РОССИЙСКАЯ АКАДЕМИЯ НАУК

МАТЕМАТИЧЕСКИЙ ИНСТИТУТ ИМ. В.А. СТЕКЛОВА
РОССИЙСКОЙ АКАДЕМИИ НАУК

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

ВЫПУСК 3

Том 62

ИЮЛЬ—АВГУСТ—СЕНТЯБРЬ

Журнал основан А. Н. Колмогоровым в январе 1956 года
Выходит 4 раза в год

МОСКВА · 2017

Журнал осуществляет свою деятельность под руководством Отделения математических наук РАН

Адрес редакции: 119991, Москва, ул. Губкина, 8, комн. 216;
тел. +7 (499) 941-01-81; e-mail: tvp@mi.ras.ru;
http://www.mathnet.ru/tvp

Заведующая редакцией М. В. Хатунцева

Учредители: Российская академия наук,
Математический институт им. В. А. Стеклова Российской академии наук

Издатель: Математический институт им. В. А. Стеклова Российской академии наук
119991 Москва, ул. Губкина, 8

Отпечатано в ФГУП «Издательство «Наука» (Типография «Наука»),
121009 Москва, Шубинский пер., 6

Подписано к печати 27.07.2017. Формат бумаги 70 × 100 1/16.
Цифровая печать. Усл.печ.л. 17,1. Усл.кр.-отт. 1,9 тыс. Бум.л. 6,6.
Уч.-изд.л. 20,9. Тираж 112 экз. Зак. 1266.

© 2017 г. ***AFENDRAS G.***, ***PAPADATOS N.****

A FACTORIAL MOMENT DISTANCE AND AN APPLICATION TO THE MATCHING PROBLEM

В этой заметке мы вводим понятие расстояния факториальных моментов для неотрицательных целочисленных случайных величин и сравниваем его с расстоянием по вариации. Кроме этого, мы изучаем скорость сходимости в классической задаче о составлении пар и в обобщенной задаче о составлении пар.

Ключевые слова и фразы: расстояние факториальных моментов, расстояние по вариации, задача о составлении пар.

DOI: https://doi.org/10.4213/tvp5125

1. Introduction. Let $\boldsymbol{\pi}_n = (\pi_n(1), \ldots, \pi_n(n))$ be a random permutation of $T_n = \{1, \ldots, n\}$, in the sense that $\boldsymbol{\pi}_n$ is uniformly distributed over $n!$ permutations of T_n. A number j is a fixed point of $\boldsymbol{\pi}_n$ if $\pi_n(j) = j$. Denote by Z_n the total number of fixed points of $\boldsymbol{\pi}_n$,

$$Z_n = \sum_{j=1}^{n} \mathbf{I}\{\pi_n(j) = j\},$$

where $\mathbf{I}$ stands for the indicator function. The study of Z_n corresponds to the famous *matching problem*, introduced by de Montmort in 1708 [5]. Obviously, Z_n can take the values $0, 1, \ldots, n-2, n$, and its exact distribution, using standard combinatorial arguments, is found to be

$$\mathbf{P}(Z_n = j) = \frac{1}{j!} \sum_{k=0}^{n-j} \frac{(-1)^k}{k!}, \qquad j = 0, 1, \ldots, n-2, n.$$

It is obvious that Z_n converges in law to Z, where Z is the standard Poisson distribution, Poi(1). Furthermore, the Poisson approximation is very accurate even for small n (evidence of this may be found in [1]). Bounds on the error of the Poisson approximation in the matching problem, especially

*Department of Biostatistics, The State University of New York at Buffalo, Buffalo, NY, USA; e-mail: gafendra@buffalo.edu

**Department of Mathematics, University of Athens, Panepistemiopolis, Greece; e-mail: npapadat@math.uoa.gr

[1)] This work was partially supported by the University of Athens' Research Fund under grant 70/4/5637. Research has been co-financed by the European Union (European Social Fund — ESF) and Greek National Funds through the Operational Program «Education and Lifelong Learning» of the National Strategic Reference Framework (NSRF) — Research Funding Program: ARISTEIA, grant 4357.

concerning the total variation distance, are also well known. Recall that the total variation distance of any two random variables X_1 and X_2 is defined as

$$d_{\rm tv}(X_1, X_2) = \sup_{A \in \mathscr{B}(\mathbf{R})} |\mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A)|,$$

where $\mathscr{B}(\mathbf{R})$ is the Borel σ-algebra of $\mathbf{R}$. An appealing result is given by Diaconis [6], who proved that $d_{\rm tv}(Z_n, Z) \leqslant 2^n/n!$. This bound has been improved by DasGupta (see [3], [4]):

$$d_{\rm tv}(Z_n, Z) \leqslant \frac{2^n}{(n+1)!}. \tag{1.1}$$

It can be seen that $d_{\rm tv}(Z_n, Z) \sim 2^n/(n+1)!$, where $a_n \sim b_n$ means that $\lim_n(a_n/b_n) = 1$; for a proof of a more general result see Theorem 3.2. Therefore, the bound (1.1) is of the correct order.

