International Journal of Statistics and Probability; Vol. 8, No. 6; November 2019
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Partitioning Problems Arising From Independent Shifted-Geometric and Exponential Samples With Unequal Intensities

Thierry E. Huillet

Correspondence: Thierry E. Huillet, Laboratoire de Physique Théorique et Modélisation Université de Cergy-Pontoise CNRS UMR-8089 Site de Saint Martin 2 avenue Adolphe-Chauvin 95302 Cergy-Pontoise, France.
E-mail: Thierry.Huillet@u-cergy.fr

Received: September 2, 2019 Accepted: October 9, 2019 Online Published: October 13, 2019
doi:10.5539/ijsp.v8n6p31 URL: https://doi.org/10.5539/ijsp.v8n6p31

Abstract

Two problems dealing with the random skewed splitting of some population into J different types are considered.

In a first discrete setup, the sizes of the sub-populations come from independent shifted-geometric with unequal characteristics. Various $J \to \infty$ asymptotics of the induced occupancies are investigated: the total population size, the number of unfilled types, the index of consecutive filled types, the maximum number of individuals in some state and the index of the type(s) achieving this maximum. Equivalently, this problem is amenable to the classical one of assigning indistinguishable particles (Bosons) at J sites, in some random allocation problem.

In a second parallel setup in the continuum, we consider a large population of say J 'stars', the intensities of which have independent exponential distributions with unequal inverse temperatures. Stars are being observed only if their intensities exceed some threshold value. Depending on the choice of the inverse temperatures, we investigate the energy partitioning among stars, the total energy emitted by the observed stars, the number of the observable stars and the energy and index of the star emitting the most.

Keywords: sum and maximum, independent shifted-geometric/exponential distributions, discrete/continuous partitioning, combinatorial probability

1. Introduction

Consider the partitioning of some population the individuals of which can be of J different types or states. We assume that the sizes of the type-j sub-populations ($j = 1, ..., J$) have independent shifted-geometric distributions with unequal success probabilities. Depending on these probabilities, we envisage various asymptotics for the occupancy distributions, including total population size and the number of unfilled states. Other statistical quantities of interest such as: the index of the consecutive filled states, the maximum number of particles in some state and the index of the site(s) achieving this maximum are also investigated. One of the asymptotics we chiefly focus on is $J \to \infty$.

A toy variant of the latter model is also investigated in the continuum which is shown to be amenable to a quite similar treatment; it deals with a population of say J 'stars' the intensities of which have independent exponential distributions with unequal inverse temperatures that can be observed or not depending on whether the intensities exceed or not some threshold value. Of parallel interest then is the energy partitioning among stars, the total energy emitted by the observed stars, the number of the observable stars, the energy and index of the star emitting the most. And the way all these quantities depend on the choice of the inverse temperatures. Some examples are detailed and the limit $J \to \infty$ is also investigated in this context. This second aspect of the partitioning problem in the continuum seems to be new.

Let us summarize our results and sketch the organization of the manuscript: motivated by examples from physics, we have studied specific partitioning problems, thereby contributing to general probability theory and discrete mathematics. Considering a population with J different sub-populations (or states) whose sizes G_j are independent (shifted-)geometrically distributed, in general with different parameters, we have studied various distributions of interest under several asymptotic regimes. First, in Section 2.1, the distribution of total population size $X_J = \sum_{j=1}^{J} G_j$, and the joint distributions of relative population sizes $G_j/\mathbf{E}(X_J)$ were investigated, also asymptotically, the latter e.g. for J fixed and large population size. A condensation phenomenon was highlighted. Section 2.2 considers the number of non-empty states and constrained occupancies problems. In Section 2.3 asymptotics where $J \to \infty$ are discussed, with a $\{0-1\}$-law distinguishing if the series of parameters converges or not. Further, the first non-empty state and site indices till consecutive records (Section 2.4) and the size and index of the most filled state (Sections 2.5 and 2.6) have been addressed. Section 2.7 gives some concrete examples of parameter sequences. Section 3 transfers the theory to the continuous analog with independent (shifted-)exponential distributions with an illustrative (non-exhaustive) physical image from astronomy.

2. The Discrete Model: A Skewed Bose Occupancy Problem

Let J be some integer denoting the number of species (or types) of some population. It can also be the number of sites at which indistinguishable particles (Bosons) will be assigned to, in some random allocation problem, (Kolchin, 1986). Let $\alpha_j \in (0,1)$ and $\overline{\alpha}_j = 1-\alpha_j$, $j = 1, ..., J$. Consider independent geometric random variables (rv's) G_j^+ each with success parameter $\overline{\alpha}_j$ and let $G_j = G_j^+ - 1$. We will rather deal with this shifted geometric rv G_j, standing for the number of particles assigned to site j (the number of individuals of type j, possibly 0) in some grand-canonical allocation problem. Thus $G_j^+ = 1 + G_j$, which is also $G_j \mid G_j \geq 1$, will stand for the number of particles assigned to site j given this site is occupied. In such an occupancy model, the total number of particles in the system, namely $X_J = \sum_{j=1}^J G_j$, is random. And the model allows for unoccupied states. A related partitioning model with this opportunity was recently developed in (Huillet, 2018). The unskewed case where $\alpha_j = \alpha$, $j = 1, ..., J$, with the G_j's independent and identically distributed (iid) is a simpler issue that will also be briefly dealt with in passing. The simpler sum $\sum_{j=1}^J G_j^+$ was considered in (Sen & Balakrishnan, 1999) in a reliability problem.

2.1 Skewed Occupancy Distributions

For any power series $\phi(z) = \sum_{i\geq 1} \phi_i z^i$, we shall denote $\left[z^i\right]\phi(z) = \phi_i$, the z^i−coefficient of $\phi(z)$.

The joint probability generating function (pgf) of the G_j's is:

$$\mathbf{E}\left(\prod_{j=1}^J z_j^{G_j}\right) = \prod_{j=1}^J \frac{1-\alpha_j}{1-\alpha_j z_j},\ z_j \in [0,1],$$

so that for any non-negative integers $i_1, ..., i_J$, the species occupancy distribution reads

$$\begin{gathered}\mathbf{P}(G_1 = i_1, ..., G_J = i_J) = \\ \left[z_1^{i_1} ... z_J^{i_J}\right] \prod_{j=1}^J \mathbf{E}\left(z_j^{G_j}\right) = \prod_{j=1}^J \left(1-\alpha_j\right) \prod_{j=1}^J \alpha_j^{i_j}.\end{gathered} \tag{1}$$

• A special asymptotic regime:

Let $g_j := \mathbf{E}\left(G_j\right) = \alpha_j / \left(1-\alpha_j\right)$ (:= means equals by definition) and $x_J := \mathbf{E}(X_J) = \sum_{j=1}^J g_j$, the mean number of particles in the system. Suppose $\alpha_j = \delta_j / \left(\varepsilon + \delta_j\right)$ for some sequence $\delta_j > 0$ and $\varepsilon > 0$ small, in such a way that $\alpha_j \to 1^-$ as $\varepsilon \to 0$ for all $j = 1, ..., J$. Then both $g_j = \varepsilon^{-1}\delta_j$ and $x_J = \varepsilon^{-1}\sum_{j=1}^J \delta_j \to \infty$ as $\varepsilon \to 0$ while $g_j / x_J = \delta_j / \sum_{j=1}^J \delta_j$ is well-defined and independent of ε. For this particular choice of α_j, when ε approaches 0 ($x_J \to \infty$), for all fixed J, we have

$$\begin{aligned}\mathbf{E}\left(\prod_{j=1}^J z_j^{G_j/x_J}\right) &= \prod_{j=1}^J \frac{1-\alpha_j}{1-\alpha_j z_j^{1/x_J}} \sim \prod_{j=1}^J \frac{1-\alpha_j}{1-\alpha_j - \frac{\alpha_j}{x_J}\log z_j} \\ &= \prod_{j=1}^J \left(1 - \frac{g_j}{x_J}\log z_j\right)^{-1} \sim \prod_{j=1}^J \left(1 + \frac{g_j}{x_J}\log z_j\right) \\ &\sim \prod_{j=1}^J e^{\left(g_j/x_J\right)\log z_j} = \prod_{j=1}^J z_j^{g_j/x_J}.\end{aligned}$$

In the above, $\sim$ means that the ratio of the quantities appearing to the left and right of this symbol tends to 1 as $x_J \to \infty$, to the dominant order.

proposition 1. *Under the above conditions on the g_j's,*

$$x_J^{-1}(G_1, ..., G_J) \underset{x_J\to\infty}{\overset{\mathbf{P}}{\to}} (g_1/x_J, ..., g_J/x_J). \tag{2}$$

The pgf of X_J itself is

$$\phi_J(z) := \mathbf{E}\left(z^{X_J}\right) = \prod_{j=1}^J \frac{1-\alpha_j}{1-\alpha_j z},\ z \in [0,1].$$

As such, the rv X_J also interprets as the number of successes before the $J-$th failure in an inhomogeneous Bernoulli process, with probabilities $1-\alpha_i$ of successes on each trial (as an extended version of the negative binomial pgf).

