UDC 519.21
DOI: https://doi.org/10.17721/1812-5409.2024/2.3

Andrii ILIENKO, PhD (Phys. & Math.), Assoc. Prof.
ORCID ID: 0000-0003-4828-0788
e-mail: andrii.ilienko@unibe.ch
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

Viktoriia STAMATIIEVA, PhD Student, Assist.
ORCID ID: 0000-0003-2721-8985
e-mail: stamatieva56@gmail.com
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

LIMIT THEOREMS IN THE GENERALIZED BIRTHDAY PROBLEM

In the authors' previous papers, a novel approach to the classical coupon collector's and birthday problems was introduced, based on the theory of random point measures and their vague convergence. In this paper, we further develop this approach using the birthday problem as a case study, and apply it to establish new limit theorems for a range of nontrivial characteristics of the model.

More specifically, we define a point process on a metric space consisting of a countable number of copies of the real line. The atoms of this process correspond to the arrival times of new objects. Owing to the structure of the constructed process, each atom is naturally associated with an integer that indicates the number of times an object of that class has arrived. We demonstrate that, as the number of classes tends to infinity, these processes converge vaguely to a certain Poisson point process on the same space.

The application of the continuous mapping theorem to the proven convergence further yields distributional limit theorems for various characteristics of the model. The essentially infinite-dimensional nature of the involved point processes leads to corresponding infinite-dimensional limit theorems for these characteristics, fully revealing their asymptotic structure.

K e y w o r d s: ***generalized birthday problem, random point measures, vague convergence, poissonization, Poisson point processes.***

AMS 2020 classification: 60C05, 60G55.

1. Introduction

The birthday problem is a classical example in combinatorial probability theory. First formalized in 1939 by R. von Mises, it has found numerous applications in fields such as statistics, cryptography, and the modeling of random processes, among others, and has served as the basis for many generalizations. In its original form, the problem involves finding the probability that, in a randomly chosen group of people, at least two will share the same birthday.

In this paper, we consider one of its generalizations. Suppose there are n distinct classes of objects that arrive independently and sequentially at discrete time points, each with equal probability $\frac{1}{n}$. Let $T_r^{(n)}$, $r \geqslant 1$, denote the first time at which one of the classes appears for the $(r+1)$-th time. Note that in the classical formulation, the objects are people, $n = 365$ (the number of days in a non-leap year), and $r = 1$ (corresponding to a single match with someone already present).

This problem is closely related to another classical problem known as the coupon collector's problem. If the classes of objects are interpreted as different types of coupons, then $T_r^{(n)}$ represents the first time at which the collector obtains a coupon of a certain type for the $(r+1)$-th time (this can also be referred to as the beginning of the $(r+1)$-th collection).

This problem can also be considered in the context of random allocations, where the arrivals of objects of different classes are interpreted as the placement of items into cells. Extensive literature is devoted to such allocations; see, e.g., the classical monographs (Kolchin, Sevast'yanov, & Chistyakov, 1976) and (Johnson & Kotz, 1977).

In (Ilienko, 2019) and (Ilienko & Stamatiieva, 2021), the authors proposed a new approach to the coupon collector's and birthday problems, based on the theory of vague convergence of point processes. More specifically, the arrival times of different classes of objects can be interpreted as the atoms of a certain random point measure. By using well-known criteria for the convergence of such measures, it is possible to establish their vague convergence as $n \to \infty$ to a certain limiting measure, which turns out to be Poisson. Further application of the continuous mapping theorem to this convergence allows one to derive the limiting distributions of a number of interesting and important characteristics in these problems.

This paper can be viewed as a far-reaching and essentially infinite-dimensional generalization of the results from (Ilienko & Stamatiieva, 2021), where the aforementioned approach was applied to individual collections. Here, we will consider all collections simultaneously and construct the corresponding point process on the union of a countable number of copies of the real line. This will enable the study of characteristics associated with different values of r simultaneously, such as the relationships between the starting times of different collections.

The proposed approach is effective for determining the limiting distributions of various characteristics but is not designed to address the asymptotics of their expected values. The latter requires completely different techniques. In the recent paper by the second author (Stamatiieva, 2023), an asymptotic expansion for $\mathbb{E}T_r^{(n)}$ was derived using Laplace's asymptotic method. For $r = 1$, this expansion is well-known in the literature as the Ramanujan-Watson-Knuth expansion (Flajolet, Grabner, Kirschenhofer, & Prodinger, 1995).