Consider now the sets of discrete random variables

$$\mathscr{D}_n := \{X \colon \mathbf{P}(X \in \{0, 1, \ldots, n\}) = 1\},$$
$$\mathscr{D}_\infty := \{X \colon \mathbf{P}(X \in \{0, 1, \ldots\}) = 1\}.$$

Since the first n moments of Z_n and Z are identical and $Z_n \in \mathscr{D}_n$, $Z \in \mathscr{D}_\infty$, one might think that

$$\inf_{X \in \mathscr{D}_n} \{d_{\rm tv}(X, Z)\} \sim d_{\rm tv}(Z_n, Z) \sim \frac{2^n}{(n+1)!}. \tag{1.2}$$

However, (1.2) is not true. In fact,

$$\min_{X \in \mathscr{D}_n} \{d_{\rm tv}(X, Z)\} = 1 - e^{-1} \sum_{j=0}^{n} \frac{1}{j!} \sim \frac{e^{-1}}{(n+1)!}. \tag{1.3}$$

Indeed, for any $X_1, X_2 \in \mathscr{D}_\infty$ with probability mass functions p_1 and p_2, the total variation distance can be expressed as

$$d_{\rm tv}(X_1, X_2) = \frac{1}{2} \sum_{j=0}^{\infty} |p_1(j) - p_2(j)| = \sum_{j=0}^{\infty} (p_1(j) - p_2(j))^+, \tag{1.4}$$

where $x^+ = \max\{x, 0\}$. Thus, for any $X_1 \in \mathscr{D}_n$ (so that $p_1(j) = 0$ for all $j > n$), we get

$$\begin{aligned} d_{\rm tv}(X_1, X_2) &= \sum_{j=0}^{n} (p_1(j) - p_2(j))^+ \geqslant \sum_{j=0}^{n} (p_1(j) - p_2(j)) \\ &= 1 - \sum_{j=0}^{n} p_2(j) = \mathbf{P}(X_2 > n), \end{aligned}$$

with equality if and only if $p_1(j) \geqslant p_2(j)$, $j = 0, 1, \ldots, n$. Applying the preceding inequality to $p_2(j) = \mathbf{P}(Z = j) = e^{-1}/j!$ we get the equality in (1.3), and the minimum is attained by any random variable $X \in \mathscr{D}_n$ with $\mathbf{P}(X = j) \geqslant e^{-1}/j!$, $j = 0, 1, \ldots, n$. Furthermore, the well-known Cauchy remainder in the Taylor expansion reads as

$$f(x) - \sum_{j=0}^{n} \frac{f^{(j)}(0)}{j!} x^j = \frac{1}{n!} \int_0^x (x-y)^n f^{(n+1)}(y)\, dy. \qquad (1.5)$$

Applying (1.5) to $f(x) = e^x$ we get the expression

$$1 - e^{-1} \sum_{j=0}^{n} \frac{1}{j!} = e^{-1}\left(e - \sum_{j=0}^{n} \frac{1}{j!}\right) = \frac{e^{-1}}{n!} \int_0^1 (1-y)^n e^y\, dy,$$

and by the obvious inequalities $1 < e^y < 1 + (e-1)y$, $0 < y < 1$, we have

$$\frac{1}{n+1} < \int_0^1 (1-y)^n e^y\, dy < \frac{1}{n+1}\left(1 + \frac{e-1}{n+2}\right).$$

It follows that

$$\frac{e^{-1}}{(n+1)!} < \min_{X \in \mathscr{D}_n} \{d_{\mathrm{tv}}(X, Z)\} = 1 - e^{-1} \sum_{j=0}^{n} \frac{1}{j!} < \frac{e^{-1}}{(n+1)!}\left(1 + \frac{e-1}{n+2}\right),$$

and, therefore, $\min_{X \in \mathscr{D}_n}\{d_{\mathrm{tv}}(X, Z)\} \sim e^{-1}/(n+1)!$.

In the present note we introduce and study a class of factorial moment distances, $\{d_\alpha, \alpha > 0\}$. These metrics are designed to capture the discrepancy among discrete distributions with finite moment generating function in a neighborhood of zero and, in addition, they satisfy the desirable property $\min_{X \in \mathscr{D}_n}\{d_\alpha(X, Z)\} = d_\alpha(Z_n, Z)$. In Section 3 we study the rate of convergence in a generalized matching problem, and we present closed form expansions and sharp inequalities for the factorial moment distance and the variational distance.

