

POISSON APPROXIMATION IN A POISSON LIMIT THEOREM INSPIRED BY COUPON COLLECTING

ANNA PÓSFAL,* *University of Szeged*

Abstract

In this paper we refine a Poisson limit theorem of Gnedenko and Kolmogorov (1954): we determine the error order of a Poisson approximation for sums of asymptotically negligible integer-valued random variables that converge in distribution to the Poisson law. As an application of our results, we investigate the case of the coupon collector's problem when the distribution of the collector's waiting time is asymptotically Poisson.

Keywords: Poisson limit theorem; Poisson approximation; total variation distance; coupon collecting

2000 Mathematics Subject Classification: Primary 60F05; 62E17; 62E20

1. Introduction

In this paper we are concerned with a Poisson approximation for the distribution of sums of asymptotically negligible integer-valued random variables. We complement the classical Poisson convergence theorem of Gnedenko and Kolmogorov [4], in the setting of triangular arrays, with error bounds that are expressed in terms of the total variation distance, which is defined as

$$d_{TV}(X, Y) := \sup_{A \subset \mathbb{Z}_+} |P(X \in A) - P(Y \in A)| \quad (1)$$

for any two random variables X and Y that map into $\mathbb{Z}_+ := \{0, 1, \dots\}$.

For each n , we approximate the distribution of the n th row sum with a Poisson distribution whose mean λ_n is defined only in terms of the distributions of the random variables in the n th row. We do not assume the existence of moments, as is the case in analogous results proved by Barbour and Hall [1], and our lower bounds are much simpler in form to theirs, being of precisely the same form, up to a constant, as our upper bounds, provided that the means λ_n are bounded away from ∞ .

We then continue with an application of these results to the coupon collector's problem. A collector samples with replacement a set of $n \in \mathbb{N} := \{1, 2, \dots\}$ distinct coupons so that the draws are independent and at each time any one of the n coupons is drawn with the same probability $1/n$. For an integer $m_n \in \{0, 1, \dots, n-1\}$ that depends on n , sampling is repeated until the first time, W_{n,m_n} , that the collector has collected $n - m_n$ distinct coupons. Baum and Billingsley [3] proved (using the method of characteristic functions) that if, as $n \rightarrow \infty$,

$$m_n \rightarrow \infty \quad \text{and} \quad \frac{n - m_n}{\sqrt{n}} \rightarrow \sqrt{2\lambda} \quad \text{for some } \lambda > 0 \text{ constant,} \quad (2)$$

then $W_{n,m_n} - (n - m_n)$ converges in distribution to the Poisson law with mean λ .

Received 5 February 2009; revision received 6 March 2009.

* Postal address: Analysis and Stochastics Research Group of the Hungarian Academy of Sciences, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged 6720, Hungary. Email address: posfai@math.u-szeged.hu

We express this problem as a special case of the Poisson limit theorem above, and immediately obtain the corresponding Poisson approximation results. An even stronger result can be proved in this special case: due to the combinatorial structure of the problem, we can explicitly determine the first-order term in the error of the approximation. We finish by stating this result.

Throughout, all asymptotic relations are meant as $n \rightarrow \infty$.

2. Poisson approximation in a Poisson limit theorem

Gnedenko and Kolmogorov [4, p. 132] gave necessary and sufficient conditions for sums of independent infinitesimal random variables to converge to the Poisson law. In the case of nonnegative integer-valued random variables, their limit theorem can be stated as follows.

Theorem 1. *Let $\{Y_{n1}, Y_{n2}, \dots, Y_{nr_n}\}_{n \in \mathbb{N}}$ be a triangular array of row-wise independent, non-negative integer-valued random variables such that*

$$\min_{1 \leq k \leq r_n} P(Y_{nk} = 0) \rightarrow 1 \quad \text{as } n \rightarrow \infty, \tag{3}$$

$$\sum_{k=1}^{r_n} P(Y_{nk} \geq 1) \rightarrow \lambda \quad \text{as } n \rightarrow \infty \ (\lambda > 0 \text{ constant}), \tag{4}$$

$$\sum_{k=1}^{r_n} P(Y_{nk} \geq 2) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{5}$$

Then

$$Y_n := \sum_{k=1}^{r_n} Y_{nk} \xrightarrow{D} N_\lambda \quad \text{as } n \rightarrow \infty,$$

where N_λ is a Poisson random variable with parameter λ and ' \xrightarrow{D} ' denotes convergence in distribution.

