Applied Probability Trust (9 November 2024)

SUPPLEMENT TO: ASYMPTOTICS OF THE ALLELE FREQUENCY SPECTRUM AND THE NUMBER OF ALLELES

ROSS A. MALLER,* *The Australian National University*

SOUDABEH SHEMEHSAVAR,** *Murdoch University and University of Tehran*

Abstract

We give here the proofs of Lemmas 3.1 and 4.1 in the paper. An extra lemma relating to the covariance matrix $\mathbf{Q}_J$ is in the Appendix.

2000 Mathematics Subject Classification: Primary 60G51, 60G52, 60G55
Secondary 60G57, 62E20, 62P10

1. Proofs of Lemmas

For convenience, we restate the lemmas with the numbering as in the paper. Equation references in the format "Eq. ( )" herein refer to the numbering in the paper.

Lemma 1 With the substitutions $k = \lfloor xn^\alpha \rfloor$, $m_j = \lfloor q_j + y_j/n^{\alpha/2} \rfloor k$, $k' = k - m_+$, we have the following limiting behaviours as $n \to \infty$:

$$\frac{\Gamma(r+k)(1-q_{+n}(\lambda n))^{k'}}{k'! \prod_{j=1}^J m_j!} \prod_{j=1}^J (q_{jn}(\lambda n))^{m_j} \sim \frac{(xn^\alpha)^{r-J/2-1} e^{-\frac{x}{2}\mathbf{y}^T \mathbf{Q}_J^{-1} \mathbf{y}}}{\sqrt{(2\pi)^J \det(\mathbf{Q}_J)}}; \tag{24}$$

$$\lim_{n\to\infty} \frac{(\lambda n)^{\alpha k}}{\Psi(\lambda n)^k} \Big(\sum_{\ell=1}^n F_\ell(\lambda n) \Big)^k = e^{-x(\lambda^{-\alpha} \vee 1)/\Gamma(1-\alpha)}; \tag{25}$$

$$\frac{1}{\Gamma(r)\Psi(\lambda n)^r} \sim \frac{1}{\Gamma(r)(\lambda n)^{r\alpha}\Gamma^r(1-\alpha)}; \tag{26}$$

$$\lim_{n\to\infty} n\mathbf{P}\Big(\sum_{i=1}^{k-m_+} X_{in}^{(J)}(\lambda n) = n - m_{++} \Big) = f_{Y_x(\lambda)}(1). \tag{27}$$

* Postal address: Ross Maller, Research School of Finance, Actuarial Studies and Statistics, Australian National University, Canberra, ACT, 0200, Australia.
* Email address: Ross.Maller@anu.edu.au
** Postal address: Soudabeh Shemehsavar, College of Science, Technology, Engineering and Mathematics, and Centre for Healthy Ageing, Health Future Institute, Murdoch University, and School of Mathematics, Statistics & Computer Sciences, University of Tehran (corresponding author).
** Email address: Soudabeh.Shemehsavar@murdoch.edu.au

Proof of Lemma 1: To prove (24), consider the first factor, in which we let $y_{jn} = y_j/n^{\alpha/2}$, $h_{jn} = q_j + y_{jn}$ and $h_{+n} = \sum_{j=1}^J h_{jn}$; then replace m_j by $h_{jn}k$, recall $k' = k - m_+ = (1-h_{+n})k$, and calculate

$$\frac{\Gamma(r+k)}{k'!\prod_{j=1}^J m_j!} \sim \frac{\sqrt{2\pi(r+k-1)}\,(r+k-1)^{r+k-1}e^{-(r+k-1)}}{\sqrt{2\pi(1-h_{+n})k}\,((1-h_{+n})k)^{(1-h_{+n})k}e^{-(1-h_{+n})k}} \times \frac{1}{\prod_{j=1}^J \sqrt{2\pi h_{jn}k}\,(h_{jn}k)^{h_{jn}k}e^{-h_{jn}k}}, \tag{1.1}$$

using Stirling's formula for the factorials. Some calculations show this is asymptotic to

$$\frac{k^{r-J/2-1}}{(2\pi)^{J/2}\sqrt{(1-h_{+n})\prod_{j=1}^J h_{jn}}} \times \frac{1}{(1-h_{+n})^{(1-h_{+n})k}\prod_{j=1}^J h_{jn}^{h_{jn}k}}. \tag{1.2}$$

Recall the q_{jn} and q_{+n} defined in Eq. (19). They depend on λ, but this was suppressed in the notation there. We have reinserted the factor λn in (24) and in what follows, for emphasis. Throughout this proof, $\lambda > 0$ is held fixed, so the convergence is pointwise in λ as $n \to \infty$. That this is enough to get the required convergence in Theorem 3.1 is shown in the argument from Eq. (29) – Eq. (30) in the paper, which avoids the need for checking a dominated convergence condition to take the limit as $n \to \infty$ through the integral with respect to λ.

