1}^{t} \mathbb{I}(D_r = w).$$

Thus, it follows from Theorem 27 that

$$E(V_w^t) = \mathrm{E}\left(\sum_{r=1}^{t} \mathbb{I}(D_r = w)\right)$$

$$= \frac{1}{2^{t-1}t!}\sum_{n=1}^{t} \frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\frac{\theta\lambda}{1-\lambda}}\mathrm{E}\left(\sum_{r=1}^{t} \mathbb{I}(Z_r = w)\right)\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

where $Z_1, Z_2, \dots, Z_t$ are independent and $Z_j \sim \text{Poisson}\left(\theta\lambda^j\right)$. Hence

$$\mathrm{E}\left(\sum_{r=1}^{t} \mathbb{I}(Z_r = w)\right) = \sum_{r=1}^{t} \mathrm{E}\big(\mathbb{I}(Z_r = w)\big) = \sum_{r=1}^{t} P(Z_r = w) = \sum_{r=1}^{t} \frac{e^{-\theta\lambda^r}(\theta\lambda^r)^w}{w!}.$$

From here we have

$$\mathrm{E}(V_w^t) = \frac{1}{2^{t-1}t!}\sum_{n=1}^{t} \frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(\mathrm{e}^{\frac{\theta\lambda}{1-\lambda}}\sum_{r=1}^{t} \frac{e^{-\theta\lambda^r}(\theta\lambda^r)^w}{w!}\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{2^{t-1}t!\,w!}\sum_{n=1}^{t}\sum_{r=1}^{t} \frac{d^t}{d\lambda^t}\frac{d^n}{d\theta^n}\left(e^{\theta\left(\frac{\lambda}{1-\lambda}-\lambda^r\right)}\theta^w\lambda^{rw}\right)\Bigg|_{\substack{\lambda=0\\ \theta=0}}$$

$$= \frac{1}{2^{t-1} t!\, w!} \sum_{n=1}^{t} \sum_{r=1}^{t} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{\lambda}{1-\lambda} - \lambda^r \right)^i \theta^{i+w} \lambda^{rw} \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{1}{2^{t-1} t!\, w!} \sum_{n=1}^{t} \sum_{r=1}^{t} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\sum_{i=0}^{\infty} \sum_{j=0}^{i} \frac{1}{i!} \left((-1)^{i-j} \binom{i}{j} \right.\right.$$

$$\left.\left. \times \left(\frac{\lambda}{1-\lambda} \right)^j \lambda^{r(i-j)} \right) \theta^{i+w} \lambda^{rw} \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{1}{2^{t-1} t!\, w!} \sum_{n=1}^{t} \sum_{r=1}^{t} \frac{d^t}{d\lambda^t} \frac{d^n}{d\theta^n} \left(\sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{\infty} (-1)^{i-j} \frac{1}{i!} \binom{i}{j} \right.$$

$$\left. \times \binom{j+k-1}{k} \theta^{i+w} \lambda^{j+k+r(i+w-j)} \right) \Bigg|_{\substack{\lambda=0 \\ \theta=0}}$$

$$= \frac{1}{2^{t-1} t!\, w!} \sum_{n=1}^{t} \sum_{r=1}^{t} \sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{\infty} \left((-1)^{i-j} \frac{t!\, n!}{i!} \binom{i}{j} \binom{j+k-1}{k} \right.$$

$$\left. \times\ \mathbb{I}(n = i + w) \mathbb{I}\big(t = j + k + r(i + w - j)\big) \right)$$

$$= \frac{1}{2^{t-1}} \sum_{n=1}^{t} \sum_{r=1}^{t} \sum_{j=0}^{n-w} (-1)^{n-w-j} \binom{n}{w} \binom{n-w}{j} \binom{t - r(n-j) - 1}{t - r(n-j) - j} \mathbb{I}(n - w \geq 0)$$

$$= \frac{1}{2^{t-1}} \sum_{n=\max\{1,w\}}^{t} \sum_{r=1}^{t} \sum_{j=0}^{n-w} (-1)^{n-w-j} \binom{n}{w} \binom{n-w}{j} \binom{t - r(n-j) - 1}{t - r(n-j) - j}.$$

Proof of Part B. V_0^t equals the number of part sizes with multiplicity 0 in a random composition of t. That is, V_0^t equals the number of part sizes that do not occur in a random composition of t. Thus $t - V_0^t$ equals the number of part sizes which do occur in a random composition of t. Therefore, $E(t - V_0^t) = t - E(V_0^t)$ equals the expected number of part sizes which occur in a random composition of t. To finish the proof we only need to substitute the result in Part A for $E(V_0^t)$.