2. Notation and main results

Let $Y_{i,r}^{(n)}$, $i \in \mathbb{N}_n = \{1, \ldots, n\}$, $r, n \in \mathbb{N}$, denote the arrival time of the $(r+1)$-th object of class i. Then $Y_{i,r}^{(n)} \sim \mathsf{NegBin}\left(r+1, \frac{1}{n}\right)$, where NegBin stands for the negative binomial distribution representing the number of trials up to and including the $(r+1)$-th success:

$$\mathbb{P}\left\{Y_{i,r}^{(n)}=k\right\}=\binom{k-1}{r}\left(\frac{1}{n}\right)^{r+1}\left(1-\frac{1}{n}\right)^{k-r-1},$$

where $k \geqslant r+1$.

For fixed values of r and n, the random variables $Y_{i,r}^{(n)}$, $i \in \mathbb{N}_n$, are identically distributed but dependent, at least because $Y_{i,r}^{(n)} \neq Y_{i',r}^{(n)}$ for $i \neq i'$.

Let us also consider the Poissonized scheme, in which objects arrive not at discrete time points but according to a Poisson process with intensity 1. Let $Z_{i,r}^{(n)}$, $i \in \mathbb{N}_n$, $r, n \in \mathbb{N}$, denote the arrival time of the $(r+1)$-th object of class i in this Poissonized scheme. Then, by the thinning theorem for Poisson processes (see, e.g., Theorem 5.8 in (Last & Penrose, 2017)), $Z_{i,r}^{(n)}$ are independent and gamma-distributed, $Z_{i,r}^{(n)} \sim \Gamma\left(r+1, \frac{1}{n}\right)$. For any fixed value of n, the sequences $\left(Y_{i,r}^{(n)}\right)$ and $\left(Z_{i,r}^{(n)}\right)$ are defined on the same probability space and are related as follows:

$$Z_{i,r}^{(n)}=\sum_{j=1}^{Y_{i,r}^{(n)}} E_j, \quad i \in \mathbb{N}_n, \quad r \geqslant 1,$$

where E_j, $j \in \mathbb{N}$, are independent $\mathrm{Exp}(1)$-distributed random variables. Moreover, the sequences $\left(Y_{i,r}^{(n)},\ 1 \leqslant i \leqslant n\right)$ and $(E_j,\ j \in \mathbb{N})$ are independent, as the former sequence is determined by the object classes while the latter defines the intervals between the arrival times.

Consider the point process $\hat{H}^{(n)}, n \in \mathbb{N}$, on the metric space $\mathbb{X}=\bigcup_{r=1}^{\infty}\{r\} \times \mathbb{R}$, defined as follows:

$$\hat{H}^{(n)}\left(\bigcup_{r=1}^{\infty}(\{r\} \times B_r)\right)=\sum_{r=1}^{\infty} \widehat{\eta}_r^{(n)}\left(B_r\right), \qquad B_r \in \mathfrak{B}(\mathbb{R}), \quad r \in \mathbb{N}, \tag{1}$$

where

$$\widehat{\eta}_r^{(n)}=\sum_{i=1}^{n} \delta_{\widehat{\psi}_r^{(n)}\left(Z_{i,r}^{(n)}\right)},$$

and

$$\widehat{\psi}_r^{(n)}(x)=\frac{x}{n^{\frac{r}{r+1}}}, \quad x \in \mathbb{R}. \tag{2}$$

Here, δ_a stands for the unit mass at a.

We also consider the analogue of $\hat{H}^{(n)}$ in the classical (non-Poissonized) setting starting from $r=2$:

$$\widehat{\Xi}^{(n)}\left(\bigcup_{r=2}^{\infty}(\{r\} \times B_r)\right)=\sum_{r=2}^{\infty} \widehat{\xi}_r^{(n)}\left(B_r\right),$$

where

$$\widehat{\xi}_r^{(n)}=\sum_{i=1}^{n} \delta_{\widehat{\psi}_r^{(n)}\left(Y_{i,r}^{(n)}\right)}.$$

Recall that the vague topology on the space of locally finite measures is generated by the integration maps $\nu \mapsto \int_{\mathbb{X}} f\, d\nu$ for all continuous functions f with bounded support; see, e.g., Chapter 4 in (Kallenberg, 2017) for a general exposition.

