RATE OF POISSON CONVERGENCE IN EQUIPROBABLE ALLOCATIONS

by

Wei–Liem Loh
Purdue University

Technical Report #89-04

Department of Statistics
Purdue University

February 1989

AMS 1980 subject classifications. Primary 60F05, Secondary 60C05.

Key words and phrases: Chen–Stein method, Poisson distribution, Equiprobable allocations.

*

RATE OF POISSON CONVERGENCE IN EQUIPROBABLE ALLOCATIONS

BY WEI-LIEM LOH

Purdue University

Consider the situation of ν balls and k urns. We assume that each of the ν balls gets into the ith urn with equal probability independently of the other balls. Let us denote by N_r the number of urns containing exactly r balls. In this paper we shall use the Chen-Stein method to obtain conditions whereby the law of N_r, or some appropriate function of N_r, converges to a Poisson distribution. This is quantified in terms of upper and lower bounds on the rate of convergence in terms of the total variation distance.

1 Introduction

Consider the situation of ν balls and k urns. We assume that each of the ν balls gets into the ith urn with equal probability independently of the other balls. Let us denote by N_r the number of urns containing exactly r balls. In this paper we shall use the Chen-Stein method (see Chen (1975), Stein (1970), (1986)) to obtain conditions whereby the law of N_r, or some suitably normalized function of N_r, converges to a Poisson distribution. This is quantified in terms of upper and lower bounds on the rate of convergence in terms of the total variation distance d which is defined as

$$d(\mu_1, \mu_2) \equiv \sup_{A \in Z^+} | \mu_1(A) - \mu_2(A) |,$$

where μ_1, μ_2 are arbitrary probability measures on Z^+, the set of non-negative integers. More precisely, in Section 2 we shall compute explicit non-asymptotic upper bounds on the rates of Poisson convergence of

1. N_r, $r \geq 0$, as $\nu/k \to \infty$,
2. $(N_0 - k + \nu)$, $[(\nu - N_1)/2]$ and N_r, $r \geq 2$, as $\nu/k \to 0$. Here $[.]$ denotes the greatest integer function.

In Section 3, we construct lower bounds on the rates of Poisson convergence of

1. N_r, $r \geq 0$, as $\nu/k \to \infty$,

2. $(N_0 - k + \nu)$ and N_r, $r \geq 2$, as $\nu/k \to 0$.

Under very weak conditions, an example being the expectation of the statistic of interest tending to a nonzero finite limit, the lower bounds are comparable to that of the corresponding upper bounds.

For simplicity of notation, we write $\mathcal{P}_\lambda$ as the Poisson distribution with mean λ and

$$\mathcal{P}_\lambda h = e^{-\lambda} \sum_{w=0}^{\infty} h(w) \frac{\lambda^w}{w!},$$

where h is a bounded real-valued function defined on Z^+. Furthermore for any random variable X, we denote the law of X by $\mathcal{L}(X)$. Kolchin, Sevastyanov and Chistyakov (1978) give a very comprehensive survey of this subject and related areas. We first state two preliminary results.

Proposition 1 (Poisson Identity) *In order that the random variable W taking values in Z^+ has a Poisson distribution with mean λ, it is necessary and sufficient that , for all bounded functions $f : Z^+ \to R$,*

$$E\{\lambda f(W+1) - W f(W)\} = 0.$$

This proposition was observed by Chen (1975). The proof can be found in Stein (1986).

Proposition 2 *Let $A \subseteq Z^+$ and I_A denote the indicator function of A. There exists a function $f : Z^+ \to R$ satisfying*

$$\lambda f(w+1) - w f(w) = I_A(w) - \mathcal{P}_\lambda I_A,$$

for all $w \in Z^+$ such that

$$\begin{aligned} \sup_{w \in Z^+} |\, f(w) \,| &\leq 1 \wedge 1.4\lambda^{-1/2}, \\ \sup_{w \in Z^+} |\, f(w+1) - f(w) \,| &\leq \lambda^{-1}(1 - e^{-\lambda}). \end{aligned}$$

Proposition 2 was proved in Barbour and Eagleson (1983).

2 Upper Bounds

In this section, we shall compute upper bounds on the rates of convergence of the distributions of

1. N_r, $r \geq 0$, to appropriate Poisson distributions as $\nu/k \to \infty$.

2. $N_0 - k + \nu$, $[(\nu - N_1)/2]$, N_r, $r \geq 2$, to appropriate Poisson distributions as $\nu/k \to 0$. Here $[x]$ denotes the greatest integer less than or equal to x.