When dealing with the joint distribution of $\left(X_J, G_j; j=1,...,J\right)$, we have

$$\begin{aligned}
\mathbf{E}\left(z^{X_J}\prod_{j=1}^{J} z_j^{G_j}\right) &= \prod_{j=1}^{J}\frac{1-\alpha_j}{1-\alpha_j z z_j}\\
\mathbf{E}\left(\prod_{j=1}^{J} z_j^{G_j}\mid X_J=i\right) &= \frac{1}{\mathbf{P}\left(X_J=i\right)}\left[z^i\right]\prod_{j=1}^{J}\frac{1-\alpha_j}{1-\alpha_j z z_j}\\
&= \frac{\left[z^i\right]\prod_{j=1}^{J}\left(1-\alpha_j z z_j\right)^{-1}}{\left[z^i\right]\prod_{j=1}^{J}\left(1-\alpha_j z\right)^{-1}}.
\end{aligned}$$

As a result, for any non-negative integers $i_1,...,i_J$, summing to i, with

$$Z_{J,i} := \left[z^i\right]\prod_{j=1}^{J}\left(1-\alpha_j z\right)^{-1} = \sum_{\substack{\sum_{j=1}^{J} i_j = i\\ i_j\geq 0,\ j=1,...,J}}\prod_{j=1}^{J}\alpha_j^{i_j},$$

an ordinary Bell polynomial (Comtet, 1970),

$$\mathbf{P}\left(G_1=i_1,...,G_J=i_J\mid X_J=i\right)=\prod_{j=1}^{J}\alpha_j^{i_j}/Z_{J,i} \tag{3}$$

also gives the canonical occupancy distribution given the total number of particles is i. Note that defining $s_j := \alpha_j/\sum_{j=1}^{J}\alpha_j$, $j=1,...,J$, summing to 1, with

$$\overline{Z}_{J,i} = Z_{J,i}/\left(\sum_{j=1}^{J}\alpha_j\right)^i$$

$$\mathbf{P}\left(G_1=i_1,...,G_J=i_J\mid X_J=i\right)=\prod_{j=1}^{J}s_j^{i_j}/\overline{Z}_{J,i},$$

interpreting as a $i-$Bose sampling procedure of the unit interval $[0,1]$ split into J parts of unequal sizes s_j.

• A condensation phenomenon:

Suppose $\alpha_1>\alpha_2\geq ...\geq\alpha_J$. Let us show that, when the number of types J is held fixed, the particles tend to concentrate on the ground state (which is the unique type with largest α_j, here chosen without loss of generality (w.l.o.g) as α_1), when the number of particles i increases. We denote $\left(G_1,...,G_J\mid X_J=i\right)=:\left(G_j\left(i\right);\ j=1,...,J\right)$

proposition 2. *With J kept fixed, as the number of particles i grows, we have the convergence in distribution*

$$\left(G_2\left(i\right),...,G_J\left(i\right)\right)\overset{d}{\underset{i\to\infty}{\to}}\left(\mathcal{G}_2,...,\mathcal{G}_J\right) \tag{4}$$

where the $\mathcal{G}_j$s are independent rv's which are geometrically distributed with success parameters $1-\alpha_j/\alpha_1,\ j=2,...,J$. Consequently,

$$\left(G_1\left(i\right)/i,...,G_J\left(i\right)/i\right)\overset{d}{\underset{i\uparrow\infty}{\to}}\left(1,0,...,0\right). \tag{5}$$

Proof: Let $[J]:=\{1,...,J\}$. Developing the product partition function $Z_{J,i}$ into a sum of J rational fractions, extracting its coefficient of z^i, we easily get (after obvious identification of the coefficients)

$$Z_{J,i}=\sum_{j=1}^{J}W_j\alpha_j^i \text{ where } W_j := \prod_{k\in[J]\setminus\{j\}}\left(1-\frac{\alpha_k}{\alpha_j}\right)^{-1}.$$

Isolate the ground state term and factorize α_1. Then

$$Z_{J,i} = \alpha_1^i \left(W_1 + \sum_{j=2}^{J} W_j a_j^i \right)$$

where $a_j := \alpha_j / \alpha_1 < 1$, $j = 2, ..., J$. With $i_1 + ... + i_J = i$, we want to estimate the joint law of the occupancies $G_j(i)$ which is

$$\mathbf{P}\left(G_j(i) = i_j;\ j = 1, ..., J\right) = \frac{1}{Z_{J,i}} \prod_{j=1}^{J} \alpha_j^{i_j}.$$

Since $i_1 = i - (i_2 + ... + i_J)$, using the expression of $Z_{J,i}$, for all i_j, $j = 2, ..., J$, obeying $\sum_{j=2}^{J} i_j < i$, the joint occupancy distribution of all states but the ground state reads

$$\mathbf{P}\left(G_j(i) = i_j;\ j = 2, ..., J\right) = \frac{\prod_{j=2}^{J} a_j^{i_j}}{W_1 \left(1 + \sum_{j=2}^{J} \frac{W_j}{W_1} a_j^i\right)}$$

$$= \frac{\prod_{j=2}^{J} a_j^{i_j} \left(1 - a_j\right)}{1 + \sum_{j=2}^{J} \frac{W_j}{W_1} a_j^i}, \text{ while using } W_1 = \prod_{k=2}^{J} \left(1 - a_j\right)^{-1}.$$

The term $\sum_{j=2}^{J} \frac{W_j}{W_1} a_j^i$ goes to 0 exponentially fast with i getting large and, since $a_J < ... < a_3 < a_2 < 1$, it has the dominant term $\varepsilon(i) := a_2^i W_2 / W_1 < 0$. As $i \to \infty$, we therefore expect

$$\mathbf{P}\left(G_j(i) = i_j;\ j = 2, ..., J\right) \sim (1 - \varepsilon(i)) \prod_{j=2}^{J} \left\{a_j^{i_j} \left(1 - a_j\right)\right\}. \tag{6}$$

When i gets large therefore, a good approximation of the joint occupancies of all ordered states but the ground state is the one of geometrically distributed finite random variables with normalized success probabilities $1 - a_j$. In other words, the probabilities of $G_j(i)/i$; $j = 2, ..., J$ all concentrate to 0 and therefore all the probability mass goes to the ground state ($j = 1$). This is the content of the statement displayed in Eq. (5). This fact is reminiscent of the Bose-Einstein condensation. □

Coming back to X_J, whenever the α_js are all distinct, the decomposition of $\phi_J(z)$ into simple fractions also gives it as a weighted sum of the elementary geometrics with negative or positive weights w_j summing to 1:

$$\phi_J(z) = \sum_{j=1}^{J} w_j \frac{\overline{\alpha}_j}{1 - \alpha_j z},\ w_j = \prod_{k \in [J] \setminus \{j\}} \frac{\overline{\alpha}_k \alpha_j}{\alpha_j - \alpha_k},$$

not a probability mixture. If the α_js are sorted in descending order, the w_j alternate in sign starting with $w_1 > 0$.

While expanding directly the product giving $\phi_J(z)$, we get the odd explicit expression

$$\mathbf{P}(X_J = i) = \left[z^i\right] \phi_J(z) = Z_{J,i} \cdot \prod_{j=1}^{J} \left(1 - \alpha_j\right).$$

The process $(X_J; J \geq 0)$, with $X_0 = 0$, has the structure of a process with independent and non-stationary increments so that

$$\mathbf{P}(X_J = i) = \sum_{j=0}^{i} \mathbf{P}(X_{J-1} = j)\, \mathbf{P}(G_J = i - j).$$

Hopefully however, by recurrence it holds:

proposition 3. *The probability mass function (pmf)* $\mathbf{P}(X_J = i)$ *can be obtained by the recurrence*

$$\mathbf{P}(X_J = i) = \alpha_J \mathbf{P}(X_J = i - 1) + (1 - \alpha_J)\, \mathbf{P}(X_{J-1} = i),\ i, J \geq 1 \tag{7}$$

with boundary conditions: $\mathbf{P}(X_J = 0) = \prod_{j=1}^{J} \left(1 - \alpha_j\right)$ *for all* $J \geq 1$ *and* $\mathbf{P}(X_0 = i) = 0$ *for all* $i \geq 1$.

This can be useful to recursively generate such probabilities on a lap-top because to produce $\mathbf{P}(X_J = i)$ only $\mathbf{P}(X_J = i-1)$ and $\mathbf{P}(X_{J-1} = i)$ are needed from previous computations and not the whole sequence $\mathbf{P}(X_{J-1} = j)$, $j = 0, ..., i$.

Let $[J]_i := \Gamma(J+i)/\Gamma(J)$ and $[{J \atop i}] := \frac{[J]_i}{i!} := \frac{(J+i-1)!}{(J-1)!i!}$. If $\alpha_j = \alpha$ for all j (the iid case), $\mathbf{P}(X_J = i) = (1-\alpha)^J \alpha^i [{J \atop i}]$ (the negative binomial probability mass function) obeys the recurrence $\mathbf{P}(X_J = i) = \alpha \mathbf{P}(X_J = i-1) + (1-\alpha)\mathbf{P}(X_{J-1} = i)$, as a result of $[{J \atop i}] = [{J \atop i-1}] + [{J-1 \atop i}]$.

Note that, as a result of $\mathbf{P}(X_J = i) = Z_{J,i} \cdot \prod_{j=1}^{J} \left(1-\alpha_j\right)$, $Z_{J,i}$itself obeys the recurrence $Z_{J,i} = \alpha_J Z_{J,i-1} + Z_{J-1,i}$, $i, J \geq 1$, with boundary conditions $Z_{J,0} = 1$ for all $J \geq 1$ and $Z_{0,i} = 0$ for all $i \geq 1$.

