ALEA, Lat. Am. J. Probab. Math. Stat. 21, 1475–1493 (2024)
DOI: 10.30757/ALEA.v21-55

Records in the Infinite Occupancy Scheme

Zakaria Derbazi, Alexander Gnedin and Alexander Marynych

School of Mathematical Sciences, Queen Mary University of London, E1 4NS London, United Kingdom
E-mail address: z.derbazi@qmul.ac.uk

School of Mathematical Sciences, Queen Mary University of London, E1 4NS London, United Kingdom
E-mail address: a.gnedin@qmul.ac.uk

Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine
E-mail address: marynych@unicyb.kiev.ua

Abstract. We consider the classic infinite occupancy scheme, where balls are thrown in boxes independently, with probability p_j of hitting box j. Each time a box receives its first ball we speak of a *record* and, more generally, call an r-*record* every event when a box receives its rth ball. Assuming that the sequence (p_j) is not decaying too fast, we show that after many balls have been thrown, the suitably scaled point process of r-record times is approximately Poisson. The joint convergence of r-record processes is argued under a condition of regular variation.

1. Introduction

In the infinite occupancy scheme first systematically studied by Karlin Karlin (1967), balls are allocated independently to an infinite series of boxes, with fixed probability p_j of hitting the jth box for each ball. There is extensive literature on asymptotic properties of the random partition associated with the allocation of a large number of balls. The most explored features include the number of boxes occupied by at least one ball, and the counts of boxes occupied r times. See Barbour and Gnedin (2009); Ben-Hamou et al. (2017); Gnedin et al. (2007); Hwang and Janson (2008) for development and many references therein. Much less attention has been devoted to the evolutionary aspects of the partition, seen as a random process when balls are thrown successively one at a time.

Recently functional Gaussian limits were shown for the mentioned statistics under Karlin's condition of regular variation on (p_j) Chebunin and Zuyev (2022); Durieu and Wang (2016). However, the contracted time scale employed for such approximation turns too rough to apprehend a short term pattern of newly 'discovered' boxes that get hit for the first time. The latter aspect is of interest for statistical applications akin to the new species search problem Bunge and Fitzpatrick (1993).

Received by the editors March 18th, 2024; accepted June 22nd, 2024.

2010 *Mathematics Subject Classification.* Primary: 60C05; Secondary: 60F05, 60G55.

Key words and phrases. Karlin's infinite occupancy scheme; Poisson approximation; Poissonisation; point processes; records.

Following the terminology from Griffiths and Spanò (2007), in this paper we call a *record* each occupancy event when a new box is hit and, more generally, call an *r-record* each event when a box is hit for the rth time. The class of models considered here has the property that the number of boxes occupied r times approaches ∞ as more balls are thrown, hence the time lag between two consecutive r-records is relatively small as compared with the total elapsed time. This covers the case of regularly varying p_j's but excludes probabilities with exponential decay.

We introduce a local time scale to ensure that r-records occur at about a constant rate, and obtain the Poisson approximation for the point processes of r-record times. The joint convergence of records of different types to independent Poisson processes is shown under the condition of regular variation.

A similar Poisson approximation has been obtained for the familiar coupon-collectors problem in Ilienko (2019), which is the occupancy scheme with finitely many equiprobable boxes. Closely related work connecting record processes to the queueing theory appeared on the Ewens sampling model Gnedin and Stark (2023), where the probabilities (p_j) are themselves random, chosen from the Poisson-Dirichlet/GEM distribution. A characteristic property of the Ewens sampling model is that the indicators of records are independent Nacu (2006).

Following Karlin's approach Karlin (1967) and the setting in much of the subsequent work, we first focus on a continuous time occupancy scheme where balls are thrown at epochs of a unit Poisson process. This has the advantage over the discrete time scheme in that the arrivals to distinct boxes occur according to independent Poisson processes. In Section 8, we proceed with de-Poissonisation to obtain the Poisson approximation for records in the traditional model, where the discrete time variable coincides with the number of balls thrown. In Section 9, we build upon the exchangeability features to develop a different approach to the discrete time model. In the last section we consider occupancy with random (p_j), where a mixed Poisson approximation to records is appropriate.

2. Poissonised setup

Suppose in the first instance that the balls labelled $1, 2, \ldots$ are thrown at epochs of a unit rate Poisson process $P = (P(t),\ t \geq 0)$. By the marking theorem, box j receives balls according to a Poisson process $P_j := (P_j(t),\ t \geq 0),\ j \geq 1$, with rate p_j, so the processes P_j are independent and $P = \sum_{j\geq 1} P_j$.

Let $K_r(t) := |\{j : P_j(t) = r\}|,\ r \geq 1$, denote the number of r-tons, that is boxes containing exactly r balls at time t, and let $K(t) := \sum_{r\geq 1} K_r(t)$ be the number of boxes occupied by at least one ball. The vector $(K_1(t), K_2(t), \ldots)$ encodes a random integer partition (possibly empty) induced by the allocation of the first $P(t)$ balls.

We call *record* (time) any jump time of the process $K := (K(t),\ t \geq 0)$. For this and other nondecreasing counting processes, we shall use the common convention to denote by the same symbol both the process and the counting measure on Borel subsets of $\mathbb{R}_+$, thus writing $K((u,t]) = K(t) - K(u), 0 \leq u < t$. The counting process of r-tons $K_r := (K_r(t),\ t \geq 0)$ has jumps ± 1. We call *r-record* (time) any jump time when K_r increments by $+1$, hence K_{r-1} falls by -1 if $r > 1$. Let

$$\beta_{jr} := \inf\{t \geq 0 : P_j(t) = r\}, \quad j \geq 1, \quad r \geq 1,$$

be the r-record time when box j receives its r-th ball. These are well defined random variables since the number of balls in each box grows to infinity, governed by the law of large numbers. For j fixed, the counting measure $\sum_{r\geq 1} \delta_{\beta_{jr}}$ is the Poisson random measure P_j of arrivals to box j, where δ_t denotes the unit mass at t.

Proposition 2.1. *For $\gamma \in [0,1]$ and $t \geq 0$, it holds that*

$$\begin{aligned} \mathbb{E}[K(\gamma t) \,|\, K(t)] &\geq \gamma K(t), \\ \mathbb{E}[K_r(\gamma t) \,|\, K_s(t)] &\geq \binom{s}{r} \gamma^r (1-\gamma)^{s-r} K_s(t), \;\; s \geq r \geq 1. \end{aligned}$$

Therefore $(K(t)/t,\ t \geq 0)$ and $(K_r(t)/t^r,\ t \geq 0)$ are reverse submartingales.

Proof: By the order statistic property of the Poisson process, given that at time t box j has s balls, the number of arrivals to the box by time γt has the Binomial(s, γ) distribution, regardless of p_j. This implies the second inequality by the following estimates

$$\begin{aligned} \mathbb{E}[K_r(\gamma t) \,|\, K_s(t)] &= \sum_{j \geq 1} \mathbb{P}[P_j(\gamma t) = r \,|\, K_s(t)] \geq \sum_{j \geq 1} \mathbb{P}[P_j(\gamma t) = r, P_j(t) = s \,|\, K_s(t)] \\ &= \sum_{j \geq 1} \mathbb{P}[P_j(\gamma t) = r \,|\, P_j(t) = s] \mathbb{P}[P_j(t) = s \,|\, K_s(t)] \\ &= \binom{s}{r} \gamma^r (1-\gamma)^{s-r} \sum_{j \geq 1} \mathbb{P}[P_j(t) = s \,|\, K_s(t)] = \binom{s}{r} \gamma^r (1-\gamma)^{s-r} K_s(t). \end{aligned}$$

The first inequality follows along the same lines by noting that if a box contains s balls at time t, then it was nonempty at time γt with probability at least γ, for every $s \geq 1$. □

We are interested in features of the point process of r-records $B_r := \sum_{j \geq 1} \delta_{\beta_{jr}}$. Note that the records are the same as 1-records, i.e. $B_1([0,\, t]) = K(t)$ for all $t \geq 0$, because K jumps when an empty box receives its first ball and becomes a singleton.

The sum of all B_r's is the Poisson process P, but the processes B_r themselves are not Poisson, rather possess a repulsion property known as the negative association; see Last and Szekli (2019) for background.

Proposition 2.2. *Each record process $B_r, r \geq 1$, is negatively associated.*

Proof: For every (disjoint) partition $A_1, \dots, A_m$ of $[0, \infty)$, the vector $(\mathbb{1}_{\{\beta_{jr} \in A_1\}}, \dots, \mathbb{1}_{\{\beta_{jr} \in A_m\}})$ is negatively associated, see part (a) in Section 3.1 of Joag-Dev and Proschan (1983). Thus, the one-point process $\delta_{\beta_{jr}}$ is negatively associated by definition. These are independent in j by the independence of arrivals to boxes. Hence B_r is negatively associated as being a sum of the negatively associated processes. □

In particular, the increments of B_r over two disjoint intervals are negatively correlated.

The processes B_r and B_s, $r \neq s$, are not independent. For instance, the first ball thrown after time $t = 1$ is a 2-record with zero probability if $\{B_1(1) = B_2(1) = 0\}$, and with positive probability if $\{B_1(1) > 0, B_2(2) = 0\}$.