2. The factorial moment distance. We start with the following observation: For the random variables Z and Z_n,

$$\mathbf{E}(Z)_k = 1 \quad \text{and} \quad \mathbf{E}(Z_n)_k = \mathbf{I}_{\{k \leqslant n\}}, \qquad k = 0, 1, \ldots, \qquad (2.1)$$

where $\mathbf{E}(X)_k$ denotes the k-th order descending factorial moment of X (for each $x \in \mathbf{R}$, $(x)_0 = 1$ and $(x)_k = x(x-1)\cdots(x-k+1)$, $k = 1, 2\ldots$). For a proof of a more general result see Lemma 3.1.

The factorial moment distance will be defined in a suitable sub-class of discrete random variables, as follows. For each $t \geqslant 0$, we define

$$\mathscr{X}(t) := \{X \in \mathscr{D}_\infty \colon \text{there exists } t' > t \text{ such that } P_X(1+t') < \infty\}, \quad (2.2)$$

where $P_X(u) = \mathbf{E}u^X$ is the probability generating function of X. Also, we define

$$\mathscr{X}(\infty) := \bigcap_{t\in[0,\infty)} \mathscr{X}(t) = \{X \in \mathscr{D}_\infty \colon P_X(1+t') < \infty \text{ for any } t' > 0\}. \quad (2.3)$$

Note that if $X \in \mathscr{D}_n$ for some n, then $X \in \mathscr{X}(t)$ for each $t \in [0,\infty]$; therefore, each $\mathscr{X}(t)$ is nonempty. For $0 \leqslant t_1 < t_2 \leqslant \infty$, it is obvious that $\mathscr{X}(t_2) \subset \mathscr{X}(t_1)$; that is, the family $\{\mathscr{X}(t), 0 \leqslant t \leqslant \infty\}$ is decreasing in t.

If $X \in \mathscr{X}(0)$, then there exists a $t' > 0$ such that $P_X(1+t') < \infty$, i.e., $\mathbf{E}e^{\theta X} < \infty$, where $\theta = \ln(1+t') > 0$. Since X is nonnegative, $\mathbf{E}e^{\theta X} < \infty$ implies that $\mathbf{E}e^{uX} < \infty$ for all $u \in (-\theta,\theta)$, which means that X has finite moment generating function at a neighborhood of zero. Therefore, X has finite moments of any order and its probability mass function is characterized by its moments; equivalently, X has finite descending factorial moment of any order and its probability mass function is characterized by these moments. This enables the following definition.

Definition 2.1. (a) Let $X_1, X_2 \in \mathscr{X}(0)$. For $\alpha > 0$, we define the *factorial moment distance of order* α of X_1, X_2 by

$$d_\alpha(X_1,X_2) := \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!} |\mathbf{E}(X_1)_k - \mathbf{E}(X_2)_k|. \quad (2.4)$$

(b) Let $X \in \mathscr{X}(0)$ and $\{X_n\}_{n=1}^\infty \subset \mathscr{X}(0)$. We say that X_n *converges in factorial moment distance of order* α to X, in symbols $X_n \xrightarrow{\alpha} X$, if

$$d_\alpha(X_n, X) \to 0 \quad \text{as} \quad n \to \infty.$$

One can easily check that the function $d_\alpha \colon \mathscr{X}(0) \times \mathscr{X}(0) \to [0,\infty]$ is a distance. Obviously, $X_n \xrightarrow{\alpha} X$ implies that the moments of X_n converge to the corresponding moments of X. Since every $X \in \mathscr{X}(0)$ is characterized by its moments, it follows that d_α convergence (for any $\alpha > 0$) is stronger than the convergence in law; the later is equivalent to the convergence in total variation (see Wang [8]). Of course, the converse is not true even in $\mathscr{X}(\infty)$. For example, consider the random variable X with $\mathbf{P}(X=0) = 1$, and the sequence of random variables $\{X_n\}_{n=1}^\infty$, where each X_n has probability mass function

$$p_n(j) = \begin{cases} 1 - \dfrac{1}{n}, & j = 0, \\ \dfrac{1}{n}, & j = n. \end{cases}$$

It is obvious that $\{X, X_1, X_2, \dots\} \subset \mathscr{X}(\infty)$, and the total variation distance is

$$d_{\mathrm{tv}}(X_n, X) = \frac{1}{n} \to 0 \quad \text{as} \quad n \to \infty.$$

Moreover, since $\mathbf{E}(X)_k = 0$ and $\mathbf{E}(X_n)_k = (n-1)_{k-1}\mathbf{I}\{k \leqslant n\}$ for all $k = 1, 2, \dots$, the d_α distance does not converge to zero:

$$d_\alpha(X_n, X) = \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!}(n-1)_{k-1}\mathbf{I}\{k \leqslant n\}$$
$$> \frac{\alpha}{2}(n-1)\mathbf{I}\{2 \leqslant n\} \to \infty \quad \text{as} \quad n \to \infty.$$