Continuing, we multiply the expression in (1.2) by the q-terms in (24) to get

$$\begin{aligned}&\frac{\Gamma(r+k)}{k'!\prod_{j=1}^J m_j!}(1-q_{+n}(\lambda n))^{k-m_+}\prod_{j=1}^J (q_{jn}(\lambda n))^{m_j}\\ &\sim \frac{k^{r-J/2-1}}{\sqrt{(2\pi)^J(1-h_{+n})\prod_{j=1}^J h_{jn}}} \times \Big(\frac{1-q_{+n}(\lambda n)}{1-h_{+n}}\Big)^{(1-h_{+n})k}\prod_{j=1}^J\Big(\frac{(q_{jn}(\lambda n))}{h_{jn}}\Big)^{h_{jn}k},\end{aligned} \tag{1.3}$$

where, recall, $q_{+n}(\lambda n) = \sum_{j=1}^J q_{jn}(\lambda n)$. Note from Eq. (3.12) that, as $n \to \infty$, for $1 \le j \le J$,

$$q_{jn}(\lambda n) = \frac{F_j(\lambda n)}{\sum_{\ell=1}^n F_\ell(\lambda n)} \to \frac{F_j(\infty)}{\sum_{\ell=1}^\infty F_\ell(\infty)} = \frac{\Gamma(j-\alpha)/j!}{\sum_{\ell=1}^\infty \Gamma(\ell-\alpha)/\ell!} = \frac{\alpha\Gamma(j-\alpha)}{j!\Gamma(1-\alpha)} = q_j, \tag{1.4}$$

and $1 - q_{+n}(\lambda n) \to 1 - q_+$, where $q_+ = \sum_{j=1}^J q_j$. Also $h_{jn} = q_j + y_{jn} \to q_j$. Now write

$$\Big(\frac{1-q_{+n}(\lambda n)}{1-h_{+n}}\Big)^{(1-h_{+n})k} \times \prod_{j=1}^J\Big(\frac{q_{jn}(\lambda n)}{h_{jn}}\Big)^{h_{jn}k}$$

$$
\begin{aligned}
&= \Big(1 + \frac{\sum_{j=1}^J (y_{jn} + q_j - q_{jn}(\lambda n))}{1 - \sum_{j=1}^J (q_j + y_{jn})}\Big)^{(1-\sum_{j=1}^J (q_j+y_{jn}))k} \\
&\times \prod_{j=1}^J \Big(1 - \frac{y_{jn} + q_j - q_{jn}(\lambda n)}{q_j + y_{jn}}\Big)^{(q_j+y_{jn})k} \\
&= \exp\Big[\Big(1 - \sum_{j=1}^J (q_j + y_{jn})\Big) k \log\Big(1 + \frac{\sum_{j=1}^J (y_{jn} + q_j - q_{jn}(\lambda n))}{1 - \sum_{j=1}^J (q_j + y_{jn})}\Big) \\
&+ \sum_{j=1}^J (q_j + y_{jn}) k \log\Big(1 - \frac{y_{jn} + q_j - q_{jn}(\lambda n)}{q_j + y_{jn}}\Big)\Big]. \qquad (1.5)
\end{aligned}
$$

Since $y_{jn} \to 0$ we can expand the RHS of (1.5) using $\log(1+z) = z - z^2/2 - \cdots$ and $\log(1-z) = -z - z^2/2 - \cdots$ for small z. The first order terms cancel and in the exponent we're left with

$$
-\frac{k}{2}\Big(\sum_{j=1}^J \frac{(y_{jn} + q_j - q_{jn}(\lambda n))^2}{q_j + y_{jn}} + \frac{\big(\sum_{j=1}^J (y_{jn} + q_j - q_{jn}(\lambda n))\big)^2}{1 - \sum_{j=1}^J (q_j + y_{jn})}\Big). \qquad (1.6)
$$

To proceed we want to estimate the $q_j - q_{jn}(\lambda n)$ term, and for this, noting (1.4), we need to estimate $F_j(\lambda n)$ and $\sum_{\ell=1}^n F_\ell(\lambda n)$. For the first, just note that

$$
F_j(\infty) - F_j(\lambda n) = \frac{\alpha}{j!}\int_{\lambda n}^\infty z^{-\alpha-1} e^{-z} \mathrm{d}z = \frac{\alpha}{j!} o(n^{-\alpha}), \text{ as } n \to \infty. \qquad (1.7)
$$

For the second, recall Eq. (2.7) and write

$$
\sum_{\ell=1}^n F_\ell(\lambda n) = \alpha \sum_{\ell=1}^n \frac{\Gamma(\ell - \alpha)}{\ell!} - \alpha \sum_{\ell=1}^n \int_{\lambda n}^\infty \frac{z^\ell}{\ell!} z^{-\alpha-1} e^{-z} \mathrm{d}z. \qquad (1.8)
$$

The first term on the RHS of (1.8) is

$$
\begin{aligned}
&\alpha\Big(\sum_{\ell=1}^\infty - \sum_{\ell>n}\Big)\int_0^\infty \frac{z^\ell}{\ell!} z^{-\alpha-1} e^{-z} \mathrm{d}z \\
&= \alpha \int_0^\infty (1 - e^{-z}) z^{-\alpha-1} \mathrm{d}z - \alpha \int_0^\infty \sum_{\ell>n} \frac{z^\ell}{\ell!} z^{-\alpha-1} e^{-z} \mathrm{d}z \qquad (1.9)
\end{aligned}
$$

Integrating by parts, we see that the first term on the RHS of (1.9) equals $\Gamma(1-\alpha)$. The second term equals

$$
n^{-\alpha}\int_0^\infty \sum_{\ell>n} \frac{(nz)^\ell}{\ell!} \alpha z^{-\alpha-1} e^{-nz} \mathrm{d}z = n^{-\alpha}\int_0^\infty \mathbf{P}\Big(\frac{\text{Poisson}(nz)}{nz} > \frac{1}{z}\Big)\alpha z^{-\alpha-1} \mathrm{d}z,
$$

with "Poisson$(\cdot)$" denoting a Poisson rv with the indicated mean. Noting that, by Markov's inequality, $\mathbf{P}\big(\text{Poisson}(nz) > n\big) \le nz/n = z$, we can apply dominated convergence and the weak law of large numbers to see that

$$\lim_{n\to\infty}\int_0^1 \mathbf{P}\Big(\frac{\text{Poisson}(nz)}{nz} > \frac{1}{z}\Big)\alpha z^{-\alpha-1}\mathrm{d}z = 0, \tag{1.10}$$

and similarly

$$\lim_{n\to\infty}\int_1^\infty \mathbf{P}\Big(\frac{\text{Poisson}(nz)}{nz} > \frac{1}{z}\Big)\alpha z^{-\alpha-1}\mathrm{d}z = \int_1^\infty \alpha z^{-\alpha-1}\mathrm{d}z = 1.$$

