MARTINGALES IN THE OK CORRAL

J. F. C. KINGMAN

1. *Introduction*

In the model of the OK Corral formulated by Williams and McIlroy [**2**]: 'Two lines of gunmen face each other, there being initially m on one side, n on the other. Each person involved is a hopeless shot, but keeps firing at the enemy until either he himself is killed or there is no one left on the other side.' They are interested in the number S of survivors when the shooting ceases, and the surprising result, for which they give both numerical and heuristic evidence, is that when $m = n$,

$$\mathbf{E}\{S\} \sim 2\,(3)^{-1/4}\pi^{-1/2}\Gamma(3/4)\,m^{3/4} \tag{1.1}$$

as $m \to \infty$ (where $\mathbf{E}$ denotes expectation).

It is the occurrence of this curious power of m, rather than any application to real gunfights, which makes the Williams–McIlroy process of interest. The purpose of this paper is to give an essentially elementary proof of the fact that if m and n are not too different, then $S/m^{3/4}$ has a simple asymptotic distribution that leads at once to (1.1).

Let X_t and Y_t denote the number of gunmen in the two gangs left standing after t deaths ($t = 0, 1, 2, \ldots$), so that

$$X_0 = m, \quad Y_0 = n \quad \text{and} \quad X_t + Y_t = N - t, \quad \text{where } N = m + n. \tag{1.2}$$

The precise formulation in [**2**] is equivalent to saying that (X_t, Y_t) is a Markov chain with states (x, y), where x and y are non-negative integers, not both zero, and the only non-zero transition probabilities are given by

$$\begin{aligned} \mathbf{P}\{X_{t+1} = X_t - 1,\, Y_{t+1} = Y_t \,|\, X_t, Y_t\} &= Y_t/(N-t), \\ \mathbf{P}\{X_{t+1} = X_t,\, Y_{t+1} = Y_t - 1 \,|\, X_t, Y_t\} &= X_t/(N-t) \end{aligned} \tag{1.3}$$

(where $\mathbf{P}$ denotes probability).

The states $(s, 0)$ and $(0, s)$ are absorbing, and the process ends in either $(S, 0)$ or $(0, S)$. The distribution of S depends only on m and n.

Let g be any function on the positive integers, and write

$$f(m, n) = \mathbf{E}\{g(S)\}. \tag{1.4}$$

The backward Kolmogorov equation is

$$(m+n)f(m, n) = nf(m-1, n) + mf(m, n-1) \quad (m, n \geqslant 1), \tag{1.5}$$

with boundary conditions

$$f(m, 0) = f(0, m) = g(m). \tag{1.6}$$

By induction on $N = m+n$, (1.5) and (1.6) determine f uniquely in terms of g, and it is clear that

$$f(m, n) = f(n, m). \tag{1.7}$$

Received 28 September 1998.

1991 *Mathematics Subject Classification* 60F05.

Bull. London Math. Soc. 31 (1999) 601–606

Conversely, if f is any solution of (1.5) and (1.7), define g by (1.6), and let f^* denote the left-hand side of (1.4). Then f^* satisfies (1.5) and (1.6), and therefore $f^* = f$. In other words, any solution f of (1.5) and (1.7) leads to an equation of the form

$$\mathbf{E}\{f(S,0)\} = f(m,n). \tag{1.8}$$

This is, of course, a martingale stopping identity, since (1.5) is the condition for $f(X_t, Y_t)$ to be a martingale.

There are also solutions of (1.5) that do not satisfy (1.7), and for those the stopping identity is a little more complicated. If the random variable χ is defined to be 1 if the process ends in $(S,0)$ and 0 if it ends in $(0,S)$, then (1.8) must be replaced by

$$\mathbf{E}\{f(S,0)\,\chi + f(0,S)\,(1-\chi)\} = f(m,n). \tag{1.9}$$

2. *Even polynomial martingales*

The simplest non-trivial solution of (1.5) is easily found to be

$$f(m,n) = (N+1)\,(m-n) = (m+1/2)^2 - (n+1/2)^2, \tag{2.1}$$

but this does not satisfy (1.7). The simplest even solution is (again checked by substitution into (1.5))

$$f(m,n) = (N+1)\,(N+2)\,\{N+3(m-n)^2\}. \tag{2.2}$$

If this is substituted into (1.4), it gives, when $m = n$, the stopping identity

$$\mathbf{E}\{S(S+1)\,(S+2)\,(3S+1)\} = N(N+1)\,(N+2), \tag{2.3}$$

from which it is easy to show that

$$\mathbf{E}\{S^4\} \sim N^3/3 \tag{2.4}$$

as $N \to \infty$. This is clearly consistent with, though it does not prove, an asymptotic increase of S with the $3/4$ power of N or $m = N/2$.