8.24 Proof of Example 30

Proof. By Theorem 1

$$\mathrm{E}\left(\left(C_{(j:n)}\right)^r\right) = \frac{1}{\binom{n+t-1}{t} t!} \frac{d^t}{dp^t}\left(\left(\frac{1}{1-p}\right)^n \mathrm{E}\left(\left(Z_{(j:n)}\right)^r\right)\right)\Bigg|_{p=0}$$

where $Z_1, \dots, Z_n$ are independent random variables such that $Z_j \sim \text{geometric}(1-p)$ for $j = 1, \dots, n$.

$$P\left(Z_{(j:n)} > z\right) = P(\text{at least } n-j+1 \text{ of } Z_1, \dots, Z_n > z)$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} P(Z_1 > z, \dots, Z_k > z)$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(P(Z_1 > z)\right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(\sum_{i=z+1}^{\infty} P(Z_1 = i)\right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left(\sum_{i=z+1}^{\infty} (1-p)p^i \right)^k$$

$$= \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} p^{k(z+1)}.$$

Therefore,

$$\mathrm{E}\left(\left(Z_{(j:n)}\right)^r\right) = \sum_{z=0}^{\infty} P\left(Z_{(j:n)} > z\right)\left((z+1)^r - z^r\right)$$

$$= \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \binom{k-1}{(n-j+1)-1} \binom{n}{k} p^{k(z+1)} \left((z+1)^r - z^r\right).$$

Thus,

$$\mathrm{E}\left(\left(C_{(j:n)}\right)^r\right) = \frac{1}{\binom{n+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^n E\left(\left(Z_{(j:n)}\right)^r\right) \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{n+t-1}{t} t!} \frac{d^t}{dp^t} \left(\left(\frac{1}{1-p} \right)^n \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} (-1)^{k-(n-j+1)} \right.$$

$$\left. \times \binom{k-1}{(n-j+1)-1} \binom{n}{k} \left((z+1)^r - z^r\right) p^{k(z+1)} \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{n+t-1}{t} t!} \frac{d^t}{dp^t} \left(\sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} (-1)^{k-(n-j+1)} \binom{n+i-1}{i} \right.$$

$$\left. \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^r - z^r\right) p^{i+k(z+1)} \right) \Bigg|_{p=0}$$

$$= \frac{1}{\binom{n+t-1}{t} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} \left((-1)^{k-(n-j+1)} \binom{n+i-1}{i} \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^r - z^r \right) \left(\left(\frac{d^t}{dp^t} p^{i+k(z+1)} \right) \Big|_{p=0} \right) \right)$$

$$= \frac{1}{\binom{n+t-1}{t} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \sum_{i=0}^{\infty} \left((-1)^{k-(n-j+1)} \binom{n+i-1}{i} \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^r - z^r \right) t!\, \mathbb{I}\big(t = i + k(z+1)\big) \right)$$

$$= \frac{1}{\binom{n+t-1}{t} t!} \sum_{z=0}^{\infty} \sum_{k=n-j+1}^{n} \left((-1)^{k-(n-j+1)} \binom{n + \big(t - k(z+1)\big) - 1}{t - k(z+1)} \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^r - z^r \right) t!\ \ \mathbb{I}(t - k(z+1) \geq 0) \right)$$

$$= \frac{1}{\binom{n+t-1}{t}} \sum_{k=n-j+1}^{n} \sum_{z=0}^{\lfloor \frac{t-k}{k} \rfloor} \left((-1)^{k-n+j-1} \binom{n+t-k(z+1)-1}{t-k(z+1)} \times \binom{k-1}{n-j} \binom{n}{k} \left((z+1)^r - z^r \right) \right).$$

