Discrete Applied Mathematics 12 (1985) 31-40
North-Holland

VOTES AND A HALF-BINOMIAL

J.S. FRAME and Dennis C. GILLILAND
Dept. of Mathematics and Dept. of Statistics and Probability, Michigan State University, Wells Hall, East Lansing, MI 48824, USA

Received 17 August 1984
Revised 6 November 1984

In two party elections with popular vote ratio p/q, $\frac{1}{2} \le p = 1-q$, a theoretical model suggests replacing the so-called MacMahon cube law approximation $(p/q)^3$, for the ratio P/Q of candidates elected, by the ratio $f_k(p)/f_k(q)$ of the two half sums in the binomial expansion of $(p+q)^{2k+1}$ for some k. This ratio is nearly $(p/q)^3$ when $k=6$. The success probability $g_k(p) = p^a/(p^a+q^a)$ for the power law $(p/q)^a \doteq P/Q$ is shown to so closely approximate $f_k(p) = \sum_0^k \binom{2k+1}{r} p^{2k+1-r} q^r$, if we choose $a = a_k = (2k+1)!/4^k k! k!$, that $1 \le f_k(p)/g_k(p) \le 1.01884086$ for $k \ge 1$ if $\frac{1}{2} \le p \le 1$. Computationally, we avoid large binomial coefficients in computing $f_k(p)$ for $k > 22$ by expressing $2f_k(p) - 1$ as the sum $(p-q) \sum_0^k (4pq)^s a_s/(2s+1)$, whose terms decrease by the factors $(4pq)(1-1/2s)$. Setting $K = 4k+3$, we compute a_k for large k using a continued fraction $\pi a_k^2 = K + 1^2/(2K + 3^2/(2K + 5^2/(2K + \cdots)))$ derived from the ratio of π to the finite Wallis product approximation.

1. Power laws

Consider an election involving two parties, Party A and Party B, with Party A having a majority of voters. Suppose that the voters are divided into districts and that a party wins a district by obtaining a majority of votes within that district. There is considerable interest among political scientists in the relationship between P, the proportion of districts won by Party A, and p, the proportion of votes for this party over all districts. With random districts of size $2k+1$, the expectation of P is the sum $f_k(p)$ of half the terms of the binomial expansion of $(p+q)^{2k+1}$. We find that $f_k(p)$ is closely approximated by a power law probability $g_k(p) = p^a/(p^a+q^a)$ with exponent $a = a_k = (3/2)(5/4)\cdots(2k+1)/2k$; specifically, $1 \le f_k(p)/g_k(p) \le 1.01884086$ for $\frac{1}{2} \le p \le 1$.

James Parker Smith testified concerning the relationship of P and p before the Royal Commission on Systems of Election in Great Britain on 19 May 1909. Having been elected four times to the House of Commons and having an interest in questions related to proportional representation, his opinion was offered in favor of changing from single-member districts to a system with proportional representation. One of the reasons given was the fact that the system with single-member districts tended to exaggerate in House seats the proportion carried nationwide by the majority party.

Smith presented the empirical evidence on the relationship of P and p for the six

House of Commons elections that took place between 1895 and 1906. This evidence showed that P exceeded p in all but the one case where p was virtually $\frac{1}{2}$.

However, Smith went beyond this. In an apparent effort to bolster his case or to describe the phenomenon under consideration, he presented an analogy whereby the electorate is thought of as a large box of red and blue marbles 'all stirred up' (Smith, 1910, [5, Par. 1253]). Here a district is created by taking a shovelful of marbles and is won by the color in the majority in the shovelful. A second district is created by taking a second shovelful, etc. After introducing this random partition model, Smith dropped the name of mathematician MacMahon, with the implication that MacMahon had analyzed the model and had reached a conclusion. Smith stated '... I have had the help in this of my friend Major MacMahon, who is one of the leading mathematicians of the day, and he gives me this as a formula: that if the electors are in ratio p to q, then the members will be at least in the ratio p^3 to q^3. ...'