It follows that

$$\alpha\sum_{\ell>n}\frac{\Gamma(\ell-\alpha)}{\ell!} = n^{-\alpha} + o(n^{-\alpha}), \tag{1.11}$$

and consequently

$$\alpha\sum_{\ell=1}^{n}\frac{\Gamma(\ell-\alpha)}{\ell!} = \Gamma(1-\alpha) - n^{-\alpha} + o(n^{-\alpha}). \tag{1.12}$$

Returning to (1.8), we next estimate

$$\alpha\sum_{\ell=1}^{n}\int_{\lambda n}^\infty \frac{z^\ell}{\ell!}z^{-\alpha-1}e^{-z}\mathrm{d}z = \alpha\Big(\sum_{\ell=1}^{\infty}-\sum_{\ell>n}\Big)\int_{\lambda n}^\infty \frac{z^\ell}{\ell!}z^{-\alpha-1}e^{-z}\mathrm{d}z$$

$$= \frac{\alpha}{n^\alpha}\int_\lambda^\infty (1-e^{-nz})z^{-\alpha-1}\mathrm{d}z - \frac{1}{n^\alpha}\int_\lambda^\infty \mathbf{P}\Big(\frac{\text{Poisson}(nz)}{nz} > \frac{1}{z}\Big)\alpha z^{-\alpha-1}\mathrm{d}z.$$

The first term on the RHS here equals $\lambda^{-\alpha}n^{-\alpha} - o(n^{-\alpha})$. In the second term, whenever $\lambda < 1$ the component of the integral over $[\lambda, 1]$ is $o(n^{-\alpha})$ by (1.10), and the component over $[1,\infty]$ is $1 + o(1)$, and when $\lambda > 1$, the integral equals $\lambda^{-\alpha} + o(1)$. Thus this second term is $\lambda^{-\alpha}n^{-\alpha}\mathbf{1}_{\{\lambda>1\}} + n^{-\alpha}\mathbf{1}_{\{\lambda\le1\}} + o(n^{-\alpha})$. Subtracting this from the first term we obtain

$$\alpha\sum_{\ell=1}^{n}\int_{\lambda n}^\infty \frac{z^\ell}{\ell!}z^{-\alpha-1}e^{-z}\mathrm{d}z = \lambda^{-\alpha}n^{-\alpha}\mathbf{1}_{\{\lambda\le1\}} - n^{-\alpha}\mathbf{1}_{\{\lambda\le1\}} + o(n^{-\alpha}),$$

and subtracting this from the RHS of (1.8), keeping in mind (1.12), we conclude

$$\begin{aligned}\sum_{\ell=1}^{n}F_\ell(\lambda n) &= \Gamma(1-\alpha) - n^{-\alpha} - \lambda^{-\alpha}n^{-\alpha}\mathbf{1}_{\{\lambda\le1\}} + n^{-\alpha}\mathbf{1}_{\{\lambda\le1\}} + o(n^{-\alpha}) \\ &= \Gamma(1-\alpha) - n^{-\alpha}\big(\lambda^{-\alpha}\vee 1\big) + o(n^{-\alpha}). \end{aligned}\tag{1.13}$$

(1.7), (1.12) and (1.13) show that $F_j(\infty) - F_j(\lambda n)$ and $\sum_{\ell=1}^n F_\ell(\lambda n) - \sum_{\ell=1}^\infty F_\ell(\infty)$ $= \sum_{\ell=1}^n F_\ell(\lambda n) - \Gamma(1-\alpha)/\alpha$ are $O(n^{-\alpha})$ as $n\to\infty$, and then, referring to (1.4), we can prove that $q_j - q_{jn}(\lambda n) = O(n^{-\alpha})$. Using this, and recalling $k = \lfloor xn^\alpha \rfloor$ and $y_{jn} = y_j/n^{\alpha/2}$, we see that (1.6) simplifies to

$$-\frac{1}{2}xn^\alpha\Big(\frac{1}{n^\alpha}\sum_{j=1}^{J}\frac{y_j^2}{q_j}+\frac{(\sum_{j=1}^{J}y_j)^2}{n^\alpha(1-q_+)}\Big)+o(1)=-\frac{x}{2}\Big(\sum_{j=1}^{J}\frac{y_j^2}{q_j}+\frac{(\sum_{j=1}^{J}y_j)^2}{1-q_+}\Big)+o(1). \quad (1.14)$$

Let $Q_1 = 1$, and when $J \geq 2$, set $Q_j = (1-\sum_{i=1,i\neq j}^{J} q_i)\prod_{i=1,i\neq j}^{J} q_i$, $2 \leq j \leq J$. Let $Q = \prod_{j=1}^{J} q_j$, and let $\mathbf{Q}_J$ be the $J \times J$ matrix whose inverse is

$$\mathbf{Q}_J^{-1} = \frac{1}{(1-\sum_{j=1}^{J} q_j)\prod_{j=1}^{J} q_j}\begin{bmatrix} Q_1 & Q & \cdots & Q \\ Q & Q_2 & \cdots & Q \\ \vdots & \vdots & \ddots & \vdots \\ Q & Q & \cdots & Q_J \end{bmatrix}. \quad (1.15)$$