2.2 Number of Filled States and Constrained Occupancies

Let $B_j = \mathbf{1}\left(G_j > 0\right)$ indicate whether site j is filled or free of particles (species j has or not at least one representative in the population). It is a Bernoulli$\left(\alpha_j\right)$rv, so with $\mathbf{P}\left(B_j = 1\right) = \alpha_j$.The total number of filled sites in the system is $P_J = \sum_{j=1}^{J} B_j$, with pgf

$$\varphi_J(z_0) := \mathbf{E}\left(z_0^{P_J}\right) = \prod_{j=1}^{J} \left(1 - \alpha_j (1 - z_0)\right).$$

With $g_j = \mathbf{E}\left(G_j\right) = \alpha_j / \left(1-\alpha_j\right) > 0$, for all $p \in \{0, ..., J\}$, we have the odd Fermi-Dirac like expressions

$$\begin{aligned}
\mathbf{P}(P_J = p) &= \left[z_0^p\right] \varphi_J(z_0) = \prod_{j=1}^{J} \left(1-\alpha_j\right) \cdot \sum_{1 \leq j_1 < ... < j_p \leq J} \prod_{q=1}^{p} g_{j_q} \\
&= \prod_{j=1}^{J} \left(1-\alpha_j\right) \cdot \sum_{\substack{\sum_{j=1}^{J} i_j = p \\ i_j \in \{0,1\},\ j=1,...,J}} \prod_{j=1}^{J} g_j^{i_j}.
\end{aligned}$$

proposition 4. *The probabilities $\mathbf{P}(P_J = p)$, $p = 0, ..., J$, can be generated by using the recurrence*

$$\mathbf{P}(P_J = p) = \alpha_J \mathbf{P}(P_{J-1} = p-1) + (1-\alpha_J)\mathbf{P}(P_{J-1} = p),\ J \geq p \geq 1, \tag{8}$$

with boundary conditions: $\mathbf{P}(P_J = 0) = \prod_{j=1}^{J} \left(1-\alpha_j\right)$ *for all* $J \geq 1$ *and* $\mathbf{P}(P_0 = p) = \delta_{p,0}$, $p \geq 0$.

This translates the fact that $(P_J; J \geq 0)$, with $P_0 \overset{d}{\sim} \delta_0$, has the structure of a time-inhomogeneous Markov chain with one-step transition matrix from step $j-1$ to j, $j \geq 1$, say $\mathcal{P}_{j-1,j} = \left[\mathcal{P}_{j-1,j}(p,q)\right]$, $p, q \geq 0$, given by:

$$\mathcal{P}_{j-1,j}(p,p) = 1-\alpha_j, \mathcal{P}_{j-1}(p,p+1) = \alpha_j \text{ and } \mathcal{P}_{j-1}(p,q) = 0 \text{ if } q \neq \{p, p+1\}.$$

Note that if $\alpha_j = \alpha$ for all j (the iid case), $\mathbf{P}(P_J = p) = \binom{J}{p}\alpha^p (1-\alpha)^{J-p}$ (the binomial distribution), known to satisfy the recurrence $\mathbf{P}(P_J = p) = \alpha \mathbf{P}(P_{J-1} = p-1) + (1-\alpha)\mathbf{P}(P_{J-1} = p)$ as a result of $\binom{J}{p} = \binom{J-1}{p-1} + \binom{J-1}{p}$, (Pascal's triangle identity for binomial coefficients).

We note that

$$\mathbf{E}(P_J) = \sum_{j=1}^{J} \alpha_j \text{ and } \sigma^2(P_J) = \sum_{j=1}^{J} \alpha_j \left(1-\alpha_j\right) = \mathbf{E}(P_J) - \sum_{j=1}^{J} \alpha_j^2.$$

while

$$\mathbf{E}(X_J) = \sum_{j=1}^{J} g_j = \sum_{j=1}^{J} \frac{\alpha_j}{1-\alpha_j} \text{ and } \sigma^2(X_J) = \sum_{j=1}^{J} \frac{\alpha_j}{\left(1-\alpha_j\right)^2} = \mathbf{E}(X_J) + \sum_{j=1}^{J} g_j^2.$$

We also have

$$\mathbf{E}\left(z_0^{B_j} z^{B_j G_j}\right) = 1 - \alpha_j + \alpha_j z_0 \mathbf{E}\left(z^{G_j^+}\right) = 1 - \alpha_j + \alpha_j z_0 \frac{z\left(1-\alpha_j\right)}{1-\alpha_j z}$$

so that

$$\begin{aligned}
\mathbf{E}\left(z_0^{P_J} \prod_{j=1}^{J} z_j^{B_j G_j}\right) &= \prod_{j=1}^{J} \left(1 - \alpha_j + \alpha_j z_0 \frac{z_j\left(1-\alpha_j\right)}{1-\alpha_j z_j}\right) \\
&= \prod_{j=1}^{J} \left(1-\alpha_j\right) \prod_{j=1}^{J} \left(1 + \alpha_j z_0 \frac{z_j}{1-\alpha_j z_j}\right).
\end{aligned}$$

Plugging $z_j = z$, $j = 1, ..., J$, observing $X_J = \sum_{j=1}^{J} B_j G_j$, we get

proposition 5.

$$\mathbf{E}\left(z_0^{P_J} z^{X_J}\right) = \prod_{j=1}^{J} \left(1 - \alpha_j + \alpha_j z_0 \frac{z\left(1-\alpha_j\right)}{1-\alpha_j z}\right) = \prod_{j=1}^{J} \left(1-\alpha_j\right) \prod_{j=1}^{J} \frac{1-\alpha_j z\left(1-z_0\right)}{1-\alpha_j z} \tag{9}$$

gives the joint pgf of (P_J, X_J).

For all fixed sequence $1 \leq j_1 < ... < j_p \leq J$ *of sites and all effective occupancies* $i_1, ..., i_p \geq 1$ *of these sites,*

$$\mathbf{P}\left(G_{j_1} = i_1, ..., G_{j_p} = i_p; P_J = p\right) =$$

$$\left[z_{j_1}^{i_1} ... z_{j_p}^{i_p}\right] \prod_{q=1}^{p} \mathbf{E}\left(z_{j_q}^{G_{j_q}^{+}}\right) = \prod_{j=1}^{J} \left(1-\alpha_j\right) \prod_{q=1}^{p} \alpha_{j_q}^{i_q - 1}.$$

Remark: Considering the joint probability

$$\mathbf{P}\left(X_J = i; P_J = p\right) = \left[z_0^p z^i\right] \mathbf{E}\left(z_0^{P_J} z^{X_J}\right),$$

we get the conditional probabilities

$$\mathbf{P}\left(X_J = i \mid P_J = p\right) = \frac{\left[z_0^p z^i\right] \mathbf{E}\left(z_0^{P_J} z^{X_J}\right)}{\left[z_0^p\right] \mathbf{E}\left(z_0^{P_J}\right)}$$

$$\mathbf{P}\left(P_J = p \mid X_J = i\right) = \frac{\left[z_0^p z^i\right] \mathbf{E}\left(z_0^{P_J} z^{X_J}\right)}{\left[z^i\right] \mathbf{E}\left(z^{X_J}\right)}.$$

2.3 Letting the Number of Types $J \to \infty$

We now run into the infinitely many species (types, sites) problem: $J \to \infty$.

proposition 6. *As $J \to \infty$, consider $X := \sum_{j=1}^{\infty} G_j$. The rv X is either ∞ with probability 1 or it is $< \infty$ with probability 1.*

Proof: (Hewitt-Savage $\{0, 1\}$ –law) Indeed, suppose $X = \infty$ ($X < \infty$) with probability $1 - p$ (p). Then the same holds true with $X^{Odd} = \sum_{j=0}^{\infty} G_{2j+1}$ and $X^{Even} = \sum_{j=1}^{\infty} G_{2j}$. But then

$$p = \mathbf{P}(X < \infty) = \mathbf{P}\left(X^{Odd} < \infty\right) \mathbf{P}\left(X^{Even} < \infty\right) = p^2$$

which is only possible if $p \in \{0, 1\}$. □

We now have

$$\mathbf{P}(X = 0) = \prod_{j \geq 1} \left(1 - \alpha_j\right)$$

which is > 0, together with $\mathbf{P}(X = i)$ for all $i \geq 1$, if and only if (iff): $\sum_{j \geq 1} \alpha_j < \infty$, because then

$$\mathbf{P}(X = i) = \mathbf{P}(X = 0) \left[z^i\right] \prod_{j \geq 1} \left(1 - \alpha_j z\right)^{-1} > 0.$$

proposition 7. *Iff $\sum_{j \geq 1} \alpha_j < \infty$, $X < \infty$ with probability 1, otherwise $X = \infty$ with probability 1.*

As $J \to \infty$, consider now $P = \sum_{j=1}^{\infty} B_j$. This rv is either ∞ with probability 1 or it is $< \infty$ with probability 1. We have

$$\mathbf{P}(P = 0) = \prod_{j \geq 1} \left(1 - \alpha_j\right)$$

which is > 0, together with $\mathbf{P}(P = p)$ for all $p \geq 1$, iff $\sum_{j \geq 1} \alpha_j < \infty$, because then

$$\mathbf{P}(P = p) = \mathbf{P}(P = 0) \left[z_0^p\right] \prod_{j \geq 1} \left(1 + \frac{\alpha_j}{1 - \alpha_j} z_0\right) > 0.$$

Thus,

proposition 8. *Concomitantly with X, iff* $\sum_{j\geq 1}\alpha_j<\infty$*,* $P<\infty$ *with probability* 1*, otherwise* $P=\infty$ *with probability* 1.