The equality in the relationship $P/Q \geq (p/q)^3$ has become known as MacMahon's cube law for election results, although there may be no reference to it in MacMahon's writings (Stanton, 1980, [6, p. 163]). The inadequacy of this law both as a descriptor across elections in general, and in terms of its lack of theoretical underpinnings is well-documented by Tufte (1973). However, there is still interest in the law and the mystery surrounding its original motivation as evidenced by a literature including Kendall and Stuart (1950), Gudgin and Taylor (1979), Stanton (1980) and Gilliland (1985).

The box of marbles analogy with marbles representing *individual* voters and random partitioning does not lead to the cube law (cf. Kendall and Stuart (1950)). For example, suppose $p=0.51$ and that districts have about 10,000 voters. Normal approximation to the binomial leads to the result that the expected proportion P of districts carried by the majority color (party) is approximately 0.98.

Since $P/Q \geq (p/q)^a$ is equivalent to $P \geq p^a/(p^a+q^a)$, an investigation of the cube law and the box of marbles analogy can concern the relationship between the half-binomial probability P and $p^3/(p^3+q^3)$. Stanton (1980) proposes a 'small shovelful theory'. This theory is supported by his demonstration that the half-binomial probability P is well-approximated by $p^3/(p^3+q^3)$ provided that the number of marbles per shovelful is about 13. A case for the applicability of this fact to the election process is made by regarding districts as being made-up of 13 equal-size blocks of voters, each block being made-up of voters of one party, and by regarding a marble as a single *block* of voters.

Stanton's small shovelful model stimulated the research reported in this paper. With $n=2k+1$, we investigate the half-binomial probabilities,

$$f_k(p)=\sum_{r=0}^{k}\binom{n}{r}p^{n-r}q^r, \quad \tfrac{1}{2}\leq p\leq 1,\ k\geq 0,\ n=2k+1 \tag{1.1}$$

and demonstrate that for each $k \geq 0$ there exists an exponent $a>0$ depending on k such that the power law

$$g_k(p)=\frac{p^a}{p^a+q^a}, \quad \tfrac{1}{2}\le p\le 1 \tag{1.2}$$

is an excellent approximation to the polynomial (1.1). (If $n=2k$ and a tie vote in a district is broken by a coin toss, then the probability that the majority party carries the district can be shown to be equal to $f_{k-1}(p)$ so that we will consider only odd-size districts and $n=2k+1$. Furthermore, $f_k(q)=1-f_k(p)$ and $g_k(q)=1-g_k(p)$ and we consider only the domain $\frac{1}{2}\le p\le 1$.) Our results establish an entire family of MacMahon type power laws including the cube law as a special case.

Note that $f_k(p)=g_k(p)=p$ at $p=\frac{1}{2}$ and $p=1$ for all $k\ge 0$. Since $g_k'(\frac{1}{2})=a$, we take

$$a=a_k=f_k'(\tfrac{1}{2}) \tag{1.3}$$

to make f_k and g_k nearly equal. From (1.1) we have

$$f_k'(p)=\sum_{r=0}^{k} np^{2k}\left[\binom{2k}{r}\left(\frac{q}{p}\right)^r-\binom{2k}{r-1}\left(\frac{q}{p}\right)^{r-1}\right]=n\binom{2k}{k}(pq)^k. \tag{1.4}$$

Thus, from (1.3) and (1.4) we have $a_0=1$ and

$$a_k=f_k'(\tfrac{1}{2})=\frac{(2k+1)!}{4^k k!k!}=\prod_{s=1}^{k}\frac{2s+1}{2s}, \quad k\ge 1. \tag{1.5}$$

This also could be derived from the representation of $f_k(p)$ as the incomplete Beta function

$$f_k(p)=\frac{(2k+1)!}{k!k!}\int_0^p x^k(1-x)^k\,\mathrm{d}x \tag{1.6}$$

(cf. Feller [1, Chapter VI, (10.9)]). Thus, a_k is seen to be the ratio of the largest coefficient $\binom{n}{k}$ in the half-binomial to the average coefficient $4^k/(k+1)$. Values for $k\le 7$ are

k	a_k	k	a_k
0	1	4	2.4609375
1	1.5	5	2.70703135
2	1.875	6	2.9326171875
3	2.1875	7	3.14208984375