Then the RHS of (1.14) can be written in the form $-x\mathbf{y}^T\mathbf{Q}_J^{-1}\mathbf{y/2}$, so the limit of the LHS of (1.5) is $\exp(-x\mathbf{y}^T\mathbf{Q}_J^{-1}\mathbf{y/2})$, and thus (1.3) is asymptotic to

$$\frac{k^{r-J/2-1}e^{-\frac{x}{2}\mathbf{y}^T\mathbf{Q}_J^{-1}\mathbf{y}}}{\sqrt{(2\pi)^J(1-\sum_{j=1}^{J} q_j)\prod_{j=1}^{J} q_j}} = \frac{k^{r-J/2-1}e^{-\frac{x}{2}\mathbf{y}^T\mathbf{Q}_J^{-1}\mathbf{y}}}{\sqrt{(2\pi)^J\det(\mathbf{Q}_J)}}. \quad (1.16)$$

In the Appendix to this Supplement, Lemma 2.1 verifies that $\mathbf{Q}_J$, the inverse of the matrix $\mathbf{Q}_J^{-1}$ in (1.15), is indeed the $J \times J$ matrix with diagonal elements $q_i(1-q_i)$ and off-diagonal elements $-q_iq_j$, $1 \leq i \neq j \leq J$ appearing in Eq. (3.2), and calculates its determinant. So (24) is proved.

To prove (25): from (6) and $\int_0^\infty(1-e^{-z})\alpha z^{-\alpha-1}\mathrm{d}z = \Gamma(1-\alpha)$ we see that

$$\begin{aligned}\Psi(\lambda n) &= 1+(\lambda n)^\alpha\int_0^{\lambda n}(1-e^{-z})\alpha z^{-\alpha-1}\mathrm{d}z \\ &= 1+(\lambda n)^\alpha\Gamma(1-\alpha)-(\lambda n)^\alpha\int_{\lambda n}^{\infty}(1-e^{-z})\alpha z^{-\alpha-1}\mathrm{d}z, \end{aligned} \quad (1.17)$$

in which, integrating by parts,

$$(\lambda n)^\alpha\int_{\lambda n}^{\infty}(1-e^{-z})\alpha z^{-\alpha-1}\mathrm{d}z = 1-e^{-\lambda n}+(\lambda n)^\alpha\int_{\lambda n}^{\infty}e^{-z}z^{-\alpha}\mathrm{d}z = 1+O(e^{-\lambda n}).$$

So $\Psi(\lambda n) = (\lambda n)^\alpha\Gamma(1-\alpha)+o(n^{-\alpha})$, and consequently

$$\frac{(\lambda n)^\alpha}{\Psi(\lambda n)} = \frac{1}{\Gamma(1-\alpha)}+o(n^{-\alpha}). \quad (1.18)$$

Combining (1.13) with (1.18) and recalling $k = \lfloor xn^\alpha \rfloor$, we obtain

$$\Big(\frac{(\lambda n)^\alpha}{\Psi(\lambda n)}\sum_{\ell=1}^{n}F_\ell(\lambda n)\Big)^k \;=\; \Big(1-\frac{(\lambda^{-\alpha}\vee 1)/\Gamma(1-\alpha)}{n^\alpha}+o(n^{-\alpha})\Big)^{\lfloor xn^\alpha \rfloor}, \quad (1.19)$$

and this tends to $e^{-x(\lambda^{-\alpha}\vee 1)/\Gamma(1-\alpha)}$, as $n\to\infty$. So we have proved (25).

To prove (26), just use (1.18) to get

$$\frac{1}{\Gamma(r)\Psi(\lambda n)^r}\sim\frac{1}{\Gamma(r)(\lambda n)^{\alpha r}\Gamma^r(1-\alpha)}. \tag{1.20}$$

This leaves just the fourth factor to deal with. To prove (27) we modify the local limit result obtained in [1], Prop. 3.1, where we dealt with a sequence (X_{in}) only slightly different to $(X_{in}^{(J)})$; namely, there we had $\mathbf{P}(X_{1n}(\lambda)=j)=p_{jn}(\lambda)$, $1\le j\le n$, and $\mathbf{p}_n(\lambda)=(p_{jn}(\lambda))_{1\le j\le n}$, rather than the expressions in the paper, and with the relevant summation being over $1\le\ell\le n$ rather than $J+1\le\ell\le n$. We indicate just the main modifications needed.