Two cases arise:

• **The case** $\sum_{j\geq 1}\alpha_j<\infty$**,** where both $X,P<\infty$ with probability 1.

This occurs when $\alpha_j\to 0$ faster than j^{-1} with the sequence of partial sums $\sum_{j=1}^{J}\alpha_j$ being bounded.

What can be said about the shape of the distribution of P and X?

- Concerning the rv P:

$$\varphi(z_0):=\mathbf{E}\left(z_0^P\right)=\prod_{j\geq 1}\left(1-\alpha_j(1-z_0)\right)$$

whose convergence radius $\sup(z_0>0:\varphi(z_0)<\infty)=+\infty$ ($\varphi(z_0)$ is an entire function). The rv P has all its moments finite. The mean and variance are already known. For instance, the third central moment is

$$\mathbf{E}\left[(P-\mathbf{E}P)^3\right]=\sum_{j\geq 1}\alpha_j\left(1-\alpha_j\right)^2-\sum_{j\geq 1}\alpha_j^2\left(1-\alpha_j\right)<\infty.$$

Having only real zeros located at $1-1/\alpha_j<0$, $\varphi(z_0)$ is not the pgf of an infinitely divisible rv P (cf. Proposition I.2.8 of (Steutel & van Harn, 2004)). The coefficients $\mathbf{P}(X=i)$ form an infinite Pòlya sequence, (Pitman, 1997).

- Concerning the rv X: its pgf

$$\phi(z)=\mathbf{E}\left(z^X\right)=\prod_{j\geq 1}\frac{1-\alpha_j}{1-\alpha_j z},\ z\in[0,1].$$

has convergence radius $z_c=\min_{j\geq 1}\left(1/\alpha_j\right)>1$, showing that $z_c^k\mathbf{P}(X=k)\to$Constant as $k\to\infty$.

proposition 9. *If* $X<\infty$*, the rv X is self-decomposable, in particular infinitely divisible (compound Poisson).*

Proof: Indeed,

$$R(z):=\log\phi(z)'=\sum_{j\geq 1}\frac{\alpha_j}{1-\alpha_j z}=\sum_{k\geq 0}z^k\left(\sum_{j\geq 1}\alpha_j^{k+1}\right)$$

with canonical sequence

$$r_k:=\left[z^k\right]R(z)=\sum_{j\geq 1}\alpha_j^{k+1}>0$$

obeying $r_k-r_{k-1}=-\left(1-\alpha_j\right)\sum_{j\geq 1}\alpha_j^k<0$, $k\geq 1$. We conclude (see Theorem V.4.13 of (Steutel & van Harn, 2004))

$$R(z)=r_0\frac{1-h(z)}{1-z}\text{ and }\phi(z)=e^{-r_0\int_z^1\frac{1-h(z')}{1-z'}dz'}$$

where $h(z)=r_0^{-1}\sum_{k\geq 1}z^k(r_{k-1}-r_k)$ is an absolutely monotone pgf obeying $h(0)=0$. □

Remark: If $\alpha_j\to 0$, there exists j_0 such that $\alpha_j<1/2$ (else $1/\left(1-\alpha_j\right)<2$) for all $j>j_0$. Therefore, for any k,

$$\sum_{j\geq 1}\alpha_j<\infty\Rightarrow\sum_{j\geq 1}\frac{\alpha_j}{\left(1-\alpha_j\right)^k}<\sum_{j=1}^{j_0}\frac{\alpha_j}{\left(1-\alpha_j\right)^k}+2^k\sum_{j>j_0}\alpha_j<\infty,$$

showing that if $X<\infty$ then X has all its moments finite. And similarly for P.

Examples: This includes the case $\alpha_j=e^{-\beta j^{\alpha}}$, $\alpha,\beta>0$ with α_j decreasing to 0 faster than j^{-1}. The nature of $\sum_{j\geq 1}\frac{\alpha_j}{1-\alpha_j}$ is the one of $I=\int_1^{\infty}\frac{e^{-\beta x^{\alpha}}}{1-e^{-\beta x^{\alpha}}}dx$ with ($\zeta(\sigma)=\sum_{k\geq 1}k^{-\sigma}$)

$$I<\zeta(1/\alpha)\Gamma(1+1/\alpha)\beta^{-1/\alpha}<\infty\text{ if }\alpha<1.$$

The case $\alpha_j = \alpha^j$, $\alpha \in (0,1)$ and also $\alpha_j = j^{-\alpha} \log^{-\beta}(j+1)$, $\alpha > 1, \beta \in \mathbf{R}$ or $\alpha = 1, \beta > 1$ are also included.

• **The case** $\sum_{j\geq 1} \alpha_j = \infty$**,** where both $X, P = \infty$ with probability 1.

The question here is: is there a way to scale (P_J, X_J) so as to obtain proper weak limits as $J \to \infty$? In such a situation, $\mathbf{E}(P_J) \to \infty$ as $J \to \infty$.

- Concerning the rv P: If both $\mathbf{E}(P_J)$ and $\sigma^2(P_J)$ tend to ∞ as $J \to \infty$, then $\sigma(P_J)/\mathbf{E}(P_J) < 1/\sigma(P_J) \to 0$ and

$$\frac{P_J - \mathbf{E}(P_J)}{\sigma(P_J)} \overset{d}{\to} \mathcal{N}(0,1) \text{ as } J \to \infty.$$

Examples:

This occurs if $\alpha_j \to 0$ slower than j^{-1}, for example $\alpha_j \sim j^{-\alpha}$ ($\alpha \in (0,1)$) or $\alpha_j \sim j^{-1}\log^{-\beta} j$ ($\beta \leq 1$). But also if $\alpha_j = \alpha$ (the iid case) with $\mathbf{E}(P_J) = \alpha J$ and $\sigma^2(P_J) = J\alpha(1-\alpha)$.

The condition $\sigma(P_J)/\mathbf{E}(P_J) \to 0$ is also met if $\alpha_j \to 1$.

For example, if $\alpha_j = 1 - \alpha^j$ ($\alpha \in (0,1)$), $\mathbf{E}(P_J) \to \infty$ as $J \to \infty$ while $\sigma^2(P_J) = \sum_{j=1}^J \alpha^j \left(1 - \alpha^j\right) \to \alpha/\left(1-\alpha^2\right)$. In this case also, $\sigma(P_J)/\mathbf{E}(P_J) \to 0$. If $\alpha_j \sim 1 - \lambda j^{-1}$, $\mathbf{E}(P_J) \to \infty$ as $J \to \infty$ while $\sigma^2(P_J) \to \lambda \sum_{j\geq 1} j^{-1}\left(1 - \lambda j^{-1}\right) = \infty$ with $\sigma(P_J)/\mathbf{E}(P_J) \to 0$.

Choosing $\alpha_j \sim 1 - \lambda j^{-\alpha}$ with $\alpha > 1$, $\mathbf{E}(P_J) \to \infty$ while $\sigma^2(P_J) \to \lambda \sum_{j\geq 1} j^{-\alpha}(1 - \lambda j^{-\alpha}) < \infty$, still with $\sigma(P_J)/\mathbf{E}(P_J) \to 0$.

- Concerning the rv X: Under the condition $\sum_{j\geq 1} \alpha_j = \infty$, both $\mathbf{E}(X_J)$ and $\sigma^2(X_J)$ tend to ∞ as $J \to \infty$.

If $\mu_j > 0$ with $\sum_{j=1}^J \mu_j = 1$, by Jensen inequality, for all $x_j > 0$

$$\left(\sum_{j=1}^J \mu_j g_j\right)^2 < \sum_{j=1}^J \mu_j x_j^2.$$

Choosing $\mu_j = \alpha_j/\mathbf{E}(P_J)$ and $x_j = 1/\left(1 - \alpha_j\right)$ yields $\sigma(X_J)/\mathbf{E}(X_J) > 1/\mathbf{E}(P_J)^{1/2}$ with $\mathbf{E}(P_J) \to \infty$. On the other hand, from the relation between l_2 and l_1 norms

$$\sigma^2(X_J) = \sum_{j=1}^J \frac{\alpha_j}{\left(1-\alpha_j\right)^2} = \sum_{j=1}^J \frac{\alpha_j}{1-\alpha_j}\left(1 + \frac{\alpha_j}{1-\alpha_j}\right) < \mathbf{E}(X_J)(1 + \mathbf{E}(X_J)),$$

showing that $\sigma(X_J)/\mathbf{E}(X_J) < (1 + 1/\mathbf{E}(X_J))^{1/2}$ with $\mathbf{E}(X_J) \to \infty$.