The difference between vague convergence and the more common weak convergence lies in the class of test functions: for weak convergence, this class includes all bounded continuous functions, whereas for vague convergence it is restricted to those with bounded support. In our case, using vague convergence allows us to avoid unnecessary consideration of the behavior of measures at infinity.

Denote by $\xrightarrow{vd}$ the distributional convergence of random point measures with underlying vague topology.

Theorem 1. *Let H be a Poisson point process on $\mathbb{X}$ with intensity measure λ defined by*

$$\lambda\left(\bigcup_{r=1}^{\infty}(\{r\} \times B_r)\right)=\sum_{r=1}^{\infty} \frac{1}{r!} \int_{B_r} x^r \cdot \mathbb{I}\{x \geqslant 0\}\, dx.$$

Then $\hat{H}^{(n)} \xrightarrow{vd} H$ as $n \to \infty$.

Remark 1. By the superposition theorem for Poisson processes (see, e.g., Theorem 3.3 in (Last & Penrose, 2017)), the additive structure of λ in Theorem 1 implies that the different levels $H(\{r\} \times \cdot)$ of the limiting process H are independent.

Remark 2. The different levels $H(\{r\} \times \cdot)$ can be interpreted as follows (see Remark 3.2 in (Ilienko, 2019)). Let ζ be a homogeneous Poisson point process with unit intensity. Define

$$h(x)=((r+1)!\, x)^{\frac{1}{r+1}}, \quad x \geqslant 0.$$

Then $H\left(\{r\}\times\cdot\right)=\sum\limits_{x\in\text{supp }\zeta}\delta_{h(x)}$ in distribution.

Indeed, by the transformation theorem for Poisson processes (see, e.g., Theorem 5.1 in (Last & Penrose, 2017)), applying the function h to ζ results in the process described in Theorem 1, since

$$h\left(\text{Leb}\right)[a,b]=\left(\text{Leb}\circ h^{-1}\right)[a,b]=\text{Leb}\left[\frac{a^{r+1}}{(r+1)!},\frac{b^{r+1}}{(r+1)!}\right]=\int\limits_a^b\frac{x^r}{r!}dx.$$

The following theorem is an analogue of Theorem 1 in the non-Poissonized setting. Due to technical reasons related to the depoissonization procedure, we can prove it only for $r\geqslant 2$.

Theorem 2. *Let H' be a Poisson point process on $\mathbb{X}$ with intensity measure λ given by*

$$\lambda\left(\bigcup_{r=2}^{\infty}\left(\{r\}\times B_r\right)\right)=\sum_{r=2}^{\infty}\frac{1}{r!}\int\limits_{B_r}x^r\cdot\mathbb{I}\left\{x\geqslant 0\right\}dx.$$

Then $\widehat{\Xi}^{(n)}\overset{vd}{\to}H'$ as $n\to\infty$.

3. Proofs

Proof of Theorem 1. By Theorem 4.18 in (Kallenberg, 2017), it suffices to show that

$$\lim_{n\to\infty}\mathbb{P}\left\{\hat{H}^{(n)}(U)=0\right\}=\mathbb{P}\left\{H\left(U\right)=0\right\},\tag{3}$$

$$\lim_{n\to\infty}\mathbb{E}\hat{H}^{(n)}\left(U\right)=\mathbb{E}H\left(U\right)\tag{4}$$

for each $U\in\mathcal{U}=\left\{\bigcup_{r=1}^{s}\left(\{r\}\times B_r\right),s\in\mathbb{N},B_r\in\mathfrak{B}\left(\mathbb{R}\right)\right\}$.

By (1), we have

$$\mathbb{P}\left\{\hat{H}^{(n)}(U)=0\right\}=\mathbb{P}\left\{\widehat{\eta}_r^{(n)}\left(B_r\right)=0,\ \ \forall r=1,\ldots,s\right\}.$$

For $1\leqslant r_1<r_2<\ldots<r_m\leqslant s$, define

$$\hat{P}^{(n)}_{r_1,\ldots,r_m}=\mathbb{P}\left\{\widehat{\psi}^{(n)}_{r_1}\left(Z^{(n)}_{i,r_1}\right)\in B_{r_1},\ldots,\widehat{\psi}^{(n)}_{r_m}\left(Z^{(n)}_{i,r_m}\right)\in B_{r_m}\right\},$$

which does not depend on i, since the variables $Z^{(n)}_{i,r}$ are independent and identically distributed for different i. Using the inclusion-exclusion formula, we obtain

$$\mathbb{P}\left\{\hat{H}^{(n)}(U)=0\right\}=\left(1-\sum_{1\leqslant r_1\leqslant s}\hat{P}^{(n)}_{r_1}+\sum_{1\leqslant r_1<r_2\leqslant s}\hat{P}^{(n)}_{r_1,r_2}-\ldots+(-1)^{s+1}\hat{P}^{(n)}_{1,\ldots,s}\right)^n.$$