Let X_i denote the number of balls in urn i. For integers r and s, we define $N_r^{(s)}$ to be a random variable having the conditional distribution of N_r given that $X_1 = s$. As Stein (1986) observed, for s fixed, N_r and $N_r^{(s)}$, $r = 0, \ldots, \nu$, can be defined on a common probability space in the following way. Uniformly distribute ν balls independently among k urns. This determines N_r, $r = 0, \ldots, \nu$. Now choose s balls and an urn at random, uniformly distributed over all possible choices, independent of N_r, $r = 0, \ldots, \nu$. We denote this urn as urn I. Remove the balls from urn I and distribute these balls uniformly among the remaining urns. Now put the s balls originally selected into urn I. This determines $N_r^{(s)}$, $r = 0, \ldots, \nu$.

Lemma 1 *With N_r and $N_r^{(s)}$ defined on the above common probability space, we have*

$$
\begin{aligned}
& E \mid N_r + I_{\{r=s\}} - N_r^{(s)} \mid \\
\leq\; & \frac{1}{r!}\left(\frac{\nu}{k}\right)^r \left(1 - \frac{1}{k}\right)^{\nu - r} \Big\{1 + \frac{\nu}{k} + \frac{rsk}{\nu} + s\left(1 - \frac{1}{k}\right)^{-s} \\
& + r\left(1 - \frac{1}{k}\right)^{1-r} + \frac{r^2}{(\nu + r - 1)}\left(\frac{\nu + r - 1}{\nu}\right)^r \Big\}.
\end{aligned}
$$

PROOF. Given $X_1, \ldots, X_k$, we define for $1 \leq i_1 < \ldots < i_j \leq s$, $0 \leq n \leq \nu$,

$$
Y_{i_1,\ldots,i_j}^{(n)} = \begin{cases} 1 & \text{if out of the } s \text{ balls picked, the } i_1\text{th,} \\ & \ldots, i_j\text{th balls come from the same} \\ & \text{urn originally containing } n \text{ balls,} \\ 0 & \text{otherwise.} \end{cases}
$$

Let I be uniformly distributed over the integers $1, \ldots, k$ and is independent of the $Y_{i_1,\ldots,i_j}^{(n)}$'s and X_i's. Given $I, X_1, \ldots, X_k$, we define for $1 \leq i_1 < \ldots <$

$i_j \le X_I$, $0 \le n \le \nu$,

$$Z^{(n)}_{i_1,\ldots,i_j} = \begin{cases} 1 & \text{if out of the } X_I \text{ balls in urn } I\text{, the } i_1\text{th,} \\ & \ldots\text{, } i_j\text{th balls are placed in the same} \\ & \text{urn originally containing } n \text{ balls,} \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} & E(N_r^{(s)} - N_r - I_{\{r=s\}})_+ \\ \le\ & E\{\sum_{m=1}^{s} \sum_{1\le i_1<\ldots<i_m\le s} Y^{(r+m)}_{i_1,\ldots,i_m} + \sum_{m=1}^{r\wedge X_I} \sum_{1\le i_1<\ldots<i_m\le X_I} Z^{(r-m)}_{i_1,\ldots,i_m}\} \\ \le\ & E\{\sum_{m=1}^{s} \frac{s!(r+m)!(\nu-m)!}{(s-m)!m!r!\nu!} I_{\{\nu\ge r+m\}} N_{r+m} \\ & + \sum_{m=1}^{r\wedge X_I} \frac{X_I!}{m!(X_I-m)!}(\frac{1}{k-1})^m N_{r-m}\} \\ \le\ & E\{\sum_{m=1}^{s\wedge(\nu-r)} \frac{s!}{(s-m)!m!r!}(\frac{\nu}{k})^r(1-\frac{1}{k})^{\nu-r-1}(\frac{1}{k-1})^{m-1} \\ & + \sum_{m=1}^{r} \frac{1}{k}\sum_{j=1}^{k} \frac{X_j!}{m!(X_j-m)!} I_{\{X_j\ge m\}}(\frac{1}{k-1})^m \sum_{i=1}^{k} I_{\{X_i=r-m\}}\}. \end{aligned}$$

On simplification, we get

$$\begin{aligned} E(N_r^{(s)} - N_r - I_{\{r=s\}})_+ \ \le\ & \frac{1}{r!}(\frac{\nu}{k})^r(1-\frac{1}{k})^{\nu-r}\{s(1-\frac{1}{k})^{-s} \\ & + \frac{r^2}{\nu+r-1}(\frac{\nu+r-1}{\nu})^r + r(1-\frac{1}{k})^{1-r}\}. \end{aligned} \tag{1}$$

Similarly, we have

$$\begin{aligned} E(N_r + I_{\{r=s\}} - N_r^{(s)})_+ \ \le\ & E\{\sum_{i=1}^{s} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}}\} \\ \le\ & \frac{1}{r!}(\frac{\nu}{k})^r(1-\frac{1}{k})^{\nu-r}(1+\frac{rsk}{\nu}+\frac{\nu}{k}). \end{aligned} \tag{2}$$