If $\sigma(X_J)/\mathbf{E}(X_J) \to 0$ and

$$\frac{\max_{j=1,\dots,J} \sigma^2\left(G_j\right)}{\sigma^2(X_J)} \to 0 \text{ as } J \to \infty,$$

then Lindeberg criterion is fulfilled, (Billingsley, 2012), and one expects

$$\frac{X_J - \mathbf{E}(X_J)}{\sigma(X_J)} \overset{d}{\to} \mathcal{N}(0,1) \text{ as } J \to \infty. \tag{10}$$

Both the B_j's and the G_j's are independent in L_2. By Kolmogorov strong law:

If $\sum_{j\geq 1} j^{-2}\sigma^2\left(B_j\right) = \sum_{j\geq 1} j^{-2}\alpha_j\left(1 - \alpha_j\right) < \infty$, then

$$\lim_{J\to\infty} \frac{1}{J}(P_J - \mathbf{E}(P_J)) = 0 \text{ almost surely.}$$

If $\sum_{j\geq 1} j^{-2}\sigma^2\left(G_j\right) = \sum_{j\geq 1} j^{-2}\alpha_j\left(1 - \alpha_j\right)^{-2} < \infty$, then

$$\lim_{J\to\infty} \frac{1}{J}(X_J - \mathbf{E}(X_J)) = 0 \text{ almost surely.}$$

2.4 Firstly Occupied Site Index and Site Indices Till Consecutive Records

Let $\mathcal{J}_J = \inf\left(j \in [J] : B_j = 1\right)$ be the index of the first filled site when the number of sites is J, finite. With the empty product being 1, we have

$$\mathbf{P}\left(\mathcal{J}_J > j\right) = \prod_{k=1}^{j} \left(1 - \alpha_k\right), \, j = 0, ..., J - 1, \mathbf{P}\left(\mathcal{J}_J > J\right) = 0$$

with $\mathbf{P}\left(\mathcal{J}_J = J\right) = \prod_{k=1}^{J-1} \left(1 - \alpha_k\right)$. Letting $J \to \infty$ and $\mathcal{J} = \inf\left(j \geq 1 : B_j = 1\right)$, it holds

$$\mathbf{P}\left(\mathcal{J} > j\right) = \prod_{k=1}^{j} \left(1 - \alpha_k\right) = \mathbf{P}\left(P_j = 0\right), \, j \geq 0. \tag{11}$$

- If $\sum_{k\geq 1} \alpha_k < \infty : \mathbf{P}\left(\mathcal{J} = \infty\right) = \prod_{k\geq 1} \left(1 - \alpha_k\right) > 0$. There is a positive probability that $\mathcal{J} = \infty$.

- If $\sum_{k\geq 1} \alpha_k = \infty$, then $\mathcal{J} < \infty$ with probability 1.

proposition 10. *More generally, with $p \geq 1$, define $\mathcal{J}^{(p)} = \inf\left(j \geq 1 : P_j = \sum_{k=1}^{j} B_k = p\right) \geq p$ be the first index when P_j equals p. Note $\mathcal{J}^{(1)} = \mathcal{J}$. Then,*

$$P_j = \sum_{p\geq 1} \mathbf{1}\left(\mathcal{J}^{(p)} \leq j\right)$$

and

$$\mathbf{P}\left(\mathcal{J}^{(p)} > j\right) = \mathbf{P}\left(P_j \leq p - 1\right).$$

proposition 11. *This probability can be generated by using the recurrence on $\mathbf{P}\left(P_J = p\right)$, leading to*

$$\mathbf{P}\left(\mathcal{J}^{(p)} > j\right) = \alpha_j \mathbf{P}\left(\mathcal{J}^{(p-1)} > j - 1\right) + \left(1 - \alpha_j\right) \mathbf{P}\left(\mathcal{J}^{(p)} > j - 1\right), \tag{12}$$

with boundary conditions: $\mathbf{P}\left(\mathcal{J}^{(1)} > j\right) = \prod_{k=1}^{j} \left(1 - \alpha_k\right)$ for all $j \geq 0$ and $\mathbf{P}\left(\mathcal{J}^{(0)} > j\right) = 0$, $j \geq 0$.

Note that $\mathcal{J}^{(p)} - \mathcal{J}^{(p-1)}$ ($\mathcal{J}^{(0)} = 0$) is the number of sites between consecutive filled states, with

$$\mathbf{P}\left(\mathcal{J}^{(p)} - \mathcal{J}^{(p-1)} = j\right) = \sum_{l\geq p-1} \mathbf{P}\left(\mathcal{J}^{(p-1)} = l\right) \mathbf{P}\left(\mathcal{J}^{(p)} - l = j \mid \mathcal{J}^{(p-1)} = l\right)$$

and

$$\mathbf{P}\left(\mathcal{J}^{(p)} - l > j \mid \mathcal{J}^{(p-1)} = l\right) = \prod_{k=1}^{j} \left(1 - \alpha_{k+l}\right).$$

2.5 Number of Particles in the Most Filled Site

The number of particles in the most filled site is $X_J^* = \max_{j=1}^{J} G_j$ with probability distribution function (pdf)

$$\mathbf{P}\left(X_J^* \leq i\right) = \prod_{j=1}^{J} \left(1 - \alpha_j^{i+1}\right), \, i \geq 0. \tag{13}$$

As $J \to \infty$, consider $X^* = \max_{j\geq 1} G_j$, so with

$$\mathbf{P}\left(X^* \leq i\right) = \prod_{j\geq 1} \left(1 - \alpha_j^{i+1}\right), \, i \geq 0.$$

- If $\sum_{j\geq 1} \alpha_j < \infty$, then $\sum_{j\geq 1} \alpha_j^{i+1} < \infty$ for all $i \geq 0$ showing that $\mathbf{P}\left(X^* \leq i\right) > 0$ for all $i \geq 0$: thus $X^* < \infty$ with probability 1, with support $\{0, ..., \infty\}$. Note in particular

$$\mathbf{P}\left(X^* = 0\right) = \prod_{j\geq 1} \left(1 - \alpha_j\right) > 0$$

corresponding to the event $X = 0$. This occurs for example when $\alpha_j = e^{-\beta(j)}$ where $\beta(j) = j^{\alpha} \log^{\beta}(j+1)$, $\alpha > 0, \beta \in \mathbb{R}$ or $\beta(j) = \alpha^{-j}$, $\alpha \in (0,1)$. In all such cases, $\alpha_j \to 0$ as $j \to \infty$.

- If $\sum_{j\geq 1} \alpha_j = \infty$, then $\mathbf{P}(X^* = 0) = 0$. Let then

$$i^* = \sup\left(i \geq 0 : \sum_{j\geq 1} \alpha_j^{i+1} = \infty\right) \in \{0, ..., \infty\}. \tag{14}$$

Two cases arise:

1/ $i^* < \infty$ then $\mathbf{P}(X^* \leq i) = 0$ for all $i \leq i^*$ and $\mathbf{P}(X^* \leq i) > 0$ for all $i > i^*$: we conclude that $X^* < \infty$ with probability 1, with support $\{i^* + 1, ..., \infty\}$. In such a situation, necessarily, $\alpha_j \to 0$ as $j \to \infty$. This occurs for example when $\alpha_j = e^{-\beta(j)}$ where $\beta(j) = \beta \log j, \beta \in (0,1]$, ($\sum_{j\geq 1} \alpha_j = \infty$ with $i^* = \lfloor 1/\beta - 1 \rfloor$).

2/ $i^* = \infty$: then $X^* = \infty$ with probability 1. This occurs in the trivial iid case $\alpha_j = \alpha$ for all $j \geq 1$ but also for all sequences $\alpha_j \to 1$ as $j \to \infty$. There are examples with $\alpha_j \to 0$ as $j \to \infty$ and $i^* = \infty$; think of $\alpha_j = e^{-\beta(j)}$ where $\beta(j) = \beta \log\log(j+1) \to \infty$.

proposition 12. *The rv X^* is also either $< \infty$ or it is ∞ with probability 1. It is finite iff $\sum_{j\geq 1} \alpha_j < \infty$ or $\sum_{j\geq 1} \alpha_j = \infty$ and $i^* < \infty$.*

- In cases $\sum_{j\geq 1} \alpha_j < \infty$ (formally with $i^* = -1$) or $\sum_{j\geq 1} \alpha_j = \infty$ and $i^* < \infty$, $X^* < \infty$ with probability 1.

What can be said about the shape of the distribution of X^* which is:

$$\mathbf{P}(X^* \leq i) = \prod_{j\geq 1}\left(1 - \alpha_j^{i+1}\right), \, i \geq i^* + 1.$$

Using $1 - \prod_{j=1}^{J}\left(1 - \alpha_j^{i+1}\right) \leq \sum_{j=1}^{J} \alpha_j^{i+1}$, by letting $J \to \infty$

$$\mathbf{E}[X^*] = \sum_{i\geq 0}\left[1 - \prod_{j\geq 1}\left(1 - \alpha_j^{i+1}\right)\right] \leq i^* + 1 + \sum_{j\geq 1} \frac{\alpha_j^{i_*+2}}{1 - \alpha_j} < \infty.$$

The mean but also the positive integral moments of X^* are finite.

- The case $\sum_{j\geq 1} \alpha_j = \infty$ and $i^* = \infty$ where $X^* = \infty$ with probability 1. The question here is: is there a way to scale X_J^* so as to obtain a proper weak limit as $J \to \infty$?

There is a partial answer to this question in (Doumas & Papanicolaou 2014) where some use was made of an analogy with a previous work on the coupon collector problem, (Doumas & Papanicolaou 2012).

proposition 13. *(Doumas & Papanicolaou 2014). If $\alpha_j = e^{-1/\lambda(j)}$ where $\lambda(j)$ is such that the function $(x \geq 1) \to \lambda(x) > 0$ is in class $\mathcal{S}$ of increasing functions, then, with $\rho(x) = -\log(\lambda'(x)/\lambda(x)) > 0$*

$$\frac{X_J^* - \lambda(J)(\rho(J) - \log\rho(J))}{\lambda(J)} \stackrel{d}{\to} G \text{ as } J \to \infty, \tag{15}$$

where G is a Gumbel rv obeying $\mathbf{P}(G \leq x) = e^{-e^{-x}}$, $x \in \mathbb{R}$.

