

**A NOTE ON THE HIGHER MOMENTS OF THE RANDOM VARIABLE  $T$  ASSOCIATED WITH THE NUMBER OF RETURNS OF A SIMPLE RANDOM WALK**

WALTER KATZENBEISSER\* and  
WOLFGANG PANNY,\* *Wirtschaftsuniversität Wien*

Let  $X_k, k = 1, 2, \dots$  be independent and identically distributed random variables with

$$P\{X_k = 1\} = P\{X_k = -1\} = \frac{1}{2}.$$

Consider the simple random walk

$$S_n = \sum_{k=1}^n X_k, n = 1, 2, \dots, 2n, \text{ with } S_0 = 0 \text{ and } S_{2n} = 0,$$

i.e. a simple random walk starting at 0 and leading to 0 after  $2n$  steps. To this random walk the random variable

$$T = [\text{number of visits to the origin}],$$

is associated. This paper primarily deals with the asymptotic behaviour of the moments of  $T$ . This investigation can be motivated by the following applications: (i) McGilchrist and Woodyer (1975) used  $T$  for a distribution-free CUSUM procedure. (ii) Katzenbeisser and Hackl (1985) examined the applicability of  $T$  as test statistic, where the Kolmogorov–Smirnov two-sample test was considered as an alternative. The distribution of  $T$  may be given by (cf. e.g. Dwass (1967)):

$$P\{T > t\} = 2^t \binom{2n-t}{n-t} \binom{2n}{n}^{-1}, \quad t = 1, 2, \dots, n.$$

The moments of  $T$  can be derived by summation by parts:

$$E\{T^s\} = \sum_{t=0}^n ((t+1)^s - t^s) 2^t \binom{2n-t}{n-t} \binom{2n}{n}^{-1}.$$

For  $s = 1, 2$  ‘closed-form’ expressions are given by Katzenbeisser and Hackl (1985):

$$E\{T\} = 2^{2n} \binom{2n}{n}^{-1},$$

$$E\{T^2\} = 4n + 2 - 2^{2n} \binom{2n}{n}^{-1}.$$

For higher moments, however, there seems to be no easy way to obtain closed-form expressions. The asymptotic behaviour of the moments of  $T$  is given by the following result.

---

Received 5 July 1985; revision received 27 September 1985.

\* Postal address: Wirtschaftsuniversität Wien, Institut für Statistik, A-1090 Wien, Augasse 2–6, Austria.

*Theorem.* Let  $n \rightarrow \infty$ , then we have:

$$E\{T\} = \sqrt{\pi n}^{\frac{1}{2}} + \frac{1}{8}\sqrt{\pi n}^{-\frac{1}{2}} + O(n^{-1}),$$

$$E\{T^2\} = 4n - \sqrt{\pi n}^{\frac{1}{2}} + 2 + O(n^{-\frac{1}{2}}).$$

For  $s \geq 3$ , the following general formula applies:

$$E\{T^s\} = \Gamma\left(\frac{s+2}{2}\right)(4n)^{s/2} - \binom{s}{2}\Gamma\left(\frac{s+1}{2}\right)(4n)^{(s-1)/2}$$

$$+ \frac{s}{24}(3s^3 - 8s^2 + 21s - 10)\Gamma\left(\frac{s}{2}\right)(4n)^{(s-2)/2} + O(n^{(s-3)/2}).$$

*Proof.* By Stirling's approximation one obtains after some manipulations

$$(1) \quad 2^{t-2n} \binom{2n-t}{n-t} = \sqrt{\frac{2}{(2n-t)\pi}} \left[ 1 - \frac{1}{48(2n-t)} \left( 12 - \frac{24t^2}{2n-t} + \frac{4t^4}{(2n-t)^2} \right) \right]$$

$$\times \exp \left\{ -\frac{t^2}{2(2n-t)} \right\} + o(n^{-2}), \quad \text{as } n \rightarrow \infty.$$