First, we claim Prop. 3.1 of [1] remains true if $X_{in}(\lambda n)$ therein is replaced by $X_{in}^{(J)}(\lambda n)$, the denominator $\Gamma(1-\alpha)$ in Eq. (3.10) of [1] is replaced by $\Gamma(1-\alpha)\big(1-q_+\big)$, and corresponding modifications are made in Eq. (3.11) and Eq. (3.12) of [1]. To check this, note that $p_{jn}(\lambda)$ in the proof of Prop. 3.1 of [1] is replaced by $p_{jn}^{(J)}(\lambda)$, so in Eq. (3.10) of [1], once $hn>3$, we need only replace the denominator by

$$\begin{aligned}\sum_{j=J+1}^{n}F_j(\lambda n)&=\sum_{j=1}^{n}F_j(\lambda n)-\sum_{j=1}^{J}F_j(\lambda n)\\ &=\Gamma(1-\alpha)-\sum_{j=1}^{J}\Gamma(j-\alpha)/j!+o(1)=\Gamma(1-\alpha)\big(1-q_+\big)+o(1).\end{aligned}$$

Here, note that $\lim_{\to\infty}F_j(\lambda n)=\alpha\Gamma(j-\alpha)/j!=\Gamma(1-\alpha)q_j$ by Eq. (7) and Eq. (10), apply (1.13) and recall that $q_+=\sum_{j=1}^{J}q_j$. The rest of the proof of Part (a) of Prop. 3.1 of [1] remains the same so the only modification necessary is to replace the denominator $\Gamma(1-\alpha)$ by $\Gamma(1-\alpha)(1-q_+)$. Likewise the proofs of Part (b) and (c) of the proposition remain valid after corresponding modifications.

We used these results in the proof of Prop. 3.2 of [1] to show that the limiting distribution of $n^{-1}\sum_{i=1}^{\lfloor xn^{\alpha}\rfloor}X_{in}(\lambda n)$ is an infinitely divisible distribution with characteristic exponent $\pi_\lambda(\mathrm{d}y)$ given by Eq. (2.7) of [1]. For the present situation, $\pi_\lambda(\mathrm{d}y)$ is modified to $\pi_\lambda^{(J)}(\mathrm{d}y)$ just by replacing the denominator $\Gamma(1-\alpha)$ in Eq. (2.7) of [1] with $\Gamma(1-\alpha)(1-q_+)$. Thus, letting $\phi_{\lambda n}^{(J)}(\nu):=\mathbf{E}\exp\big(\mathrm{i}\nu X_{1n}^{(J)}(\lambda n)\big)$, $\nu\in\mathbb{R}$, and recalling that $k'=k-m_+\sim\lfloor xn^{\alpha}(1-q_+)\rfloor$, we have

$$\lim_{n\to\infty}\big(\phi_{\lambda n}^{(J)}(\nu/n')\big)^{k'}\quad=\quad\lim_{n\to\infty}\mathbf{E}\exp\Big(\frac{\mathrm{i}\nu}{n'}\sum_{i=1}^{\lfloor xn^{\alpha}(1-q_+)\rfloor}X_{in}^{(J)}(\lambda n)\Big)$$

$$= \quad e^{-x(1-q_+)\int_{\mathbb{R}\setminus\{0\}}(e^{\mathrm{i}\nu y}-1)\Pi_\lambda^{(J)}(\mathrm{d}y)}$$

$$= \quad e^{-x\int_{\mathbb{R}\setminus\{0\}}(e^{\mathrm{i}\nu y}-1)\Pi_\lambda(\mathrm{d}y)} = \mathbf{E}(e^{\mathrm{i}\nu Y_x(\lambda)}).$$

Note that the factor of $(1-q_+)$ cancels and (27) follows as in [1]. This completes the proof of Lemma 3.1. □

Lemma 2 With the substitutions $k = \lfloor xn^\alpha \rfloor$, $m_j = \lfloor q_j + y_j/n^{\alpha/2} \rfloor k$, $k' = k - m_+$, we have the following limiting behaviour as $n \to \infty$:

$$\frac{(k-1)!(1-q_{+n})^{k'}}{k'!\prod_{j=1}^J m_j!}\prod_{j=1}^J q_{jn}^{m_j} \sim \frac{(xn^\alpha)^{-J/2-1}e^{-\frac{x}{2}\mathbf{y}^T\mathbf{Q}_J^{-1}\mathbf{y}}}{\sqrt{(2\pi)^J\det(\mathbf{Q}_J)}}; \tag{1.21}$$

$$\lim_{n\to\infty}\Big(\frac{\alpha}{\Gamma(1-\alpha)}\sum_{\ell=1}^n \frac{\Gamma(\ell-\alpha)}{\ell!}\Big)^{\lfloor xn^\alpha \rfloor} = e^{-x/\Gamma(1-\alpha)}; \tag{1.22}$$

$$\frac{\Gamma(n)\Gamma(\theta/\alpha+k)\Gamma(\theta+1)}{\alpha k'!\Gamma(\theta/\alpha+1)\Gamma(n+\theta)} \sim \frac{x^{\theta/\alpha}(k-1)!\Gamma(\theta+1)}{\alpha k'!\Gamma(\theta/\alpha+1)}; \tag{1.23}$$

$$\lim_{n\to\infty} n\mathbf{P}\Big(\sum_{i=1}^{k'} X_{in}^{(J)} = n'\Big) = f_{Y_x(1)}(1). \tag{1.24}$$

Proof of Lemma 2: As in the proof of Lemma 1 we let $y_{jn} = y_j/n^{\alpha/2}$, $h_{jn} = q_j + y_{jn}$, replace m_j by $h_{jn}k$, and in (1.21) consider first the factor

$$\frac{(k-1)!}{k'!\prod_{j=1}^J m_j!} = \frac{(k-1)!}{((1-h_{+n})k)!\prod_{j=1}^J (h_{jn}k)!}$$
$$\sim \frac{k^{-J/2-1}}{\sqrt{(2\pi)^J(1-h_{+n})\prod_{j=1}^J h_{jn}}} \times \frac{1}{(1-h_{+n})^{(1-h_{+n})k}\prod_{j=1}^J h_{jn}^{h_{jn}k}}.$$

(see (1.1)). Multiply this by the q-terms in (1.21) (again we reinsert the λn argument) to get

$$\frac{(k-1)!}{k'!\prod_{j=1}^J m_j!}(1-q_{+n}(\lambda n))^{k'}\prod_{j=1}^J (q_{jn}(\lambda n))^{m_j}$$
$$\sim \frac{k^{-J/2-1}}{\sqrt{(2\pi)^J(1-h_{+n})\prod_{j=1}^J h_{jn}}}\Big(\frac{1-q_{+n}(\lambda n))}{1-h_{+n}}\Big)^{(1-h_{+n})k}\prod_{j=1}^J\Big(\frac{q_{jn}(\lambda n)}{h_{jn}}\Big)^{h_{jn}k}.$$

Very similar working as in (1.5)–(1.16) then gives the asymptotic in (1.21) for this, and similar working as in (1.12) and (1.19) proves (1.22).