The lemma follows immediately from (1) and (2). □

Theorem 1 *Let $\mathcal{L}(N_r)$ denote the law of N_r and $\lambda = E(N_r)$. Then*

$$
\begin{aligned}
d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \;\le\; & \frac{1}{r!}(1-e^{-\lambda})(\frac{\nu}{k})^r(1-\frac{1}{k})^{\nu-2r}\{1+2r \\
& +\frac{r^2}{\nu}(\frac{\nu+r-1}{\nu})^{r-1}+\frac{r^2k}{\nu}+\frac{\nu}{k}\}.
\end{aligned}
$$

PROOF. We observe from Proposition 2 that for each $A \subseteq Z^+$, there exists an f such that

$$
\begin{aligned}
& E\{I_A(N_r) - \mathcal{P}_\lambda I_A\} \\
= \; & E\{\lambda f(N_r+1) - \sum_{i=1}^{k} I_{\{X_i=r\}} f(N_r)\} \\
= \; & E\{\lambda f(N_r+1) - \sum_{i=1}^{k} I_{\{X_i=r\}} E(f(N_r) \mid X_i = r)\} \\
= \; & E\{\lambda(f(N_r+1) - f(N_r^{(r)}))\}.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
& \mid E\{I_A(N_r) - \mathcal{P}_\lambda I_A\} \mid \\
\le \; & \lambda(\sup_{w\in Z^+} \mid f(w+1) - f(w) \mid)E \mid N_r + 1 - N_r^{(r)} \mid \\
\le \; & \frac{1}{r!}(1-e^{-\lambda})(\frac{\nu}{k})^r(1-\frac{1}{k})^{\nu-2r}\{1+2r \\
& +\frac{r^2}{\nu}(\frac{\nu+r-1}{\nu})^{r-1}+\frac{r^2k}{\nu}+\frac{\nu}{k}\}.
\end{aligned}
$$

The second inequality uses Proposition 2 and Lemma 1. □
Here are two immediate corollaries.

Corollary 1 *Let $\lambda = EN_r$ where r is a fixed integer greater than or equal to 2. Then $d(\mathcal{L}(N_r), \mathcal{P}_\lambda)$ tends to 0 as $\nu/k \to 0$. More precisely,*

$$
d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \le \frac{r^2}{r!}(1-e^{-\lambda})(\frac{\nu}{k})^{r-1}(1+o(1)).
$$

Corollary 2 *Let $\lambda = EN_r$ where r is a fixed integer greater than or equal to 0. Then $d(\mathcal{L}(N_r), \mathcal{P}_\lambda)$ tends to 0 as $\nu/k \to \infty$. More precisely,*

$$
d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \le \frac{1}{r!}(1-e^{-\lambda})(\frac{\nu}{k})^{r+1}(1-\frac{1}{k})^{\nu}(1+o(1)).
$$

Next we consider the asymptotic distribution of $N_0 - k + \nu$ as $\nu/k \to 0$. To do so, it would be convenient to have the following lemma which follows directly from Lemma 1.

Lemma 2 *Let $\lambda = E(N_0) - k + \nu$. With N_r's and $N_r^{(s)}$'s defined on the common probability space described at the beginning of this section, we have*

$$\begin{aligned}
&\sum_{r=2}^{\nu}(r-1)E \mid N_r + I_{\{r=s\}} - N_r^{(s)} \mid \\
\leq\ & \frac{\nu}{k}(1-\frac{1}{k})^{\nu-3}e^{2\nu/(k-1)}\{s + \frac{\nu}{k}(\frac{3}{2} + \frac{s}{2}(1-\frac{1}{k})^{-s} + \frac{8}{\nu}) + (\frac{\nu}{k})^2(\frac{1}{2} + \frac{8}{\nu})\}.
\end{aligned}$$

Theorem 2 *Let $W = N_0 - k + \nu$ and $\lambda = E(W)$. Then*

$$\begin{aligned}
&d(\mathcal{L}(W), \mathcal{P}_\lambda) \\
\leq\ & \frac{k}{\lambda}(1-e^{-\lambda})(\frac{\nu}{k})^3(1-\frac{1}{k})^{\nu-7}e^{3\nu k/(k-1)^2}(\frac{4}{3} + \frac{5\nu}{4k} + \frac{\nu^2}{4k^2} + \frac{4}{k} + \frac{4\nu}{k^2}).
\end{aligned}$$

Thus $d(\mathcal{L}(W), \mathcal{P}_\lambda)$ tends to 0 as $\nu/k \to 0$. More precisely,

$$d(\mathcal{L}(W), \mathcal{P}_\lambda) \leq \frac{8}{3}(1-e^{-\lambda})\frac{\nu}{k}(1+o(1)).$$