The important thing about (2.3), however, is that it is the first of an infinite family of identities using solutions to (1.5) which are polynomials symmetric in m and n.

THEOREM 1. *For every positive integer K, there are solutions of* (1.5) *of the form*

$$f(m,n) = \sum_{k=0}^{K} a_k(N)\,(m-n)^{2k}, \tag{2.5}$$

where each a_k is a polynomial in N with leading term

$$(2K)!\,N^{3K-k}/6^{K-k}(2k)!\,(K-k)!. \tag{2.6}$$

Proof. Substituting (2.5) into (1.5), we obtain

$$\begin{aligned} N\sum_{k=0}^{K} a_k(N)\,(m-n)^{2k} &= n\sum_{k=0}^{K} a_k(N-1)\,(m-1-n)^{2k} + m\sum_{k=0}^{K} a_k(N-1)\,(m-n+1)^{2k} \\ &= \sum_{k=0}^{K} a_k(N-1)\sum_{i=0}^{2k}\binom{2k}{i}(m-n)^i\,\{n(-1)^{2k-i}+m\} \\ &= \sum_{k=0}^{K} a_k(N-1)\left\{\sum_{j=0}^{k}\binom{2k}{2j}(m-n)^{2j}\,N \right. \\ &\qquad \left. + \sum_{j=1}^{k}\binom{2k}{2j-1}(m-n)^{2j-1}(m-n)\right\}, \end{aligned}$$

https://doi.org/10.1112/S0024609399006098 Published online by Cambridge University Press

and this will be true for all m and n if, for $j = 0,1,2,\ldots,K$,

$$Na_j(N) = \sum_{k=j}^{K} \left\{ N\binom{2k}{2j} + \binom{2k}{2j-1} \right\} a_k(N-1). \tag{2.7}$$

Equations (2.7) can be solved by downwards induction on j, starting with

$$Na_K(N) = (N+2K)\,a_K(N-1).$$

This is satisfied by

$$a_K(N) = (N+1)(N+2)\ldots(N+2K) = (N+1)^{(2K)}, \tag{2.8}$$

using a standard notation for rising factorials. Now write

$$a_j(N) = (N+1)^{(2j)}\,b_j(N),$$

so that (2.7) becomes

$$b_j(N) - b_j(N-1) = \sum_{k=j+1}^{K} \left\{ N\binom{2k}{2j} + \binom{2k}{2j-1} \right\} (N+2j+1)^{(2k-2j-1)} b_k(N-1).$$

This is satisfied if

$$b_j(N) = \sum_{r=0}^{N-1} \sum_{k=j+1}^{K} \left\{ (r+1)\binom{2k}{2j} + \binom{2k}{2j-1} \right\} (r+2j+2)^{(2k-2j-1)}\,b_k(r). \tag{2.9}$$

It now follows by downwards induction on j that b_j is a polynomial with leading term

$$(2K)!\,N^{3(K-j)}/6^{K-j}(2j)!\,(K-j)!,$$

where we use the fact that if P is a polynomial with leading term pN^n, then

$$\sum_{r=0}^{N-1} P(r)$$

is a polynomial with leading term $pN^{n+1}/(n+1)$. This then shows that a_j is a polynomial with leading term (2.6).

Note that it is not asserted that (2.5) is unique. The proof has made a natural choice among a range of possibilities, but any two choices differ by a linear combination of the polynomials for smaller values of K. The solutions of (1.5) which are, for fixed N, even polynomials in $m-n$ of degree $\leqslant 2K$ form a vector space of dimension K.