This formula holds for all  $t$ . Moreover, it is well known (cf. Feller (1968), p. 180) that  $\binom{2n-t}{n} 2^{t-2n}$  becomes exponentially small for  $t \geq n^{1/2+\epsilon}$ ,  $\epsilon > 0$ , as  $n \rightarrow \infty$ . Expanding the powers of  $(1-t/(2n))$  occurring in (1) yields

$$2^{t-2n} \binom{2n-t}{n-t} = \frac{1}{\sqrt{n\pi}} \left[ 1 + \frac{t}{4n} + \frac{1}{8n} \left( -1 + \frac{7t^2}{4n} - \frac{t^4}{12n^2} \right) \right]$$

$$\times \exp \left\{ -\frac{t^2}{2(2n-t)} \right\} + O(n^{-2}),$$

where  $t = O(n^{1/2+\epsilon})$ ,  $\epsilon$  sufficiently small. Hence,

$$2^t \binom{2n-t}{n-t} \binom{2n}{n}^{-1} = \left[ 1 + \frac{t}{4n} + \frac{1}{32n} \left( \frac{7t^2}{n} - \frac{t^4}{3n^2} \right) \right] \exp \left\{ -\frac{t^2}{2(2n-t)} \right\} + O(n^{-3/2}).$$

Since

$$\exp \{-t^2/(2(2n-t))\} = \left[ 1 - \frac{t}{8n} + \frac{1}{16n} \left( \frac{-t^4}{n^2} + \frac{t^6}{8n^3} \right) \right]$$

$$\times \exp \left\{ -\frac{t^2}{4n} \right\} + O(n^{-\frac{3}{2}}),$$

one obtains

$$(2) \quad 2^t \binom{2n-t}{n-t} \binom{2n}{n}^{-1} = \left[ 1 + \frac{t}{4n} \left( 1 - \frac{t^2}{2n} \right) + \frac{1}{16n} \left( \frac{7t^2}{2n} - \frac{5t^4}{3n^2} + \frac{1t^6}{8n^3} \right) \right]$$

$$\times \exp \left\{ -\frac{t^2}{4n} \right\} + O(n^{-\frac{3}{2}}),$$

where again,  $t = O(n^{1/2+\epsilon})$ .

It will be convenient to write

$$E\{T^s\} = 1 + S_s.$$

Expanding  $(t+1)^s - t^s$  according to the binomial theorem one gets

$$S_s = \sum_{0 \leq \lambda \leq s-1} \binom{s}{\lambda} \sum_{1 \leq t \leq n} t^\lambda 2^t \binom{2n-t}{n-t} \binom{2n}{n}^{-1}.$$

The last expression suggests the use of the functions

$$(3) \quad r_\lambda(n) = \sum_{t \geq 1} t^\lambda \exp \left\{ -\frac{t^2}{n} \right\} \quad \lambda = 0, 1, \dots,$$

which were introduced in Katzenbeisser and Panny (1984), p. 268. By means of these functions  $S_s$  can be rewritten as

$$(4) \quad S_s = \sum_{0 \leq \lambda \leq s-1} \binom{s}{\lambda} \left[ r_\lambda + \frac{1}{4n} \left( r_{\lambda+1} - \frac{1}{2n} r_{\lambda+3} \right) + \frac{1}{16n^2} \left( \frac{7}{2} r_{\lambda+2} - \frac{5}{3n} r_{\lambda+4} + \frac{1}{8n^2} r_{\lambda+6} \right) \right] + O(n^{-(s-3)/2}),$$

as  $n \rightarrow \infty$ . For the sake of brevity the arguments  $4n$  of the  $r_\lambda$ -functions are omitted in (4). To prove the error term a ‘look ahead’ argument is used. Following the method applied in the derivation of (2) the accuracy of the approximation could be improved by using more terms in the underlying expansion. Proceeding this way the  $O(n^{-3/2})$ -term in (2) can be refined to

$$(5) \quad \frac{1}{n^2} \left( \frac{-t}{16} + \frac{25t^3}{128n} - \frac{71t^5}{768n^2} + \frac{17t^7}{1536n^3} - \frac{t^9}{3072n^4} \right) \exp \left\{ -\frac{t^2}{4n} \right\} + O(n^{-2}).$$

Apart from the  $O$ -term this produces an error of  $O(n^{-(s-3)/2})$ , since the order of magnitude of the  $r_\lambda(n)$ -functions is  $O(n^{(\lambda+1)/2})$ . This is most conveniently obtained by estimating the right-hand side of (3) by means of the  $\Gamma$ -function.