3. The Bernstein function

Without loss of generality, we assume that the boxes are labelled by decreasing popularity, that is $p_1 \geq p_2 \geq \cdots > 0$, $\sum_{j \geq 1} p_j = 1$. The probabilities (p_j) are conveniently encoded into the infinite counting measure $\nu := \sum_{j \geq 1} \delta_{p_j}$ on $[0, 1]$. This allows one to write sums over the boxes in the form of integrals,

$$\sum_{j \geq 1} f(p_j) = \int_{[0,\,1]} f(x) \nu(\mathrm{d}x).$$

A dual way to parametrise the model is to use the *Bernstein function*

$$\Phi(t) := \int_{[0,\,1]} (1 - e^{-tx}) \nu(\mathrm{d}x), \quad t \geq 0, \tag{3.1}$$

which uniquely determines ν and is important for the analysis. In the context of this integral representation, ν is sometimes called the Lévy measure Schilling et al. (2012). The tilted measure $\nu_1(x) := x\,\nu(\mathrm{d}x)$ is normalised and can be interpreted as the probability distribution of the popularity of the box hit by ball 1.

The Bernstein function has a transparent probabilistic meaning as the expected number of boxes occupied by time t, that is $\Phi(t) = \mathbb{E}[K(t)]$. Furthermore, the expected counts of r-tons $\Phi_r(t) := \mathbb{E}[K_r(t)]$ are expressible via the derivatives of Φ as

$$\Phi_r(t) \quad = \quad \frac{t^r(-1)^{r+1}}{r!}\Phi^{(r)}(t) = \frac{t^r}{r!}\int_{[0,1]} e^{-tx}x^r\nu(\mathrm{d}x).$$

The formulas imply the recursion

$$\Phi_{r+1}(t) = \frac{r}{r+1}\Phi_r(t) - \frac{t}{r+1}\Phi_r'(t). \tag{3.2}$$

Formulas for the variance and large-t asymptotics are found in Barbour and Gnedin (2009); Bogachev et al. (2008); Gnedin et al. (2007); Hwang and Janson (2008); Karlin (1967).

In accord with Proposition 2.1, we have the monotonicity $\Phi(t)/t\downarrow$, $\Phi_r(t)/t^r\downarrow$.

Proposition 3.1. *The functions Φ_r satisfy for $0 \le \gamma \le 1$ and $t \ge 0$ the inequalities*

$$\Phi_r(\gamma t) \quad \ge \quad \binom{s}{r}\gamma^r(1-\gamma)^{s-r}\Phi_s(t), \; 1 \le r \le s, \tag{3.3}$$

$$\Phi_r(\gamma t) \quad \le \quad \gamma^r\left(\frac{t^r}{r!}\right)^{1-\gamma}\Phi_r^\gamma(t). \tag{3.4}$$

Proof: The first inequality follows from Proposition 2.1. The second follows from Lemma 10.2 in Appendix. □

We note in passing that a sharp constant in (3.3) is obtained by replacing the binomial coefficient with the factor $\frac{s!}{r!}\left(\frac{e}{s-r}\right)^{s-r}$; see Eq. (4.4) in Barbour and Gnedin (2009).

Proposition 3.2. *The point process of r-records has the intensity measure*

$$\mathbb{P}[B_r(\mathrm{d}t) = 1] = \frac{r\,\Phi_r(t)}{t}\mathrm{d}t, \quad t \ge 0.$$

Proof: By the order statistic property of the Poisson process, if box j contains r balls at time $t+\mathrm{d}t$, the latest of them arrived within the interval $[t, t+\mathrm{d}t]$ with probability $r\mathrm{d}t/t$, regardless of p_j. Given $K_r(t+\mathrm{d}t) = k$, the probability that a r-record occurs in the interval is $rk\mathrm{d}t/t$. The intensity formula follows by taking the expectation. □

The Bernstein function is concave, subadditive, satisfies $\Phi(t)\uparrow\infty$ and $\Phi(t) \ll t$ as $t\to\infty$, where $f(t) \ll g(t)$ means that $\lim_{t\to\infty} f(t)/g(t) = 0$; see Schilling et al. (2012). In contrast, the expected number of r-tons can be less regular. Thus, as $t\to\infty$ the function $\Phi_r(t)$ may stay bounded, converge to ∞ or oscillate between a finite level and ∞; see Barbour and Gnedin (2009); Bogachev et al. (2008) for classification of the modes of behaviour. By considering the r-records, our first and foremost assumption will be that $\Phi_r(t)\to\infty$, which is equivalent to $\mathrm{Var}[K_r(t)]\to\infty$. It is known, see Barbour and Gnedin (2009), that if $\Phi_r(t)\to\infty$ holds for some r, then also for all $r' \le r$; and a sufficient condition for this is $p_j^r \ll \sum_{i\ge j+1} p_i^r,\; j\to\infty$. This excludes the light-tailed distributions (p_j) like Poisson or geometric.

In the case $\limsup \Phi_r(t) < \infty$, an approximation to the process of r-records can be sought on the contraction scale $\theta t,\; \theta\in[0,1]$, similar to Gnedin and Stark (2023) or to the well known Poisson limit for record times in the extreme-value theory Resnick (2008). But this case falls outside the scope of the present note.

4. Poisson approximation to r-records

Assuming $\Phi_r(t) \to \infty$, we aim at a local Poisson approximation for B_r. The strategy is to fix some initial time t_0, called in the sequel *lower cutoff* and treated as a large parameter, while introducing an auxiliary temporal variable $\theta > 0$ to measure the size of a properly scaled time window. These are related via a time change

$$t = t_0 + f(t_0)\,\theta, \tag{4.1}$$

where f is a scaling function defined by (4.5) below and satisfying $f(t_0) \ll t_0$. Denote for shorthand $h := f(t_0)\,\theta$. The focus is then on

$$\widehat{B}_r(\theta) := B_r((t_0, t_0 + h]) = B_r((t_0, t_0 + f(t_0)\,\theta]), \tag{4.2}$$

which is the number of r-records arriving within the time window $h \ll t_0$. The point process $\widehat{B}_r$ as a component of P is an instance of 'normalised ordered thinning' Böker and Serfozo (1983).

By independence across the boxes, $\widehat{B}_r(\theta)$ has the Poisson-binomial distribution with success probabilities

$$q_j := \mathbb{P}[\beta_{jr} \in (t_0, t_0 + h]] = \int_{t_0}^{t_0+h} e^{-tp_j} \frac{(tp_j)^{r-1}}{(r-1)!}\, p_j \mathrm{d}t, \tag{4.3}$$

and expectation

$$\lambda_r := \mathbb{E}[\widehat{B}_r(\theta)] = \sum_{j\geq 1} q_j = \int_{t_0}^{t_0+h} \frac{r\Phi_r(t)}{t} \mathrm{d}t, \tag{4.4}$$

as follows from Proposition 3.2.

We proceed with the scaling function

$$f(t_0) = \frac{t_0}{r\Phi_r(t_0)}. \tag{4.5}$$

The value (4.5) is the mean inter-arrival time in a homogeneous Poisson process with the rate equal to the instantaneous rate of B_r at time t_0.

Application of Theorems 1.C(i) and 2.M from Barbour et al. (1992) yields the following estimate of the total variation distance to the Poiss(θ) distribution,

$$d_{\mathrm{TV}}(\widehat{B}_r(\theta), \mathrm{Poiss}(\theta)) \leq \frac{1 - e^{-\lambda_r}}{\lambda_r} \sum_{j\geq 1} q_j^2 + |\lambda_r - \theta|. \tag{4.6}$$

The first part is the seminal Barbour-Eagleson bound on the total variation distance between $\widehat{B}_r(\theta)$ and Poiss(λ_r). The second part $|\lambda_r - \theta|$ is the bound for the distance between two Poisson distributions, see Lemma 1 in Ruzankin (2004), which appears as the interpolation error caused by adopting parameter θ in place of the genuine mean λ_r.

To estimate the first part of the approximation error (4.6), start with the inequality

$$q_j < \frac{e^{-t_0 p_j} p_j^r}{(r-1)!} \int_{t_0}^{t_0+h} t^{r-1} \mathrm{d}t < \frac{e^{-t_0 p_j} p_j^r}{r!}\, t_0^r\, \frac{2rh}{t_0} = \frac{e^{-t_0 p_j}(p_j t_0)^r}{r!} \frac{2\theta}{\Phi_r(t_0)},$$

which holds for all sufficiently large t_0. Squaring this and summing over the boxes, after some bookkeeping and application of (3.3) to $\Phi_{2r}(2t_0)/\Phi_r(t)$ we obtain

$$\sum_{j\geq 1} q_j^2 < 4\theta^2 2^{-2r} \binom{2r}{r} \frac{\Phi_{2r}(2t_0)}{\Phi_r^2(t_0)} < \frac{4\theta^2}{\Phi_r(t_0)}.$$

Quite satisfactorily, we see that $d_{\mathrm{TV}}(\widehat{B}_r(\theta), \mathrm{Poiss}(\lambda_r))$ tends to zero under the sole assumption $\Phi_r(t) \to \infty$ as $t \to \infty$.