We will compute $\lim\limits_{n\to\infty}\ln\mathbb{P}\left\{\hat{H}^{(n)}\left(U\right)=0\right\}$. To do this, we use the asymptotic equivalence $\ln\left(1+\alpha\right)\sim\alpha,\ \alpha\to 0$. The fact that $\alpha\to 0$ in our case follows from the arguments below. We have

$$\lim_{n\to\infty}\ln\mathbb{P}\left\{\hat{H}^{(n)}\left(U\right)=0\right\}=-\sum_{1\leqslant r_1\leqslant s}\lim_{n\to\infty}\left(n\cdot\hat{P}^{(n)}_{r_1}\right)+\sum_{1\leqslant r_1<r_2\leqslant s}\lim_{n\to\infty}\left(n\cdot\hat{P}^{(n)}_{r_1,r_2}\right)-\ldots+(-1)^s\lim_{n\to\infty}\left(n\cdot\hat{P}^{(n)}_{1,\ldots,s}\right).$$

To prove (3), it suffices to show that the limit in the first sum equals $\frac{1}{r_1!}\int_{B_{r_1}}x^{r_1}\cdot\mathbb{I}\{x\geqslant 0\}\,dx$, while the limit in the second sum equals zero. Since all subsequent sums are bounded from above by the second one, this will prove the claim.

Since $Z^{(n)}_{i,r}\sim\Gamma\left(r+1;\frac{1}{n}\right)$, and $Z^{(n)}_{i,2}-Z^{(n)}_{i,1},\ldots,Z^{(n)}_{i,s}-Z^{(n)}_{i,s-1}$ are independent $\text{Exp}\left(\frac{1}{n}\right)$-distributed random variables, the densities $f^{(n)}_r(x)$ and $f^{(n)}_{r_1,r_2}(x,y)$ of $Z^{(n)}_{i,r}$ and $\left(Z^{(n)}_{i,r_1},Z^{(n)}_{i,r_2}\right)$, respectively, have the form:

$$f^{(n)}_r\left(x\right)=\frac{1}{n^{r+1}r!}x^r\exp\left(-\frac{x}{n}\right)\mathbb{I}\left\{x\geqslant 0\right\},$$

$$f^{(n)}_{r_1,r_2}\left(x,y\right)=\frac{x^{r_1}\cdot\left(y-x\right)^{r_2-r_1-1}}{n^{r_2+1}\cdot r_1!\left(r_2-r_1-1\right)!}\exp\left(-\frac{y}{n}\right)\mathbb{I}\left\{0\leqslant x\leqslant y\right\}.$$

Then, the density of $\widehat{\psi}^{(n)}_{r_1}\left(Z^{(n)}_{i,r_1}\right)$ is given by

$$\hat{f}^{(n)}_{r_1}(x)=n^{\frac{r_1}{r_1+1}}\cdot f^{(n)}_{i,r_1}\left(n^{\frac{r_1}{r_1+1}}\cdot x\right)=n^{\frac{r_1}{r_1+1}}\cdot\frac{1}{n^{r_1+1}r_1!}\cdot\left(n^{\frac{r_1}{r_1+1}}\cdot x\right)^{r_1}\times$$
$$\times\exp\left(-\frac{n^{\frac{r_1}{r_1+1}}\cdot x}{n}\right)\mathbb{I}\{x\geqslant 0\}=\frac{1}{r_1!\,n}\cdot x^{r_1}\cdot\exp\left(-\frac{x}{n^{\frac{1}{r_1+1}}}\right)\mathbb{I}\{x\geqslant 0\},$$

and the density of $\left(\widehat{\psi}_{r_1}^{(n)}\left(Z_{i,r_1}^{(n)}\right), \widehat{\psi}_{r_2}^{(n)}\left(Z_{i,r_2}^{(n)}\right)\right)$ is