Remark 2.1. Let $X \in \mathscr{D}_n \setminus \{X\colon X \in \mathscr{D}_n,\ X \stackrel{\mathrm{d}}{=} Z_n\}$. It is obvious that $\mathbf{E}(X)_k = 0$ for all $k > n$, and we can find an index $k \in \{1, \dots, n\}$ such that $\mathbf{E}(X)_k \neq 1$. From (2.1) and (2.4) we see that $d_\alpha(X, Z) > d_\alpha(Z_n, Z)$. Hence,

$$\inf_{X \in \mathscr{D}_n} \{d_\alpha(X, Z)\} = d_\alpha(Z_n, Z) \quad \text{for all} \quad \alpha > 0.$$

Proposition 2.1. *Let* $0 < \alpha_1 < \alpha_2$ *and* $X_1, X_2 \in \mathscr{X}(0)$. *Then*
(a) $d_{\alpha_1}(X_1, X_2) \leqslant d_{\alpha_2}(X_1, X_2)$;
(b) *we cannot find a constant* $C = C(\alpha_1, \alpha_2) < 1$ *such that for all any random variables* $X_1, X_2 \in \mathscr{X}(0)$, $d_{\alpha_1}(X_1, X_2) \leqslant C d_{\alpha_2}(X_1, X_2)$.

Proof. (a) is obvious. To see (b), it suffices to consider X_1 and X_2 with $\mathbf{P}(X_1 = 0) = \mathbf{P}(X_2 = 1) = 1$. Then $d_\alpha(X_1, X_2) = 1$ for every $\alpha > 0$. The proposition is proved.

From (a) of the preceding proposition, $X_n \stackrel{\alpha_2}{\to} X$ implies $X_n \stackrel{\alpha_1}{\to} X$ for every $\alpha_1 < \alpha_2$. In the sequel we shall show that for any $\alpha \geqslant 2$, the inequality $d_{\mathrm{tv}}(X_n, X) \leqslant d_\alpha(X_n, X)$ holds true, provided $\{X, X_1, X_2, \dots\} \subseteq \mathscr{X}(1)$. To this end, we shall make use of the following «*moment inversion*» formula. Though this result is widely used (see, e.g., [2, p. 49]), we provide a short proof, specifying a condition under which the formula is valid.

Lemma 2.1. *If* $X \in \mathscr{X}(1)$, *then its probability mass function* p *can be written as*

$$p(j) = \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!}\binom{k}{j}\mathbf{E}(X)_k, \qquad j = 0, 1, \dots. \tag{2.5}$$

Proof. By the assumption $X \in \mathscr{X}(1)$, we can find a number $t' > 1$ such that $\mathbf{E}(1+t')^X = \sum_{j=0}^{\infty}(1+t')^j p(j) < \infty$. Since X is nonnegative, its probability generating function admits a Taylor expansion around 0 with radius of convergence $R \geqslant 1 + t' > 2$, i.e., $P(u) = \sum_{j=0}^{\infty} u^j p(j) \in \mathbf{R}$, $|u| < R$. It is well known that $\frac{d^k}{du^k}P(u)\big|_{u=1} = \mathbf{E}(X)_k$, and since P admits a Taylor expansion around 1 with radius of convergence $R' \geqslant t' > 1$, we have

$$P(u) = \sum_{k=0}^{\infty} \frac{\mathbf{E}(X)_k}{k!}(u-1)^k, \qquad |u-1| < R'.$$

Using the preceding expansion and the fact that $0 \in (1-R', 1+R')$ we get

$$p(j) = \frac{1}{j!} \cdot \frac{d^j}{du^j} P(u)\Big|_{u=0} = \frac{1}{j!} \sum_{k=j}^{\infty} \frac{(u-1)^{k-j}}{(k-j)!} \mathbf{E}(X)_k \Big|_{u=0}$$
$$= \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{j!\,(k-j)!} \mathbf{E}(X)_k,$$

completing the proof.

It should be noted that the condition $X \in \mathscr{X}(t)$ for some $t \in [0,1)$ is not sufficient for (2.5). As an example, consider the geometric random variable X with probability mass function $p(j) = 2^{-j-1}$, $j = 0, 1, \ldots$. It is clear that $X \notin \mathscr{X}(1)$, but $X \in \mathscr{X}(t)$ for each $t \in [0,1)$. The factorial moments of X are $\mathbf{E}(X)_k = k!$, $k = 0, 1, \ldots$, and the right-hand side of (2.5), $\sum_{k=j}^{\infty} (-1)^{k-j} \binom{k}{j}$, is a nonconvergent series.

