Asymptotic Normality of the Generalized Eulerian Numbers

Lane Clark
Department of Mathematics
Southern Illinois University at Carbondale
Carbondale, IL 62901

ABSTRACT. Let $A_m(n,k)$ denote the number of permutations of $\{1,\ldots,n\}$ with exactly k rises of size at least m. We show that, for each positive integer m, the $A_m(n,k)$ are asymptotically normal.

For positive integers m and n, we say a permutation $\sigma = (\sigma(1),\ldots,\sigma(n))$ of $[n] = \{1,\ldots,n\}$ has an **m-rise at i** provided $\sigma(i+1)-\sigma(i) \geq m$. Let $\mathcal{A}_m(n,k)$ denote the set of permutations of $[n]$ with exactly k m-rises and set $A_m(n,k) = |\mathcal{A}_m(n,k)|$. Hence, the $A_1(n,k)$ are the classical Eulerian numbers (before shifting). Comtet [1; pp. 240–246] gives recurrence relations for the $A_1(n,k)$, as well as, recurrence relations and generating functions for the shifted numbers $B_1(n,k) = A_1(n,k-1)$. Magagnosc [5] introduced the $A_m(n,k)$ and gave several recurrence relations for them. (Many of the results therein actually refer to the shifted numbers $B_m(n,k) = A_m(n,k-1)$.) David and Barton [2; pp. 150–154] showed that the $A_1(n,k)$ are asymptotically normal by computing cumulants. In this paper we show that, for each positive integer m, the $A_m(n,k)$ are also asymptotically normal. Our proof was inspired by the proof of Harper [3] of a similar result for the Stirling numbers of the second kind.

Our first result gives a recurrence relation for the $A_m(n,k)$. (For a different proof of a recurrence relation for the $B_m(n,k)$ see [5].)

Lemma 1. *For $n \geq 3$ and $k \geq 1$ with $n+1 \geq m+k$,*

$$A_m(n,k) = (k+m)A_m(n-1,k) + (n-k-m+1)A_m(n-1,k-1). \quad (1)$$

Proof: Let $\sigma \in \mathcal{A}_m(n-1,k)$ with m-rises at $1 \leq i_1 < i_2 < \cdots < i_k \leq n-2$. Now $\sigma(i_j+1)-\sigma(i_j) \geq m$ so that all $\sigma(i_j) \leq n-1-m$ and, hence,

$k \le n-1-m$. Placing an n at the beginning; after any of $n-m+1, \ldots, n-1$; or after any of $\sigma(i_1), \ldots, \sigma(i_k)$ gives $k+m$ distinct $\sigma' \in \mathcal{A}_m(n,k)$. Next, let $\sigma \in \mathcal{A}_m(n-1, k-1)$ with m-rises at $1 \le i_1 < i_2 < \cdots < i_{k-1} \le n-2$. Again, all $\sigma(i_j) \le n-1-m$. Placing an n after any $\sigma(i) \le n-m$ with $i \notin \{i_1, \ldots, i_{k-1}\}$ gives $n-k-m+1$ distinct $\sigma'' \in \mathcal{A}_m(n,k)$. As deleting n from $\sigma \in \mathcal{A}_m(n,k)$ gives a permutation σ''' of $[n-1]$ satisfying precisely one of the above conditions, our result follows. (Similarly for $k = 1, n-m$; $k = n+1-m$ is trivial.) □

Remark. As a consequence of the first part of the proof of Lemma 1, we have the following

$$A_m(n,k) = 0 \quad \text{for} \quad m+k \ge n+1 \quad \text{with} \quad n \ge 2, k \ge 1.$$

Hence, with $A_m(1,0) = 1$,

$$A_m(n,0) = n! \quad \text{for} \quad 1 \le n \le m$$

while the first construction in Lemma 1 is valid for $k = 0$ and $n \ge m+1$ so that

$$A_m(n,0) = mA_m(n-1,0) \quad \text{for} \quad n \ge m+1.$$

Let

$$P_n(x) = P_{n,m}(x) = \sum_{k=0}^{n} A_m(n,k)x^k.$$

Hence, for $1 \le n \le m$,

$$P_n(x) = n!$$

and, for $n \ge m+1$,

$$P_n(x) = \sum_{k=0}^{n-m} A_m(n,k)x^k$$

since $A_m(n,k) = 0$ for $k \ge n-m+1$ (≥ 2). For $n \ge m+1$ with $n \ge 3$, (1) implies

$$P_n(x) = [(n-m)x + m]P_{n-1}(x) + (x - x^2)P'_{n-1}(x) \tag{2}$$

which is correct for $n = 2$ and $m = 1$ as well. Note that for $n \ge 1$,

$$P_n(1) = n!.$$

Lemma 2. *For* $n \ge m+1$,

$$P_n(x) = m!x^{n-m} + \sum_{k=1}^{n-m-1} m!a_k x^k + m!m^{n-m} \in \mathbb{Z}^+[x]$$

and $P_n(x)$ has $n-m$ distinct negative real roots.