9 Ready Reference

Definition for Factorial Moments

$\boldsymbol{k^{th}}$ **Rising Factorial Moment**

$$a^{[k]} = a(a+1)\cdots(a+k-1)$$

$\boldsymbol{k^{th}}$ **Falling Factorial Moment**

$$a_{[k]} = a(a-1)\cdots(a-k+1).$$

Identity Connecting Falling and Rising Factorial Moments

$$(x+c)_{[k]} = \sum_{j=0}^{k} \frac{k!}{j!} \binom{c-j}{k-j} x^{[j]}$$

$$\mathrm{E}\left((n-X)_{(k)}\right) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (n-j)_{(k-j)} \, \mathrm{E}\left(X_{(j)}\right).$$

However, a multivariate generalization of Charles Jordan's inversion formula (see Takács [128] for references) allows us to express joint probabilities in terms of joint descending factorial moments.

In particular, under conditions given in Takács [128], Brandt, Brandt and Sulanke [15], and Lenard [98], we have that for general discrete random variables defined on $\{0, 1, \dots\}$

$$P(X_1 = x_1, \dots, X_n = x_n)$$

$$= \sum_{j_1 = x_1}^{\infty} \cdots \sum_{j_n = x_n}^{\infty} (-1)^{(j_1 + \dots + j_n) - (x_1 + \dots + x_n)} \binom{j_1}{x_1} \cdots \binom{j_n}{x_n}$$

$$\times \frac{1}{(j_1! \cdots j_n!)} \mathrm{E}\Big((X_1)_{(j_1)} \cdots (X_n)_{(j_n)}\Big).$$

A sufficient condition for this representation to hold is the existence of some $c < \infty$ such that $P(X_1 \leq c, \dots, X_n \leq c) = 1$.

Theorem 36 (k^{th} Falling Factorial Moment of the Sum of Indicator Variables)

Consider an experiment with probability space $(\Omega, \mathcal{A}, P)$ and suppose $A_1, \dots, A_n$ are all events within $\mathcal{A}$. Suppose the experiment is performed and let $\omega \in \Omega$ be the outcome of this experiment. Define

$$X = \text{number of events among } A_1, \dots, A_n \text{ that } \omega \text{ is an element of}$$
$$= \mathbb{I}(\omega \in A_1) + \cdots + \mathbb{I}(\omega \in A_n).$$

Then,

$$\frac{1}{r!}\,\mathrm{E}\big(X_{[r]}\big) = \sum_{(j_1, \dots, j_r) \in \mathbb{C}_r} P\big(A_{j_1} \cap \cdots \cap A_{j_r}\big)$$

where $\mathbb{C}_r$ is the set of all r subsets of $\{1, 2, \dots, n\}$.

Theorem 37 (Generalized Principle of Inclusion-Exclusion for Probabilities)

Suppose $A_1,\ldots,A_n$ are sets within a universal set Ω.

Define :

$$\mathbb{H}_m = \left\{x \in \Omega \mid x \text{ is an element of } \underline{\text{exactly}}\ m \text{ of the } n \text{ sets } A_1,\ldots,A_n\right\}$$

$$\mathbb{H}_{\geq m} = \left\{x \in \Omega \mid x \text{ is an element of } \underline{\text{at least}}\ m \text{ of the } n \text{ sets } A_1,\ldots,A_n\right\}$$

$$\mathbb{H}_{\leq m} = \left\{x \in \Omega \mid x \text{ is an element of } \underline{\text{at most}}\ m \text{of the } n \text{ sets } A_1,\ldots,A_n\right\}.$$

Define

$$P(\mathbb{S}_k) = \begin{cases} \displaystyle\sum \cdots \sum_{(j_1,\ldots,j_k)\in\mathbb{C}_k} P\left(A_{j_1} \cap \cdots \cap A_{j_k}\right) & 1 \leq k \leq n \\ 1 & k = 0 \end{cases}$$

where $\mathbb{C}_k$ is the set of all subsets of $\{1,\ldots,n\}$ with k elements. Then,

$$P(\mathbb{H}_m) = \sum_{k=m}^{n} (-1)^{k-m} \binom{k}{m} P(\mathbb{S}_k)$$

and

$$P(\mathbb{H}_{\geq m}) = \sum_{k=m}^{n} (-1)^{k-m} \binom{k-1}{m-1} P(\mathbb{S}_k)$$

and

$$P(\mathbb{H}_{\leq m}) = 1 - P(\mathbb{H}_{\geq m+1}).$$

In the event $P\left(A_{j_1} \cap \cdots \cap A_{j_k}\right)$ is the same for all $(j_1, \ldots, j_k) \in \mathbb{C}_k$, then

$$P(\mathbb{S}_k) = \begin{cases} \binom{n}{k} P(A_1 \cap \cdots \cap A_k) & 1 \leq k \leq n \\ 1 & k = 0. \end{cases}$$