(1.7)

From the Wallis product $\pi_k=(4k+2)/a_k^2$ we shall derive in Section 3 the continued fraction representation

$$\pi a_k^2=K+\frac{1^2|}{|2K}+\frac{3^2|}{|2K}+\frac{5^2|}{|2K}+\frac{7^2|}{|2K}+\frac{9^2|}{|2K}+\cdots, \quad K=4k+3. \tag{1.8}$$

The first two approximants yield inequalities

$$1/(8k+6.2)<\pi a_k^2-(4k+3)<1/(8k+6), \quad k\ge 5 \tag{1.9}$$

from which $3.1420891 < a_7 < 3.1420917$, thus demonstrating that (1.9) yields six figure accuracy for $k \geq 7$.

With the choice of exponent $a = a_k$ in $g_k(p)$ we denote the relative error function by

$$h_k(p) = \frac{f_k(p)}{g_k(p)} - 1, \quad \tfrac{1}{2} \leq p \leq 1. \tag{1.10}$$

Since $h_0(p) \equiv 0$, we assume $k \geq 1$ and seek first to find the extreme values of $h_k(p)$ for each k. In Section 2 we find, for selected $k \leq 960$, that

$$0 \leq h_k(p) \leq 0.01884086, \quad \tfrac{1}{2} \leq p \leq 1. \tag{1.11}$$

In fact, we prove that the lower bound holds for all k and give an analysis to suggest that the upper bound holds for all k as well.

It is easy to verify $\lim_{k \to \infty} h_k(p) = 0$, $\frac{1}{2} \leq p \leq 1$. However, the pointwise convergence is not uniform in p. Consider for fixed c, $0 \leq c \leq 1$, random variables X_k distributed binomial with $n = 2k+1$ trials and chance of success parameter $p_k = \frac{1}{2} + c/a_k$, and let

$$Y_k = \frac{X_k - np_k}{\sqrt{np_k q_k}}, \quad k \geq 3. \tag{1.12}$$

By Stirling's formula (or (1.9)), $a_k^2 \sim 4k/\pi$, and it follows that the sequence of random variables Y_k converges in distribution to the standard normal distribution, from which

$$\lim_k f_k(p_k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c\sqrt{2\pi}} \exp(-\tfrac{1}{2}t^2)\,\mathrm{d}t. \tag{1.13}$$

It is easy to verify

$$\lim_k g_k(p_k) = \mathrm{e}^{2c}/(\mathrm{e}^{2c} + \mathrm{e}^{-2c}). \tag{1.14}$$

The maximum difference RHS (1.13) − RHS (1.14) is attained at a c-value approximately equal to 0.69 at which the difference is approximately $0.95815 - 0.94048 = 0.01767$. It follows from (1.10) that

$$\lim_k h_k(p_k) \doteq 0.95815/0.94048 - 1 = 0.0188. \tag{1.15}$$

From (1.11) we see that the maximum value of $h_k(p)$, $\frac{1}{2} \leq p \leq 1$, is nearly attained at p_k.

2. Bounds on the relative error

Extreme values of the relative error function $h_k(p)$ in (1.10) occur at the end points of the interval $\frac{1}{2} \leq p \leq 1$ under consideration, where $h_k(p) = 0$, and at any interior points where $h_k'(p)=0$. It is convenient to introduce the following notation:

$$x=p/q=e^{2u}, \quad y=4pq=\operatorname{sech}^2 u, \quad \upsilon=a_k u, \quad w=x^{a_k}=e^{2\upsilon}, \tag{2.1}$$

$$G_k(p)=1/g_k(p)=1+(q/p)^{a_k}=1+e^{-2\upsilon}=1+w^{-1}. \tag{2.2}$$

Theorem 2.1. *The relative error function $h_k(p)$, $\frac{1}{2}\le p\le 1$, achieves its minimum value 0 at $p=\frac{1}{2}$ and $p=1$, is positive on $\frac{1}{2}<p<1$ with a unique maximum occurring at some point p_k.*