3. *The fair gunfight*

Substituting (2.5) into (1.4), we have the identity

$$\mathbf{E}\left\{ \sum_{k=0}^{K} a_k(S)\,S^{2k} \right\} = \sum_{k=0}^{K} a_k(N)\,(m-n)^{2k} \tag{3.1}$$

for $K = 1,2,3,\ldots$. This is enough to derive the asymptotic distribution of S when N is large, and the next theorem deals with the case when m and n are equal, or nearly so.

https://doi.org/10.1112/S0024609399006098 Published online by Cambridge University Press

THEOREM 2. *If $m, n \to \infty$ in such a way that*

$$m - n = o(\sqrt{N}), \tag{3.2}$$

then the distribution of

$$3S^4/2N^3 \tag{3.3}$$

converges to the gamma distribution with parameter 1/2, *and*

$$\mathbf{E}\{(3S^4/2N^3)^\nu\} \to \Gamma(\nu + 1/2)/\Gamma(1/2) \tag{3.4}$$

for any $\nu > 0$. For any fixed $\lambda > 0$,

$$\mathbf{P}\{S \leqslant \lambda N^{3/4}\} \to 2\,\Phi(\lambda^2 \sqrt{3}) - 1, \tag{3.5}$$

where Φ is the standard normal distribution function.

To prove the theorem, use (2.6) and (3.2) to show that the term in k on the right of (3.1) is

$$a_k(N)\,(m-n)^{2k} = O(N^{3K-k})\,o(N^k) = o(N^{3K})$$

for $k \geqslant 1$, while the term in $k = 0$ is

$$a_0(N) \sim (2K)!\,N^{3K}/6^K K!.$$

Hence

$$\mathbf{E}\left\{\sum_{k=0}^{K} a_k(S)\,S^{2k}\right\} \sim (2K)!\,N^{3K}/6^K K!.$$

The left-hand side is

$$\mathbf{E}\left\{S^{2K}(S+1)^{(2K)} + \sum_{k=0}^{K-1} a_k(S)\,S^{2k}\right\} = \mathbf{E}\{S^{4K}\} + O[\mathbf{E}\{S^{4K}\}^{(4K-1)/4K}]$$

by (2.6) and Hölder's inequality. Hence

$$\mathbf{E}\{S^{4K}\} \sim (2K)!\,N^{3K}/6^K K! \tag{3.6}$$

as $N \to \infty$. If R denotes the random variable (3.3), this means that

$$\mathbf{E}\{R^K\} \to (2K)!/2^{2K}K! = \Gamma(K + 1/2)/\Gamma(1/2) \tag{3.7}$$

for $K = 1, 2, 3, \ldots$. Since $\Gamma(K+\alpha)/\Gamma(\alpha)$ is the Kth moment of the gamma distribution with parameter α, and of no other distribution, this means that R has the limiting distribution with density

$$x^{-1/2}\,e^{-x}/\Gamma(1/2), \tag{3.8}$$

and [**1**, Section 11.4] that all the moments converge, giving (3.4).

For fixed $\lambda > 0$,

$$\begin{aligned}
\mathbf{P}\{S \leqslant \lambda N^{3/4}\} &= \mathbf{P}\{R \leqslant 3\lambda^4/2\} \\
&\to \int_0^{3\lambda^4/2} x^{-1/2}\,e^{-x}/\Gamma(1/2)\,dx \\
&= \int_0^{\lambda^2\sqrt{3}} (2/\pi)^{1/2}\,e^{-y^2/2}\,dy \\
&= 2\Phi(\lambda^2\sqrt{3}) - 1,
\end{aligned}$$

and the proof is complete.

Setting $\nu = 1/4$ in (3.4) gives the Williams–McIlroy result (1.1).

4. *Unfair fights*

The identity (3.1) contains useful information not only when m and n are effectively equal, but also when they are quite different. For instance, if $m-n$ is large compared with $\sqrt{N}$, it implies that

$$\mathbf{E}\{S^{4K}\} \sim N^{2K}(m-n)^{2K}, \tag{4.1}$$

for $K = 1, 2, 3, \ldots$. From this it follows that, with probability 1,

$$S \sim \sqrt{\{N(m-n)\}} = \sqrt{|m^2-n^2|}, \tag{4.2}$$

which accords with the deterministic approximation in [**2**].