Theorem 38 (Factorial Moments in Terms of Cumulative Distribution)

Let X be a discrete random variable defined on $\mathbb{S} \subseteq \{0,1, \ldots\}$. Then

$$\mu_{[r]} = r \sum_{n=0}^{\infty} P(X > n) n_{[r-1]}.$$

In the special case $r = 1$ we have the familiar result

$$\mu = \sum_{n=0}^{\infty} P(X > n).$$

Proof.

$$\mu_{[r]} = E(X_{[r]}) = \sum_{n=0}^{\infty} P(X = n) n_{[r]}$$

$$= \sum_{n=0}^{\infty} P(X = n) \left(\frac{d^r}{dt^r} t^n \bigg|_{t=1} \right)$$

$$= \left(\frac{d^r}{dt^r} \sum_{n=0}^{\infty} P(X = n) t^n \right) \Bigg|_{t=1}$$

$$= \left(\frac{d^r}{dt^r} g(t) \right) \Bigg|_{t=1} \text{ with } g(t) = \sum_{n=0}^{\infty} P(X = n) t^n.$$

Now define $h(t) = \sum_{n=0}^{\infty} P(X > n) t^n$. We note that

$$1 + (t-1)h(t)$$

$$= 1 + (t-1)\left(\sum_{n=0}^{\infty} P(X > n) t^n \right)$$

$$= \big(1 - P(X > 0)\big) t^0 + \big(P(X > 0) - P(X > 1)\big) t^1$$

$$+ \ (P(X > 1) - P(X > 2)) t^2 + \cdots$$

$$= (P(X = 0)) t^0 + (P(X = 1)) t^1 + (P(X = 2)) t^2 + \cdots$$

$$= g(t).$$

Therefore,

$$\frac{d^r}{dt^r} g(t) = \frac{d^r}{dt^r} \big(1 + (t-1)h(t)\big)$$

$$\frac{d^r}{dt^r} g(t) = \frac{d^r}{dt^r} \big((t-1)h(t)\big)$$

$$\frac{d^r}{dt^r}g(t) = \sum_{j=0}^{r} \binom{r}{j} \left(\frac{d^j}{dt^j}(t-1)\right)\left(\frac{d^{r-j}}{dt^{r-j}}h(t)\right)$$

[Leibniz's Product Rule for Differentiation]

$$\frac{d^r}{dt^r}g(t) = \binom{r}{0}\left((t-1)\right)\left(\frac{d^r}{dt^r}h(t)\right) + \binom{r}{1}\left(\frac{d^{r-1}}{dt^{r-1}}h(t)\right) + 0.$$

Hence,

$$\begin{aligned}
\mathrm{E}(X_{[r]}) &= \left(\frac{d^r}{dt^r}g(t)\right)\bigg|_{t=1} \\
&= r\left(\frac{d^{r-1}}{dt^{r-1}}h(t)\right)\bigg|_{t=1} \\
&= r\left(\frac{d^{r-1}}{dt^{r-1}}\sum_{n=0}^{\infty} P(X>n)t^n\right)\bigg|_{t=1} \\
&= r\sum_{n=0}^{\infty} P(X>n)\left(\frac{d^{r-1}}{dt^{r-1}}t^n\right)\bigg|_{t=1} \\
&= r\sum_{n=0}^{\infty} P(X>n)n_{[r-1]}.
\end{aligned}$$

Theorem 39 (Raw Moments in Terms of Cumulative Distribution)