Proof: By (2), both $P_{m+1}(x) = m!x + m \cdot m!$ and $P_{m+2}(x) = m!x^2 + m!(3m+1)x + m!m^2$ have the desired property for $m \geq 1$. For $n \geq m+2$, (2) implies

$$\begin{aligned}
P_{n+1}(x) = {} & \big[(n-m+1)x + m\big]\left[m!x^{n-m} + \sum_{k=1}^{n-m-1} m!a_k x^k + m!m^{n-m}\right] \\
& + (x - x^2)\left[m!(n-m)x^{n-m-1} + \sum_{k=1}^{n-m-1} m!ka_k x^{k-1}\right] \\
= {} & m!x^{n-m+1} + m!\big[2a_{n-m-1} + n\big]x^{n-m} \\
& + \sum_{k=2}^{n-m-1} m!\big[(n-m-k+2)a_{k-1} + (m+k)a_k\big]x^k \\
& + m!\big[(m+1)a_1 + (n-m+1)m^{n-m}\big]x + m!m^{n-m+1} \in \mathbb{Z}^+[x].
\end{aligned}$$

Suppose $P_n(x)$ has distinct real roots $z_1 < z_2 < \cdots < z_{n-m} < 0$. Then $P_n'(z_1), \ldots, P_n'(z_{n-m})$ alternate signs with $P_n'(z_{n-m}) > 0$. From (2), $P_{n+1}(z_i) = (z_i - z_i^2)P_n'(z_i)$ while $z_i - z_i^2 < 0$ and, hence, $P_{n+1}(z_1), \ldots, P_{n+1}(z_{n-m})$ alternate signs with $P_{n+1}(z_{n-m}) < 0$ and $P_{n+1}(0) = m!m^{n-m+1}$. Consequently, $P_{n+1}(x)$ has $n-m$ distinct real roots in $(z_1, 0)$. For $n-m$ even, $P_n'(z_1) < 0$ so that $P_{n+1}(z_1) > 0$ while $\lim_{x\to-\infty} P_{n+1}(x) = -\infty$ and, hence, $P_{n+1}(x)$ has a real root in $(-\infty, z_1)$. For $n-m$ odd, $P_n'(z_1) > 0$ so that $P_{n+1}(z_1) < 0$ while $\lim_{x\to-\infty} P_{n+1}(x) = \infty$ and, hence, $P_{n+1}(x)$ has a real root in $(-\infty, z_1)$. □

In what follows we assume $n \geq m+1$. First, for $n \geq m+1$, (2) implies

$$\begin{aligned}
P_n'(x) = {} & (n-m)P_{n-1}(x) + \big[(n-m-2)x + m + 1\big]P_{n-1}'(x) \\
& + (x - x^2)P_{n-1}''(x)
\end{aligned} \tag{3}$$

so that

$$P_n'(1) = (n-m)P_{n-1}(1) + (n-1)P_{n-1}'(1)$$

which upon iteration using the values of $P_k(1)$, as well as, $P_{m+1}'(1) = m!$ gives

$$P_n'(1) = (n-1)!\binom{n-m+1}{2}.$$

Next, for $n \geq m+2$, (3) implies

$$P_n''(x) = 2(n-m-1)P_{n-1}'(x) + \big[(n-m-4)x + m + 2\big]P_{n-1}''(x) + (x - x^2)P_{n-1}'''(x)$$

so that

$$P_n''(1) = 2(n-m-1)P_{n-1}'(1) + (n-2)P_{n-1}''(1)$$

which upon iteration using the values of $P_k'(1)$, as well as, $P_{m+2}''(1) = 2m!$ gives

$$P_n''(1) = (n-2)!\binom{n-m+1}{3}\frac{3n-3m-2}{2}$$

which is correct for $n = m+1$ as well.

Let $-r_{n,1} < \cdots < -r_{n,n-m} < 0$ be the roots of $P_n(x)$. Now Lemma 2 implies

$$P_n(x) = m!\prod_{j=1}^{n-m}(x + r_{n,j}) \tag{4}$$

so that

$$\frac{P_n'(x)}{P_n(x)} = \sum_{j=1}^{n-m}(x + r_{n,j})^{-1}$$

and, hence,

$$\frac{P_n(x)P_n''(x) - [P_n'(x)]^2}{[P_n(x)]^2} = \frac{\mathrm{d}}{\mathrm{dx}}\left[\frac{P_n'(x)}{P_n(x)}\right] = -\sum_{j=1}^{n-m}(x + r_{n,j})^{-2}.$$

We now introduce a triangular array of row independent random variables $X_{n,j}$ taking on the values $0, 1$, with

$$P(X_{n,j} = 0) = r_{n,j}(1 + r_{n,j})^{-1}, \quad P(X_{n,j} = 1) = (1 + r_{n,j})^{-1}.$$