Theorem 2.1. *If $X_1, X_2 \in \mathscr{X}(1)$, then $d_{\mathrm{tv}}(X_1, X_2) \leqslant d_\alpha(X_1, X_2)$ for each $\alpha \geqslant 2$.*

Proof. In view of Proposition 2.1, (a), it is enough to prove the desired result for $\alpha = 2$. By (1.4) and (2.5) we get

$$d_{\mathrm{tv}}(X_1, X_2) = \frac{1}{2} \sum_{j=0}^{\infty} \Bigg| \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!} \binom{k}{j} (\mathbf{E}(X_1)_k - \mathbf{E}(X_2)_k) \Bigg|$$
$$\leqslant \frac{1}{2} \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} \frac{1}{k!} \binom{k}{j} |\mathbf{E}(X_1)_k - \mathbf{E}(X_2)_k|.$$

Interchanging the order of summation according to Tonelli's theorem, we have

$$d_{\mathrm{tv}}(X_1, X_2) \leqslant \frac{1}{2} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{k!} \binom{k}{j} |\mathbf{E}(X_1)_k - \mathbf{E}(X_2)_k|$$
$$= \sum_{k=0}^{\infty} \frac{2^{k-1}}{k!} |\mathbf{E}(X_1)_k - \mathbf{E}(X_2)_k|.$$

The proof is completed by the fact that $\mathbf{E}(X_1)_0 = \mathbf{E}(X_2)_0 = 1$.

Theorem 2.1 quantifies the fact that for any $\alpha \geqslant 2$, the d_α convergence (in $\mathscr{X}(1)$) implies the convergence in total variation, and provides convenient bounds for the rate of the total variation convergence. However, we note that such convenient bounds do not hold for $\alpha < 2$. In fact, for given $\alpha \in (0,2)$ and $t \geqslant 0$, we cannot find a finite constant $C = C(\alpha, t) > 0$ such that $d_{\mathrm{tv}}(X_1, X_2) \leqslant C d_\alpha(X_1, X_2)$ for all $X_1, X_2 \in \mathscr{X}(t)$ (see Remark 3.1).

3. An application to a generalized matching problem. Consider the classical matching problem where, now, we record only a proportion of the matches, due to a random censoring mechanism. The censoring mechanism decides independently to every individual match. Specifically, when a particular match occurs, the mechanism counts this match with probability λ, independently of the other matches, and ignores this match with probability $1-\lambda$, where $0<\lambda\leqslant 1$. We are now interested on the number $Z_n(\lambda)$ of the counted matches. The case $\lambda=1$ corresponds to the classical matching problem, where all coincidences are recorded, so that $Z_n=Z_n(1)$.

The probabilistic formulation is as follows: Let $\boldsymbol{\pi}_n=(\pi_n(1),\ldots,\pi_n(n))$ be a random permutation of $\{1,\ldots,n\}$, as in the introduction. Let also $J_1(\lambda),\ldots,J_n(\lambda)$ be independent and identically distributed Bernoulli(λ) random variables, independent of $\boldsymbol{\pi}_n$. The number $Z_n(\lambda)$ of the recorded coincidences can be written as

$$Z_n(\lambda)=\sum_{i=1}^{n}J_i(\lambda)\mathbf{I}\{\pi_n(i)=i\}.$$

Let $A_i=\{J_i(\lambda)=1\}$, $B_i=\{\pi_n(i)=i\}$, $E_i=A_i\cap B_i$, $i=1,\ldots,n$. Then $Z_n(\lambda)$ presents the number of the events E's that will occur and, by standard combinatorial arguments,

$$\begin{aligned}\mathbf{P}(Z_n(\lambda)=j)&=\mathbf{P}(\text{exactly } j \text{ among } E_1,\ldots,E_n \text{ occur})\\&=\sum_{i=j}^{n}(-1)^{i-j}\binom{i}{j}S_{i,n},\end{aligned}$$

where

$$S_{0,n}=1,\quad S_{i,n}=\sum_{1\leqslant k_1<\cdots<k_i\leqslant n}\mathbf{P}(E_{k_1}\cap\cdots\cap E_{k_i}),\qquad i=1,\ldots,n.$$

Since the A's are independent of the B's, we have

$$\mathbf{P}(E_{k_1}\cap\cdots\cap E_{k_i})=\mathbf{P}(A_{k_1}\cap\cdots\cap A_{k_i})\mathbf{P}(B_{k_1}\cap\cdots\cap B_{k_i})=\lambda^i\frac{(n-i)!}{n!},$$

so that

$$S_{i,n}=\binom{n}{i}\lambda^i\frac{(n-i)!}{n!}=\frac{\lambda^i}{i!},\qquad i=0,1,\ldots,n.$$

Therefore, the probability mass function of $Z_n(\lambda)$ is given by

$$\begin{aligned}p_{n;\lambda}(j):=\mathbf{P}(Z_n(\lambda)=j)&=\frac{1}{j!}\sum_{i=j}^{n}(-1)^{i-j}\frac{\lambda^i}{(i-j)!}\\&=\frac{\lambda^j}{j!}\sum_{i=0}^{n-j}\frac{(-\lambda)^i}{i!},\qquad j=0,1,\ldots,n.\end{aligned}\tag{3.1}$$

The generalized matching distribution (3.1) has been introduced by Niermann [7], who showed that $p_{n;\lambda}$ is a proper probability mass function for all $\lambda \in (0,1]$; however, Niermann did not give a probabilistic interpretation to the probability mass function $p_{n;\lambda}$, and derived its properties analytically.