PROOF. We observe that $W = \sum_{r=2}^{\nu}(r-1)N_r$. It follows from Proposition 2 that for each $A \subseteq Z^+$, there exists an f such that

$$\begin{aligned}
&E\{I_A(W) - \mathcal{P}_\lambda I_A\} \\
=\ & E\{\lambda f(W+1) - \sum_{r=2}^{\nu}(r-1)N_r f(W)\} \\
=\ & E\{\lambda f(W+1) - \sum_{r=2}^{\nu}(r-1)\sum_{i=1}^{k} I_{\{X_i=r\}} f(W)\} \\
=\ & \sum_{r=2}^{\nu}(r-1)E(N_r)\{E(f(W+1)) - E(f(W) \mid X_1 = r)\}.
\end{aligned}$$

Consequently, it follows from Lemma 2 and Proposition 2 that

$$\begin{aligned}
&\mid EI_A(W) - \mathcal{P}_\lambda I_A \mid \\
\leq\ & \lambda^{-1}(1-e^{-\lambda})\sum_{r=2}^{\nu}(r-1)E(N_r)\{r-2
\end{aligned}$$

$$+\sum_{s=2}^{\nu}(s-1)E\mid N_s+I_{\{s=r\}}-N_s^{(r)}\mid\}$$
$$\leq \lambda^{-1}(1-e^{-\lambda})\{\sum_{r=3}^{\nu}(r-2)(r-1)E(N_r)$$
$$+\frac{\nu}{k}(1-\frac{1}{k})^{\nu-3}e^{2\nu/(k-1)}\sum_{r=2}^{\nu}(r-1)E(N_r)$$
$$\times\{r+\frac{\nu}{k}(\frac{3}{2}+\frac{r}{2}(1-\frac{1}{k})^{-r}+\frac{8}{\nu})+(\frac{\nu}{k})^2(\frac{1}{2}+\frac{8}{\nu})\}\}.$$

Hence we conclude, after a straightforward calculation, that

$$d(\mathcal{L}(W),\mathcal{P}_\lambda)$$
$$\leq \frac{k}{\lambda}(1-e^{-\lambda})(\frac{\nu}{k})^3(1-\frac{1}{k})^{\nu-7}e^{3\nu k/(k-1)^2}(\frac{4}{3}+\frac{5\nu}{4k}+\frac{\nu^2}{4k^2}+\frac{4}{k}+\frac{4\nu}{k^2}).$$

This completes the proof. □

Finally we consider the asymptotic distribution of $[(\nu-N_1)/2]$ as $\nu/k\to 0$. Here $[x]$ denotes the greatest integer less than or equal to x.

Lemma 3 *With the N_r's and $N_r^{(2)}$'s defined on the common probability space described at the beginning of this section, we have*

$$E\mid[\frac{1}{2}\sum_{r=3}^{\nu}rN_r]-[\frac{1}{2}\sum_{r=3}^{\nu}rN_r^{(2)}]\mid$$
$$\leq e^{2\nu/(k-1)}(1-\frac{1}{k})^{\nu-5}(\frac{\nu}{k})^2(3+\frac{3\nu}{k}+\frac{36}{k}+\frac{\nu^2}{2k^2}).$$

PROOF. We observe that

$$E\mid[\frac{1}{2}\sum_{r=3}^{\nu}rN_r]-[\frac{1}{2}\sum_{r=3}^{\nu}rN_r^{(2)}]\mid$$
$$= E\{\mid[\frac{1}{2}\sum_{r=3}^{\nu}rN_r]-[\frac{1}{2}\sum_{r=3}^{\nu}rN_r^{(2)}]\mid\mid\mid\sum_{r=3}^{\nu}(rN_r-rN_r^{(2)})\mid\geq 1\}$$
$$\times P\{\mid\sum_{r=3}^{\nu}r(N_r-N_r^{(2)})\mid\geq 1\}$$
$$\leq \{E\{\mid\frac{1}{2}\sum_{r=3}^{\nu}r(N_r-N_r^{(2)})\mid\mid\mid\sum_{r=3}^{\nu}r(N_r-N_r^{(2)})\mid\geq 1\}+\frac{1}{2}\}$$

$$\times P\{|\sum_{r=3}^{\nu} r(N_r - N_r^{(2)})| \geq 1\}$$
$$\leq \quad E\sum_{r=3}^{\nu} r \mid N_r - N_r^{(2)} \mid .$$

The last inequality uses Markov's inequality. By a calculation similar to that of Lemma 2, we arrive at

$$\begin{aligned} & E \mid [\frac{1}{2}\sum_{r=3}^{\nu} rN_r] - [\frac{1}{2}\sum_{r=3}^{\nu} rN_r^{(2)}] \mid \\ \leq \quad & e^{2\nu/(k-1)}(1-\frac{1}{k})^{\nu-5}(\frac{\nu}{k})^2(3+\frac{3\nu}{k}+\frac{36}{k}+\frac{\nu^2}{2k^2}). \end{aligned}$$