We shall refine the obvious approximation of the Y_n s that the limit theorem suggests by approximating the distribution of each of the Y_n random variables not with the limiting Poisson distribution, but with a Poisson distribution that has a suitably chosen parameter that depends on n , namely by the distribution of $N_{\lambda_n} \sim \text{Poisson}(\lambda_n)$, where

$$\lambda_n = \sum_{k=1}^{r_n} P(Y_{nk} \geq 1).$$

Theorem 2. (The upper bound.) *We have*

$$d_{TV}(Y_n, N_{\lambda_n}) \leq \sum_{k=1}^{r_n} [P(Y_{nk} \geq 2) + P(Y_{nk} \geq 1)^2].$$

Proof. The proof follows the same argument as in [2, p. 181]. For each $k = 1, 2, \dots, r_n$, $n \in \mathbb{N}$, we define the random variable

$$I_{nk} := \begin{cases} 0 & \text{if } Y_{nk} = 0, \\ 1 & \text{if } Y_{nk} \geq 1. \end{cases}$$

Thus, for each $n \in \mathbb{N}$, $I_n := \sum_{k=1}^{r_n} I_{nk}$ is a sum of independent Bernoulli random variables with success probabilities $q_{nk} := P(Y_{nk} \geq 1)$, $k = 1, 2, \dots, r_n$. By Le Cam's inequality [5],

$$d_{TV}(I_n, N_{\lambda_n}) \leq \sum_{k=1}^{r_n} q_{nk}^2 = \sum_{k=1}^{r_n} P(Y_{nk} \geq 1)^2.$$

Also, for any two random variables X and Y defined on the same probability space, the coupling inequality (see, for example, [6, p. 12]) says that

$$d_{TV}(X, Y) \leq P(X \neq Y);$$

hence, we have

$$\begin{aligned} d_{TV}(Y_n, I_n) &\leq P\left(\sum_{k=1}^{r_n} Y_{nk} \neq \sum_{k=1}^{r_n} I_{nk}\right) \\ &= P\left(\bigcup_{k=1}^{r_n} \{Y_{nk} \neq I_{nk}\}\right) \\ &\leq \sum_{k=1}^{r_n} P(Y_{nk} \neq I_{nk}) \\ &= \sum_{k=1}^{r_n} P(Y_{nk} \geq 2). \end{aligned}$$

Putting these two bounds together in

$$d_{TV}(Y_n, N_{\lambda_n}) \leq d_{TV}(Y_n, I_n) + d_{TV}(I_n, N_{\lambda_n}),$$

the assertion of the theorem follows.

Theorem 3. (The lower bound.) *For all n such that $\min_{1 \leq k \leq r_n} P(Y_{nk} = 0) \geq \frac{3}{4}$,*

$$d_{TV}(Y_n, N_{\lambda_n}) \geq \frac{1}{10} \left(\prod_{k=1}^{r_n} P(Y_{nk} = 0) \right) \sum_{k=1}^{r_n} [P(Y_{nk} \geq 2) + P(Y_{nk} \geq 1)^2].$$

Before turning to the proof of Theorem 3, we prove a simple result that will be needed later.

Proposition 1. *If $0 \leq y_i \leq x_i \leq 1$ for all $i = 1, 2, \dots, n$, $n \in \mathbb{N}$, then*

$$\left(\prod_{i=1}^n y_i\right) \sum_{i=1}^n (x_i - y_i) \leq \prod_{i=1}^n x_i - \prod_{i=1}^n y_i \leq \sum_{i=1}^n (x_i - y_i).$$

Proof. Defining $y_0 := 1$, we can write the difference of the two products in the form of a telescopic sum; thus,

$$\begin{aligned} \prod_{i=1}^n x_i - \prod_{i=1}^n y_i &= \sum_{k=1}^n [y_1 \cdots y_{k-1} x_k \cdots x_n - y_1 \cdots y_k x_{k+1} \cdots x_n] \\ &= \sum_{k=1}^n (x_k - y_k)(y_1 \cdots y_{k-1} x_{k+1} \cdots x_n). \end{aligned}$$

Owing to our assumption on the y_i s, the last expression can be bounded from above and from below by

$$y_1 \cdots y_n \sum_{k=1}^n (x_k - y_k) \leq \sum_{k=1}^n (x_k - y_k)(y_1 \cdots y_{k-1} x_{k+1} \cdots x_n) \leq \sum_{k=1}^n (x_k - y_k),$$

and the assertion follows.