Exploring $|\lambda_r - \theta|$ turns more intricate. From (4.4) and monotonicity

$$\lambda_r \le \frac{r\Phi_r(t_0)}{t_0^r}\int_{t_0}^{t_0+h} t^{r-1}\mathrm{d}r = \frac{\Phi_r(t_0)}{t_0^r}\left((t_0+h)^r - t_0^r\right)$$
$$= \Phi_r(t_0)\left(\frac{hr}{t_0} + \frac{r(r-1)h^2}{2t_0^2} + \cdots\right) \le \theta + \frac{(r-1)}{2r}\frac{\theta^2}{\Phi_r(t_0)} + \cdots,$$

where the remaining terms are of the smaller order as $t_0 \to \infty$. In the case $r = 1$ the estimate simplifies as $\lambda_1 \le \theta$. Similarly, a lower bound becomes

$$\lambda_r \ge \frac{\Phi_r(t_0+h)}{(t_0+h)^r}\left((t_0+h)^r - t_0^r\right)$$
$$= \frac{\Phi_r(t_0+h)}{\Phi_r(t_0)}\frac{t_0}{t_0+h}\theta - \frac{(r-1)}{2r}\frac{t_0^2}{(t_0+h)^2}\frac{\Phi_r(t_0+h)}{\Phi_r^2(t_0)}\theta^2 + \cdots.$$

To achieve $|\lambda_r - \theta| \to 0$, it will be sufficient to fulfil

$$\frac{\Phi_r(t_0+h)}{\Phi_r(t_0)} \to 1 \quad \text{as} \quad t_0 \to \infty, \quad \text{for} \quad h = \frac{t_0}{r\Phi_r(t_0)}\theta. \tag{4.7}$$

Expanding $\Phi_r(t_0+h) = \Phi_r(t_0) + \Phi_r'(t_0+uh)h,\ u \in [0,1]$, after some plain manipulations using $\Phi_r(t) \to \infty$ and monotonicity, we show that a sufficient condition for (4.7) to hold for each $\theta > 0$ is

$$\frac{\Phi_{r+1}(t)}{\Phi_r^2(t)} \to 0 \quad \text{as} \quad t \to \infty. \tag{4.8}$$

In Section 6, we shall discuss regularity conditions that imply (4.8).

The next result shows that, with probability close to one, there are no boxes that receive both rth and $(r+1)$st balls within the time window $[t_0, t_0+h]$.

Lemma 4.1. *If $\Phi_r(t) \to \infty$ and either (4.7) or (4.8) holds, then*

$$\sum_{j\ge1}\mathbb{P}[t_0 \le \beta_{jr} < \beta_{j,r+1} \le t_0+h] \to 0 \quad \text{as} \quad t_0 \to \infty.$$

Proof: The generic term is the probability of the event that box j receives rth record and some other balls, hence

$$\sum_{j\ge1}\mathbb{P}[t_0 \le \beta_{jr} < \beta_{j,r+1} \le t_0+h] = \sum_{j\ge1}\int_{t_0}^{t_0+h} e^{-tp_j}\frac{(tp_j)^{r-1}}{(r-1)!}p_j\left(1 - e^{-(t_0+h-t)p_j}\right)\mathrm{d}t$$
$$= \lambda_r - \sum_{j\ge1}\int_{t_0}^{t_0+h} e^{-(t_0+h)p_j}\frac{(tp_j)^{r-1}}{(r-1)!}p_j\mathrm{d}t = \lambda_r - \frac{\Phi_r(t_0+h)}{(t_0+h)^r}\left((t_0+h)^r - {t_0}^r\right).$$

Here, both terms converge to θ as before. □

We shall tacitly use two equivalent approaches to the functional convergence of point processes on $\mathbb{R}_+ := [0,\infty)$; see Section 3 in Resnick (2008) or Section 11.1 in Daley and Vere-Jones (2008). Recall that we identify $\widehat{B}_r$ with a random element of the space $M_p(\mathbb{R}_+)$ of locally finite point measures on $\mathbb{R}_+$ by writing $\widehat{B}_r(A) = B_r(t_0+f(t_0)(A))$ for a Borel $A \subset \mathbb{R}_+$. We endow the space $M_p(\mathbb{R}_+)$ with the topology of vague convergence. According to Lemma 11.1.XI in Daley and Vere-Jones (2008), the weak convergence on $M_p(\mathbb{R}_+)$ is equivalent to the weak convergence of the corresponding cumulative processes in the Skorokhod space $D(\mathbb{R}_+)$ of càdlàg functions endowed with the J_1-topology. Thus, the random measures $\widehat{B}_r$ converge in distribution on $M_p(\mathbb{R}_+)$ if and only if the random processes $(\widehat{B}_r(\theta), \theta \ge 0)$ converge in distribution on $D(\mathbb{R}_+)$. Furthermore, by Theorem 11.1.VII in Daley and

Vere-Jones (2008), both types of convergence are equivalent to the convergence of finite-dimensional distributions of $(\widehat{B}_r(\theta), \theta \geq 0)$ at the continuity points of the limit.

We give next the main result of this paper.

Theorem 4.2. *If $\Phi_r(t) \to \infty$ as $t \to \infty$ and either of the conditions (4.7) or (4.8) holds, then as $t_0 \to \infty$ the process $(\widehat{B}_r(\theta),\ \theta \geq 0)$ converges in distribution to a Poisson process of unit rate.*

Proof: We first prove that for given $\theta > 0$ the restriction of $\widehat{B}_r$ on $[0, \theta]$ converges to a Poisson process of unit rate on this interval. To this end, it suffices to show that the number and positions of atoms of $\widehat{B}_r$ on $[0, \theta]$ converge to the number and positions of atoms of the unit rate Poisson process on $[0, \theta]$. The number of points of $\widehat{B}_r$ on $[0, \theta]$ converges to Poiss(θ), because both parts of the error (4.6) approach 0. For the convergence of positions recall a familiar re-statement of the order statistic property: the Poisson process on $[0, \theta]$ can be characterised as a mixed binomial process, whose number of points has Poiss(θ) distribution; see Section 3 in Kallenberg (2017). Let A be the event that no box receives the rth and $(r+1)$st balls within the same time window $[t_0, t_0 + h]$. Given A, the process $\widehat{B}_r$ restricted to $[0, \theta]$ is mixed binomial, because a single arrival to a box within the time window is uniformly distributed there, and under our scaling (4.1), (4.5) the arrival becomes uniformly distributed on $[0, \theta]$. By Lemma 4.1 $\mathbb{P}[A] \to 1$ as $t_0 \to \infty$, which together with the convergence of $\widehat{B}_r(\theta)$ implies convergence to the Poisson process on $[0, \theta]$. Since this holds for every θ, the proof is complete. □

5. A conditional limit theorem

So far we have been concerned with the features of the process of r-records averaged over the history prior to the cutoff time t_0. However, the online observer, being aware of the history, will see different patterns depending on the allocation at time t_0. Two extreme cases are accumulation of all balls in just one box and the situation when each arrival before t_0 hits a different box. In this section, we show that typically the past allocation does not impact the future. That is, with high probability the same Poisson approximation works even conditionally on the full history of the occupancy process.

Let $\mathcal{F}_{t_0} := \sigma(P_j(t), j \geq 1, t \leq t_0)$ be the sigma-algebra generated by the occupancy counts up to and including time t_0.

Proposition 5.1. *Under the assumptions of* Theorem 4.2*, as $t_0 \to \infty$ the conditional distribution of $(\widehat{B}_r(\theta), \theta \geq 0)$ given $\mathcal{F}_{t_0}$ converges in probability to the distribution of a Poisson process of unit rate.*

Proof: According to Lemma 4.1

$$\widehat{B}_r(\theta) = \sum_{j \geq 1} \mathbb{1}_{\{\beta_{jr} \in (t_0, t_0 + f(t_0)\theta\}} = \sum_{j \geq 1} \mathbb{1}_{\{P_j(t_0 + f(t_0)\theta) \geq r, P_j(t_0) < r\}}$$
$$= \sum_{j \geq 1} \mathbb{1}_{\{P_j(t_0 + f(t_0)\theta) \geq r, P_j(t_0) = r-1\}} + \varepsilon,$$

where $\varepsilon = \varepsilon(\theta)$ converges to zero in probability locally uniformly with respect to θ, as $t_0 \to \infty$. Denote by $P'_j(t) := P_j(t + t_0) - P_j(t_0)$, $j \geq 1$, $t \geq 0$ and note that $(P'_j, j \geq 1)$ is independent of $\mathcal{F}_{t_0}$. Thus we need to check that the conditional distribution of the process

$$\theta \mapsto \sum_{j : P_j(t_0) = r-1} \mathbb{1}_{\{P'_j(f(t_0)\theta) \geq 1\}}, \tag{5.1}$$

given $\mathcal{F}_{t_0}$, converges to the desired Poisson limit. Note that for fixed θ, given $\mathcal{F}_{t_0}$, the variable (5.1) is a sum of independent Bernoulli variables. Arguing in the same way as in the proof of Theorem 4.2, we see that it suffices to prove that the conditional distribution of (5.1) converges in probability to Poiss(θ), for every fixed $\theta > 0$. Put

$$E_{t_0,r} := \mathbb{E}\left[\sum_{j:P_j(t_0)=r-1} \mathbb{1}_{\{P'_j(f(t_0)\theta)\geq 1\}}\Big|\mathcal{F}_{t_0}\right] = \sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})\mathbb{1}_{\{P_j(t_0)=r-1\}},$$

and note that

$$\mathbb{E}[E_{t_0,r}] = \sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!}$$
$$= \theta - \sum_{j\geq 1}(p_j f(t_0)\theta - 1 + e^{-p_j f(t_0)\theta})e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!}.$$

Using an elementary inequality $x - 1 + e^{-x} \leq x(1-e^{-x})$, $x \geq 0$, we see that

$$\begin{aligned}
0 &\leq \sum_{j\geq 1}(p_j f(t_0)\theta - 1 + e^{-p_j f(t_0)\theta})e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!} \\
&\leq \sum_{j\geq 1} p_j f(t_0)\theta(1-e^{-p_j f(t_0)\theta})e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!} \\
&= \frac{r f(t_0)\theta}{t_0}\left(\Phi_r(t_0) - \frac{t_0^r}{(t_0+f(t_0)\theta)^r}\Phi_r(t_0+f(t_0)\theta)\right) \qquad (5.2)\\
&= 1 - \frac{t_0^r}{(t_0+f(t_0)\theta)^r}\frac{\Phi_r(t_0+f(t_0)\theta)}{\Phi_r(t_0)} \to 0,
\end{aligned}$$

by the assumption (4.7). Similarly, using $(1-e^{-x})^2 \leq x(1-e^{-x})$, $x \geq 0$, we obtain from (5.2)

$$\mathrm{Var}[E_{t_0,r}] \leq \sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})^2 e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!}$$
$$\leq \sum_{j\geq 1} p_j f(t_0)\theta(1-e^{-p_j f(t_0)\theta})e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!} \to 0. \qquad (5.3)$$