For (1.23) just note that

$$\frac{\Gamma(n)\Gamma(\theta/\alpha+k)\Gamma(\theta+1)}{\alpha k'!\Gamma(\theta/\alpha+1)\Gamma(n+\theta)} \sim \frac{n^{-\theta}k^{\theta/\alpha}(k-1)!\Gamma(\theta+1)}{\alpha k'!\Gamma(\theta/\alpha+1)} \sim \frac{x^{\theta/\alpha}(k-1)!\Gamma(\theta+1)}{\alpha k'!\Gamma(\theta/\alpha+1)}.$$

Finally in (1.24) is the factor involving the X_{in}. Using Fourier inversion as in Eq. (5.13)

$$n\mathbf{P}\Big(\sum_{i=1}^{k'} X_{in}^{(J)} = n'\Big) = \frac{n}{2\pi n'}\int_{-n'\pi}^{n'\pi} e^{-\mathrm{i}\nu}\big(\phi_n^{(J)}(\nu/n')\big)^{k'}\mathrm{d}\nu, \tag{1.25}$$

where $\phi_n^{(J)}(\nu) := \mathbf{E}\big(\exp(\mathrm{i}\nu X_{1n}^{(J)})\big)$, $\nu\in\mathbb{R}$. For this we have

$$\begin{aligned}\phi_n^{(J)}(\nu/n) &= \frac{\sum_{j=J+1}^n e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)/j!}{\sum_{\ell=J+1}^n \Gamma(\ell-\alpha)/\ell!}\\ &= \frac{\sum_{j=1}^n e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)/j! - \sum_{j=1}^J e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)/j!}{\sum_{\ell=1}^n \Gamma(\ell-\alpha)/\ell! - \sum_{\ell=1}^J \Gamma(\ell-\alpha)/\ell!}.\end{aligned} \tag{1.26}$$

(For notational simplicity, in (1.26) and what follows, we replace n' by n, which is irrelevant asymptotically.) In the numerator of (1.26) we can replace the exponentials in the second summation by $1+O(1/n)$, so that (1.26) becomes

$$\phi_n^{(J)}(\nu/n) = \frac{\sum_{j=1}^n e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)/j! - \sum_{j=1}^J \Gamma(j-\alpha)/j!}{\sum_{\ell=1}^n \Gamma(\ell-\alpha)/\ell! - \sum_{\ell=1}^J \Gamma(\ell-\alpha)/\ell!} + O(1/n). \tag{1.27}$$

We will next show that

$$\sum_{j=1}^n \frac{e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)}{j!} = \frac{\Gamma(1-\alpha)}{\alpha} + n^{-\alpha}\int_0^1 (e^{\mathrm{i}\nu z}-1)z^{-\alpha-1}\mathrm{d}z - \frac{n^{-\alpha}}{\alpha} + o(n^{-\alpha}). \tag{1.28}$$

To prove (1.28), first consider

$$\begin{aligned}\sum_{j=1}^\infty \frac{e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)}{j!} &= \int_0^\infty \sum_{j=1}^\infty \frac{e^{\mathrm{i}\nu j/n}z^j}{j!}e^{-z}z^{-\alpha-1}\mathrm{d}z = \int_0^\infty \big(e^{ze^{\mathrm{i}\nu/n}}-1\big)e^{-z}z^{-\alpha-1}\mathrm{d}z\\ &= \int_0^\infty (e^z-1)e^{-z}z^{-\alpha-1}\mathrm{d}z + \int_0^\infty \big(e^{ze^{\mathrm{i}\nu/n}}-e^z\big)e^{-z}z^{-\alpha-1}\mathrm{d}z\\ &= \int_0^\infty (1-e^{-z})z^{-\alpha-1}\mathrm{d}z + n^{-\alpha}\int_0^\infty \big(e^{nz(e^{\mathrm{i}\nu/n}-1)}-1\big)z^{-\alpha-1}\mathrm{d}z\\ &= \Gamma(1-\alpha)/\alpha + n^{-\alpha}\int_0^\infty (e^{\mathrm{i}\nu z}-1)z^{-\alpha-1}\mathrm{d}z + o(n^{-\alpha}).\end{aligned} \tag{1.29}$$

(We integrated by parts to get the first term on the RHS of (1.29).) Next, towards (1.28), we show

$$\sum_{j>n} \frac{e^{\mathrm{i}\nu j/n}\Gamma(j-\alpha)}{j!} = \frac{n^{-\alpha}}{\alpha} + n^{-\alpha}\int_1^\infty (e^{\mathrm{i}\nu z}-1)z^{-\alpha-1}\mathrm{d}z + o(n^{-\alpha}). \tag{1.30}$$