Proof of Theorem 3. We introduce the notation $P(Y_{nk} = 0) = p_{nk}$ and $P(Y_{nk} = 1) = (1 - p_{nk})\tilde{p}_{nk}$, $k = 1, 2, \dots, r_n$; thus, $\lambda_n = \sum_{k=1}^{r_n} (1 - p_{nk})$, $n \in \mathbb{N}$. We are going to prove the theorem by approximating the following elementary lower bound for the total variation distance of the distributions considered:

$$d_{TV}(Y_n, N_{\lambda_n}) \geq \frac{1}{2}|P(Y_n = 0) - P(N_{\lambda_n} = 0)| + \frac{1}{2}|P(Y_n = 1) - P(N_{\lambda_n} = 1)|, \tag{6}$$

which can be justified by taking $A = \{0\}$ and $A = \{1\}$ in (1).

We start by bounding the difference of the point probabilities at 0. Since

$$P(N_{\lambda_n} = 0) = \exp\{-\lambda_n\} = \prod_{k=1}^{r_n} \exp\{-(1 - p_{nk})\}, \tag{7}$$

$$P(Y_n = 0) = \prod_{k=1}^{r_n} p_{nk}, \tag{8}$$

and $\exp\{-(1 - p_{nk})\} \geq p_{nk}$ for all $k = 1, 2, \dots, r_n$, $n \in \mathbb{N}$, applying the proposition above yields

$$|P(N_{\lambda_n} = 0) - P(Y_n = 0)| \geq \left(\prod_{k=1}^{r_n} p_{nk}\right) \sum_{k=1}^{r_n} [\exp\{-(1 - p_{nk})\} - p_{nk}].$$

Since $1 - p_{nk} \leq 1$, we have

$$\exp\{-(1 - p_{nk})\} \geq 1 - (1 - p_{nk}) + \frac{1}{2}(1 - p_{nk})^2 - \frac{1}{6}(1 - p_{nk})^3 \geq p_{nk} + \frac{1}{3}(1 - p_{nk})^2$$

for $k = 1, 2, \dots, r_n$, $n \in \mathbb{N}$, which yields

$$|P(N_{\lambda_n} = 0) - P(Y_n = 0)| \geq \frac{1}{3} \left(\prod_{k=1}^{r_n} p_{nk}\right) \sum_{k=1}^{r_n} (1 - p_{nk})^2. \tag{9}$$

This inequality implies the assertion of the theorem in the case when $3 \sum_{k=1}^{r_n} (1 - p_{nk})^2 \geq 2 \sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk})$, because we can bound $\frac{3}{5}$ ths of the sum in the display above using this assumption. In fact, in this case we obtain a better bound than the one we aimed at. Otherwise, if $2 \sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) \geq 3 \sum_{k=1}^{r_n} (1 - p_{nk})^2$, we need to examine the point probabilities at 1 too to improve our current bound.

We have

$$P(N_{\lambda_n} = 1) = \lambda_n \exp\{-\lambda_n\} = \sum_{k=1}^{r_n} (1 - p_{nk}) \exp\left\{-\sum_{k=1}^{r_n} (1 - p_{nk})\right\},$$

and since, for an arbitrary $n \in \mathbb{N}$, $Y_n = 1$ if and only if, for $k = 1, 2, \dots, r_n$, exactly one of the Y_{nk} s is 1 and the rest are 0,

$$P(Y_n = 1) = \left(\prod_{k=1}^{r_n} p_{nk} \right) \sum_{k=1}^{r_n} \frac{(1 - p_{nk}) \tilde{p}_{nk}}{p_{nk}}.$$