Thus, $E_{t_0,r}$ converges to θ in probability as $t_0 \to \infty$ by Chebyshev's inequality. In view of (4.6), given $\mathcal{F}_{t_0}$,

$$d_{\mathrm{TV}}\left(\sum_{j:P_j(t_0)=r-1}\mathbb{1}_{\{P'_j(f(t_0)\theta)\geq 1\}}, \mathrm{Poiss}(\theta)\right)$$
$$\leq \frac{1-e^{-E_{t_0,r}}}{E_{t_0,r}}\sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})^2\mathbb{1}_{\{P_j(t_0)=r-1\}} + |E_{t_0,r}-\theta|.$$

The right-hand side converges to zero in probability by Markov's inequality, since

$$\mathbb{E}\left[\sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})^2\mathbb{1}_{\{P_j(t_0)=r-1\}}\right] = \sum_{j\geq 1}(1-e^{-p_j f(t_0)\theta})^2 e^{-p_j t_0}\frac{(p_j t_0)^{r-1}}{(r-1)!} \to 0,$$

as in (5.3). $\square$

6. Regularity and growth

If $\Phi_r'(t) \geq 0$ then $\Phi_{r+1}(t) \leq \frac{r}{r+1}\Phi_r(t)$ by (3.2), hence (4.8) is always true if $t \to \infty$ along a sequence of increase points of Φ_r (subject to the only condition $\Phi_r(t) \to \infty$). However, we could neither verify (4.8) in full generality nor construct a counter-example in terms of (p_j) or Φ. In this section, we give various conditions to ensure (4.8) or directly (4.7), hence the Poisson convergence of the process of r-records, in accord with Theorem 4.2.

Example 6.1 (Regular variation). Karlin's condition of regular variation Karlin (1967) reads as

$$\nu[x, 1] \sim x^{-\alpha}\ell(1/x), \quad x \to 0+, \tag{6.1}$$

where $0 \leq \alpha \leq 1$ is the *index* and ℓ is some function slowly varying at ∞. In the *proper* case $0 < \alpha < 1$ this implies that, asymptotically, the expectations of the r-ton counts only differ by constant factors:

$$\Phi(t) \sim \Gamma(1-\alpha)\, t^{\alpha}\ell(t) \quad \text{and} \quad \Phi_r(t) \sim \frac{\alpha\Gamma(r-\alpha)}{r!\Gamma(1-\alpha)}\,\Phi(t). \tag{6.2}$$

In the case of *rapid* variation $\alpha = 1$ the slowly varying factor must satisfy $\ell(t) \to 0$ as $t \to \infty$ (to agree with $\sum_{j\geq 1} p_j = 1$), in which case the asymptotic formulas are

$$\Phi(t) \sim t\ell_1(t), \quad \Phi_1(t) \sim \Phi(t) \quad \text{and} \quad \Phi_r(t) \sim \frac{1}{r(r-1)} t\ell(t) \quad \text{for} \quad r \geq 2, \tag{6.3}$$

where $\ell_1 \gg \ell$ is another slowly varying function, thus $\Phi_1 \gg \Phi_r$ for $r \geq 2$.

Speaking of the case of *slow* variation $\alpha = 0$ we shall mean a slightly stronger condition

$$\nu_1[0, x] \sim x\ell_0(1/x), \quad x \to 0+, \tag{6.4}$$

with slowly varying $\ell_0(t) \to \infty$, $t \to \infty$; then (6.1) holds with some $\ell \gg \ell_0$, and

$$\Phi_r(t) \sim \frac{1}{r}\ell_0(t), \quad \Phi(t) \sim \ell(t), \quad t \to \infty, \tag{6.5}$$

so the Φ_r's are of the same order while $\Phi(t) \gg \Phi_r(t)$ for all $r \geq 1$. Note that the mean number $r\Phi_r(t)$ of balls contained in r-ton boxes is asymptotic to the same function $\ell_0(t)$ regardless of r. Summarizing, we conclude that (4.7) holds under (6.1) if $0 < \alpha \leq 1$ or under (6.4) if $\alpha = 0$. See Gnedin et al. (2007) for conditions of slow variation expressed in terms of (p_j) or ν.

Example 6.2 (A weaker form of regular variation). Under the regular variation (understood as (6.4) if $\alpha = 0$) we obviously have

$$\liminf_{\gamma\downarrow 1} \liminf_{t\to\infty} \frac{\Phi_r(\gamma t)}{\Phi_r(t)} \geq 1, \tag{6.6}$$

which in turn implies (since $\Phi_r(t)/t^r$ decreases) the desired (4.7) for all $\theta > 0$. Condition (6.6) itself is well known in the Tauberian theory; see p. 19 in Bingham et al. (1989). If it holds we have

$$\lim_{\gamma\to 1} \limsup_{t\to\infty} |\Phi_r(\gamma t)/\Phi_r(t) - 1| = 0.$$

The latter asymptotic condition defines the class of *pseudo-regularly varying* functions treated in a recent monograph Buldygin et al. (2018).

Example 6.3 (Slow decrease). Assuming $\Phi_r(t) \to \infty$, it is easy to check that

$$\frac{\Phi_{r+1}(t)}{\Phi_r^2(t)} \to 0 \quad \Longleftrightarrow \quad \left(\frac{t}{\Phi_r(t)}\right)' \to 0. \tag{6.7}$$

Writing

$$\frac{1}{\Phi_r(t)} = \frac{1}{t}\int_a^t \left(\frac{u}{\Phi_r(u)}\right)' \mathrm{d}u + o(1),$$

we have the left-hand side converging to 0, which only forces the Cesáro summability, hence a priori does not guarantee convergence of the integrand. Thus, some extra assumption to limit the variability of Φ_r seems inevitable.

In the classic text Hardy (1949, Section 6.2) a function g is called *slowly decreasing* if

$$\liminf(g(\gamma t) - g(t)) \geq 0$$

for $t \to \infty$, $\gamma \geq 1$ and $\gamma \to 1$. A sufficient condition for this is $g'(t) > -c/t$. Applying Hardy (1949) Theorem 68, condition (4.8) holds if any of the involved functions Φ_r and Φ_{r+1} is slowly decreasing.

Example 6.4 (A bounded ratio). Given $\Phi_r(t) \to \infty$, a sufficient condition for (4.8) is

$$\limsup_{t\to\infty} \frac{\Phi_{r+1}(t)}{\Phi_r(t)} < \infty. \tag{6.8}$$

Now suppose that for some $1 < \gamma_1 < \gamma_2$

$$\liminf_{t\to\infty} \sup_{\gamma\in[\gamma_1,\gamma_2]} \frac{\Phi_{r+1}(\gamma t)}{\Phi_{r+1}(t)} > 0. \tag{6.9}$$

Then from (3.3) for any such γ

$$\frac{\Phi_{r+1}(\gamma t)}{\Phi_r(t)} \leq \frac{\gamma^{r+1}}{(r+1)(\gamma-1)} \leq \frac{\gamma_2^{r+1}}{(r+1)(\gamma_1-1)},$$

whence choosing ε less than the limes inferior in (6.9), for large enough t

$$\frac{\Phi_{r+1}(t)}{\Phi_r(t)} < \frac{\gamma_2^{r+1}}{\varepsilon(r+1)(\gamma_1-1)},$$

so (6.8) is implied by (6.9). By the same token, (6.8) also follows if condition (6.9) is imposed on Φ_r instead of Φ_{r+1}.

Example 6.5 (Superlogarithmic growth). Finally, we show that the superlogarithmic growth condition

$$\Phi_r(t) \gg \log t \tag{6.10}$$

is sufficient for (4.7). Indeed, from (3.3) and (3.4) we may estimate (4.8) as

$$\begin{aligned}\frac{\Phi_{r+1}(t)}{\Phi_r^2(t)} &\leq \frac{\Phi_r(\gamma t)}{(r+1)\gamma^r(1-\gamma)\Phi_r^2(t)} \\ &\leq \frac{2\,t^{r(1-\gamma)}\,\Phi_r^{\gamma}(t)}{(r!)^{1-\gamma}(r+1)(1-\gamma)\,\Phi_r^2(t)} \leq \frac{t^{r(1-\gamma)}}{(1-\gamma)\Phi_r^{2-\gamma}(t)}.\end{aligned}$$

If $\Phi_r(t) = g(t)\log t$ with some $g(t) \to \infty$, with the choice of the parameter $\gamma = 1 - (\log t)^{-1}$ the estimate becomes

$$\frac{\Phi_{r+1}(t)}{\Phi_r^2(t)} = \frac{e^r}{g(t)}\{g(t)\log t\}^{-1/\log t} \to 0.$$

A similar check shows that the bound does not converge to 0 whenever $\limsup g(t) < \infty$. This only means that the bound is inconclusive, because $\Phi_{r+1}(t)/\Phi_r^2(t) \to 0$ may still hold for Φ_r arbitrarily slowly growing, e.g. under the condition of slow variation (6.4). In the view of discussion in Example 6.3, it looks unexpected that the growth rate (6.10) assumes the role of a Tauberian condition.