The contribution of the  $O(n^{-2})$ -term is only  $O(n^{-(s-4)/2+\epsilon})$ . This is easily seen by recalling that for  $t > n^{1/2+\epsilon}$  the error in (2) actually becomes exponentially small. Consequently, its contribution may be neglected for  $\epsilon$  sufficiently small. The error produced by extending the range of summation to infinity by use of the  $r_\lambda$ -functions is only  $O(n^{-m})$  for all  $m \geq 0$  and covered by the given  $O$ -term.

In Katzenbeisser and Panny (1984) asymptotic series for the  $r_\lambda$ -functions,  $n \rightarrow \infty$ , are given (cf. Lemma 2, p. 269). They were derived by means of the so called  $\Gamma$ -function approach (cf. de Bruijn et al. (1972)). Making use of them we get the desired result.

For the variance ( $\sigma^2$ ) and skewness ( $\gamma$ ) of  $T$  one obtains, for example,

$$\sigma^2 = (4 - \pi)n - \sqrt{\pi}n^{1/2} + \frac{1}{4}(8 - \pi) + O(n^{-1/2}),$$

$$\gamma = \frac{2\pi^{1/2}(\pi - 3)}{(4 - \pi)^{3/2}} - \frac{3(16 - 5\pi)}{(4 - \pi)^{3/2}} n^{-1/2} + \frac{\pi^{1/2}(19\pi^2 - 65\pi + 16)}{4(4 - \pi)^{7/2}} n^{-1} + O(n^{-3/2}).$$

By the same method the theorem could be extended to achieve an error term as small as we please. The following table summarizes computational results, where exact (ex.) and approximate (app.) values are compared.

$n$	$\mu$		$\sigma^2$		$\gamma$	
	ex.	app.	ex.	app.	ex.	app.
10	5.6755	5.6751	4.11	4.19	0.1683	0.1738
20	7.9763	7.9762	10.40	10.46	0.3167	0.3184
30	9.7487	9.7486	17.21	17.26	0.3788	0.3797
40	11.2451	11.2450	24.30	24.34	0.4148	0.4153
50	12.5645	12.5645	31.57	31.60	0.4389	0.4393
100	17.7467	17.7467	69.31	69.33	0.4975	0.4976

To conclude, we should like to mention further that the distribution of  $T/(2\sqrt{n})$  tends to a standard Weibull distribution with parameter 2, as  $n \rightarrow \infty$  (Johnson and Kotz (1970), p. 250). This is easily seen from Equation (2).

### References

DE BRUIJN, N. G., KNUTH, D. E. AND RICE, S. O. (1972) The average height of planted plane trees. In *Graph Theory and Computing*, ed. R. C. Read, Academic Press, New York, 15–22.

DWASS, M. (1967) Simple random walk and rank order statistics. *Ann. Math. Statist.* **38**, 1042–1053.

FELLER, W. (1968) *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd edn. Wiley, New York.

JOHNSON, N. L. AND KOTZ, S. (1970) *Continuous Univariate Distributions-1*. Houghton Mifflin, Boston.

KATZENBEISSER, W. AND HACKL, P. (1985) An alternative to the Kolmogorov-Smirnov two-sample test. *Commun. Statist. Theory Methods*. To appear.

KATZENBEISSER, W. AND PANNY, W. (1984) Asymptotic results on the maximal deviation of simple random walks. *Stoch. Proc. Appl.* **18**, 263–275.

MCGILCHRIST, C. A. AND WOODYER, K. D. (1975) Note on a distribution-free CUSUM technique. *Technometrics* **17**, 321–325.