Proof. Since

$$G_k'(p)=a_k\left(\frac{q}{p}\right)^{a_k-1}\frac{d(q/p)}{dp}=-\frac{a_k}{pq}\left(\frac{q}{p}\right)^{a_k}=-\frac{4a_k}{yw}, \tag{2.3}$$

and, since $f_k'(p)=a_k y^k$ by (1.4) and $df_k/d\upsilon=\frac{1}{2}y^{k+1}$, we have

$$h_k'(p)=a_k y^k G_k(p)-4f_k(p)a_k/yw=H_k(\upsilon)a_k/yw \tag{2.4}$$

where

$$\begin{aligned} H_k(\upsilon)&=y^{k+1}(w+1)-4f_k(p)\\ &=[\operatorname{sech}^{2k+2}(\upsilon/a_k)](e^{2\upsilon}+1)-4f_k(p) \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} \frac{dH_k}{d\upsilon}&=2y^{k+1}\left[-\frac{(k+1)}{a_k}\tanh\left(\frac{\upsilon}{a_k}\right)(e^{2\upsilon}+1)+e^{2\upsilon}-1\right]\\ &=2y^{k+1}(w+1)\left[\tanh\upsilon-\frac{(k+1)}{a_k^2}\left(a_k\tanh\frac{\upsilon}{a_k}\right)\right]. \end{aligned} \tag{2.6}$$

At $p=\frac{1}{2}$, we have $u=\upsilon=0$, $x=y=w=1$ so that $H_k=H_k'=0$.

For $k\ge 1$, the function $\tanh\upsilon/(a_k\tanh(\upsilon/a_k))$ decreases monotonically in $\upsilon>0$ from 1 to $1/a_k\le 2/3$ and passes once through the value $(k+1)/a_k^2$, which is between $\pi/4$ and $8/9$, when $\upsilon=\upsilon_k^*$, say. For large k, υ_k^* is near the root $\upsilon^*=0.9305436$ of $\tanh\upsilon^*=\pi\upsilon^*/4$. For smaller $k\ge 1$, we have $H_k'(\upsilon_k^*)=0$ if

$$\upsilon_k^*\doteq 0.93054-1/(18.72k+15)>0.901. \tag{2.7}$$

Thus, $H_k'(\upsilon)$ is positive for $0<\upsilon<\upsilon_k^*$ and negative for $\upsilon>\upsilon_k^*$ where $0.901<\upsilon_k^*<0.931$. Hence, $H_k(\upsilon)=0$ has at most one positive root. Since $H_k(\upsilon)$ is negative (near -4) for large υ, it follows that $H_k(\upsilon)=0$ does have a (unique) positive root $\upsilon_k>0.901$. The corresponding unique root p_k of $h_k'(p)=0$, $\frac{1}{2}<p<1$, is

$$p_k=\tfrac{1}{2}(1+\tanh u_k), \qquad u_k=\upsilon_k/a_k=\tfrac{1}{2}\ln x_k. \quad \square \tag{2.8}$$

The value of $h_k(p)$ at its maximum point for fixed k is M_{k-1} where

$$\begin{aligned} M_k &= f_k(p_k)G_k(p_k) \\ &= y_k^{k+1}(w_k+1)(w_k^{-1}+1)/4 \\ &= \operatorname{sech}^{2k+2}(\upsilon_k/a_k)\cosh^2\upsilon_k \end{aligned} \tag{2.9}$$

where $f_k(p_k)$ was found by setting $H_k(\upsilon_k)=0$ in (2.5).

Numerical work was done to determine M_k values. For each k we estimate the root υ_k of $H_k(\upsilon)=0$, or the corresponding $w_k = e^{2\upsilon_k}$, solve for p_k, x_k, y_k, u_k, evaluate H_k, and correct by Newton's method.