This completes the proof. □

Theorem 3 *Let $V = [(\nu - N_1)/2]$ and $\lambda = E(N_2)$. Then*

$$\begin{aligned} & d(\mathcal{L}(V), \mathcal{P}_\lambda) \\ \leq \quad & (1 \wedge \frac{5}{7}\lambda^{1/2})(1-\frac{1}{\nu})^{-1/2}(1-\frac{1}{k})^{(\nu-4)/2}e^{\nu/(k-1)}\frac{\nu^2}{2k^{3/2}} \\ & + e^{2\nu/(k-1)}(1-e^{-\lambda})(1-\frac{1}{k})^{\nu-5}\frac{\nu}{k}(2+\frac{11\nu}{2k}+\frac{7\nu^2}{2k^2}+\frac{4}{k}+\frac{36\nu}{k^2}+\frac{\nu^3}{2k^3}). \end{aligned}$$

This implies that $d(\mathcal{L}(V), \mathcal{P}_\lambda)$ tends to 0 whenever $\nu^4/k^3 \to 0$.

PROOF. We observe that $V = N_2 + [\sum_{r=3}^{\nu} rN_r/2]$. It follows from Proposition 2 that for each $A \subseteq Z^+$, there exists an f such that

$$\begin{aligned} & E\{I_A(V) - \mathcal{P}_\lambda I_A\} \\ = \quad & E\{\lambda f(V+1) - \sum_{i=1}^{k} I_{\{X_i=2\}} f(V) - [\frac{1}{2}\sum_{r=3}^{\nu} rN_r]f(V)\} \\ = \quad & E\{\lambda(f(V+1) - f(V \mid X_1 = 2)) - [\frac{1}{2}\sum_{r=3}^{\nu} rN_r]f(V)\}. \end{aligned}$$

Now it follows from Proposition 2 that

$$\begin{aligned} & \mid E\{I_A(V) - \mathcal{P}_\lambda I_A\} \mid \\ \leq \quad & (1-e^{-\lambda})E\{\mid N_2 + 1 - N_2^{(2)} \mid + \mid [\frac{1}{2}\sum_{r=3}^{\nu} rN_r] - [\frac{1}{2}\sum_{r=3}^{\nu} rN_r^{(2)}] \mid\} \\ & + \frac{1}{2}\sum_{r=3}^{\nu} rE(N_r)(1 \wedge 1.4\lambda^{-1/2}). \end{aligned}$$

Using Lemma 3, we conclude after some computation that

$$\begin{aligned} & d(\mathcal{L}(V), \mathcal{P}_\lambda) \\ \leq \ & (1 \wedge \frac{5}{7}\lambda^{1/2})(1-\frac{1}{\nu})^{-1/2}(1-\frac{1}{k})^{(\nu-4)/2}e^{\nu/(k-1)}\frac{\nu^2}{2k^{3/2}} \\ & + e^{2\nu/(k-1)}(1-e^{-\lambda})(1-\frac{1}{k})^{\nu-5}\frac{\nu}{k}(2+\frac{11\nu}{2k}+\frac{7\nu^2}{2k^2}+\frac{4}{k}+\frac{36\nu}{k^2}+\frac{\nu^3}{2k^3}). \end{aligned}$$

This clearly implies that if $\nu^4/k^3 \to 0$, then $d(\mathcal{L}(V), \mathcal{P}_\lambda)$ tends to 0. □

3 Lower Bounds

In this section we shall adopt the ingenious adaptation of the Chen-Stein method of Barbour and Hall (1984) to arrive at the lower bounds on the rates of Poisson convergence of the distributions of

1. N_r, $r \geq 0$, as $\nu/k \to \infty$,
2. $(N_0 - k + \nu)$ and N_r, $r \geq 2$, as $\nu/k \to 0$.

Lemma 4 *Let $f(m) = (m-\lambda)e^{-(m-\lambda)^2/(\theta\lambda)}$, $m \in Z^+$, where λ, θ are positive constants. Then*

$$\begin{aligned} & E\{f(N_r+1) - f(N_r^{(r)})\} \\ \geq \ & E\{N_r + 1 - N_r^{(r)} - \{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+/(\theta\lambda)\}. \end{aligned}$$