Since

$$\lim_{n\to\infty} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!} = e^{-\lambda}$$

for any fixed j, we see that $p_{n;\lambda}$ converges pointwise to the probability mass function of $Z(\lambda)$, where $Z(\lambda)$ is a Poisson random variable with mean λ, Poi(λ). Interestingly enough, the Poisson approximation is extremely accurate; numerical results are shown in Niermann's work. Also, Niermann proved that $\mathbf{E}Z_n(\lambda) = \operatorname{Var} Z_n(\lambda) = \lambda$ for all $n \geqslant 2$ and $\lambda \in (0,1]$. In fact, the following general result shows that the first n moments of $Z_n(\lambda)$ and $Z(\lambda)$ are identical, giving some light to the amazing accuracy of the Poisson approximation.

Lemma 3.1. *For any* $\lambda \in (0,1]$, $\mathbf{E}\big(Z_n(\lambda)\big)_k = \lambda^k \mathbf{I}\{k \leqslant n\}$, $k = 1, 2, \ldots$.

Proof. For $k > n$ the relation is obvious, since $Z_n(\lambda) \in \mathscr{D}_n$. For $k = 1, \ldots, n-1$,

$$\begin{aligned}\mathbf{E}\big(Z_n(\lambda)\big)_k &= \sum_{j=k}^{n} \frac{\lambda^j}{(j-k)!} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!} \\ &= \lambda^k \sum_{r=0}^{n-k} \frac{\lambda^r}{r!} \sum_{i=0}^{(n-k)-r} \frac{(-\lambda)^i}{i!} = \lambda^k \sum_{r=0}^{n-k} p_{n-k;\lambda}(r),\end{aligned}$$

and, since $p_{n-k;\lambda}$ is a probability mass function supported on $\{0, 1, \ldots, n-k\}$, we get the desired result. For $k = n$, $\mathbf{E}\big(Z_n(\lambda)\big)_n = n!\, p_{n;\lambda}(n) = \lambda^n$, completing the proof.

Corollary 3.1. *For any* $\lambda \in (0,1]$ *and* $\alpha > 0$,

$$\inf_{X\in\mathscr{D}_n} \big\{d_\alpha\big(X, Z(\lambda)\big)\big\} = d_\alpha\big(Z_n(\lambda), Z(\lambda)\big).$$

Thus, for $\lambda \in (0,1]$, $Z_n(\lambda)$ minimizes the factorial moment distance from $Z(\lambda)$ over all random variables supported in a subset of $\{0, 1, \ldots, n\}$. Using (2.5) it is easily verified that $Z_n(\lambda)$ is unique. Moreover, it is worth pointing out that for $\lambda > 1$, we cannot find a random variable $X \in \mathscr{D}_n$ such that $\mathbf{E}(X)_k = \lambda^k \mathbf{I}\{k \leqslant n\}$ for all k. Indeed, since $\mathscr{D}_n \subset \mathscr{X}(\infty) \subset \mathscr{X}(1)$, assuming $X \in \mathscr{D}_n$ and $\mathbf{E}(X)_k = \lambda^k \mathbf{I}\{k \leqslant n\}$, we get from (2.5) that

$$0 \leqslant \mathbf{P}(X = n-1) = \frac{\lambda^{n-1}(1-\lambda)}{(n-1)!},$$

which implies that $\lambda \leqslant 1$. Therefore, finding $\inf_{X\in\mathscr{D}_n}\big\{d_\alpha\big(X, Z(\lambda)\big)\big\}$ for $\lambda > 1$ seems to be a rather difficult task.

We now evaluate some exact and asymptotic results for the factorial moment distance and the total variation distance between $Z_n(\lambda)$ and $Z(\lambda)$ when $\lambda \in (0,1]$.