For $k>22$, some binomial coefficients in (1.1) have more than 13 digits and direct computation may involve multiple precision. In order to accurately compute $f_k(p)$, we use a series representation in powers of $y=4pq$ where all coefficients are less than 1. This representation follows by noting that

$$\begin{aligned} f_s(p)-f_{s-1}(p) &= \sum_{r=0}^{s}\left[\binom{2s-1}{r}+2\binom{2s-1}{r-1}+\binom{2s-1}{r-2}\right]p^{2s+1-r}q^r \\ &\quad -\sum_{r=0}^{s-1}\binom{2s-1}{r}(p^2+2pq+q^2)p^{2s-1-r}q^r \\ &= \binom{2s-1}{s}p^{s+1}q^s+0-\binom{2s-1}{s-1}p^sq^{s+1} \\ &= \tfrac{1}{2}(p-q)\binom{2s}{s}(pq)^s = \tfrac{1}{2}\tanh u\binom{-\frac{1}{2}}{s}(-y)^s. \end{aligned} \tag{2.10}$$

We sum the differences and use the fact that $2f_0(p)=2p$ to obtain

Theorem 2.2. *The half-binomial probability $f_k(p)$ satisfies*

$$2f_k(p)-1=(p-q)\sum_{s=0}^{k}c_s \quad \text{where } c_0=1, \quad c_s=c_{s-1}y(2s-1)/2s. \tag{2.11}$$

Since $y=1-(p-q)^2$, we see that $2f_k(p)-1$ is the truncated expansion of $(p-q)(1-y)^{-1/2}$ which increases with k from 0 to 1.

To solve $H_k(\upsilon)=0$ for its positive root υ_k and obtain the related quantities w_k, x_k, y_k, u_k, p_k we need a good initial estimate of υ_k and an approximate value of H_k' to use in applying Newton's method to correct the estimate. To avoid the factor y^{k+1} in $dH_k/d\upsilon$ in (2.6), which tends to 0 for large k, we find instead the root of $H_k/y^{k+1}=Z_k(w)$ whose derivative changes very slowly with k. We set

$$Z_k(w)=H_k/y^{k+1}=w+1-4f_k/y^{k+1}. \tag{2.12}$$

Noting that $df_k/dw=-y^{k+1}/4w$ we find

$$\begin{aligned} Z_k'(w_k) &= 1-w_k^{-1}-(4f_k/y^{k+1})(k+1)\tanh u_k/(a_kw_k) \\ &= (1+w_k^{-1})[\tanh\upsilon_k-(k+1)(\tanh(\upsilon_k/a_k))/a_k]. \end{aligned} \tag{2.13}$$

Starting with the exact root $x_1 = p_1/q_1 = (\sqrt{2}+1)^2$ we find $w_1 = x_1^{a_1} = (\sqrt{2}+1)^3 = 14.07107$. Calculations show that w_k increases very slowly with increasing k and that good first approximations for w_k and $v_k = a_k u_k$ are

$$w_k \doteq 14.375 - 0.35/k, \qquad v_k = \tfrac{1}{2} \ln w_k \doteq 1.33275 - 0.0122/k. \tag{2.14a}$$

The function $Z'(w_k)$ changes very little with k and can be approximated in (2.13) by $Z_k'(w_k) = -0.19 + (k+1)/4$. We correct w_k by adding $\Delta w = -Z_k/Z_k'$.

To obtain more accurate values of w_k for $k > 120$, we set $k = 960/z$ for $z = 6, 5, 4, 3, 2, 1$ and used a computer displaying 16 figure accuracy to obtain improved approximations of a first estimate for w_k. We found

$$w_k \doteq 14.3752461 - 0.3454/(k - 0.7), \tag{2.14b}$$

which for $k > 120$ requires a first correction Δw with $|\Delta w| \leq 10^{-5}$. An improved estimate for Δw was found to be

$$\Delta w \doteq -Z_k(w)/Z_k'(w) \doteq (5.28759 + 6.92/k) Z_k(w). \tag{2.15}$$

A comparison of initial estimates using (2.14a) and (2.14b) with final computed values are shown in the following table.