$$\hat{f}_{r_1,r_2}^{(n)}(x,y) = n^{\frac{r_1}{r_1+1}} \cdot n^{\frac{r_2}{r_2+1}} \cdot f_{r_1,r_2}^{(n)}\left(n^{\frac{r_1}{r_1+1}} \cdot x, n^{\frac{r_2}{r_2+1}} \cdot y\right) =$$

$$= \frac{n^{\frac{r_1}{r_1+1}} \cdot n^{\frac{r_2}{r_2+1}}}{n^{r_2+1} \cdot r_1!\,(r_2-r_1-1)!} \cdot \left(n^{\frac{r_1}{r_1+1}} \cdot x\right)^{r_1} \left(n^{\frac{r_2}{r_2+1}} \cdot y - n^{\frac{r_1}{r_1+1}} \cdot x\right)^{r_2-r_1-1} \times$$

$$\times \exp\left(-\frac{n^{\frac{r_2}{r_2+1}} \cdot y}{n}\right) \mathbb{I}\left\{0 \leqslant n^{\frac{r_1}{r_1+1}} \cdot x \leqslant n^{\frac{r_2}{r_2+1}} \cdot y\right\} = \frac{n^{\frac{r_1+1}{r_2+1}-2}}{r_1!\,(r_2-r_1-1)!} \cdot x^{r_1} \cdot \left(y - n^{\frac{r_1}{r_1+1}-\frac{r_2}{r_2+1}} \cdot x\right)^{r_2-r_1-1} \times$$

$$\times \exp\left(-\frac{y}{n^{\frac{1}{r_2+1}}}\right) \mathbb{I}\left\{0 \leqslant n^{\frac{r_1}{r_1+1}} \cdot x \leqslant n^{\frac{r_2}{r_2+1}} \cdot y\right\}.$$

Thus,

$$\hat{f}_{r_1}^{(n)}(x) = \frac{1}{r_1!\,n} \cdot x^{r_1} \cdot \exp\left(-\frac{x}{n^{\frac{1}{r_1+1}}}\right) \mathbb{I}\{x \geqslant 0\},$$

$$\hat{f}_{r_1,r_2}^{(n)}(x,y) = \frac{n^{\frac{r_1+1}{r_2+1}-2}}{r_1!\,(r_2-r_1-1)!} \cdot x^{r_1}\left(y - n^{\frac{r_1}{r_1+1}-\frac{r_2}{r_2+1}} \cdot x\right)^{r_2-r_1-1} \exp\left(-\frac{y}{n^{\frac{1}{r_2+1}}}\right) \mathbb{I}\left\{0 \leqslant n^{\frac{r_1}{r_1+1}} \cdot x \leqslant n^{\frac{r_2}{r_2+1}} \cdot y\right\}.$$

Then,

$$n \cdot \hat{P}_{r_1}^{(n)} = n \cdot \int\limits_{B_{r_1}} \hat{f}_{r_1}^{(n)}(x)\,dx = \frac{1}{r_1!} \int\limits_{B_{r_1}} x^{r_1} \cdot \exp\left(-\frac{x}{n^{\frac{1}{r_1+1}}}\right) \mathbb{I}\{x \geqslant 0\}\,dx \underset{n\to\infty}{\longrightarrow} \frac{1}{r_1!} \int\limits_{B_{r_1}} x^{r_1} \cdot \mathbb{I}\{x \geqslant 0\}\,dx.$$

Let $\alpha = \inf B_{r_1}$ and $\beta = \inf B_{r_2}$. Then we have

$$n \cdot \hat{P}_{r_1,r_2}^{(n)} \leqslant n \cdot \int\limits_{\alpha}^{+\infty} \int\limits_{\beta}^{+\infty} \hat{f}_{r_1,r_2}^{(n)}(x,y)\,dx\,dy =$$

$$= \frac{n^{\frac{r_1+1}{r_2+1}-1}}{r_1!\,(r_2-r_1-1)!} \int\limits_{\alpha}^{+\infty} \int\limits_{\beta}^{+\infty} x^{r_1}\left(y - n^{\frac{r_1}{r_1+1}-\frac{r_2}{r_2+1}} \cdot x\right)^{r_2-r_1-1} \exp\left(-\frac{y}{n^{\frac{1}{r_2+1}}}\right) \mathbb{I}\left\{0 \leqslant n^{\frac{r_1}{r_1+1}} \cdot x \leqslant n^{\frac{r_2}{r_2+1}} \cdot y\right\} dx\,dy.$$

Since $r_1 < r_2$, we have $\frac{r_1+1}{r_2+1} - 1 < 0$ and $\frac{r_1}{r_1+1} - \frac{r_2}{r_2+1} < 0$, which implies that the last expression vanishes as $n \to \infty$. This proves (3).