Let X be a discrete random variable defined on $\mathbb{S} \subseteq \{0,1,\ldots\}$. Then

$$\mathrm{E}(X^r) = \sum_{n=0}^{\infty} \Big((n+1)^r - n^r\Big) P(X > n).$$

In the special case $r = 1$ we have the familiar result

$$\mu = \sum_{n=0}^{\infty} P(X > n).$$

10 Probability Distributions: Definitions and Properties

Negative Binomial

•	$P(X = x) = \binom{x+m-1}{m-1}\theta^m(1-\theta)^x \mathbb{I}(x \in \{0,1,\dots\})$ where $m \in \{1,2,\dots\}$ and $0 < \theta < 1$.
•	We write this as "$X \sim$ negative binomial(m,θ)".
•	$\mathrm{E}(X_{[r]}) = \left(\frac{1-\theta}{\theta}\right)^r m^{[r]}$
•	If $X_j \sim$ negative binomial(m_j,θ), $j = 1,2$ and if X_1 and X_2 are independent, then $(X_1 + X_2) \sim$ negative binomial$(m_1 + m_2, \theta)$.

Geometric

•	Special case $m = 1$ of the negative binomial distribution.

α – shifted negative binomial

•	$P(X = x) = \binom{m+(x-\alpha)-1}{m-1}\theta^m(1-\theta)^{x-\alpha} \mathbb{I}(x \in \{\alpha,\alpha+1,\dots\})$ where $m \in \{1,2,\dots\}$, $\alpha \in \{0,1,2,\dots\}$ and $0 < \theta < 1$.
•	We write this as "$X \sim \alpha -$ shifted negative binomial(m,θ)."
•	$\mathrm{E}((X-\alpha)_{[r]}) = \left(\frac{1-\theta}{\theta}\right)^r m^{[r]}$
•	If $X \sim$negative binomial(m,θ) then $(X+\alpha) \sim \alpha -$ shifted negative binomial(m,θ).
•	If $X_j \sim \alpha -$ shifted negative binomial(m_j,θ), $j = 1,2$ and if X_1 and X_2 are independent, then $(X_1 + X_2) \sim (\alpha_1 + \alpha_2) -$ shifted negative binomial$(m_1 + m_2, \theta)$.

Poisson

•	$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}\mathbb{I}(x\in\{0,1,\dots\}), \lambda>0.$
•	We write this as "X~Poisson(λ)."
•	$\mathrm{E}(X_{[r]})=\lambda^r$
•	If X_j~ Poisson (λ_j), $j=1,2$ and if X_1 and X_2 are independent, then $(X_1+X_2)\sim$ Poisson$(\lambda_1+\lambda_2)$.

Dirichlet

•	$f(x_1,x_2,\dots,x_n)=\frac{\Gamma(m_1+m_2+\cdots+m_n)}{\Gamma(m_1)\cdots\Gamma(m_n)}x_1^{m_1-1}\cdots x_n^{m_n-1},$
	for all points $(x_1,x_2,\dots,x_n)$ such that $x_1+x_2+\cdots+x_n=1$, $0\le x_j\le 1$ for $j=1,2,\dots,n,$
	and $m_j\in\{\quad\}, j=1,2,\dots,n.$
•	We write this as "$(X_1,\dots,X_n)\sim$ Dirichlet$(m_1,\dots,m_n)$".
•	$\mathrm{E}(X_1^{r_1}X_2^{r_2}\cdots X_n^{r_n})=\frac{m_1^{[r_1]}m_2^{[r_2]}\cdots m_n^{[r_n]}}{(m_1+m_2+\cdots+m_n)^{[r_1+r_2+\cdots+r_n]}}$
•	(Aggregation Property) Let $(X_1,\dots,X_n)\sim$ Dirichlet$(m_1,\dots,m_n)$ and let sets $S_1,\dots,S_r$ be a
	partition of $\{1,2,\dots,n\}$, then
	$\left(\sum_{i\in S_1}X_i,\dots,\sum_{i\in S_r}X_i\right)\sim$ Dirichlet$\left(\sum_{i\in S_1}m_i,\dots,\sum_{i\in S_r}m_i\right)$

Gamma

•	$$f_X(x) = \frac{\lambda^m}{\Gamma(m)} x^{m-1} \mathrm{e}^{-\lambda x} \mathbb{I}(x > 0), \lambda > 0$$
	where $m \in \{\quad\}$.
•	We write this as "$X \sim \text{Gamma}(m, \lambda)$".
•	$\mathrm{E}(X^r) = \dfrac{\Gamma(m+r)}{\lambda^r \Gamma(m)}$
•	If $X_j \sim \text{Gamma}\big(m_j, \lambda\big)$, $j = 1,2$ and if X_1 and X_2 are independent, then $(X_1 + X_2) \sim \text{Gamma}(m_1 + m_2, \lambda)$.