Write the LHS of (1.30) as

$$\sum_{j>n} \frac{\Gamma(j-\alpha)}{j!} + \sum_{j>n} \left(e^{\mathrm{i}\nu j/n} - 1\right) \frac{\Gamma(j-\alpha)}{j!}. \tag{1.31}$$

The first term in (1.31) equals $n^{-\alpha}/\alpha + o(n^{-\alpha})$, by (1.11). To deal with the second term, define rvs Y_n with

$$P(Y_n = j) = \frac{\Gamma(j-\alpha)}{j! \sum_{\ell>n} \Gamma(\ell-\alpha)/\ell!},\ j > n, \tag{1.32}$$

and note that $Y_n > n$ w.p.1. Take $u > 1$ and calculate

$$P(Y_n/n \le u) = P(Y_n \le nu) = P(n < Y_n \le nu)$$

$$= \sum_{n<j\le nu} \frac{\Gamma(j-\alpha)}{j!} \Big/ \sum_{j>n} \frac{\Gamma(j-\alpha)}{j!} = \alpha n^{\alpha}(1+o(1)) \sum_{n<j\le nu} \frac{\Gamma(j-\alpha)}{j!}, \tag{1.33}$$

where we used (1.11) to estimate the denominator. Then note that

$$\alpha n^{\alpha} \sum_{n<j\le nu} \frac{\Gamma(j-\alpha)}{j!} = \alpha \int_0^\infty \sum_{n<j\le nu} \frac{(nz)^j}{j!} e^{-nz} z^{-\alpha-1} \mathrm{d}z$$

$$= \alpha \int_0^\infty P\big(n < \text{Poisson}(nz) \le nu\big) z^{-\alpha-1} \mathrm{d}z. \tag{1.34}$$

This expression equals

$$\int_0^\infty P\Big(\frac{n(1-z)}{\sqrt{nz}} < \frac{\text{Poisson}(nz) - nz}{\sqrt{nz}} \le \frac{n(u-z)}{\sqrt{nz}}\Big) \alpha z^{-\alpha-1} \mathrm{d}z,$$

and when $0 < z < 1$, $P(\text{Poisson}(nz) > n) \le z$, so we can apply dominated convergence to see that the last expression converges to

$$\int_0^\infty P\big(-\infty < N(0,1) < \infty\big) \mathbf{1}_{\{1<z\le u\}} \alpha z^{-\alpha-1} \mathrm{d}z = 1 - u^{-\alpha}. \tag{1.35}$$

Consequently $Y_n/n \xrightarrow{\mathrm{D}} Y$, where $P(Y \le u) = (1-u^{-\alpha})\mathbf{1}_{\{u>1\}}$ and so

$$E\big(e^{\mathrm{i}\nu Y_n/n} - 1\big) \to \int_1^\infty e^{\mathrm{i}\nu u} \alpha u^{-\alpha-1} \mathrm{d}u - 1 = \int_1^\infty \big(e^{\mathrm{i}\nu u} - 1\big) \alpha u^{-\alpha-1} \mathrm{d}u. \tag{1.36}$$

Hence we have for the second term in (1.31)

$$\sum_{j>n} \left(e^{\mathrm{i}\nu j/n} - 1\right) \frac{\Gamma(j-\alpha)}{j!} = \frac{n^{-\alpha}}{\alpha} \Big(1 + E\big(e^{\mathrm{i}\nu Y_n/n} - 1\big)\Big) + o(n^{-\alpha})$$

$$= \frac{n^{-\alpha}}{\alpha} + n^{-\alpha}\int_1^\infty \big(e^{\mathrm{i}\nu z}-1\big)z^{-\alpha-1}\mathrm{d}z + o(n^{-\alpha}), \tag{1.37}$$

proving (1.30). Subtracting (1.30) from (1.29) gives gives (1.28).

Taking account of (1.29) we can write (1.27) as

$$\begin{aligned}
&\phi_n^{(J)}(\nu/n)\\
&= \frac{\Gamma(1-\alpha)/\alpha + n^{-\alpha}\int_0^1(e^{\mathrm{i}\nu z}-1)z^{-\alpha-1}\mathrm{d}z - n^{-\alpha}/\alpha - \sum_{j=1}^J \Gamma(j-\alpha)/j!}{\Gamma(1-\alpha)/\alpha - n^{-\alpha}/\alpha - \sum_{\ell=1}^J\Gamma(\ell-\alpha)/\ell!} + o(n^{-\alpha})\\
&= 1 + \frac{n^{-\alpha}\int_0^1(e^{\mathrm{i}\nu z}-1)z^{-\alpha-1}\mathrm{d}z}{\Gamma(1-\alpha)/\alpha - n^{-\alpha}/\alpha - \sum_{\ell=1}^J\Gamma(\ell-\alpha)/\ell!} + o(n^{-\alpha}),
\end{aligned} \tag{1.38}$$

where we obtained the denominator by putting $\nu = 0$ in (1.27). The denominator in (1.38) is

$$\begin{aligned}
&\Gamma(1-\alpha)/\alpha - n^{-\alpha}/\alpha - \sum_{\ell=1}^J \Gamma(\ell-\alpha)/\ell! + o(n^{-\alpha})\\
&= \Gamma(1-\alpha)(1-q_+)/\alpha - n^{-\alpha}/\alpha + o(n^{-\alpha}),
\end{aligned} \tag{1.39}$$

and so we obtain

$$\phi_n^{(J)}(\nu/n) = 1 + \frac{\int_0^1(e^{\mathrm{i}\nu z}-1)\alpha z^{-\alpha-1}\mathrm{d}z/\Gamma(1-\alpha)}{n^\alpha(1-q_+) - 1/\Gamma(1-\alpha)} + o(n^{-\alpha}). \tag{1.40}$$