7. Multivariate processes of records

For the joint convergence of r-record processes to Poisson processes with constant rates, one needs a common scaling function f in (4.1) to serve different types of records. But this is only possible under a condition of regular variation. Indeed, note that $L_r(t) := \Phi_r(t)/(t^r/r!)$ is the Laplace transform of the measure $x^r\nu(\mathrm{d}x)$ on $[0,1]$, hence relation (3.2) between Φ_r and Φ_{r+1} becomes

$$L_r(t) = \int_t^\infty L_{r+1}(u)\mathrm{d}u,$$

and therefore

$$\frac{\Phi_{r+1}(t)}{\Phi_r(t)} = \frac{tL_{r+1}(t)}{(r+1)\int_t^\infty L_{r+1}(u)\mathrm{d}u}.$$

If the ratio converges as $t \to \infty$ to some constant, say $(r-\alpha)/(r+1)$, then by Karamata's theorem, see Theorem 1.6.1 in Bingham et al. (1989), L_{r+1} is regularly varying with index $-(r+1)+\alpha$, thus both functions Φ_{r+1} and Φ_r are regularly varying with index α.

Theorem 7.1. *Fix $m \geq 2$. Suppose Φ is regularly varying with index $0 \leq \alpha < 1$ (respectively, $\alpha = 1$). Then with the scaling function $f(t_0)$ the collection of processes $(\widehat{B}_1, \dots, \widehat{B}_m)$ (respectively, $(\widehat{B}_2, \dots, \widehat{B}_m)$) converge jointly in distribution, as $t_0 \to \infty$, to independent homogeneous Poisson processes with rates $\rho_1, \dots, \rho_m$ (respectively, $\rho_2, \dots, \rho_m$), where*

$$\begin{aligned}
&\text{for } 0<\alpha<1: && f(t_0) = \frac{t_0}{\Phi(t_0)} \text{ and } && \rho_r = \frac{\alpha\Gamma(r-\alpha)}{(r-1)!\,\Gamma(1-\alpha)}, && r \geq 1;\\
&\text{for } \alpha=1: && f(t_0) = \frac{t_0}{\Phi_2(t_0)} \text{ and } && \rho_r = \frac{2}{r-1}, && r \geq 2;\\
&\text{for } \alpha=0 \text{ under (6.4)}: && f(t_0) = \frac{t_0}{\Phi_1(t_0)} \text{ and } && \rho_r = 1, && r \geq 1.
\end{aligned}$$

Proof: The rates are clear from (6.2), (6.3) and (6.5). The marginal convergence of the scaled r-record processes follows from Theorem 4.2. For the joint convergence, we need to show that for each r the processes $\widehat{B}_1, \dots, \widehat{B}_r$ converge jointly to a multivariate Poisson process. To that end, we apply Corollary 11.2.VII in Daley and Vere-Jones (2008). With the marginal convergence at hand, the convergence of intensity measures, see Eq. (11.2.11) in Daley and Vere-Jones (2008), holds automatically. It remains to verify the condition

$$\sum_j \mathbb{P}\left[\sum_{s=1}^r \delta_{\beta_{js}}([t_0, t_0 + f(t_0)\theta]) \geq 2\right] \to 0 \quad \text{as} \quad t_0 \to \infty,$$

which is Eq. (11.2.10) in Daley and Vere-Jones (2008). But this is satisfied by Lemma 4.1, since

$$\sum_{s=1}^r \delta_{\beta_{js}}([t_0, t_0+h]) \geq 2$$

means precisely that $t_0 \leq \beta_{js} < \beta_{j,s+1} \leq t_0 + h$, for some $1 \leq s < r$. □

A multivariate counterpart of (4.6) can be derived from estimates in Roos (1999).

In the case $\alpha = 1$ of rapid variation the scaling in Theorem 7.1 is not suitable for the process B_1, because the 1-records are then much more frequent than records of any other type $r > 1$.

8. The discrete-time model: de-Poissonisation

We turn next to the occupancy scheme, where balls are thrown at discrete times. De-Poissonisation is a folk name for methods aiming to derive properties of 'fixed-n' models from their 'Poiss(n)' counterparts. For $K(t)$ and $K_r(t)$ the de-Poissonisation relies on concentration properties of the

Poisson distribution, which enable efficient coupling where the values of the variables in both models are determined by much the same bulk of balls; see, for example, Section 6.2 in Buraczewski et al. (2021). For r-record processes such coupling can be constructed if $\Phi_r(t) \ll t^{1/2}$, but cannot in models with $\Phi_r(t) \gg t^{1/2}$ where the time window has an order smaller than the square root fluctuations. Fortunately, if $\Phi_r(t) \to \infty$ the de-Poissonisation can be universally justified due to the square root insensitivity to the choice of the lower cutoff t_0.

Denote $\mathcal{K}_n$ the total number of occupied boxes when n balls are allocated and $\mathcal{K}_{n,r}$ the number of boxes occupied by exactly r balls, $r \geq 1$. Thus,

$$K(t) = \mathcal{K}_{P(t)}, \quad K_r(t) = \mathcal{K}_{P(t),r}, \quad r \geq 1.$$

For the moments we shall use the approximate formulas

$$\mathbb{E}[\mathcal{K}_n] = \Phi(n) + o(1),\ \mathbb{E}[\mathcal{K}_{n,r}] = \Phi_r(n) + o(1),\ n \to \infty, \tag{8.1}$$

found in Bogachev et al. (2008) along with an explicit fixed-n analogue of (3.1).

We shall also need a vector of occupancy counts $(\mathcal{P}_j(n), j \geq 1)$, where $\mathcal{P}_j(n)$ is the number of balls in box j after allocating n balls. A discrete-time counterpart of the record time β_{jr} is

$$\widetilde{\beta}_{jr} := \min\{n \in \mathbb{N} : \mathcal{P}_j(n) = r\}, \quad j \geq 1, \quad r \geq 1,$$

and we have an obvious relation $\widetilde{\beta}_{jr} = P(\beta_{jr})$. Therefore, the point process of discrete-time r-records

$$\mathcal{B}_r := \sum_{j\geq 1} \delta_{\widetilde{\beta}_{jr}}, \quad r \geq 1,$$

is the random measure on $\mathbb{N}$, which may be represented as the push-forward of B_r under the random mapping $t \mapsto P(t)$, that is $\mathcal{B}_r = B_r \circ P^{-1}$. Note that B_r and P in this representation are dependent making analysis of $\mathcal{B}_r$ harder. Let $\widehat{\mathcal{B}}_r(\theta) := \mathcal{B}_r((n_0, n_0 + f(n_0)\,\theta]),\ \theta \geq 0$.

Theorem 8.1. *Fix $r \geq 1$. If $\Phi_r(t) \to \infty$ as $t \to \infty$ and either of the conditions (4.7) or (4.8) is satisfied, then with the scaling function $f(n_0) = n_0/(r\Phi_r(n_0))$, as $n_0 \to \infty$ the process $\widehat{\mathcal{B}}_r$ converges in distribution to a Poisson process of unit rate.*

To prove the theorem we need two auxiliary lemmas. Recall that we assume $p_1 \geq p_2 \geq \cdots$. For $t > 1$, let $j(t) \in \mathbb{N}$ be the unique index such that

$$p_{j(t)} > \frac{2\log t}{t} \geq p_{j(t)+1}.$$

We regard a box with $j \leq j(t_0)$ as 'popular' since for large time $t_0 = n_0$, it is likely to contain more than r balls, both in Poisson and discrete time occupancy schemes. The intuition suggests that the popular boxes make a negligible contribution to the normalised record processes. We prefer to justify this in the continuous time setting, the other case being completely analogous.

Lemma 8.2. *For $r \geq 1$,*

$$\mathbb{P}[P_j(t) \leq r \text{ for some } j \leq j(t)] \to 0 \quad \text{as} \quad t \to \infty.$$

Proof: For $j \leq j(t)$ we have $\mathbb{E}[P_j(t)] \geq 2\log t$. The assertion follows from the elementary estimate

$$\mathbb{P}[P_j(t) \leq r] \leq \frac{(r+1)}{r!}\frac{(2\log t)^r}{t^2} \quad \text{for} \quad t > e^{r/2}, \tag{8.2}$$

by observing that $j(t) \leq t/(2\log t)$ because there are at most $1/p$ boxes with probability larger p. Bound (8.2) is a consequence of the chain of estimates: for $r \leq 2\log t \leq \theta$,

$$\mathbb{P}[\text{Poiss}(\theta) \leq r] \leq (r+1)\mathbb{P}[\text{Poiss}(\theta) = r] = \frac{(r+1)}{r!}e^{-\theta}\theta^r \leq \frac{(r+1)}{r!}e^{-2\log t}(2\log t)^r,$$

where for the last inequality we used that $\theta \mapsto e^{-\theta}\theta^r$ is decreasing for $\theta \geq r$. □

By the lemma, a truncated version of the normalised r-record process,

$$\theta \mapsto \sum_{j \geq j(t_0)} \delta_{\beta_{jr}}((t_0, t_0 + f(t_0)\,\theta]), \quad \theta \geq 0, \tag{8.3}$$

for large t_0 coincides with $\widehat{B}_r$ with probability close to one.