Some elementary algebra gives

$$\begin{aligned} P(N_{\lambda_n} = 1) - P(Y_n = 1) &= \prod_{k=1}^{r_n} p_{nk} \left(\sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) - \sum_{k=1}^{r_n} \frac{(1 - p_{nk})^2 \tilde{p}_{nk}}{p_{nk}} \right) \\ &\quad + \left(\exp \left\{ - \sum_{k=1}^{r_n} (1 - p_{nk}) \right\} - \prod_{k=1}^{r_n} p_{nk} \right) \sum_{k=1}^{r_n} (1 - p_{nk}), \end{aligned}$$

where in the second term we recognize the point probabilities at 0. Using

$$\frac{\tilde{p}_{nk}}{p_{nk}} \leq \frac{1}{\min_{1 \leq k \leq r_n} p_{nk}}$$

and the fact that the difference in the point probabilities at 0 in the formula above is always positive, we obtain

$$P(N_{\lambda_n} = 1) - P(Y_n = 1) \geq \prod_{k=1}^{r_n} p_{nk} \left(\sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) - \frac{1}{\min_{1 \leq k \leq r_n} p_{nk}} \sum_{k=1}^{r_n} (1 - p_{nk})^2 \right).$$

From this, by (9) we obtain

$$\begin{aligned} &P(N_{\lambda_n} = 0) - P(Y_n = 0) + P(N_{\lambda_n} = 1) - P(Y_n = 1) \\ &\geq \left(\prod_{k=1}^{r_n} p_{nk} \right) \left(\sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) + \left[\frac{1}{3} - \frac{1}{\min_{1 \leq k \leq r_n} p_{nk}} \right] \sum_{k=1}^{r_n} (1 - p_{nk})^2 \right). \end{aligned}$$

Now $\frac{1}{3} - 1/\min_{1 \leq k \leq r_n} p_{nk} \geq -1$ in the range of n for which the assumption $\min_{1 \leq k \leq r_n} p_{nk} \geq \frac{3}{4}$ of the theorem holds; thus,

$$\begin{aligned} &P(N_{\lambda_n} = 0) - P(Y_n = 0) + P(N_{\lambda_n} = 1) - P(Y_n = 1) \\ &\geq \left(\prod_{k=1}^{r_n} p_{nk} \right) \left(\sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) - \sum_{k=1}^{r_n} (1 - p_{nk})^2 \right), \end{aligned}$$

and it can be seen that the latter bound is at most

$$\frac{1}{5} \left(\prod_{k=1}^{r_n} p_{nk} \right) \left(\sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) + \sum_{k=1}^{r_n} (1 - p_{nk})^2 \right)$$

for all n such that $2 \sum_{k=1}^{r_n} (1 - p_{nk})(1 - \tilde{p}_{nk}) \geq 3 \sum_{k=1}^{r_n} (1 - p_{nk})^2$. This together with (6) completes the proof.

Theorem 2 and Theorem 3 together state that the order of the error of our Poisson approximation for the random variables in Theorem 1 is $\sum_{k=1}^{r_n} [\mathbb{P}(Y_{nk} \notin \{0, 1\}) + \mathbb{P}(Y_{nk} \geq 1)^2]$. Barbour and Hall [1] proved similar results using Stein’s method: they approximated a sum $\sum_{j=1}^n Y_j$ of independent, nonnegative integer-valued random variables with a Poisson variable that has mean $\sum_{j=1}^n \mathbb{P}(Y_j = 1)$ or $\sum_{j=1}^n \mathbb{E}(Y_j)$. (Note that the parameter of our approximating Poisson random variable is between these two values.) Their bounds are expressed differently, and involve second moments of the random variables Y_j . Moreover, their lower bounds would yield no useful information at all in the application to be considered in the next section.

We also obtain the following result.