k	0.375-0.35/k	$w_k - 14$	k	0.52461-34.54/$(k-0.7)$	100w-1437
20	0.35750	0.35737 31	160	0.30778 64	0.30767 4234
24	0.36042	0.36042 28	192	0.34405 59	0.34398 3264
30	0.36333	0.36344 82	240	0.38027 72	0.38023 2312
40	0.36625	0.36674 64	320	0.41643 59	0.41642 0663
60	0.36917	0.36941 44	480	0.45254 66	0.45254 7590
120	0.37208	0.37234 88	960	0.48860 46	0.48861 2367

(2.16)

After computing the values $w_{960/z}$ for $z = 1, \ldots, 6$ we fit a polynomial of degree 5 to these values and setting $z = 0$ were led to the following estimates for the limits of w_k and v_k:

$$\begin{aligned} W &= \lim_k w_k \doteq 14.3752461427 \\ V &= \tfrac{1}{2} \ln W \doteq 1.33275385475. \end{aligned} \tag{2.17}$$

The extreme values M_k are given by (2.9) and they increase with k. By (2.9),

$$\ln M_k = 2 \ln \cosh v_k - (2k+2) \ln \cosh (v_k/a_k)$$

so that

$$\begin{aligned} \ln M = \lim_k \ln M_k &= 2 \ln \cosh V - \lim_k (k+1) V^2/a_k^2 \\ &= 2 \ln \cosh V - \pi V^2/4 \end{aligned} \tag{2.18}$$

where use is made of (1.9). Hence

$$M = \cosh^2 V e^{(-\pi V^2)/4} = 1.01884085746. \tag{2.19}$$

Allowing for uncertainty in the tenth significant figure for W and V, we can assert

$$0 \le h_k(p) \le 0.01884086, \quad \tfrac{1}{2} \le p \le 1, \quad k \ge 1. \tag{2.20}$$

The numerical evaluations establish the upperbound for the range $k \le 960$ and the analysis suggests the result for all k.

Our first look at power laws resulted in a positive lower bound estimate for $h_k(p)$. This proof is too lengthy to present here and is based on a representation:

$$\mathrm{e}^{a_k u} \cosh^n u \; h_k(p) = E \sinh[2u(X - EX)]. \tag{2.21}$$

Here X is a random variable with truncated binomial distribution $p(x) = 4^{-k}\binom{n}{x}$, $x = k+1, \ldots, n$. One can show that the central moments satisfy $E(X - EX)^j > 0$, $j \ge 2$, which together with the series representation of the sinh function and (2.21) establishes the lower bound in (2.20). The results on the central moments are new and of some independent interest. We do conjecture the non-negativity of all central moments for any discrete distribution on a lattice with monotone decreasing probabilities.

Of course, by choice of an exponent a in (1.2) which is somewhat larger than a_k, $\max\{|h_k(p)| : \tfrac{1}{2} \le p \le 1\}$ can be reduced below that value resulting from the exponent a_k. We have not explored methods leading to such choices.

3. Wallis product correction

Consider the Wallis product

$$\pi_k = 2 \cdot \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdots \frac{2k}{2k-1} \cdot \frac{2k}{2k+1} = \frac{4k+2}{a_k^2}, \quad k \ge 1 \tag{3.1}$$

which converges to π as $k \to \infty$. We use this to compute a_k.

Theorem 3.1. *The Wallis product correction factor is the ratio*

$$\frac{\pi}{\pi_k} = \prod_{j=k+1}^{\infty} \frac{4j^2}{4j^2 - 1} = \frac{\pi a_k^2}{4k+2}, \quad k \ge 1, \tag{3.2}$$

and it is expressible by the following continued fraction where $K = 4k+3$:

$$\frac{\pi}{\pi_k} = \frac{1}{\mathrm{K}-1}\left(K + \frac{1|}{|2K} + \frac{9|}{|2K} + \frac{25|}{|2K} + \frac{49|}{|2K} + \frac{81|}{|2K} + \frac{121|}{|2K} + \cdots\right),$$
$$K = 4k+3. \tag{3.3}$$