For $n = 1, 2, \dots$ the discrete time allocations are naturally identified with the configuration of balls in boxes at random times S_n, where $S_n := \min\{t \geq 0 : \ P(t) = n\}$ is the n-th arrival time in the Poisson process P. Consider the process

$$\widehat{B}^*_r(\theta) := \sum_{j \geq 1} \delta_{\beta_{jr}}((S_{n_0}, S_{n_0} + f(n_0)\theta]), \ \theta \geq 0,$$

which has the same window size as $\widehat{B}_r(\theta)$ but the lower cutoff $t_0 = n_0$ is replaced by the n_0th point of P. Replacing n_0 by S_{n_0} is a nontrivial step which turns possible in full generality due to a key observation from Barbour and Gnedin (2009) that the counts of balls within unpopular boxes at large times $t_0 = n_0$ are similar for both Poisson and discrete-time schemes.

Lemma 8.3. *Under conditions of* Theorem 4.2, *for $r \geq 1$, as $n_0 \to \infty$ the process $(\widehat{B}^*_r(\theta), \ \theta \geq 0)$ converges in distribution to a Poisson process with unit rate.*

Proof: Lemma 8.2 implies that it is sufficient to restrict summation over $j > j(n_0)$. Since $P(S_{n_0}) = n_0$, we may apply the total variation estimate (2.6) from Barbour and Gnedin (2009), which in our notation reads as

$$\begin{aligned} d_{\text{TV}}\,((\mathcal{P}_j(n_0), \ j > j(n_0), \ (P_j(n_0), \ j > j(n_0)) \\ = d_{\text{TV}}\,((P_j(S_{n_0}), \ j > j(n_0)), (P_j(n_0), \ j > j(n_0)) \leq \sum_{j > j(n_0)} p_j. \end{aligned} \tag{8.4}$$

Letting $n_0 \to \infty$ gives $j(n_0) \to \infty$ hence the right side approaches 0. That is to say, the occupancy numbers in unpopular boxes are likely to be the same at times n_0 and S_{n_0}. On the other hand, the r-record process after time n_0 depends on the history only through the allocation of balls at this time. The assertion now follows from the convergence of $\widehat{B}_r$. □

We are now in position to prove Theorem 8.1.

Proof: The r-record process is nondecreasing, therefore the result will follow from Lemma 8.3 by a sandwich argument provided we can justify that

$$\mathbb{P}[\widehat{B}^*_r(\theta_i - \varepsilon) \leq \widehat{B}_r(\theta_i) \leq \widehat{B}^*_r(\theta_i + \varepsilon), i = 1, \dots, m] \to 1, \ n_0 \to \infty,$$

for every fixed $m \in \mathbb{N}$, $0 \leq \theta_1 < \dots < \theta_m$ and $\varepsilon > 0$. Clearly, it suffices to consider the case $m = 1$. Note that the event

$$\Big\{P\big((S_{n_0}, S_{n_0} + f(n_0)(\theta_1 - \varepsilon)]\big) \subseteq (n_0, n_0 + f(n_0)\theta_1] \subseteq P\big((S_{n_0}, S_{n_0} + f(n_0)(\theta_1 + \varepsilon)]\big)\Big\}$$

implies the event $\{\widehat{B}^*_r(\theta_1 - \varepsilon) \leq \widehat{B}_r(\theta_1) \leq \widehat{B}^*_r(\theta_1 + \varepsilon)\}$. But the former, by monotonicity of P and the fact that $P(S_{n_0}) = n_0$, coincides with the event

$$\Big\{P(S_{n_0} + f(n_0)(\theta_1 - \varepsilon)) - P(S_{n_0}) \leq f(n_0)\theta_1 \leq P(S_{n_0} + f(n_0)(\theta_1 + \varepsilon)) - P(S_{n_0})\Big\}.$$

The probability of this event tends to one by the weak law of large numbers for $(P(S_{n_0} + t) - P(S_{n_0}), t \geq 0)$, which is again a unit rate Poisson process by the strong Markov property of P. □

A discrete-time version of Theorem 7.1 also holds true. The proof proceeds along the same lines, does not involve any new ideas and is therefore omitted.

Theorem 8.4. *Under the assumptions of* Theorem 7.1 *the random processes $\widehat{\mathcal{B}}_r$ converge jointly in distribution, as $n_0 \to \infty$, to independent homogeneous Poisson processes with rates ρ_r.*

9. The discrete-time model: the use of exchangeability

Records arriving at large times typically emerge due to unpopular boxes that rarely change their occupancy status. Therefore, at later stages t all relatively recent r-record balls, for fixed r, are likely to belong to boxes that contain exactly r balls at time t. This intuitive feature, combined with the intrinsic exchangeability of the occupancy scheme, leads to another approach to the Poisson approximation, which we sketch in the discrete time setting and under the assumption of regular variation. The aim here is to show weak convergence to a multivariate Poisson distribution of the random vector or record counts $(\mathcal{B}_s((n_0, n_1]),\ 1 \leq s \leq r)$ for a suitable time window $(n_0, n_1]$ with $n_0 \to \infty$. For simplicity we also exclude the case $\alpha = 1$ where the multivariate approximation holds for $r \geq 2$. Under multivariate Poisson distribution we understand the distribution of an integer vector with independent univariate Poisson components.

For given n_1 define $\mathcal{C}_r$ to be the random set of balls contained in the r-ton boxes present at time n_1. Formally, $\mathcal{C}_r$ is a random point measure on $\{1, \dots, n_1\}$, though we do not include n_1 in the notation. By the definition, n is an atom of $\mathcal{C}_r$ if and only if $n = \widetilde{\beta}_{jr} \leq n_1 < \widetilde{\beta}_{j,r+1}$ for some j, and by virtue of (8.1) this event has probability

$$\frac{r}{n_1}\,\mathbb{E}[\mathcal{K}_{n_1,r}] = \frac{r}{n_1}\,\Phi_r(n_1) + o\left(\frac{1}{n_1}\right),$$

which is the same for all $n \leq n_1$. To compare $\mathcal{C}_r$ with r-record counts in terms of their means, choose $n_0 < n_1$ and observe that

$$\mathbb{E}[\mathcal{C}_r((n_0, n_1])] = \frac{r(n_1 - n_0)}{n_1}\,\Phi_r(n_1) + o\left(\frac{n_1 - n_0}{n_1}\right), \tag{9.1}$$

$$\mathbb{E}[\mathcal{B}_r((n_0, n_1])] = \sum_{n=n_0+1}^{n_1} \frac{r\Phi_r(n)}{n} + o\left(\frac{n_1 - n_0}{n_1}\right). \tag{9.2}$$

Setting $n_1 = n_0 + \lfloor f(n_0)\theta \rfloor$ with the scaling function as in (4.5), we obtain under conditions of Theorem 4.2 that $\mathbb{E}[\mathcal{B}_r((n_0, n_1]) - \mathcal{C}_r((n_0, n_1])] \to 0$ as $n_0 \to \infty$ locally uniformly in $\theta \geq 0$.

The multivariate Poisson approximation to records will be justified in several steps.

Step 1: Approximation of $(\mathcal{B}_s((n_0, n_1]),\ 1 \leq s \leq r)$ by $(\mathcal{C}_s((n_0, n_1]),\ 1 \leq s \leq r)$.

The point process $\mathcal{C}_r$ is much better tractable than $\mathcal{B}_r$, because $\mathcal{C}_r$ is exchangeable, that is, invariant under re-labelling of balls $1, \dots, n_1$ by permutations. If a generic ball n at time n_1 belongs to an s-ton for some $s \leq r$, then n is also a s_1-record time for some $s_1 \leq s$. This gives pointwise inequality between measures

$$\mathcal{C}_1 + \dots + \mathcal{C}_s \leq \mathcal{B}_1 + \dots + \mathcal{B}_s, \quad s \geq 1. \tag{9.3}$$

Applying the Markov inequality, the total variation distance is estimated in terms of the means as

$$\begin{aligned}
&d_{\mathrm{TV}}((\mathcal{C}_s((n_0, n_1]), 1 \leq s \leq r), (\mathcal{B}_s((n_0, n_1]), 1 \leq s \leq r)) \\
&\quad \leq \mathbb{P}[\mathcal{C}_s((n_0, n_1]), 1 \leq s \leq r) \neq (\mathcal{B}_s((n_0, n_1]), 1 \leq s \leq r))] \\
&\quad = \mathbb{P}\left[\sum_{s=1}^{r_1} \mathcal{C}_s((n_0, n_1]) \neq \sum_{s=1}^{r_1} \mathcal{B}_s((n_0, n_1]) \ \text{ for some } \ r_1 \leq r\right] \\
&\quad \leq \sum_{r_1=1}^{r} \mathbb{P}\left[\sum_{s=1}^{r_1} (\mathcal{B}_s((n_0, n_1]) - \mathcal{C}_s((n_0, n_1])) \geq 1\right] \\
&\quad \leq \sum_{r_1=1}^{r} \mathbb{E}\left[\sum_{s=1}^{r_1} \{\mathcal{B}_s((n_0, n_1]) - \mathcal{C}_s((n_0, n_1])\}\right].
\end{aligned}$$

From (9.1), (9.2), the bound approaches zero as $n_0 \to \infty$ and $n_1 = n_0 + \lfloor \theta n_0/\Phi(n_0) \rfloor$, provided that Φ satisfies the condition of regular or slow variation with $\alpha \in [0, 1)$. For the sequel we assume this holds, and focus on approximating $(\mathcal{C}_s((n_0, n_1]),\ 1 \le s \le r)$.

Step 2: Multinomial approximation to the conditional distribution of $(\mathcal{C}_s((n_0, n_1]),\ 1 \le s \le r)$ given $\{\mathcal{K}_{n_1,s} = k_s,\ 1 \le s \le r\}$.