Let us proceed to the proof of (4). Consider $\mathbb{E}\hat{H}^{(n)}(U)$. Then,

$$\mathbb{E}\hat{H}^{(n)}(U) = \mathbb{E}\hat{H}^{(n)}\left(\bigcup_{r=1}^{s}(\{r\} \times B_r)\right) = \sum_{r=1}^{s} \mathbb{E}\widehat{\eta}_r^{(n)}(B_r).$$

Using the fact that $\widehat{\eta}_r^{(n)}(B_r) \sim \mathrm{Bin}(n, \hat{P}_r^{(n)})$, we obtain

$$\mathbb{E}\hat{H}^{(n)}(U) = \sum_{r=1}^{s} n \cdot \hat{P}_r^{(n)} \underset{n\to\infty}{\longrightarrow} \sum_{r=1}^{s} \frac{1}{r!} \int_{B_{r_1}} x^r \cdot \mathbb{I}\{x \geqslant 0\}\,dx = \mathbb{E}H(U),$$

which, together with (3), completes the proof. □

The proof of Theorem 2 is based on the depoissonization procedure which follows the same lines as in the proof of Theorem 1 in (Ilienko & Stamatiieva, 2021). Therefore, we omit it here.

4. Applications

Fix $s \geqslant 1$ and consider the following random variables:

$$U_r^{(n)} = \mathrm{card}\left\{i\colon Z_{i,r}^{(n)} < \min_{i \leqslant n} Z_{i,s}^{(n)} \cdot d_{r,s}^{(n)}\right\}, \quad r = 1, \ldots, s-1,$$

where

$$d_{r,s}^{(n)} = n^{\frac{r}{r+1}-\frac{s}{s+1}}.$$

In other words, $U_r^{(n)}$ represents the number of objects in the $(r+1)$-th collection before the starting time of the $(s+1)$-th one multiplied by the adjustment factor $d_{r,s}^{(n)}$. The necessity of multiplying by this factor is due to the fact that, without it, $U_r^{(n)}$ would grow unboundedly.

The following theorem holds.

Theorem 3. *Denote*

$$(G_r,\ r=1,\dots,s-1)=\left(N_r\left(\frac{(E\,(s+1)!)^{\frac{r+1}{s+1}}}{(r+1)!}\right),\ r=1,\dots,s-1\right),$$

where $E\sim$ Exp(1), and N_r are Poisson processes with unit intensity, independent of each other and of E. Then

$$\left(U_r^{(n)}, r=1,\dots,s-1\right)\xrightarrow{d}(G_r,\ r=1,\dots,s-1)$$

in $\mathbb{R}^{s-1}$ as $n\to\infty$.

Proof. Let

$$\hat{T}_s^{(n)}=\sup\left\{x\in\mathbb{R}\colon\ \hat{H}^{(n)}\left(\{s\}\times(0,x)\right)=0\right\}, \tag{5}$$

$$V_r^{(n)}=\hat{H}^{(n)}\left(\{r\}\times\left(0,\hat{T}_s^{(n)}\right)\right),\ r=1,\dots,s-1.$$

Thus, $\hat{T}_s^{(n)}$ is the first point of $\widehat{\eta}_r^{(n)}$, and $V_r^{(n)}$ is the number of points in $\widehat{\eta}_r^{(n)}$ to the left of $\hat{T}_s^{(n)}$.

By the definition of $\hat{H}^{(n)}$ and $\widehat{\eta}_r^{(n)}$, we obtain

$$\hat{T}_s^{(n)}=\min_{i\leqslant n}\widehat{\psi}_s^{(n)}\left(Z_{i,s}^{(n)}\right)=\widehat{\psi}_s^{(n)}\left(T_s^{(n)}\right). \tag{6}$$

Taking into account (2) and (6), $V_r^{(n)}$ can be represented as

$$V_r^{(n)}=\text{card}\left\{i\colon\ \frac{Z_{i,r}^{(n)}}{n^{\frac{r}{r+1}}}<\frac{T_s^{(n)}}{n^{\frac{s}{s+1}}}\right\}=\text{card}\left\{i\colon\ Z_{i,r}^{(n)}<T_s^{(n)}\cdot d_{r,s}^{(n)}\right\}.$$