The class $\mathcal{S}$ of functions consists of positive and strictly increasing ones in $C^2(1,\infty)$ which grow at infinity slower than exponentials but faster than positive powers of logarithms (see (Doumas & Papanicolaou 2014)). In such cases, $\lambda'(x)/\lambda(x) \to 0$ as $x \to \infty$ ($\rho(x) \to \infty$) and with

$$\mathbf{E}(X_J^*) = \lambda(J)(\rho(J) - \log\rho(J)) \text{ and } \sigma(X_J^*) = \lambda(J),$$

$\sigma\left(X_J^*\right)/\mathbf{E}\left(X_J^*\right) \to 0$ as $J \to \infty$. For such a class, $\alpha_j \to 1$ as $J \to \infty$ and the order of magnitude of X_J^* is $\mathbf{E}\left(X_J^*\right)$, with fluctuations of order $\sigma\left(X_J^*\right)$. The functions $\lambda(x) = x^{\alpha} \log^{\beta} x$, $\alpha > 0, \beta \in \mathbb{R}$ and $e^{x^{\alpha}}$, $\alpha \in (0,1)$ are in class $\mathcal{S}$. For instance, taking $\lambda(x) = x^{\alpha}$, $\alpha > 0$,

$$\mathbf{E}(X_J^*) \sim J^{\alpha} \log J \text{ while } \sigma(X_J^*) \sim J^{\alpha}$$

while taking $\lambda(x) = e^{x^\alpha}$, $\alpha \in (0,1)$,

$$\mathbf{E}(X_J^*) \sim (1-\alpha)\, e^{J^\alpha} \log J \text{ and } \sigma(X_J^*) \sim e^{J^\alpha}.$$

There are cases left behind with λ not in class $\mathcal{S}$: (i)Whenever $\alpha_j = \alpha \in (0,1)$ for all $j \geq 1$ (the G_j's are iid), there is no way to scale X_J^* so as to obtain a proper weak limit as $J \to \infty$. All that can be said is that $X_J^*/\log J$ converges in L_1 and in probability to $-1/\log\alpha$ as $J \to \infty$. See (Eisenberg, 2008) and (Wilf & Ewens 2010) for refinements. (ii)Whenever $\alpha_j = e^{-\alpha^j}$, $\alpha \in (0,1)$

$$\frac{X_J^*}{\alpha^{-J}} \xrightarrow{d} X \text{ as } J \to \infty$$

with $\mathbf{P}(X \leq t) = \prod_{j\geq 0}\left(1 - e^{\alpha^{-j}t}\right)$, $t \geq 0$, (see (Doumas & Papanicolaou 2014)). The rv X_J^* grows logarithmically (exponentially) with J in case (i) (respectively (ii)).

Remark: the authors of (Wilf & Ewens 2010) and (Doumas & Papanicolaou 2014) were motivated by the time to genomic evolution dealing with $X_J^* = \max_{j=1}^J G_j^+$ instead of $X_J^* = \max_{j=1}^J G_j$. On each site j of a gene with J sites, a clock is launched whose duration is G_j^+ (the time till a favorable allele shows up there) and their X_J^* represents the time till all sites have switched favorable, ending up evolution.

2.6 The Index of the Site Achieving the Maximum Number of Particles

Let $\mathcal{J}_J^*$ be any of the indices of the sites achieving $G_{\mathcal{J}_J^*}^* = \max_{j=1}^J G_j$. Then

$$\begin{aligned}\mathbf{P}(\mathcal{J}_J^* = j) &= \sum_{l\geq 0}\mathbf{P}\left(G_j = l\right)\prod_{k\in[J]\setminus\{j\}}\mathbf{P}(G_k \leq l)\\ &= \left(1-\alpha_j\right)\sum_{l\geq 0}\alpha_j^l\prod_{k\in[J]\setminus\{j\}}\left(1-\alpha_k^{l+1}\right)\\ &= \left(1-\alpha_j\right)\sum_{l\geq 0}\frac{\alpha_j^l}{1-\alpha_j^{l+1}}\prod_{k=1}^J\left(1-\alpha_k^{l+1}\right)\end{aligned}$$

Summing the latter expression over $j = 1, ..., J$ gives 1 in principle, leading to a combinatorial identity. If α_j is a strictly decreasing sequence, one can check that $\mathbf{P}\left(\mathcal{J}_J^* = j+1\right) < \mathbf{P}\left(\mathcal{J}_J^* = j\right)$.

For this index to be uniquely determined, the corresponding occupancy should contain at least one particle. This index is thus unique and it is j with probability

$$\begin{aligned}\mathbf{P}(\mathcal{J}_J^* = j) &= \sum_{l\geq 1}\mathbf{P}\left(G_j = l\right)\prod_{k\in[J]\setminus\{j\}}\mathbf{P}(G_k \leq l-1)\\ &= \left(1-\alpha_j\right)\sum_{l\geq 1}\alpha_j^l\prod_{k\in[J]\setminus\{j\}}\left(1-\alpha_k^{l}\right)\\ &= \left(1-\alpha_j\right)\sum_{l\geq 1}\frac{\alpha_j^l}{1-\alpha_j^{l}}\prod_{k=1}^J\left(1-\alpha_k^{l}\right).\end{aligned}$$

Summing the latter expression over $j = 1, ..., J$ gives the probability that the maximum is achieved at a single site. We also have

$$\begin{aligned}\mathbf{P}(\mathcal{J}_J^* = j, X_J^* \leq i) &= \sum_{l=0}^{i}\mathbf{P}\left(G_j = l\right)\prod_{k\in[J]\setminus\{j\}}\mathbf{P}(G_k \leq l)\\ &= \left(1-\alpha_j\right)\sum_{l=0}^{i}\alpha_j^l\prod_{k\in[J]\setminus\{j\}}\left(1-\alpha_k^{l+1}\right).\end{aligned}$$

Summing also the latter expression over $j = 1, ..., J$ gives an alternative expression of $\mathbf{P}\left(X_J^* \leq i\right) = \prod_{j=1}^J\left(1-\alpha_j^{i+1}\right)$.

Clearly,

proposition 14. *$\left(\mathcal{J}_J^*, X_J^*\right)$ are not independent.*

2.7 Special Examples

We give here some details for specific 'critical' sequences α_j.

• Special cases with $\alpha_j \to 1$:

$$\begin{aligned}\alpha_j &= 1 - \alpha/j,\ 0 < \alpha < 1\\ \alpha_j &= (j/(j+1))^{\alpha},\ \alpha > 0.\end{aligned}$$

For both sequences, to the dominant order in J:

$$\begin{aligned}\mathbf{E}(P_J) &\sim J - \alpha \log J \text{ and } \sigma(P_J) \sim \sqrt{\alpha \log J}\\ \mathbf{E}(X_J) &\sim \frac{1}{2\alpha} J^2 \text{ and } \sigma(X_J) \sim \frac{1}{\sqrt{3}\alpha} J^{3/2}\end{aligned}$$

For both sequences, $\alpha_j \sim e^{-\alpha/j}$ for large j, with $\lambda(j) = \alpha^{-1} j$ in class $\mathcal{S}$. To the dominant order, $\mathbf{E}\left(X_J^*\right) \sim \alpha^{-1} J \log J$ while $\sigma\left(X_J^*\right) \sim \alpha^{-1} J$. On average, there are $\mathbf{E}(X_J)/\mathbf{E}(P_J) \sim \frac{1}{2\alpha} J$ particles per filled site while there are order $\alpha^{-1} J \log J$ particles in the most filled state. And most states are filled.

• Special cases with $\alpha_j \to 0$:

$$\begin{aligned}\alpha_j &= \alpha/j,\ 0 < \alpha < 1\\ \alpha_j &= 1 - (j/(j+1))^{\alpha},\ \alpha > 0.\end{aligned}$$

For both sequences, to the dominant order in J:

$$\begin{aligned}\mathbf{E}(P_J) &\sim \alpha \log J \text{ and } \sigma(P_J) \sim \sqrt{\alpha \log J}\\ \mathbf{E}(X_J) &\sim \alpha^{-1} \log J \text{ and } \sigma(X_J) \sim \sqrt{\alpha^{-1} \log J}\end{aligned}$$

For both sequences, $\alpha_j \sim \alpha/j$ for large j with $i^* = 0$. Then $X_J^* \to X^* < \infty$ almost surely as $J \to \infty$. We have $\mathbf{P}(X^* = 0) = 0$. And filled states are rare with $O(1)$ particles per state on average.

3. A Variant of the Model in the Continuum

Let $c > 0$ stand for some cutoff value to be interpreted as a minimum detection limit of some sensor. Let E_j, $j = 1, ..., J$ be independent exponentially distributed random variables, each with rate parameter $\beta_j > 0$ (the reciprocal of the scale parameter $\lambda_j := \beta_j^{-1}$). Let also $E_j^+ = E_j + c \overset{d}{=} E_j \mid E_j > c$, $j = 1, ..., J$, so each with densities

$$f_j^+(\epsilon) = \beta_j e^{-\beta_j(\epsilon - c)},\ \epsilon > c.$$

Let $B_j = \mathbf{1}\left(E_j > c\right)$, so that $B_j = 1$ with probability $e^{-\beta_j c}$, $= 0$ with complementary probability. These rv's indicate whether item j has or not 'energy' E_j exceeding c.