PROOF. For simplicity of notation, let $A = N_r + 1 - N_r^{(r)}$. From the observation that

$$0 \leq 1 - (d/dw)(we^{-w^2/(\theta\lambda)}) \leq 3w^2/(\theta\lambda),$$

we get for $A \geq 0$,

$$\begin{aligned} A - f(N_r+1) + f(N_r^{(r)}) &= \int_{N_r^{(r)}-\lambda}^{N_r+1-\lambda} \{1-(d/dw)(we^{-w^2/(\theta\lambda)})\}dw \\ &\leq \int_{N_r^{(r)}-\lambda}^{N_r+1-\lambda} (3w^2/(\theta\lambda))dw \\ &= \{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}/(\theta\lambda), \end{aligned}$$

and for $A < 0$, we have

$$\begin{aligned} f(N_r+1) - f(N_r^{(r)}) &= \int_{N_r^{(r)}-\lambda}^{N_r+1-\lambda} (d/dw)(we^{-w^2/(\theta\lambda)})dw \\ &\geq A. \end{aligned}$$

This implies that

$$f(N_r+1) - f(N_r^{(r)}) \geq A - \{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+/(\theta\lambda),$$

and this completes the proof. □

Proposition 3 *Let $f(m) = (m-\lambda)e^{-(m-\lambda)^2/(\theta\lambda)}$, $m \in Z^+$, where λ, θ are positive constants. Then*

$$d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \geq (1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1}) E\{f(N_r+1) - f(N_r^{(r)})\}/2.$$

PROOF. As in Barbour and Hall (1984), we observe that

$$\begin{aligned} & 2d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \sup_j \mid \lambda f(j+1) - jf(j) \mid \\ \geq\ & E\{\lambda f(N_r+1) - N_r f(N_r)\} \\ =\ & \lambda E\{f(N_r+1) - f(N_r^{(r)})\}. \end{aligned} \tag{3}$$

Since

$$\lambda(1 \vee (2e^{-3/2} + \theta e^{-1})) \geq \sup_j \mid \lambda f(j+1) - jf(j) \mid,$$

the proposition follows from (3). □

Theorem 4 *Let $\lambda = E(N_r)$ where r is a fixed integer greater than or equal to 2. Suppose that ν, k tend to infinity in such a manner that $\nu/k \to 0$, then*

$$d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \geq \frac{er^2}{2(2e^{-1/2}+\theta)r!}(\frac{\nu}{k})^{r-1}\{1 - \frac{6}{\theta}(2\lambda^{-1}+1)\}(1+o(1)),$$

where $\theta = 6(2\lambda^{-1}+1) + \{12(2\lambda^{-1}+1)(6\lambda^{-1}+3+e^{-1/2})\}^{1/2}$.

PROOF. We observe that

$$\begin{aligned} & E(N_r + 1 - N_r^{(r)}) \\ =\ & \frac{\nu!}{(\nu-r)!r!k^{r-1}}(1-\frac{1}{k})^{\nu-r} - \frac{(\nu-r)!}{(\nu-2r)!r!(k-1)^{r-1}}(1-\frac{1}{k-1})^{\nu-2r} \\ =\ & \frac{r}{(r-1)!}(\frac{\nu}{k})^{r-1}(1+o(1)). \end{aligned}$$

Next, for $N_r + 1 - N_r^{(r)} \geq 0$, we observe that

$$\begin{aligned}
& \{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+ \\
\leq\ & 3(N_r+1-\lambda)^2(N_r+1-N_r^{(r)}) + 3\lambda(N_r+1-N_r^{(r)})^2 \\
\leq\ & 3(N_r+1-\lambda)^2(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}}) \\
& +3\lambda(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})^2.
\end{aligned}$$

Here the final inequality uses the notation of Lemma 1 and the proof of (2). Thus we conclude that

$$\begin{aligned}
& E\{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+ \\
\leq\ & 3E\{(N_r+1-\lambda)^2(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}}) \\
& +\lambda(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})^2\}.
\end{aligned}$$

By expanding and taking expectations, we get

$$E\{(N_r+1-\lambda)^2(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})\} = \frac{r^2}{r!}(\frac{\nu}{k})^{r-1}(4+\lambda)(1+o(1)),$$

and

$$\lambda E(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})^2 \leq \frac{r^2}{r!}(\frac{\nu}{k})^{r-1}\lambda(1+o(1)).$$

Now it follows from Lemma 4 and Proposition 3 that

$$\begin{aligned}
& 2d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \\
\geq\ & (1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1}) E\{N_r + 1 - N_r^{(r)} \\
& -\{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+/(\theta\lambda)\} \\
\geq\ & (1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1}) \frac{r^2}{r!}(\frac{\nu}{k})^{r-1}\{1 - \frac{6}{\theta}(2\lambda^{-1}+1)\}(1+o(1)).
\end{aligned}$$

To ensure that the right hand side of the last inequality is strictly positive, it is necessary to restrict the possible values of θ to be greater or equal to 6. Then

$$(1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1}) = e(2e^{-1/2} + \theta)^{-1}.$$