The processes $\mathcal{C}_1, \dots, \mathcal{C}_r$ do not have common points and have the following structure. Conditionally on $\mathcal{K}_{n_1,s} = k_s,\ 1 \le s \le r$, they jointly can be represented by sampling without replacement from an urn with sk_s balls of colour s (that occupy s-ton boxes at time n_1) and $n_1 - \sum_{s=1}^r sk_s$ uncoloured balls. In particular, the conditional distribution of the random vector $(\mathcal{C}_s((n_0, n_1]),\ 1 \le s \le r)$ (complemented with the uncoloured component) given $\{\mathcal{K}_{n_1,s} = k_s,\ 1 \le s \le r\}$ is a multivariate hypergeometric distribution with parameters

$$n_1 - n_0;\ \left(k_1, 2k_2, \dots, rk_r, n_1 - \sum_{s=1}^r sk_s\right).$$

By the Diaconis-Freedman bound, see Theorem 4 in Diaconis and Freedman (1980), the total variation distance between this multivariate hypergeometric distribution and its multinomial counterpart with parameters

$$n_1 - n_0;\ \left(\frac{k_1}{n_1}, \frac{2k_2}{n_2}, \dots, \frac{rk_r}{n_1}, \frac{n_1 - \sum_{s=1}^r sk_s}{n_1}\right) \tag{9.4}$$

is at most $2(r+1)(n_1 - n_0)/n_1$. Remarkably, the bound does not depend on $k_1, \dots, k_r$. According to Theorem 1 in McDonald (1980), the total variation distance between the first r components of this multinomial distribution and the r-variate Poisson distribution

$$\text{Poiss}\left(\frac{(n_1 - n_0)}{n_1}\, sk_s,\ 1 \le s \le r\right)$$

is bounded by $2(n_1 - n_0)\left(\sum_{s=1}^r \frac{sk_s}{n_1}\right)^2$ which goes to zero.

Step 3: Mixed multivariate Poisson approximation of $(\mathcal{C}_s((n_0, n_1]),\ 1 \le s \le r)$.

To eliminate conditioning we use the elementary consequence of the fact that the total variation distance derives from a norm. This fact is implicit in Diaconis and Freedman (1980), see the proof of Theorem 3 therein.

Lemma 9.1. *For two families of probability measures* $(F_\alpha), (G_\alpha)$ *the convex mixtures satisfy*

$$d_{\text{TV}}\left(\sum_\alpha a_\alpha F_\alpha, \sum_\alpha a_\alpha G_\alpha\right) \le \sum_\alpha a_\alpha\ d_{\text{TV}}(F_\alpha, G_\alpha).$$

Together with the estimates from Steps 2 and 3 this allows us to assess the approximation by a mixed multivariate Poisson distribution with random parameters:

$$d_{\text{TV}}\left((\mathcal{C}_s((n_0, n_1]),\ 1 \le s \le r),\ \text{Poiss}\left(\frac{(n_1 - n_0)}{n_1}\, s\mathcal{K}_{n_1,s},\ 1 \le s \le r\right)\right) \le \frac{(r+1)(n_1 - n_0)}{n_1} + \frac{2(n_1 - n_0)}{n_1^2}\mathbb{E}\left[\left(\sum_{s=1}^r s\mathcal{K}_{n_1,s}\right)^2\right]. \tag{9.5}$$

To proceed we need the following lemma. For the Poissonised scheme it is quite easy too see that $\text{Var}[K_r(t)] < \mathbb{E}[K_r(t)] = \Phi_r(t)$; see p. 370 in Barbour and Gnedin (2009). We are not aware of discrete time analogue of this inequality and will use instead a rougher estimate.

Lemma 9.2. *For $r \geq 1$ and $n \geq 1$,*

$$\mathrm{Var}\left[\sum_{s=1}^{r} \mathcal{K}_{n,s}\right] \leq \mathbb{E}\left[\sum_{s=1}^{r} \mathcal{K}_{n,s}\right] \quad \textit{and} \quad \mathrm{Var}\left[\mathcal{K}_{n,r}\right] \leq 4\,\mathbb{E}\left[\sum_{s=1}^{r} \mathcal{K}_{n,s}\right].$$

Proof: The array $(\mathcal{P}_j(n),\ j \geq 1)$ has multinomial distribution, hence it is negatively associated; see Section 3 in Joag-Dev and Proschan (1983). The indicators $\mathbb{1}_{\{\mathcal{P}_j(n)\leq r\}}$ are nonincreasing functions of the array, hence they are pairwise negatively correlated. Thus,

$$\mathrm{Var}\left[\sum_{s=1}^{r} \mathcal{K}_{n,s}\right] = \mathrm{Var}\left[\sum_{j\geq 1} \mathbb{1}_{\{\mathcal{P}_j(n)\leq r\}}\right] \leq \sum_{j\geq 1} \mathrm{Var}[\mathbb{1}_{\{\mathcal{P}_j(n)\leq r\}}]$$
$$\leq \sum_{j\geq 1} \mathbb{E}[\mathbb{1}_{\{\mathcal{P}_j(n)\leq r\}}] = \mathbb{E}\left[\sum_{s=1}^{r} \mathcal{K}_{n,s}\right].$$

By the virtue of $\mathcal{K}_{n,r} = \sum_{s=1}^{r} \mathcal{K}_{n,s} - \sum_{s=1}^{r-1} \mathcal{K}_{n,s}$ the second inequality follows from the first. □

By the first inequality in Lemma 9.2 the second summand on the right-hand side of (9.5) is bounded by

$$\frac{2r^2(n_1-n_0)}{n_1^2}\mathbb{E}\left[\left(\sum_{s=1}^{r} \mathcal{K}_{n_1,s}\right)^2\right] \leq \frac{2r^2(n_1-n_0)}{n_1^2}\left(\mathbb{E}\left[\sum_{s=1}^{r} \mathcal{K}_{n_1,s}\right] + \left(\sum_{s=1}^{r} \mathbb{E}\mathcal{K}_{n_1,s}\right)^2\right).$$

Since $\mathbb{E}\left[\sum_{s=1}^{r} \mathcal{K}_{n_1,s}\right] = \sum_{s=1}^{r} (\Phi_s(n_1) + o(1))$ and $n_1 = n_0 + \lfloor \theta f(n_0) \rfloor$, this tends to zero in case $\alpha \in [0, 1)$ of Theorem 7.1.

Step 4: Multivariate Poisson approximation.

Finally, we wish to replace the random parameters in (9.5) by their mean values. Applying Lemma 9.2 and Theorem 10.C from Barbour et al. (1992)

$$d_{\mathrm{TV}}\left(\mathrm{Poiss}\left(\frac{(n_1-n_0)s}{n_1}\mathcal{K}_{n_1,s},\ 1\leq s\leq r\right), \mathrm{Poiss}\left(\frac{(n_1-n_0)s}{n_1}\mathbb{E}[\mathcal{K}_{n_1,s}],\ 1\leq s\leq r\right)\right)$$
$$\leq \mathbb{E}\left[\left(\sum_{s=1}^{r} \frac{(n_1-n_0)s}{n_1}\left|\mathcal{K}_{n_1,s} - \mathbb{E}[\mathcal{K}_{n_1,s}]\right|\right)^2\right] \leq \frac{r(n_1-n_0)^2}{n_1^2}\sum_{s=1}^{r} s^2\,\mathrm{Var}[\mathcal{K}_{n_1,s}]$$
$$\leq \frac{4r^4(n_1-n_0)^2}{n_1^2}\sum_{s=1}^{r}\mathbb{E}[\mathcal{K}_{n_1,s}] = \frac{4r^4(n_1-n_0)^2}{n_1^2}\sum_{s=1}^{r}(\Phi_s(n_1)+o(1)).$$

For $n_1 = n_0 + \lfloor \theta f(n_0) \rfloor$ this bound approaches 0 in all cases of Theorem 7.1. In particular, in the proper regular variation case (i.e. with index $0 < \alpha < 1$) with the scaling function $f(n_0) = \Phi(n_0)$ the bound is of the order $O(1/\Phi(n_0))$. The same bound is valid if we approximate by $\mathrm{Poiss}(\theta\rho_s,\ 1 \leq s \leq r)$, that is, with the interpolated rate instead of the natural mean.

From these total variation bounds the analogue of Theorem 7.1 for $(\widehat{\mathcal{B}}_r(\theta),\ 1 \leq s \leq r)$ follows in case $\alpha \in [0, 1)$. In particular, in the proper regular variation case, the vector converges in distribution to $(\mathrm{Poiss}(\theta\rho_s), 1 \leq s \leq r)$.