In the spirit of the previous discrete random allocation study, we wish to consider the rv's $X_J = \sum_{j=1}^J B_j E_j$ and $P_J = \sum_{j=1}^J B_j$, together with the joint distribution of $\left(B_j E_j, j = 1, .., J; P_J\right)$, taking into account only those E_j exceeding the cutoff in a random allocation of energy process.

Note that (unless $c = 0$), $X_J < \sum_{j=1}^J E_j = \sum_{j=1}^J \left(1 - B_j\right) E_j + \sum_{j=1}^J B_j E_j$ because whenever $B_j = 0$, the associated energy E_j is $\leq c$, but not zero. We have $\mathbf{E}\left(\sum_{j=1}^J E_j\right) = \sum_{j=1}^J \beta_j^{-1}$ and $\sigma^2\left(\sum_{j=1}^J E_j\right) = \sum_{j=1}^J \beta_j^{-2}$ showing that (see (Anderson, 1991) p. 18 − 19)

$$\sum_{j\geq 1} E_j < \infty \text{ with probability 1, iff } \sum_{j\geq 1} \beta_j^{-1} < \infty. \quad (16)$$

Suppose for instance that J stars emit light the intensity of which are the E_j's. In this context, the β_j's may be related to the inverse temperatures of the stars [1]. Suppose that due to lack of perfect resolution, only those stars whose intensity exceed the cutoff c are being detected. The rv X_J is the cumulated energy detected, while $P_J \leq J$ is the number of detected stars. The rv's

$$X_J / \sum_{j=1}^J E_j \text{ and } P_J / J$$

[1] Other physical images are of course at stake.

are the fraction of the total intensity emitted which is being detected and the proportion of the detected stars.

We have

$$\mathbf{E}\left(z_0^{B_j}e^{-\omega B_j E_j}\right) = 1-e^{-\beta_j c}+e^{-\beta_j c}z_0\mathbf{E}\left(e^{-\omega E_j^+}\right) = 1-e^{-\beta_j c}+z_0\frac{\beta_j e^{-c\left(\omega+\beta_j\right)}}{\omega+\beta_j}$$

so that with $\omega_j \geq 0$

$$\mathbf{E}\left(z_0^{P_J}\prod_{j=1}^{J}e^{-\omega_j B_j E_j}\right)=\prod_{j=1}^{J}\left(1-e^{-\beta_j c}+z_0\,\frac{\beta_j e^{-c\left(\omega_j+\beta_j\right)}}{\omega_j+\beta_j}\right).$$

The joint pgf and Laplace-Stieltjes transform of (P_J, X_J) is (plugging $\omega_j = \omega$)

$$\mathbf{E}\left(z_0^{P_J}e^{-\omega X_J}\right)=\prod_{j=1}^{J}\mathbf{E}\left(z_0^{B_j}e^{-\omega B_j E_j}\right)=\prod_{j=1}^{J}\left(1-e^{-\beta_j c}+z_0\frac{\beta_j e^{-c\left(\omega+\beta_j\right)}}{\omega+\beta_j}\right)$$

As a consequence,

$$\mathbf{E}\left(P_J\right)=\sum_{j=1}^{J}e^{-\beta_j c}\text{ and }\sigma^2\left(P_J\right)=\sum_{j=1}^{J}e^{-\beta_j c}\left(1-e^{-\beta_j c}\right)=\mathbf{E}\left(P_J\right)-\sum_{j=1}^{J}e^{-2\beta_j c}.$$

while

$$\mathbf{E}\left(X_J\right)=\sum_{j=1}^{J}\frac{e^{-\beta_j c}}{\beta_j}\left(c\beta_j+1\right)\text{ and }\sigma^2\left(X_J\right)=\sum_{j=1}^{J}\frac{e^{-\beta_j c}}{\beta_j^2}\left(1+\left(c\beta_j+1\right)^2\left(1-e^{-\beta_j c}\right)\right).$$

In view of $e^{-\beta_j c}\left(c\beta_j+1\right)<1$, we have as expected $\mathbf{E}\left(X_J\right)<\mathbf{E}\left(\sum_{j=1}^{J}E_j\right)=\sum_{j=1}^{J}\beta_j^{-1}$.

Note $\sigma^2\left(X_J\right)<\sum_{j=1}^{J}\beta_j^{-2}\left(2\cosh\left(c\beta_j\right)-1\right)$ and $\sigma^2\left(\sum_{j=1}^{J}E_j\right)=\sum_{j=1}^{J}\beta_j^{-2}$.

3.1 The rv P

As $J\to\infty$, consider now $P=\sum_{j=1}^{\infty}B_j$. This rv is either ∞ with probability 1 or it is $<\infty$ with probability 1. We have

$$\varphi\left(z_0\right):=\mathbf{E}\left(z_0^{P}\right)=\prod_{j=1}^{J}\left(1-e^{-\beta_j c}+z_0e^{-\beta_j c}\right),$$

with

$$\mathbf{P}\left(P=0\right)=\prod_{j\geq 1}\left(1-e^{-\beta_j c}\right).$$

Two cases can arise:

• The case $\sum_{j\geq 1}e^{-\beta_j c}<\infty$, where $\mathbf{P}\left(P=0\right)>0$, together with $\mathbf{P}\left(P=p\right)$ for all $p\geq 1$. If so, $P<\infty$ with probability 1. This occurs when $c>\epsilon^*$ where

$$\epsilon^*=\sup\left(\epsilon\geq 0:\sum_{j\geq 1}e^{-\beta_j\epsilon}=\infty\right)\in[0,\infty], \tag{17}$$

and this requires $\epsilon^*<\infty$, itself requiring $\beta_j\to\infty$.

If $\epsilon^*=0$, the condition $c>\epsilon^*=0$ is always fulfilled. This occurs when $\beta_j\to\infty$ like $\beta_j=\beta j, \beta>0$ or faster.

A case with $0<\epsilon^*<\infty$ occurs typically when $\beta_j\to\infty$ like $\beta_j=\beta\log j, \beta>0$, with $\epsilon^*=1/\beta$.

Under the condition $P<\infty$, what can be said about the shape of the distribution of P? The pgf $\varphi\left(z_0\right)$ has convergence radius $\sup\left(z_0>0:\varphi\left(z_0\right)<\infty\right)=+\infty$ ($\varphi\left(z_0\right)$ is an entire function). The rv P has all its moments finite. The mean and variance are $\mathbf{E}\left(P\right)=\sum_{j\geq 1}e^{-\beta_j c}$ and $\sigma^2\left(P\right)=\sum_{j\geq 1}e^{-\beta_j c}\left(1-e^{-\beta_j c}\right)$, both finite. The third central moment is

$$\mathbf{E}\left[\left(P-\mathbf{E}P\right)^3\right]=\sum_{j\geq 1}e^{-\beta_j c}\left(1-e^{-\beta_j c}\right)^2-\sum_{j\geq 1}e^{-2\beta_j c}\left(1-e^{-\beta_j c}\right)<\infty.$$

Having only real zeros located at $z_j = 1 - e^{\beta_j c} < 0$, $\varphi(z_0)$ is not the pgf of an infinitely divisible rv P.

• In case $\sum_{j\geq 1} e^{-\beta_j c} = \infty$, then $P = \infty$ with probability 1. This will happen either because $\epsilon^* < \infty$ and $c \leq \epsilon^*$ or because $\epsilon^* = \infty$. An estimate of the way $P_J \to \infty$ as $J \to \infty$ can be obtained from a large J estimate of $\mathbf{E}(P_J)$ and $\sigma^2(P_J)$.

To summarize, we observe that a sharp transition phenomenon is possible:

proposition 15. *For those sequences β_j such that $0 < \epsilon^* < \infty$, depending on the threshold $c > \epsilon^*$ or $c \leq \epsilon^*$ the number P of detected stars can switch from finite to infinite with probability* 1.

3.2 The rv X

As $J \to \infty$, consider now $X = \sum_{j=1}^{\infty} B_j E_j < \sum_{j=1}^{\infty} E_j$, with

$$\mathbf{E}\left(e^{-\omega X}\right) = \prod_{j\geq 1}\left(1 - e^{-\beta_j c} + \frac{\beta_j e^{-c(\omega+\beta_j)}}{\omega + \beta_j}\right).$$

Clearly if $P < \infty$ with probability 1, so is X. Can X be finite if $P = \infty$? We have

proposition 16. *$X < \infty$ with probability* 1*iff $P < \infty$. Otherwise, $X = \infty$ with probability* 1.

Proof: Suppose $P = \infty$ with probability 1, equivalently $\sum_{j\geq 1} e^{-\beta_j c} = \mathbf{E}(P) = \infty$. Then $X = \sum_{j=1}^{\infty} B_j E_j$, with $\left(B_j E_j;\, j \geq 1\right)$ a sequence of independent rv's each with mean $\mathbf{E}\left(B_j E_j\right) = \frac{e^{-\beta_j c}}{\beta_j}\left(c\beta_j + 1\right) \geq c e^{-\beta_j c}$. So X has at least infinite mean.

If $\mathbf{P}(X = \infty) > 0$, then $\mathbf{E}(X) = \infty$. Suppose $P = \infty$ and consider now

$$\mathbf{E}\left(e^{-X}\right) = \prod_{j\geq 1}\left(1 - e^{-\beta_j c}\left(1 - \frac{\beta_j e^{-c}}{1 + \beta_j}\right)\right).$$

We have $X = \infty$ with probability 1 iff $\mathbf{E}\left(e^{-X}\right) = 0$.