Of more interest is the intermediate situation in which $(m-n)$ and $\sqrt{N}$ are comparable. Suppose for definiteness that, for some constant θ,

$$N^{-1/2}(m-n) \to \theta \quad \text{as } N \to \infty. \tag{4.3}$$

Then (3.1) and (2.6) show that

$$\begin{aligned}\mathbf{E}\{S^{4K}\} &\sim \sum_{k=0}^{K} \frac{(2K)!\,N^{3K-k}\,\theta^{2k}\,N^k}{6^{K-k}(2k)!\,(K-k)!} \\ &= N^{3K} \sum_{k=0}^{K} \frac{(2K)!\,\theta^{2k}}{6^{K-k}(2k)!\,(K-k)!}.\end{aligned} \tag{4.4}$$

This leads to a generalisation of Theorem 2, giving a limiting distribution for (3.3) which depends on the value of θ. But it is now a matter of concern to which gang the S survivors belong, which is indicated by the value of χ. It turns out to be neatest to code the pair (S, χ) by means of

$$Z = S^2(2\chi-1)/N^{3/2}, \tag{4.5}$$

which contains both S and χ, since

$$S^2 = |Z|\,N^{3/2} \tag{4.6}$$

and

$$\chi = 1 \quad \text{if and only if} \quad Z > 0. \tag{4.7}$$

Equation (4.4) gives the asymptotic values of the even moments of Z. To derive a limiting distribution for Z, it is necessary to estimate the odd moments too, and this is done by using odd solutions of (1.5), of the form

$$f(m,n) = \sum_{k=1}^{K} a_k(N)\,(m-n)^{2k-1}. \tag{4.8}$$

An argument exactly similar to that used to prove Theorem 1 shows that there are such solutions in which each a_k is a polynomial with leading term

$$(2K-1)!\,N^{3K-k-1}/6^{K-k}(2k-1)!\,(K-k)!.$$

This shows that, for $K = 1, 2, 3, \ldots$,

$$\mathbf{E}\{S^{4K-2}(2\chi-1)\} \sim N^{3K-3/2} \sum_{k=1}^{K} \frac{(2K-1)!\,\theta^{2k-1}}{6^{K-k}(2k-1)!\,(K-k)!}, \tag{4.9}$$

so that

$$\mathbf{E}\{Z^{2K-1}\} \to \sum_{k=1}^{K} \frac{(2K-1)!\,\theta^{2k-1}}{6^{K-k}(2k-1)!\,(K-k)!}. \tag{4.10}$$

Since, from (4.4),

$$\mathbf{E}\{Z^{2K}\} \to \sum_{k=0}^{K} \frac{(2K)!\,\theta^{2k}}{6^{K-k}(2k)!\,(K-k)!}, \tag{4.11}$$

we have the limiting values of all the moments of Z. The reader will recognise the right-hand sides of (4.10) and (4.11) as the moments of the normal distribution with mean θ and variance $1/3$, and of this distribution alone, so that this is the limiting distribution of Z. Thus we have proved the following theorem.

THEOREM 3. *Let* $m, n \to \infty$ *in such a way that* (4.3) *holds. Then the random variable* Z *defined by* (4.5) *has a distribution that converges to the normal distribution with mean* θ *and variance* $1/3$. *In particular*,

$$\mathbf{P}\{\chi = 1\} \to \Phi(\theta\sqrt{3}). \tag{4.12}$$

Thus the first gang can be confident of winning, in the sense of annihilating the second, if (for large m, n)

$$m - n > 2\sqrt{(m+n)},$$

for then

$$\mathbf{P}\{\chi = 1\} > 0.9997.$$

The occurrence of the random variable Z can be made less surprising by noting that the solution of (1.5) given by (2.1) yields the martingale

$$M_t = (X_t + 1/2)^2 - (Y_t + 1/2)^2, \tag{4.13}$$

whose final value when the shooting stops is

$$M = S(S+1)(2\chi - 1). \tag{4.14}$$

Thus Theorem 3 asserts the asymptotic normality of M, and can be seen as an example of the propensity of martingales to have normal limits.

Finally, it is worth noting that Theorem 3 is consistent with the generalisation to the unfair case of the diffusion argument given in [**2**]. I owe this remark, and encouragement to work on this problem, to Professor David Williams, to whom it is a pleasure to record my thanks.

References

1. M. LOEVE, *Probability theory* (van Nostrand, Princeton, NJ, 1963).
2. D. WILLIAMS and P. MCILROY, 'The OK Corral and the power of the law (a curious Poisson-kernel formula for a parabolic equation)', *Bull. London Math. Soc.* 30 (1998) 166–170.

University of Bristol
Senate House
Tyndall Avenue
Bristol BS8 1TH