10. Random frequencies

Finally, we sketch a mixed Poisson approximation for records in occupancy schemes where the frequencies (p_j) are random. For simplicity of exposition we shall consider only the classic ‘pure power laws', see Section 10 in Gnedin et al. (2007). More precisely, assume that for some $\alpha \in (0,1)$

$$\nu[x,1] \ \sim\ Dx^{-\alpha}, \quad x \to 0+ \quad \text{a.s.}, \tag{10.1}$$

where D is a strictly positive random variable sometimes called the α-diversity Pitman (2006). Intensely studied examples of (p_j) leading to (10.1) are the two-parameter Poisson-Dirichlet frequencies Pitman (2006) and their generalisations Gnedin and Pitman (2005); Gnedin et al. (2006); Griffiths and Spanò (2007). By Proposition 23 in Gnedin et al. (2007), the relation (10.1) is equivalent to

$$p_j \ \sim\ D^{1/\alpha} j^{-1/\alpha}, \quad j \to \infty \quad \text{a.s.} \tag{10.2}$$

By Karamata's theorem, see Theorem 1.6.1 in Bingham et al. (1989), either of these relations implies

$$\Phi(n_0) \ \sim\ \Gamma(1-\alpha) D n_0^{\alpha}, \quad n_0 \to \infty \quad \text{a.s.} \tag{10.3}$$

By Theorem 8.4, the processes

$$\left(\widehat{\mathcal{B}}_r\left(\left(n_0, n_0 + \theta \frac{n_0}{\Phi(n_0)}\right]\right), \theta \geq 0\right), \quad r \geq 1,$$

given (p_j), converge jointly, as $n_0 \to \infty$, to independent homogeneous Poisson processes with rates $\rho_r = \frac{\alpha\Gamma(r-\alpha)}{(r-1)!\,\Gamma(1-\alpha)}$. Combining this with (10.3) we arrive at

Proposition 10.1. *Assume either of equivalent conditions* (10.1) *or* (10.2)*. Then, as* $n_0 \to \infty$*,*

$$\left(\widehat{\mathcal{B}}_r\left(\left(n_0, n_0 + n_0^{1-\alpha}\theta\right]\right), \theta \geq 0\right),$$

converge jointly in distribution as $n_0 \to \infty$*, to mixed Poisson processes with random rates* $\frac{D\alpha\Gamma(r-\alpha)}{(r-1)!}$*. Conditionally on* D*, the limit processes are independent Poisson.*

Appendix

Lemma 10.2. *For* μ *a measure on the halfline with* $0 < \mu([0,\infty)) \leq 1$*, the Laplace transform*

$$L(t) := \int_0^\infty e^{-tx}\mu(\mathrm{d}x), \quad t \geq 0,$$

satisfies for $0 \leq \gamma \leq 1$

$$L(\gamma t) \leq\ L^{\gamma}(t).$$

Proof: Let ξ be a nonnegative random variable with Laplace transform $L(t)/L(0)$. Using Jensen's inequality,

$$\frac{L(\gamma t)}{L(0)} = \mathbb{E}[e^{-\gamma t\xi}] \leq \left(\mathbb{E}[e^{-t\xi}]\right)^{\gamma} = \left(\frac{L(t)}{L(0)}\right)^{\gamma}.$$

The assertion follows from this by noting that $L(0) = \mu([0,\infty)) \leq 1$. □

Remark on exchangeable partitions. Nacu (2006) proved that the distribution of the point process of records $\mathcal{B}_1$ restricted to $\{1,\dots,n\}$ uniquely determines the distribution of counts $(\mathcal{K}_{n,1},\dots,\mathcal{K}_{n,n})$ that encode a partition of integer n induced by the allocation of n balls. Letting n vary, by the virtue of Kingman's theory of exchangeable partitions Pitman (2006), this fact implies that the distribution of probabilities (p_j) (arranged in decreasing order) is uniquely determined by the distribution of $\mathcal{B}_1$ seen as a point process on $\mathbb{N}$.

Recall that $\mathcal{C}_r(n_1) = r\mathcal{K}_{n_1,r}$. Changing notation n_1 to n, we have from (9.3)

$$\mathcal{K}_{n,1} + 2\mathcal{K}_{n,2} + \cdots + r\mathcal{K}_{n,r} \leq \mathcal{B}_1(n) + \mathcal{B}_2(n) + \cdots + \mathcal{B}_r(n), \quad 1 \leq r \leq n.$$

Now, using this inequality, arguments similar to Nacu (2006) allow one to show that the distribution of $(\mathcal{B}_1(n),\dots,\mathcal{B}_n(n))$ uniquely determines the distribution of the partition of n. Letting n vary, we conclude that the distributions of vectors $(\mathcal{B}_1(n), \dots, \mathcal{B}_n(n))$, $n \geq 1$, offer yet another way to describe the law of the exchangeable partition of $\mathbb{N}$ induced by the allocation of infinitely many balls.

Acknowledgments

This work has been accomplished during AM's visit to Queen Mary University of London as Leverhulme Visiting Professor in July-December 2023. AM gratefully acknowledges financial support from the Leverhulme Trust. The authors sincerely thank the anonymous referee for careful reading of the manuscript and suggestions that led to improvement of the manuscript.

References

Barbour, A. D. and Gnedin, A. V. Small counts in the infinite occupancy scheme. *Electron. J. Probab.*, **14**, no. 13, 365–384 (2009). MR2480545.

Barbour, A. D., Holst, L., and Janson, S. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York (1992). ISBN 0-19-852235-5. MR1163825.

Ben-Hamou, A., Boucheron, S., and Ohannessian, M. I. Concentration inequalities in the infinite urn scheme for occupancy counts and the missing mass, with applications. *Bernoulli*, **23** (1), 249–287 (2017). MR3556773.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1989). ISBN 0-521-37943-1. MR1015093.

Bogachev, L. V., Gnedin, A. V., and Yakubovich, Y. V. On the variance of the number of occupied boxes. *Adv. in Appl. Math.*, **40** (4), 401–432 (2008). MR2412154.

Böker, F. and Serfozo, R. Ordered thinnings of point processes and random measures. *Stochastic Process. Appl.*, **15** (2), 113–132 (1983). MR702486.

Buldygin, V. V., Indlekofer, K.-H., Klesov, O. I., and Steinebach, J. G. *Pseudo-regularly varying functions and generalized renewal processes*, volume 91 of *Probability Theory and Stochastic Modelling*. Springer, Cham (2018). ISBN 978-3-319-99536-6; 978-3-319-99537-3. MR3839312.

Bunge, J. and Fitzpatrick, M. Estimating the number of species: A review. *J. Amer. Statist. Assoc.*, **88** (421), 364–373 (1993). DOI: 10.1080/01621459.1993.10594330.

Buraczewski, D., Dovgay, B., and Marynych, A. Moderate parts in regenerative compositions: the case of regular variation. *J. Math. Anal. Appl.*, **497** (1), Paper No. 124894, 20 (2021). MR4192220.

Chebunin, M. and Zuyev, S. Functional central limit theorems for occupancies and missing mass process in infinite urn models. *J. Theoret. Probab.*, **35** (1), 1–19 (2022). MR4379457.

Daley, D. J. and Vere-Jones, D. *An introduction to the theory of point processes. Vol. II General theory and structure*. Probability and its Applications (New York). Springer, New York, second edition (2008). ISBN 978-0-387-21337-8. MR2371524.

Diaconis, P. and Freedman, D. Finite exchangeable sequences. *Ann. Probab.*, **8** (4), 745–764 (1980). MR577313.

Durieu, O. and Wang, Y. From infinite urn schemes to decompositions of self-similar Gaussian process. *Electron. J. Probab.*, **21**, Paper No. 43, 23 (2016). MR3530320.

Gnedin, A., Hansen, B., and Pitman, J. Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Probab. Surv.*, **4**, 146–171 (2007). MR2318403.

Gnedin, A. and Pitman, J. Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, **325** (Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 12), 83–102, 244–245 (2005). MR2160320.

Gnedin, A., Pitman, J., and Yor, M. Asymptotic laws for compositions derived from transformed subordinators. *Ann. Probab.*, **34** (2), 468–492 (2006). MR2223948.

Gnedin, A. and Stark, D. Random permutations and queues. *Adv. in Appl. Math.*, **149**, Paper No. 102549, 26 (2023). MR4587907.

Griffiths, R. C. and Spanò, D. Record indices and age-ordered frequencies in exchangeable Gibbs partitions. *Electron. J. Probab.*, **12**, 1101–1130 (2007). MR2336601.

Hardy, G. H. *Divergent Series.* Oxford, at the Clarendon Press (1949). MR30620.

Hwang, H.-K. and Janson, S. Local limit theorems for finite and infinite urn models. *Ann. Probab.*, **36** (3), 992–1022 (2008). MR2408581.

Ilienko, A. Convergence of point processes associated with coupon collector's and Dixie cup problems. *Electron. Commun. Probab.*, **24**, Paper No. 51, 9 (2019). MR4003125.

Joag-Dev, K. and Proschan, F. Negative association of random variables, with applications. *Ann. Statist.*, **11** (1), 286–295 (1983). MR684886.

Kallenberg, O. *Random measures, theory and applications*, volume 77 of *Probability Theory and Stochastic Modelling.* Springer, Cham (2017). ISBN 978-3-319-41596-3; 978-3-319-41598-7. MR3642325.

Karlin, S. Central limit theorems for certain infinite urn schemes. *J. Math. Mech.*, **17**, 373–401 (1967). MR216548.

Last, G. and Szekli, R. On negative association of some finite point processes on general state spaces. *J. Appl. Probab.*, **56** (1), 139–152 (2019). MR3981150.

McDonald, D. R. On the Poisson approximation to the multinomial distribution. *Canad. J. Statist.*, **8** (1), 115–118 (1980). DOI: 10.2307/3314676.

Nacu, S. Increments of Random Partitions. *Combin. Probab. Comput.*, **15** (4), 589–595 (2006). DOI: 10.1017/S0963548305007455.

Pitman, J. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin (2006). ISBN 978-3-540-30990-1; 3-540-30990-X. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, 2002. MR2245368.

Resnick, S. I. *Extreme values, regular variation and point processes.* Springer Series in Operations Research and Financial Engineering. Springer, New York (2008). ISBN 978-0-387-75952-4. MR2364939.

Roos, B. On the rate of multivariate Poisson convergence. *J. Multivariate Anal.*, **69** (1), 120–134 (1999). MR1701409.

Ruzankin, P. S. On the rate of Poisson process approximation to a Bernoulli process. *J. Appl. Probab.*, **41** (1), 271–276 (2004). MR2036288.

Schilling, R. L., Song, R., and Vondraček, Z. *Bernstein functions. Theory and applications*, volume 37 of *De Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, second edition (2012). ISBN 978-3-11-025229-3; 978-3-11-026933-8. MR2978140.