Theorem 3.1. *Fix $\alpha > 0$ and $\lambda \in (0,1]$ and let $d_\alpha(n) := d_\alpha\big(Z_n(\lambda), Z(\lambda)\big)$. Then,*

$$d_\alpha(n) = \frac{\alpha^n \lambda^{n+1}}{n!} \int_0^1 (1-y)^n e^{\alpha\lambda y}\, dy. \tag{3.2}$$

Moreover, the following double inequality holds:

$$1 + \frac{\alpha\lambda}{n+2} + \frac{a^2\lambda^2}{(n+2)(n+3)} < \frac{(n+1)!}{\alpha^n\lambda^{n+1}} d_\alpha(n) < 1 + \frac{\alpha\lambda}{n+2} + \frac{a^2\lambda^2 e^{\alpha\lambda}}{(n+2)(n+3)}. \tag{3.3}$$

Hence, as $n \to \infty$,

$$\begin{gathered} d_\alpha(n) \sim \frac{\alpha^n\lambda^{n+1}}{(n+1)!} \quad \text{and, more precisely,} \\ d_\alpha(n) = \frac{\alpha^n\lambda^{n+1}}{(n+1)!}\left(1 + \frac{\alpha\lambda}{n+2} + o\Big(\frac{1}{n}\Big)\right). \end{gathered} \tag{3.4}$$

Proof. From the definition of d_α and in view of (1.5) and Lemma 3.1,

$$d_\alpha(n) = \frac{1}{\alpha}\sum_{k=n+1}^{\infty} \frac{(\alpha\lambda)^k}{k!} = \frac{1}{\alpha}\left(e^{\alpha\lambda} - \sum_{k=0}^{n} \frac{(\alpha\lambda)^k}{k!}\right) = \frac{1}{\alpha n!}\int_0^{\alpha\lambda} (\alpha\lambda - x)^n e^x\, dx,$$

and the substitution $x = \alpha\lambda y$ leads to (3.2). Now (3.3) follows from the inequalities

$$1 + \alpha\lambda y + \frac{1}{2}\alpha^2\lambda^2y^2 < e^{\alpha\lambda y} < 1 + \alpha\lambda y + \frac{1}{2}e^{\alpha\lambda}\alpha^2\lambda^2y^2, \qquad 0 < y < 1,$$

while (3.4) is obvious from (3.3). Theorem 3.1 is proved.

Theorems 2.1 and 3.1 give the following statement.

Corollary 3.2. *An upper bound for $d_{\mathrm{tv}}(Z_n, Z)$ is given by*

$$d_{\mathrm{tv}}(Z_n, Z) < \frac{2^n}{(n+1)!}\left(1 + \frac{2}{n+2} + \frac{4e^2}{(n+2)(n+3)}\right) \sim \frac{2^n}{(n+1)!}. \tag{3.5}$$

The bound in (3.2) is of the correct order, and the same is true for the better result (1.1), given by DasGupta [3], [4]. In contrast, the bound $d_{\mathrm{tv}}(Z_n, Z) \leqslant 2^n/n!$, given by Diaconis [6], is not asymptotically optimal, because $2^n/(n+1)! = o(2^n/n!)$. Thus, it is of some interest to point out that the factorial distance d_2 provides an optimal rate upper bound for the variational distance in the matching problem. The situation is similar for the generalized matching distribution, as the following result shows.

Theorem 3.2. *For any* $\lambda \in (0,1]$, *let* $d_{\mathrm{tv}}(n) := d_{\mathrm{tv}}\big(Z_n(\lambda), Z(\lambda)\big)$ *be the variational distance between* $Z_n(\lambda)$ *and* $Z(\lambda)$. *Then,*

$$d_{\mathrm{tv}}(n) = \frac{\lambda^{n+1}}{2n!}\int_0^1 [y^n + (2-y)^n]e^{-\lambda y}\,dy. \tag{3.6}$$

Moreover, the following inequalities hold:

$$\begin{aligned} d_{\mathrm{tv}}(n) &> \frac{2^n\lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2}\left(1 - \frac{1}{2^{n+1}}\right)\right), \\ d_{\mathrm{tv}}(n) &< \frac{2^n\lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2}\left(1 - \frac{1}{2^{n+1}}\right) + \frac{4\lambda^2}{(n+2)(n+3)}\left(1 - \frac{n+3}{2^{n+2}}\right)\right). \end{aligned} \tag{3.7}$$

Hence, as $n \to \infty$,

$$\begin{aligned} d_{\mathrm{tv}}(n) &\sim \frac{2^n\lambda^{n+1}}{(n+1)!} \quad \textit{and, more precisely,} \\ d_{\mathrm{tv}}(n) &= \frac{2^n\lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2} + o\left(\frac{1}{n}\right)\right). \end{aligned} \tag{3.8}$$

Proof. Clearly, (3.8) is an immediate consequence of inequalities (3.7). Moreover, the inequalities (3.7) are obtained from (3.6) and the fact that

$$1 - \lambda y < e^{-\lambda y} < 1 - \lambda y + \frac{1}{2}\lambda^2 y^2, \qquad 0 < y < 1.$$