Exponential

•	Special case $m = 1$ of the gamma distribution.
•	If $X_1, X_2, \ldots, X_n$ are independent with $X_j \sim$ exponential(λ), then $X_{(1:n)} \sim$ exponential$(n\lambda)$.

11 References

[1] J. Ahuja and E. Enneking. Convolution of independent left-truncated negative binomial variables and limiting distributions. Ann. Inst. Statist. Math., 26:265–270, 1974.

[2] J. Ahuja. Distribution of the sum of independent decapitated negative binomial variables. Ann. Math. Stat., 42:383–384, 1971.

[3] R. Arratia and S. Tavaré. The cycle structure of random permutations. Ann. Probab., 20(3):1567–1591, 1992.

[4] R. Arratia and S. Tavaré. Independent process approximations for random combinatorial structures. Adv. Math., 104(1):90–154, 1994.

[5] K. Baclawski, M. Cerasoli, and G. Rota. Introduzione Alla Probabiltà. Unione Mathematica Italiana, 1984.

[6] K. Baclawski, G. Rota, and S. Billey. An Introduction to the Theory of Probability. Department of Mathematics, MIT, second preliminary edition, 1989.

[7] P. Billingsley. Probability and Measure. John Wiley & Sons, 1979.

[8] M. L. Bizley. Probability: An Intermediate Text-Book. Cambridge University Press, 1964.

[9] D. Blackwell and J. MacQueen. Ferguson distributions via polya urn schemes. The Annals of Statistics,Vol. 1, No. 2 (Mar., 1973), pp. 353-355, 1(2):353–355, 1973.

[10] J. Buoncristiani and M. Cerasoli. Multidimensional occupancy problems with Poisson randomization. J. Math. Anal. Appl., 104:526–536, 1984.

[11] G. Casella and R. Berger. Statistical Inference. Duxbury, 2nd edition, 2002.

[12] M. Cerasoli. Poisson randomization in occupancy problems. J. Math. Anal. Appl., 94:150–165, 1983.

[13] C. Charalambides and M. Koutras. On the differences of the generalized factorials at an arbitrary point and their combinatorial applications. Discrete Math., 47:183–201, 1983.

[14] C. Charalambides. A unified derivation of occupancy and sequential occupancy distributions. In Advances in Combinatorial Methods and Applications to Probability and Statistics, pages 259–273. Birkhauser, 1997.

[15] J. Chen. Weighing the heaviest Polya urn. Operations Research Letters, 42(6-7):394–397, September 2014.

[16] S. Corteel, B. Pittel, C. Savage, and H. Wilf. On the multiplicity of parts in a random partition. Random Structures and Algorithms, 14:1999, 185-197 1999.

[17] H. A. David and H. N. Nagaraja. Order Statiistics. Wiley, 3rd edition, 2003.

[18] C. Domb. On the use of a random parameter in combinatorial problems. Proc. Phys. Soc. Sect. A., 65:305–309, 1952.

[19] M. Dwass. The distribution of linear combinations of random divisions of an interval. Trabajos de Estadistica, 12(1-2):11–17, 1961.

[20] M. Dwass. Simple random walk and rank order statistics. Ann. Math. Statist., 38:1042–1053, 1967.

[21] W. Feller. An Introduction to Probability Theory and its Applications, volume 1. John Wiley & Sons, 1950.

[22] B. Fristedt. The structure of random partitions of large integers. Trans. Amer. Math. Soc., 337:703–735, 1993.

[23] T. Gerstenkom and J. Jarzebska. The inverse Pólya distribution and its moments derived by the compounding operation. In Statistics and Probability (Visegrád, 1982), pages 73–87. Reidel, Dordrecht, 1984.

[24] W. Goh and E. Schmutz. The number of different part sizes in random integer partition. J. Combin. Theory Ser. A, 69:149–158, 1995.