Thus we conclude that

$$(4)\qquad d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \geq \frac{er^2}{2r!(2e^{-1/2}+\theta)}(\frac{\nu}{k})^{r-1}\{1-\frac{6}{\theta}(2\lambda^{-1}+1)\}(1+o(1)),$$

where the value of θ which maximizes the right hand side of (4) is given by

$$\theta = 6(2\lambda^{-1}+1) + \{12(2\lambda^{-1}+1)(6\lambda^{-1}+3+e^{-1/2})\}^{1/2}.$$

This completes the proof of the theorem. □

Theorem 5 *Let $\lambda = E(N_r)$ where r is a fixed integer greater than or equal to 0. Suppose that ν, k tend to infinity in such a way that $\nu/k \to \infty$ and $\nu/k^2 \to 0$. Then*

$$d(\mathcal{L}(N_r), \mathcal{P}_\lambda) \geq \frac{e}{2(2e^{-1/2}+\theta)r!}(\frac{\nu}{k})^{r+1}(1-\frac{1}{k})^{\nu}\{1-\frac{6}{\theta}(2\lambda^{-1}+1)\}(1+o(1)),$$

where $\theta = 6(2\lambda^{-1}+1) + \{12(2\lambda^{-1}+1)(6\lambda^{-1}+3+e^{-1/2})\}^{1/2}$.

PROOF. We observe that

$$E(N_r + 1 - N_r^{(r)}) = \frac{1}{r!}(\frac{\nu}{k})^{r+1}(1-\frac{1}{k})^{\nu}(1+o(1)).$$

As in the proof of Theorem 4, we have

$$\begin{aligned} & E\{(N_r+1-\lambda)^3 - (N_r^{(r)}-\lambda)^3\}_+ \\ \leq\ & 3E\{(N_r+1-\lambda)^2(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}}) \\ & +\lambda(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})^2\}, \end{aligned}$$

and

$$\begin{aligned} & E\{(N_r+1-\lambda)^2(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})\} \\ =\ & \frac{1}{r!}(\frac{\nu}{k})^{r+1}(1-\frac{1}{k})^{\nu}(4+\lambda)(1+o(1)), \end{aligned}$$

and

$$\lambda E(\sum_{i=1}^{r} Y_i^{(r)} + \sum_{j=1}^{X_I} Z_j^{(r)} + I_{\{X_I=r\}})^2 \le \frac{1}{r!}(\frac{\nu}{k})^{r+1}(1-\frac{1}{k})^{\nu}\lambda(1+o(1)).$$

Now the rest of the proof is almost identical to that of Theorem 4 and hence will be omitted. □

We shall now construct a lower bound on the rate of Poisson convergence of the distribution of $N_0 - k + \nu$ as $\nu/k \to 0$.

Theorem 6 *Let $W = N_0 - k + \nu$ and $\lambda = E(W)$. Suppose that ν, k tend to infinity in such a way that $\nu/k \to 0$, then*

$$d(\mathcal{L}(W), \mathcal{P}_\lambda) \ge \frac{2e}{3(2e^{-1/2}+\theta)}\frac{\nu}{k}\{1-\frac{9}{\theta}(4\lambda^{-1}+\frac{9}{8})\}(1+o(1)),$$

where $\theta = 9(4\lambda^{-1}+\frac{9}{8}) + 3\{(4\lambda^{-1}+\frac{9}{8})(36\lambda^{-1}+\frac{81}{8}+2e^{-1/2})\}^{1/2}$.

PROOF. Let f be defined as in Lemma 4. Then as in the proofs of Lemma 4 and Proposition 3, we have

$$\begin{aligned} & 2\lambda(1\vee(2e^{-3/2}+\theta e^{-1}))d(\mathcal{L}(W),\mathcal{P}_\lambda) \\ \ge\ & E\{\lambda f(W+1) - \sum_{r=2}^{\nu}(r-1)N_r E(f(W)\mid X_1=r)\} \\ \ge\ & \sum_{r=2}^{\nu}(r-1)E(N_r)E\{W+1-W^{(r)} \\ & -\{(W+1-\lambda)^3-(W^{(r)}-\lambda)^3\}_+/(\theta\lambda)\}, \end{aligned} \tag{5}$$

where $W^{(r)}$ has the condtional distribution of W given that $X_1 = r$. Also, we observe that

$$\begin{aligned} E(W+1-W^{(r)}) \ \ge\ & (1-\frac{1}{k-1})^{-r}\{2-r+\frac{r(\nu-2)}{k-1}-\frac{r\nu(\nu-1)}{2k(k-1)} \\ & -\frac{\nu(\nu-1)}{k(k-1)}-\frac{\nu(\nu-1)}{2(k-1)^3}-\frac{\nu^2(\nu-1)^2}{4k^2(k-1)^3}\}, \end{aligned}$$

and hence

$$\sum_{r=2}^{\nu}(r-1)E(N_r)E(W+1-W^{(r)}) \ge \frac{2\nu^3}{3k^2}(1+o(1)). \tag{6}$$