This occurs iff $\sum_{j\geq 1} e^{-\beta_j c}\left(1 - \frac{\beta_j e^{-c}}{1+\beta_j}\right) = \infty$. Owing to $1 - \frac{\beta_j e^{-c}}{1+\beta_j} \geq 1 - e^{-c}$, we have $\sum_{j\geq 1} e^{-\beta_j c}\left(1 - \frac{\beta_j e^{-c}}{1+\beta_j}\right) \geq \left(1 - e^{-c}\right)\sum_{j\geq 1} e^{-\beta_j c} = \infty$. So iff $P = \infty$ with probability 1, does concomitantly $X = \infty$ with probability 1. □

We obtained a sharp transition phenomenon for X as well:

proposition 17. *For those sequences β_j such that $0 < \epsilon^* < \infty$, depending on the threshold $c > \epsilon^*$ or $c \leq \epsilon^*$, the energy X of the detected stars can be either finite or infinite with probability* 1.

3.3 The rv X^*

The energy of the star emitting the most is $X_J^* = \max_{j=1}^{J} E_j$ with pdf

$$\begin{aligned}\mathbf{P}(X_J^* \leq \epsilon) &= \prod_{j=1}^{J}\left(1 - e^{-\beta_j \epsilon}\right)\\ &= 1 + \sum_{k=1}^{J}(-1)^k \sum_{1\leq j_1<\ldots<j_k\leq J} e^{-\left(\sum_{l=1}^{k} j_l\right)\epsilon},\ \epsilon > 0.\end{aligned}$$

It is detectable with probability $\mathbf{P}\left(X_J^* > c\right) = 1 - \prod_{j=1}^{J}\left(1 - e^{-\beta_j c}\right)$. The $q-$moments of X_J^* are given by

$$\mathbf{E}\left((X_J^*)^q\right) = q\int_0^{\infty} \epsilon^{q-1}\left(1 - \prod_{j=1}^{J}\left(1 - e^{-\beta_j \epsilon}\right)\right) d\epsilon,\ q > 0.$$

As $J \to \infty$, consider $X^* = \max_{j\geq 1} E_j$, so with

$$\mathbf{P}(X^* \leq \epsilon) = \prod_{j\geq 1}\left(1 - e^{-\beta_j \epsilon}\right),\ \epsilon > 0.$$

As before, consider ϵ^* as from (17). Two cases arise:

1/ $0 \leq \epsilon^* < \infty$: then $\mathbf{P}(X^* \leq \epsilon) = 0$ for all $\epsilon \leq \epsilon^*$ and $\mathbf{P}(X^* \leq \epsilon) > 0$ for all $\epsilon > \epsilon^*$: we conclude that $X^* < \infty$ with probability 1, with support (ϵ^*, ∞) . A case with $0 < \epsilon^* < \infty$ occurs when $\beta_j = \beta \log j, \beta > 0$, (with $\epsilon^* = 1/\beta$). A case with $\epsilon^* = 0$ occurs when $\beta_j = \beta j, \beta > 0$.

2/ $\epsilon^* = \infty$: then $X^* = \infty$ with probability 1. This occurs in the trivial iid case $\beta_j = \beta$ for all $j \geq 1$ but also for all sequences β_j tending to 0 as $j \to \infty$.

proposition 18. *The rv X^* is also either $< \infty$ or it is ∞ with probability 1. It is $< \infty$ with support (ϵ^*, ∞) iff $0 \leq \epsilon^* < \infty$.*

Whenever $X^* < \infty$,

$$\mathbf{P}(X^* \leq \epsilon) = \prod_{j \geq 1} \left(1 - e^{-\beta_j \epsilon}\right), \epsilon > \epsilon^*$$

and X^* has all its moments finite. Whenever $X^* = \infty$ and $\beta_j = 1/\lambda_j$ where $\lambda_j = \lambda(j)$ and $\lambda(x)$ is in class $\mathcal{S}$, X_J^* can be scaled as in the discrete setup, with a Gumbel weak limit.

3.4 The Joint Distribution of $\left(\mathcal{J}_J^*, X_J^*\right)$

With $\mathcal{J}_J^*$ be the unique index j achieving $\max_{j=1}^J E_j$, it holds

$$\begin{aligned} \mathbf{P}(\mathcal{J}_J^* = j, X_J^* \leq \epsilon) &= \int_0^\epsilon \mathbf{P}\left(E_j \in d\epsilon'\right) \prod_{k \in [J] \setminus \{j\}} \mathbf{P}(E_k \leq \epsilon') \\ &= \beta_j \int_0^\epsilon d\epsilon' e^{-\beta_j \epsilon'} \prod_{k \in [J] \setminus \{j\}} \left(1 - e^{-\beta_k \epsilon'}\right) \\ &= 1 - e^{-\beta_j \epsilon} + \sum_{k=1}^{J-1} (-1)^k \sum_{\substack{1 \leq j_1 < \ldots < j_k \leq J \\ j_l \in [J] \setminus \{j\}}} \frac{\beta_j \left(1 - e^{-(\beta_j + \sum_{l=1}^k \beta_{j_l})\epsilon}\right)}{\beta_j + \sum_{l=1}^k \beta_{j_l}}, \end{aligned}$$

resulting in $(j = 1, ..., J)$

$$\begin{aligned} \mathbf{P}(\mathcal{J}_J^* = j) &= \beta_j \int_0^\infty d\epsilon' e^{-\beta_j \epsilon'} \prod_{k \in [J] \setminus \{j\}} \left(1 - e^{-\beta_k \epsilon'}\right) \\ &= 1 + \sum_{k=1}^{J-1} (-1)^k \sum_{\substack{1 \leq j_1 < \ldots < j_k \leq J \\ j_l \in [J] \setminus \{j\}}} \frac{\beta_j}{\beta_j + \sum_{l=1}^k \beta_{j_l}}. \end{aligned}$$

Unless $\beta_j = \beta$ for all $j = 1, ..., J$, the rv's $\left(J^*, X_J^*\right)$ are not independent. If $\beta_j = \beta$ for all $j = 1, ..., J$, $\mathbf{P}\left(\mathcal{J}_J^* = j\right) = 1/J$ and

$$\begin{aligned} \mathbf{P}(\mathcal{J}_J^* = j, X_J^* \leq \epsilon) &= \beta \int_0^\epsilon d\epsilon' e^{-\beta \epsilon'} \left(1 - e^{-\beta \epsilon'}\right)^{J-1} \\ &= \frac{1}{J} \left(1 - e^{-\beta \epsilon}\right)^J = \mathbf{P}(\mathcal{J}_J^* = j) \mathbf{P}(X_J^* \leq \epsilon). \end{aligned}$$

In contrast, it can be easily be checked, proceeding similarly, that the unique index achieving $\min_{j=1}^J E_j$ has probability mass at site $j : \beta_j / \sum_{k=1}^J \beta_k$ and it is independent of the latter rv, which is exponentially distributed with rate $\sum_{k=1}^J \beta_k$.

Acknowledgments

The author acknowledges partial support from the "Chaire *Modélisation mathématique et biodiversité"* of Veolia-Ecole Polytechnique-MNHN-Fondation X. He also acknowledges support from the labex MME-DII Center of Excellence (*Modèles mathématiques et économiques de la dynamique, de l'incertitude et des interactions*, ANR-11-LABX-0023-01 project).

References

Anderson, W. J. (1991). *Continuous-Time Markov Chains: An Applications-Oriented Approach.* Springer-Verlag, New York. https://doi.org/10.1007/978-1-4612-3038-0

Billingsley, P. (2012). *Probability and Measure.* Anniversary Edition, John Wiley & Sons, Inc., Publication.

Comtet, L. (1970). *Analyse Combinatoire.* Tomes 1 et 2. Presses Universitaires de France, Paris.

Doumas, A. V., & Papanicolaou, V. G. (2012). Asymptotics of the rising moments for the coupon collector's problem. *Electronic Journal of Probability*, *18,* paper 41. https://doi.org/10.1214/EJP.v18-1746

Doumas, A. V., & Papanicolaou, V. G. (2014). The maximum of independent geometric random variables as the time for genomic evolution. *Stochastic Models*, *30*, 125-141. https://doi.org/10.1080/15326349.2014.868742

Eisenberg, B. (2008). On the expectation of the maximum of IID geometric random variables. *Statistics & Probability Letters*, *78*, 135-143. https://doi.org/10.1016/j.spl.2007.05.011

Huillet, T. E. (2018). Karlin-McGregor Mutational Occupancy Problem Revisited. *Journal of Statistical Physics*, *171*(6), 1136-1149. https://doi.org/10.1007/s10955-018-2056-3

Kolchin, V. F. (1986). *Random mappings*. Translated from the Russian. With a foreword by S. R. S. Varadhan. Translation Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York.

Pitman, J. (1997). Probabilistic Bounds on the Coefficients of Polynomials with Only Real Zeros. *Journal of Combinatorial Theory*, *77*, 279-303. https://doi.org/10.1006/jcta.1997.2747

Sen, A., & Balakrishnan, N. (1999). Convolution of geometrics and a reliability problem. *Statistics & Probability Letters*, *43*, 421-426. https://doi.org/10.1016/S0167-7152(98)00284-3

Steutel, F. W., & Van Harn, K. (2003). *Infinite divisibility of probability distributions on the real line*. CRC Press. https://doi.org/10.1201/9780203014127

Wilf, H. S., & Ewens, W. J. (2010). There's plenty of time for evolution. *PNAS*, *107*(52), 22454-22456. https://doi.org/10.1073/pnas.1016207107

Copyrights