Proof. Expressing the logarithm as an integral by writing

$$\ln r = \int_1^r s^{-1}\,\mathrm{d}s = \int_0^\infty \int_1^r \mathrm{e}^{-st}\,\mathrm{d}s\,\mathrm{d}t = \int_0^\infty (\mathrm{e}^{-t} - \mathrm{e}^{-rt})\,\mathrm{d}t/t, \tag{3.4}$$

we have

$$\ln\left(\frac{4j^2}{4j^2-1}\right)=\int_0^\infty (e^{-(2j-1)t}-2e^{-2jt}+e^{-(2j+1)t})dt/t$$

$$=\int_0^\infty \tanh(t/2)(e^{-(2j-1)t}-e^{-(2j+1)t})dt/t. \quad (3.5)$$

Summing from $j=k+1$ to ∞, and setting $t=2u$ yields

$$\ln(\pi/\pi_k)=\int_0^\infty \tanh u\, e^{-(K-1)u}du/u. \quad (3.6)$$

Noting that $\tanh u = 1-e^{-u}\operatorname{sech} u$ we have

$$\ln((K-1)/K)=\int_0^\infty (e^{-Ku}-e^{-(K-1)u})du/u,$$

$$\ln\left(\frac{\pi}{\pi_k}\cdot\frac{K-1}{K}\right)=\int_0^\infty e^{-Ku}(1-\operatorname{sech} u)du/u. \quad (3.7)$$

Expanding $1-\operatorname{sech} u$ in terms of the Euler numbers

$$E_j=1,5,61,1385,50521,2702765,\ldots$$

we express (3.7) by the asymptotic series

$$\frac{\pi}{\pi_k}\cdot\frac{K-1}{K}=\exp\left(\sum_{j=1}^{\infty}(-1)^{j-1}\frac{E_j}{2jK^{2j}}\right)$$

$$=1+1/2K^2-9/8K^4+153/16K^6-21429/2^7K^8+$$

$$+1268343/2^8K^{10}-\cdots \quad (3.8)$$

Surprisingly, this series yields a continued fraction having simple predictable numerators

$$\frac{\pi}{\pi_k}\cdot\frac{K-1}{K}=$$

$$=1+\frac{1/2K^2|}{|1}+\frac{9/4K^2|}{|1}+\frac{25/4K^2|}{|1}+\frac{49/4K^2|}{|1}+\frac{81/4K^2|}{|1}+\cdots \quad (3.9)$$

from which (3.3) follows. □

To illustrate the rapid rate of convergence of (3.3) take $k=3$ so that $\pi_3=512/175=2.925714\ldots$, $K=15$. The sixth convergent correction in (3.3) yields

$$\pi \doteq \frac{256}{1225}\left(15+\frac{1|}{|30}+\frac{9|}{|30}+\frac{25|}{|30}+\frac{49|}{|30}+\frac{81|}{|30}+\frac{121|}{|30}\right)$$

$$=3.14159265349\ldots \quad (3.10)$$

which is in error by less than 10^{-10}.

References

[1] W. Feller, An Introduction to Probability Theory and its Applications, Vol. I, 2nd ed. (Wiley, New York, 1960).

[2] D.C. Gilliland, On the MacMahon cube law for election results, J. Appl. Statistics (1985), to appear.

[3] G. Gudgin and P.J. Taylor, Seats, Votes, and the Spatial Organization of Elections (Pron Ltd., London, 1979).

[4] M.G. Kendall and A. Stuart, The law of the cubic proportion in election results, British J. Sociol. 1 (1950) 183–196.

[5] J.P. Smith, Minutes of evidence before the royal commission on systems of election (Cmd. 5352, B.P.P. XXVI (1910)) 77–86.

[6] R.G. Stanton, A result of MacMahon on electoral predictions, Ann. Discrete Math. 8 (1980) 163–167.

[7] E.R. Tufte, The relationship between seats and votes in two-party systems, Am. Pol. Sci. Review 67 (1973) 540–554.