Corollary 1. *For the rate of convergence in Theorem 1, we have the upper bound*

$$d_{TV}(Y_n, N_{\lambda_n}) \leq \sum_{k=1}^{r_n} [\mathbb{P}(Y_{nk} \geq 2) + \mathbb{P}(Y_{nk} \geq 1)^2] + \left| \sum_{k=1}^{r_n} \mathbb{P}(Y_{nk} \geq 1) - \lambda \right|, \quad n \in \mathbb{N}.$$

Proof. Since

$$d_{TV}(Y_n, N_{\lambda_n}) \leq d_{TV}(Y_n, N_{\lambda_n}) + d_{TV}(N_{\lambda_n}, N_{\lambda}),$$

the assertion follows from Theorem 2 and because, for any $N_{\nu_1} \sim \text{Poisson}(\nu_1)$ and $N_{\nu_2} \sim \text{Poisson}(\nu_2)$, where $0 < \nu_1 < \nu_2$, we have

$$d_{TV}(N_{\nu_1}, N_{\nu_2}) \leq \min\{1, \nu_2^{-1/2}\}(\nu_2 - \nu_1).$$

For reference, see, for example, [2, Remark 1.1.4].

3. Coupon collecting with an approximately Poisson distributed waiting time

We begin this section by examining how the coupon collector’s problem defined in the introduction fits in the framework of the previous section. It can be seen that the following equality in distribution holds for $\tilde{W}_{n,m_n} := W_{n,m_n} - (n - m_n)$:

$$\tilde{W}_{n,m_n} \stackrel{D}{=} \sum_{i=m_n+1}^n \tilde{X}_{n,i},$$

where the \tilde{X}_{ni} random variables are independent, and $\tilde{X}_{n,i} + 1$ has geometric distributions with success probability i/n , $i \in \{m_n + 1, \dots, n\}$, $n \in \mathbb{N}$, that is,

$$\mathbb{P}(\tilde{X}_{n,i} + 1 = j) = \left(1 - \frac{i}{n}\right)^{j-1} \frac{i}{n}, \quad j \in \mathbb{N}, i \in \{m_n + 1, \dots, n\}.$$

The triangular array $\{\tilde{X}_{n,m_n+1}, \dots, \tilde{X}_{n,n}\}_{n \in \mathbb{N}}$ satisfies the conditions of Theorem 1. The variables of the array are infinitesimal, i.e. they satisfy condition (3): for any $0 < \varepsilon < 1$,

$$\begin{aligned} \max_{m_n+1 \leq i \leq n} \mathbb{P}(\tilde{X}_{n,i} > \varepsilon) &= 1 - \min_{m_n+1 \leq i \leq n} \mathbb{P}(\tilde{X}_{n,i} = 0) \\ &= 1 - \min_{m_n+1 \leq i \leq n} \frac{i}{n} \\ &= \frac{n - m_n + 1}{n} \\ &\rightarrow 0, \end{aligned}$$

by (2); they also satisfy (4) and (5), because

$$\sum_{i=m_n+1}^n P(Y_{nk} \geq 1) = \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right) = \frac{(n - m_n)(n - m_n - 1)}{2n} \rightarrow \lambda$$

and

$$\sum_{i=m_n+1}^n P(Y_{nk} \geq 2) = \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right)^2 = \frac{(n - m_n)(n - m_n - 1/2)(n - m_n - 1)}{3n^2} \rightarrow 0,$$

by (2) again. Thus, we see that the limit theorem proved in [3] concerning the coupon collector’s problem is a special case of the Gnedeno–Kolmogorov theorem [4, p. 132]. If we apply the results of the previous section to \tilde{W}_{n,m_n} , we obtain the following result.

Corollary 2. *If $\{m_n\}_{n \in \mathbb{N}}$ is a sequence of integers that satisfies (2) then the error of the approximation of the coupon collector’s waiting time, \tilde{W}_{n,m_n} , with the Poisson random variable N_{λ_n} , which has mean $\lambda_n = \sum_{i=m_n+1}^n (1 - i/n)$, is of order $\sum_{i=m_n+1}^n (1 - i/n)^2$. In fact, for all n such that $\min_{m_n+1 \leq i \leq n} i/n \geq \frac{3}{4}$,*

$$\frac{1}{5} \left(\prod_{i=m_n+1}^n \frac{i}{n} \right) \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right)^2 \leq d_{TV}(\tilde{W}_{n,m_n}, N_{\lambda_n}) \leq 2 \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right)^2.$$

Corollary 3. *For the rate of convergence in the Poisson limit theorem concerning the coupon collector’s problem, we have the upper bound*

$$d_{TV}(\tilde{W}_{n,m_n}, N_{\lambda}) \leq 2 \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right)^2 + \left| \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right) - \lambda \right|.$$

Finally, we note that the Poisson approximation of Corollary 2 can be refined. Using the special combinatorial structure of the coupon collector’s problem, with elementary but delicate combinatorial considerations and calculations, we can derive the first asymptotic corrections of the $P(\tilde{W}_{n,m_n} = k)$, $k = 0, 1, \dots$, probabilities to the corresponding Poisson point probabilities. We state this result in the following theorem, the proof can be found in [8].