From the definitions of $U_r^{(n)}$, $V_r^{(n)}$, and $\widehat{\psi}_r$, it follows that $V_r^{(n)}=U_r^{(n)}$. Using Remarks 1, 2, and Theorem 1, we have by continuous mapping that

$$\left(U_r^{(n)},\ r=1,\dots,s-1\right)\xrightarrow{d}\left(N_r\left(\mu_{r,s}\right),\ r=1,\dots,s-1\right),\quad n\to\infty,$$

where $\mu_{r,s}=h_r^{-1}\left(h_s\left(E\right)\right)=\frac{(E(s+1)!)^{\frac{r+1}{s+1}}}{(r+1)!}$, which completes the proof. □

Theorem 4. *Let $W_r,\ r\geqslant 1$, be independent random variables following the Weibull distribution with the cdf of the form*

$$\mathbb{P}\left\{W_r\leqslant x\right\}=1-\exp\left(-\frac{x^{r+1}}{(r+1)!}\right),\quad x\geqslant 0.$$

Then

$$\left(\frac{T_r^{(n)}}{n^{\frac{r}{r+1}}},\ r\geqslant 1\right)\xrightarrow{d}(W_r,\ r\geqslant 1)\ \text{in } \mathbb{R}^{\infty}\text{ as } n\to\infty.$$

Proof. Taking into account Theorem 1, (5), (6), and Remarks 1, 2, we have

$$\left(\widehat{\psi}_r^{(n)}\left(T_r^{(n)}\right),\ r\geqslant 1\right)=\left(\hat{T}_r^{(n)},\ r\geqslant 1\right)\xrightarrow{d}\left(\left((r+1)!\,E_r\right)^{\frac{1}{r+1}},\ r\geqslant 1\right),\quad n\to\infty,$$

where E_r are independent Exp(1). To complete the proof, we now find the distribution function of $\left((r+1)!\,E_r\right)^{\frac{1}{r+1}}$:

$$\mathbb{P}\left\{\left((r+1)!\,E_r\right)^{\frac{1}{r+1}}\leqslant x\right\}=\mathbb{P}\left\{E_r\leqslant\frac{x^{r+1}}{(r+1)!}\right\}=1-\exp\left(-\frac{x^{r+1}}{(r+1)!}\right),\quad x\geqslant 0.$$

This completes the proof. □

Finally, we return to the depoissonized setting. Let

$$L_r^{(n)}=\text{card}\left\{i\colon\ Y_{i,r}^{(n)}<\min_{i\leqslant n}Y_{i,s}^{(n)}\cdot d_{r,s}^{(n)}\right\},\quad r=2,\dots,s-1,$$

$$X_r^{(n)}=\min_{i\leqslant n}Y_{i,r}^{(n)}.$$

The following theorems hold.

Theorem 5.

$$\left(L_r^{(n)}, r=2,\dots,s-1\right)\xrightarrow{d}(G_r, r=2,\dots,s-1)\ \text{in } \mathbb{R}^{s-2}\text{ as } n\to\infty,$$

where G_r are defined in Theorem 3.

Theorem 6.

$$\left(\frac{X_r^{(n)}}{n^{\frac{r}{r+1}}},\ r \geqslant 2\right) \xrightarrow{d} (W_r,\ r \geqslant 2) \text{ in } \mathbb{R}^\infty \text{ as } n \to \infty,$$

where W_r are defined in Theorem 4.

The proofs of Theorems 5 and 6 are similar to the proofs of Theorems 3 and 4, but are based on the application of Theorem 2 instead of Theorem 1.

5. Discussion and conclusions

In this paper, we investigate the generalized birthday problem from the perspective of point processes. The approach to this problem, based on studying the vague convergence of specially constructed random point measures, proves to be highly fruitful for obtaining the limiting distributions of a wide variety of model characteristics.

Authors' contribution: Andrii Ilienko — methodology; Viktoriia Stamatiieva — derivation of the main results and preparation of the paper.

Acknowledgments, funding sources. The authors thank the anonymous reviewer for the careful reading of the manuscript and the constructive comments. The first author was supported by the Swiss National Science Foundation, Grant 229505.