Next, let N_s, $N_s^{(r)}$, $s = 0, \ldots, \nu$, be defined on the common probability space described at the beginning of Section 2. Also let I, $Y_i^{(1)}$, $i = 1, \ldots, r$, be as in Lemma 1 and define on that probability space

$$\tilde{Z}_i^{(0)} = \begin{cases} 1 & \text{if the } i\text{th ball from urn } I \text{ falls into either one of} \\ & \text{the originally } N_0 \text{ empty urns or the urn} \\ & \text{where the second of } r \text{ balls is taken from,} \\ 0 & \text{otherwise.} \end{cases}$$

Then we have for $r \geq 2$,

$$\begin{aligned} (N_0 + 1 - N_0^{(r)})_+ &\leq \sum_{i=1}^{X_I} \tilde{Z}_i^{(0)} + (I_{\{X_I=0\}} - Y_1^{(1)})_+ + 1 - Y_2^{(1)} \\ &= B, \text{ say.} \end{aligned}$$

Hence writing $E(N_0) = \lambda_0$, we have

$$\begin{aligned} & E\{(W + 1 - \lambda)^3 - (W^{(r)} - \lambda)^3\}_+ \\ =\ & E\{(N_0 + 1 - \lambda_0)^3 - (N_0^{(r)} - \lambda_0)^3\}_+ \\ \leq\ & E\{\frac{9}{2}(N_0 + 1 - \lambda_0)^2 B + \frac{5}{2}B^3\}. \end{aligned} \tag{7}$$

Furthermore, by a straightforward but tedious computation, we observe that

$$\frac{9}{2}E(N_0 + 1 - B)^2 B \leq (\frac{81\nu}{2k} + \frac{27\nu^3}{4k^2})(1 + o(1)),$$

and

$$\frac{5}{2}EB^3 \leq \frac{15\nu}{2k}(1 + o(1)).$$

Thus it follows from (7) that for $r \geq 2$,

$$E\{(W + 1 - \lambda)^3 - (W^{(r)} - \lambda)^3\}_+ \leq (\frac{48\nu}{k} + \frac{27\nu^3}{4k^2})(1 + o(1)),$$

and we conclude from (5) and (6) that

$$\begin{aligned} & d(\mathcal{L}(W), \mathcal{P}_\lambda) \\ \geq\ & \frac{\nu}{2k}(1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1})\{\frac{4}{3} - \frac{1}{\theta}(48\lambda^{-1} + \frac{27}{2})\}(1 + o(1)). \end{aligned} \tag{8}$$

To ensure that the right hand side of (8) is strictly positive it is necessary that $\theta > 81/8$. Thus

$$1 \wedge (2e^{-3/2} + \theta e^{-1})^{-1} = e(2e^{-1/2} + \theta)^{-1}.$$

Hence

$$d(\mathcal{L}(W), \mathcal{P}_\lambda) \geq \frac{2\nu}{3k}(\frac{e}{2e^{-1/2}+\theta})\{1 - \frac{9}{\theta}(4\lambda^{-1} + \frac{9}{8})\}(1 + o(1)),$$

and the value of θ that maximizes the right hand side is given by

$$\theta = 9(4\lambda^{-1} + \frac{9}{8}) + 3\{(4\lambda^{-1} + \frac{9}{8})(36\lambda^{-1} + \frac{81}{8} + 2e^{-1/2})\}^{1/2}.$$

This completes the proof. □

REMARK. Due to its more complicated functional form, we have not been able to construct a lower bound on the rate of Poisson convergence of $[(\nu - N_1)/2]$ as $\nu/k \to 0$ which is comparable to the upper bound of Theorem 3.

4 Acknowledgments

I would like to thank Professor Charles M. Stein for introducing this area of probability theory to me and also to Professor Louis H. Y. Chen for his interest and encouragement over the years.

References

[1] Barbour, A. D. and Eagleson, G. K. (1983). Poisson approximation for some statistics based on exchangeable trials. *Adv. in Appl. Probab.* **15** 585-600.

[2] Barbour, A. D. and Hall, P. (1984). On the rate of Poisson convergence. *Math. Proc. Camb. Phil. Soc.* **95** 473-480.

[3] Chen, L. H. Y. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3** 534-545.

[4] Kolchin, V. F., Sevastyanov, B. A. and Chistyakov, V. P. (1978). *Random Allocations.* V. H. Winston, Washington.

[5] Stein, C. M. (1970). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 583-602. Univ. of California Press.

[6] Stein, C. M. (1986). *Approximate Computation of Expectations.* IMS, Hayward.