[25] R. L. Graham, D. E. Knuth, and O. Patashnik. Concrete Mathematics: A Foundation for Computer Science. Addison-Wesley, 1994.

[26] J. Hansen. A functional central limit theorem for random mappings. Ann. Probab., 17(1):317–332, 1989.

[27] J. Hansen. A functional central limit theorem for the Ewen's sampling formula. J. Appl. Probab., 27:28–43, 1990.

[28] J. Hansen. Order statistics for decomposable combinatorial structures. Random Structures Algorithms, 5(4):517–533, 1994.

[29] W. Hardy. A recursive method of solution for occupancy problems. SIAM J. Appl. Math., 17(1):192–197, 1969.

[30] P. Hitczenko and C. Savage. On the multiplicity of parts in a random composition of a large integer. SIAM J. Discrete Math., 18(2):418–435, 2004.

[31] P. Hitczenko and G. Stengle. Expected number of distinct part sizes in a random integer composition. Comb. Probab. Comput., 9:519–527, 2000.

[32] T. Huillet and S. Martinez. Dirichlet-Kingman partition revisited. East Journal of Theoretical Statistics, 24(1):1–33, 2008.

[33] T. Huillet. Random covering of the circle: The size of the connected components. Advances in Applied Probability, 35(3):563–582, 2003.

[34] S. Mohanty. Lattice Path Counting and Applications. Academic Press, 1979.

[35] A. M. Mood, F. A. Graybill, and D. C. Boes. Introduction to the Theory of Stati. McGraw-Hill, 3rd edition, 1974.

[36] J. Mosimann. On the compound negative multinomial distribution and correlations among inversely sampled pollen counts. Biometrika, 49:65–82, 1962.

[37] S. Nandi and S. Dutta. On associated and generalized Lah numbers and applications to discrete distributions. Fibonacci Quart., 25(2):128–136, 1987.

[38] W. Schelter. Maxima, a Computer Algebra System. http://maxima.sourceforge.net/. Version 5.28.0-2.

[39] E. Schrödinger. A combinatorial problem in counting cosmic rays. Proc. Phys. Soc. Sect. A, 64:1040–1041, 1951.

[40] L. Shepp and S. Lloyd. Ordered cycle lengths in a random permutation. Trans. Amer. Math. Soc., 121(2):340–357, 1966.

[41] J. Skellam. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. J. Roy. Statist. Soc. Ser. B., 10:257–261, 1948.

[42] K. S. Srikantan. Recurrence relations between the pdf's of order statistics and some applications. Annals of Mathematical Statistics, 33:166–177, 1962.

[43] F. W. Steutel. Random division of an interval. Statistica Neerlandica 21 (1967) nr. 3 en 4, 21(3-4):231–244, 1967.

[44] K. Suman. Functions of ordered numbers, order statistics and the principle of inclusion-exclusion. Available on request.

[45] J. W. Tukey. Moments of random group size distributions. Ann. Math. Stat., 20(4):523–539, 1949.

[46] M. Tweedie. Further results concerning expectation-inversion technique. In Classical and Contagious Discrete Distributions (Proc. Internat. Sympos., McGill Univ., Montreal, Que., 1963), pages 195–218, Calcutta, 1965. Statistical Publishing Society.

[47] T. Wallstrom. The equalization probability of the Pólya urn. The American Mathematical Monthly, 119(6):516–518, 2012.

[48] G. Watterson. Models for the logarithmic species abundance distributions. Theoret. Population Biology, 6:217–250, 1974.

[49] G. Watterson. The sampling theory of selectively neutral alleles. Advances in Appl. Probability, 6:463–488, 1974.

[50] H. S. Wilf. generatingfunctionology. A. K. Peters, 2006.

[51] H. Wilf. Three problems in combinatorial asymptotics. J. Combin. Theory Ser. A, 35:199–207, 1983.

[52] S. S. Wilks. A combinatorial test for the problem of two samples from continuous distributions. In Proc. Fourth Berkeley Symp. on Math. Stat. and Prob., 1, pages 707–717. University of California Press, 1960.

[53] S. S. Wilks. Mathematical Statistics. John Wiley & Sons, New York, 1962.

[54] Wolfram Alpha LLC. 2015. Wolfram Alpha. https://www.wolframalpha.com, 2015. (accessed July 17, 2015).