References

Flajolet, P., Grabner, P. J., Kirschenhofer, P., & Prodinger, H. (1995). On Ramanujan's Q-function. *Journal of Computational and Applied Mathematics*, 58(1), 103–116. https://doi.org/10.1016/0377-0427(93)E0258-N

Ilienko, A. (2019). Convergence of point processes associated with coupon collector's and dixie cup problems. *Electronic Communications in Probability*, 24(51), 1–9. https://doi.org/10.1214/19-ecp263

Ilienko, A., & Stamatiieva, V. (2021). A limit theorem for point processes associated with the generalized birthday problem. *Scientific Bulletin of Uzhhorod University; Series of Mathematics and Informatics*, 39(2), 38–46 [in Ukrainian]. [Ільєнко, А., & Стаматієва, В. (2021). Гранична теорема для точкових процесів, пов'язаних з узагальненою задачею про дні народження. *Науковий вісник Ужгородського університету. Серія «Математика і інформатика»*, 39(2), 38–46]. https://doi.org/10.24144/2616-7700.2021.39(2).38-46

Johnson, N. L., & Kotz, S. (1977). *Urn models and their application*. John Wiley & Sons.

Kallenberg, O. (2017). *Random measures, theory and applications* (Vol. 77). Springer, Cham. https://doi.org/10.1007/978-3-319-41598-7

Kolchin, V. F., Sevast'yanov, B. A., & Chistyakov, V. P. (1976). *Random allocations*. Nauka [in Russian]. [Колчин, В. Ф., Севастьянов, Б. А., & Чистяков, В. П. (1976). *Случайные размещения*. Наука]

Last, G., & Penrose, M. (2017). *Lectures on the poisson process*. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781316104477

Stamatiieva, V. (2023). Generalization of the asymptotic expansion of Ramanujan – Watson – Knuth. *Scientific Bulletin of Uzhhorod University; Series of Mathematics and Informatics*, 43(2), 52–61 [in Ukrainian]. [Стаматієва, В. (2023). Узагальнення асимптотичного розкладу Рамануджана – Ватсона – Кнута. *Науковий вісник Ужгородського університету. Серія «Математика і інформатика»*, 43(2), 52–61]. https://doi.org/10.24144/2616-7700.2023.43(2).52-61

Отримано редакцією журналу / Received: 02.09.24
Прорецензовано / Revised: 27.10.24
Схвалено до друку / Accepted: 28.11.24

Андрій ІЛЬЄНКО, канд. фіз.-мат. наук, доц.
ORCID ID: 0000-0003-4828-0788
e-mail: andrii.ilienko@unibe.ch
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

Вікторія СТАМАТІЄВА, асп., асист.
ORCID ID: 0000-0003-2721-8985
e-mail: stamatieva56@gmail.com
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ГРАНИЧНІ ТЕОРЕМИ В УЗАГАЛЬНЕНІЙ ЗАДАЧІ ПРО ДНІ НАРОДЖЕННЯ

У попередніх роботах авторами було запропоновано новий підхід до класичних задачі збирача купонів і задачі про дні народження, що заснований на теорії випадкових точкових мір і їх грубій збіжності. У пропонованій статті зазначений підхід розвинено на прикладі задачі про дні народження та застосовуємо його у процесі доведення нових граничних теорем для низки нетривіальних характеристик моделі.

Авторами задано точковий процес на метричному просторі, який складається зі зліченної кількості копій дійсної прямої. Атоми цього процесу відповідають моментам появи нових об'єктів. Завдяки структурі побудованого процесу кожному атому природним чином відповідає ціле число, яке вказує на кількість разів, коли об'єкт цього класу з'являвся раніше. Ми показуємо, що у разі збільшення кількості класів до нескінченності ці процеси грубо збігаються до певного точкового процесу Пуассона на тому ж просторі.

Застосування теореми про неперервне відображення до доведеної збіжності дає змогу тепер доводити граничні теореми для розподілів різноманітних характеристик моделі. По суті, нескінченновимірна природа залучених точкових процесів приводить до відповідних нескінченновимірних граничних теорем для цих характеристик, що повністю розкриває їх асимптотичну структуру.

Ключові слова: ***узагальнена задача про дні народження, випадкові точкові міри, груба збіжність, пуассонізація, точкові процеси Пуассона.***

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження (у зборі, аналізі чи інтерпретації даних, якщо це мало місце), у написанні рукопису та в рішенні про публікацію результатів.

The authors declare no conflicts of interest. The funders had no role in the design of the study (in the collection, analyses or interpretation of data if applicable), in the writing of the manuscript as well as in the decision to publish the results.