Raised to power $k' = k - m_+ \sim xn^\alpha(1-q_+)$, this converges to

$$x\int_0^1 (e^{\mathrm{i}\nu z}-1)\alpha z^{-\alpha-1}\mathrm{d}z/\Gamma(1-\alpha).$$

Note that again the factor $1-q_+$ cancels, and it follows as in [1] that

$$\lim_{n\to\infty} n\mathbf{P}\Big(\sum_{i=1}^{k'} X_{in}^{(J)} = n'\Big) = f_{Y_x(1)}(1). \tag{1.41}$$

This completes the proof of Lemma 4.1. □

2. Appendix: Inverse of $\mathbf{Q}_J$

Lemma 2.1. *The matrix $\mathbf{Q}_J$ defined as the inverse of the matrix $\mathbf{Q}_J^{-1}$ in (1.15) has diagonal elements $q_i(1-q_i)$, off-diagonal elements $-q_iq_j$, $1 \le i \ne j \le J$, and determinant*

$$(1 - \sum_{j=1}^J q_j)\prod_{j=1}^J q_j.$$

Proof of Lemma 2.1: For the inverse, we refer to [2]. Take the matrix $\mathbf{P}$ in their notation to be our matrix $\mathbf{Q}_J$ (so their vector $\mathbf{p}$ is our $(q_1, \dots, q_J)$). Let $a = 1 - \sum_{j=1}^J q_j$ and recall $Q = \prod_{j=1}^J q_j$. Then by Eq. (21) of [2] (note $J < n$, so $\mathbf{Q}_J^{-1}$ exists)

$$\mathbf{Q}_J^{-1} = \mathrm{diag}\Big(\frac{1}{q_1}, \dots, \frac{1}{q_J}\Big) + \frac{1}{1 - \sum_{j=1}^J q_j}\mathbf{1}\mathbf{1}^T = \mathrm{diag}\Big(\frac{1}{q_1}, \dots, \frac{1}{q_J}\Big) + \frac{1}{a}\mathbf{1}\mathbf{1}^{\mathbf{T}},$$

where $\mathbf{1}$ is a J-vector of 1s. The diagonal elements of the matrix on the RHS are

$$\frac{1}{q_j} + \frac{1}{a} = \frac{a + q_j}{q_j a} = \frac{(1 - \sum_{i=1, i\neq j}^J q_i)\prod_{i=1, i\neq j}^J q_i}{aQ} = \frac{Q_j}{aQ}.$$

The off-diagonal elements of the matrix are $1/a = Q/(aQ)$. This verifies that the matrix in (1.15) is indeed the inverse of the matrix $\mathbf{Q}_J$ in Theorem 3.1.

The determinant can be obtained by modifying the calculations in [2], or directly as follows. Denote by $\mathbf{A}$ the $J \times J$ matrix in (1.15) (without the premultiplying factor), let $\mathbf{D} = \mathrm{diag}(Q_1 - Q, \dots, Q_J - Q)$ and let $\mathrm{adj}(\mathbf{D})$ be the adjugate matrix of $\mathbf{D}$. Then $\mathbf{A} = \mathbf{D} + Q\mathbf{1}\mathbf{1}^{\mathbf{T}}$ and by the Sherman-Morrison formula

$$\det(\mathbf{A}) = \det(\mathbf{D}) + Q\mathbf{1}^{\mathbf{T}}\mathrm{adj}(\mathbf{D})\mathbf{1} = \prod_{\mathrm{j=1}}^{\mathrm{J}}(\mathrm{Q_j} - \mathrm{Q}) + \mathrm{Q}\sum_{\mathrm{i=1}}^{\mathrm{J}}\prod_{\mathrm{j\neq i}}(\mathrm{Q_j} - \mathrm{Q}).$$

Note that

$$Q_i = \big(1 - \sum_{1\le j\le J, j\neq i} q_j\big) \prod_{1\le j\le J, j\neq i} q_j = \frac{(a + q_i)Q}{q_i},$$

so $Q_i - Q = aQ/q_i$. Hence

$$\begin{aligned}
&\prod_{j=1}^J (Q_j - Q) + Q\sum_{i=1}^J \prod_{j\neq i}(Q_j - Q) = \prod_{j=1}^J (Q_j - Q)\Big(1 + Q\sum_{i=1}^J \frac{q_i}{aQ}\Big) \\
= \;& \prod_{j=1}^J (Q_j - Q)\Big(1 + Q\frac{1-a}{aQ}\Big) = \frac{1}{a}\prod_{j=1}^J \frac{aQ}{q_j} = \frac{a^J Q^J}{aQ} = a^{J-1}Q^{J-1}.
\end{aligned}$$

Taking into account the premultiplying factor in (1.15) gives $a^{-1}Q^{-1}$ for the determinant of the matrix in (1.15), hence $aQ = \big(1 - \sum_{j=1}^J q_j\big)\prod_{j=1}^J q_j$ as the determinant of $\mathbf{Q}$. □

References

[1] Ipsen, Y.F., Maller, R.A. & Shemehsavar, S. (2021) A generalised Dickman distribution and the number of species in a negative binomial process model. *Adv. Appl. Prob., 53, 370–399.*

[2] Tanabe, K. & Sagae, M. (1992) An exact Cholesky decomposition and the generalized inverse of the variance-covariance matrix of the multinomial distribution, with applications. *J. Roy. Statist. Soc. B (Method.)*, 54, 211–219.