Let

$$S_n = \sum_{j=1}^{n-m} X_{n,j}$$

so that

$$E(S_n) = \sum_{j=1}^{n-m}(1 + r_{n,j})^{-1} = \frac{P_n'(1)}{P_n(1)} = \frac{(n-m+1)(n-m)}{2n}$$

and

$$\begin{aligned}\sigma^2(S_n) = \mathrm{Var}(S_n) &= \sum_{j=1}^{n-m}(1 + r_{n,j})^{-1} - \sum_{j=1}^{n-m}(1 + r_{n,j})^{-2} \\ &= \frac{P_n'(1)}{P_n(1)} + \frac{P_n''(1)}{P_n(1)} - \left[\frac{P_n'(1)}{P_n(1)}\right]^2 \\ &= \frac{n^4 - n^2 - 2m(2m-1)(m-1)n + 3m^2(m-1)^2}{12n^2(n-1)}.\end{aligned}$$

From (4) we see that the coefficient $A_m(n,k)$ of x^k in $P_n(x)$ is

$$m! \sum_{\substack{K \subseteq [n-m] \\ |K|=k}} \prod_{j \notin K} r_{n,j}.$$

Now independence of the $X_{n,j}$ implies

$$\begin{aligned}
P(S_n = k) &= P(\text{Exactly } k \text{ of the } X_{n,j} = 1) \\
&= \sum_{\substack{K \subseteq [n-m] \\ |K|=k}} \prod_{j \in K} (1 + r_{n,j})^{-1} \cdot \prod_{j \notin K} r_{n,j}(1 + r_{n,j})^{-1} \\
&= \prod_{j=1}^{n-m} (1 + r_{n,j})^{-1} \cdot \sum_{\substack{K \subseteq [n-m] \\ |K|=k}} \prod_{j \notin K} r_{n,j} \\
&= \frac{m!}{P_n(1)} \cdot \frac{A_m(n,k)}{m!} \\
&= \frac{A_m(n,k)}{n!}.
\end{aligned}$$

Hence, the distribution function $F_n(x)$ of S_n satisfies

$$F_n(x) = P(S_n \leq x) = \sum_{k=0}^{\lfloor x \rfloor} \frac{A_m(n,k)}{n!}.$$

Finally, let $G_{n,j}(x)$ denote the distribution function of

$$Y_{n,j} = \frac{X_{n,j} - E(X_{n,j})}{\sigma(S_n)}$$

and $G_n(x)$ denote the distribution function of

$$T_n = \sum_{j=1}^{n-m} Y_{n,j} = \frac{S_n - E(S_n)}{\sigma(S_n)}.$$

Now $Y_{n,j}$ takes on the values $-(1 + r_{n,j})^{-1}/\sigma(S_n)$, $r_{n,j}(1 + r_{n,j})^{-1}/\sigma(S_n)$ so that $|Y_{n,j}| < \sigma^{-1}(S_n)$. For $\epsilon > 0$, $\sigma(S_n) > \epsilon^{-1}$ for all sufficiently large n, so that

$$\int_{|x| \geq \epsilon} x^2 \mathrm{d}G_{n,j} = 0$$

and, hence,

$$\sum_{j=1}^{n-m} \int_{|x| \geq \epsilon} x^2 \mathrm{d}G_{n,j} = 0.$$

By the Lindeberg-Feller Theorem (see [4; p. 295]), $G_n(x)$ converges weakly to the distribution function of a normal random variable with mean 0 and standard deviation 1. Hence, for each $x \in \mathbb{R}$,

$$\frac{1}{n!}\sum_{k=0}^{\lfloor x_n \rfloor} A_m(n,k) = F_n(x_n) = G_n(x) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{dt} \quad \text{as} \quad n \to \infty$$

where

$$x_n = x\sigma(S_n) + E(S_n).$$

We have thus proved the following result.

Theorem 3. *For each* $x \in \mathbb{R}$,

$$\frac{1}{n!}\sum_{k=0}^{\lfloor x_n \rfloor} A_m(n,k) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt \quad \text{as} \quad n \to \infty$$

where

$$x_n = x\left[\frac{n^4 - n^2 - 2m(2m-1)(m-1)n + 3m^2(m-1)^2}{12n^2(n-1)}\right]^{\frac{1}{2}} + \frac{(n-m+1)(n-m)}{2n}.$$

From these results, the reader can easily derive analogous results for the shifted numbers $B_m(n,k) = A_m(n,k-1)$. Finally observe that $\sigma^2(S_n) \to \infty$ as $n \to \infty$ for $m = m(n) = o(n)$, hence, we need not assume m to be constant in Theorem 3, merely that $m = m(n) = o(n)$.

References

[1] L. Comtet, *Advanced Combinatorics*, D. Reidel, Boston, 1974.

[2] F.N. David and D.E. Barton, *Combinatorial Chance*, Charles Griffin, London, 1962.

[3] L.H. Harper, Stirling behavior is asymptotically normal, *Ann. Math. Statist.* **38** (1967), 410–414.

[4] M. Loeve, *Probability Theory*, Van Nostrand, New York, 1963.

[5] D. Magagnosc, Recurrences and formulae in an extension of the Eulerian numbers, *Discrete Math.* **30** (1980), 265–268.