It remains to show (3.6). From (1.4) with $p_1 = p_{n;\lambda}$ and p_2 the probability mass function of $\mathrm{Poi}(\lambda)$, we get

$$\begin{aligned} d_{\mathrm{tv}}(n) &= \sum_{j=0}^{n} \frac{\lambda^j}{j!}\left[\sum_{i=0}^{n-j}\frac{(-\lambda)^i}{i!} - e^{-\lambda}\right]^+ \\ &= \sum_{j=0}^{n} \frac{\lambda^j}{j!}\left[\frac{(-1)^{n-j}}{(n-j)!}\int_0^\lambda (\lambda - x)^{n-j}e^{-x}\,dx\right]^+, \end{aligned}$$

where the integral expansion is deduced by an application of (1.5) to the function $f(\lambda) = e^{-\lambda}$. Thus,

$$\begin{aligned} d_{\mathrm{tv}}(n) &= \sum_{k=0}^{n} \frac{\lambda^{n-k}}{(n-k)!}\left[\frac{(-1)^k}{k!}\int_0^\lambda (\lambda - x)^k e^{-x}\,dx\right]^+ \\ &= \frac{1}{n!}\int_0^\lambda e^{-x}\left(\sum_{k\ \mathrm{even}}\binom{n}{k}(\lambda - x)^k\lambda^{n-k}\right)dx. \end{aligned}$$

Since

$$\sum_{k\ \mathrm{even}}\binom{n}{k}(\lambda - x)^k\lambda^{n-k} = \frac{1}{2}\big[x^n + (2\lambda - x)^n\big],$$

we obtain

$$d_{\mathrm{tv}}(n) = \frac{1}{2n!}\int_0^{\lambda} [x^n + (2\lambda - x)^n] e^{-x}\,dx,$$

and a final change of variables $x = \lambda y$ yields (3.6). Theorem 3.2 is proved.

Remark 3.1. Although the factorial moment distance d_α dominates the variational distance when $\alpha \geqslant 2$, the situation for $\alpha \in (0,2)$ is completely different. To see this, assume that for some $\alpha \in (0,2)$ and $t \geqslant 0$ we can find a finite constant $C = C(\alpha,t) > 0$ such that

$$d_{\mathrm{tv}}(X_1,X_2) \leqslant C d_\alpha(X_1,X_2) \quad \text{for all } X_1,X_2 \in \mathscr{X}(t). \tag{3.9}$$

Obviously, Z and Z_n, $n = 1,2,\ldots$, lie in $\mathscr{X}(\infty) \subset \mathscr{X}(t)$. From Theorem 3.2 we know that

$$\lim_{n\to\infty} \frac{(n+1)!}{2^n} d_{\mathrm{tv}}(Z_n,Z) = 1.$$

On the other hand, from (3.3) with $\lambda = 1$,

$$d_\alpha(Z_n,Z) < \frac{\alpha^n}{(n+1)!}\left(1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^\alpha}{(n+2)(n+3)}\right),$$

and, since $\alpha/2 < 1$, this inequality contradicts (3.9):

$$\begin{aligned} 1 &= \lim_{n\to\infty} \frac{(n+1)!}{2^n} d_{\mathrm{tv}}(Z_n,Z) \\ &\leqslant C \lim_{n\to\infty}\left(\left(\frac{\alpha}{2}\right)^n \left(1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^\alpha}{(n+2)(n+3)}\right)\right) = 0. \end{aligned}$$

Acknowledgment. The authors would like to thank an anonymous referee for the detail reading of the manuscript, leading to an improvement of its presentation. The first author would like to thank Dr. M. Markatou for financial support.

REFERENCES

1. A. D. Barbour, L. Holst, S. Janson, *Poisson approximation*, Oxford Stud. Probab., **2**, The Clarendon Press, Oxford Univ. Press, New York, 1992, x+277 pp.
2. Ch. A. Charalambides, *Combinatorial methods in discrete distributions*, Wiley Ser. Probab. Stat., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2005, xiv+415 pp.
3. A. DasGupta, *The matching problem with random decks and the Poisson approximation*, Technical report #99-01, Purdue Univ., West Lafayette, IN, 1999, 18 pp.
4. A. DasGupta, "The matching, birthday and the strong birthday problem: a contemporary review", *J. Statist. Plann. Inference*, **130**:1-2 (2005), 377–389.

5. P. R. de Montmort, *Essay d'analyse sur les jeux de hazard*, Photographic reprint of the 2nd ed., Chelsea Publishing, New York, 1980, xlii+416 pp.; 1st ed., 1708; 2nd ed., Jacques Quillau, Paris, 1713.
6. P. Diaconis, "Application of the method of moments in probability and statistics", *Moments in mathematics* (San Antonio, Tex., 1987), Proc. Sympos. Appl. Math., **37**, Amer. Math. Soc., Providence, RI, 1987, 125–142.
7. S. Niermann, "A generalization of the matching distribution", *Statist. Papers*, **40**:2 (1999), 233–238.
8. Y. H. Wang, "A compound Poisson convergence theorem", *Ann. Probab.*, **19**:1 (1991), 452–455.

Поступила в редакцию
30.X.2015
Исправленный вариант
07.III.2016