Theorem 4. *If $\{m_n\}_{n \in \mathbb{N}}$ is a sequence of nonnegative integers that satisfies (2),*

$$\lambda_n = \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right), \quad \text{and} \quad \lambda_{n,2} = \sum_{i=m_n+1}^n \left(1 - \frac{i}{n}\right)^2,$$

then

$$\begin{aligned} P(\tilde{W}_{n,m_n} = 0) &= \exp\{-\lambda_n\} - \exp\{-\lambda_n\} \frac{\lambda_{n,2}}{2} + O\left(\frac{1}{n}\right), \\ P(\tilde{W}_{n,m_n} = 1) &= \exp\{-\lambda_n\} \lambda_n - \exp\{-\lambda_n\} \lambda_n \frac{\lambda_{n,2}}{2} + O\left(\frac{1}{n}\right), \\ P(\tilde{W}_{n,m_n} = k) &= \exp\{-\lambda_n\} \frac{\lambda_n^k}{k!} + \exp\{-\lambda_n\} \left(\frac{\lambda_n^{k-2}}{(k-2)!} - \frac{\lambda_n^k}{k!} \right) \frac{\lambda_{n,2}}{2} + O\left(\frac{1}{n}\right), \quad k \geq 2. \end{aligned}$$

We note that $\lambda_{n,2} = (2\lambda_n)^{3/2} / 3\sqrt{n} + O(1/n)$.

The results of this section complete the refinement of the limit theorems proved by Baum and Billingsley [3], who determined all possible limiting distributions of the appropriately standardized waiting times of the coupon collector, which vary according to the way the sequence of m_n s behaves as $n \rightarrow \infty$. They obtained three different nondegenerate limiting distributions: a ‘Gumbel-like’ distribution when m_n is constant for all n , the rate of convergence is given in [9]; a standard normal distribution when $m_n \rightarrow \infty$ and $(n - m_n)/\sqrt{n} \rightarrow \infty$, the rate of convergence is treated in [7]; and, finally, a Poisson distribution with mean λ when m_n is as in (2).

Acknowledgements

I would like to thank Vilmos Totik for his ideas on complementing my preliminary results. I am also grateful to Andrew Barbour for his suggestions.

References

- [1] BARBOUR, A. D. AND HALL, P. (1984). On the rate of Poisson convergence. *Math. Proc. Camb. Phil. Soc.* **95**, 473–480.
- [2] BARBOUR, A. D., HOLST, L. AND JANSON, S. (1992). *Poisson Approximation* (Oxford Stud. Prob. **2**). Clarendon Press, New York.
- [3] BAUM, L. E. AND BILLINGSLEY, P. (1965). Asymptotic distributions for the coupon collector’s problem. *Ann. Math. Statist.* **36**, 1835–1839.
- [4] GNEDENKO, B. V. AND KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Cambridge, MA.
- [5] LE CAM, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* **10**, 1181–1197.
- [6] LINDVALL, T. (1992). *Lectures on the Coupling Method*. John Wiley, New York.
- [7] PÓSFAL, A. (2007). Rates of convergence for normal approximation in incomplete coupon collection. *Acta. Sci. Math.* **73**, 333–348.
- [8] PÓSFAL, A. (2009). A supplement to the paper Poisson approximation in a Poisson limit theorem inspired by coupon collecting. Preprint. Available at <http://arxiv.org/abs/0904.4924>.
- [9] PÓSFAL, A. AND CSÖRGŐ, S. (2008). Asymptotic approximations for coupon collectors. *Studia Sci. Math. Hung.* **46**, 61–96.