Sankhyā : The Indian Journal of Statistics
1978, Volume 40, Series A, Pt. 2, pp. 199–207.

A CHARACTERIZATION OF THE ARC SINE LAW*

By R. SHANTARAM
Indian Statistical Institute

SUMMARY. It is shown that if X and Y are non-discrete identically distributed independent random variables and $X+Y$ has the same distribution as XY then X follows the arc sine law.

0. Introduction

Consider the arc sine density given by ($c \neq 0$)

$$f(x) = \pi^{-1}(4c^2-x^2)^{-\frac{1}{2}}, \; |x| < 2|c|$$
$$= 0 \quad , \qquad \text{else} \qquad \ldots \quad (0.1)$$

whose odd order moments are zero and even ($2n$-th) order moment is given by $\binom{2n}{n} c^{2n}$, $n = 0, 1, 2, \ldots$. If X and Y are independent identically distributed (i.i.d.) random variables (r.v.'s) distributed according to (0. 1) then it is known (Norton, 1975) that

$$c(X+Y) \sim XY \qquad \ldots \quad (0.2)$$

where $\sim$ stands for "has the same distribution as". We say that a distribution satisfies (0.2) if i.i.d. r.v.'s X, Y following this distribution satisfy (0.2).

In attempting to characterize distributions satisfying (0.2), Norton (1978) is led to the following conjecture :

> the arc sine is the only non-discrete distribution (having all moments) satisfying (0.2).

In this note we prove this conjecture and incidentally obtain some interesting determinantal identities. Section 1 contains certain preliminary results that are needed and the conjecture is proved in Section 2.

* This work was done while the author was on sabbatical leave from the University of Michigan, Flint, USA.

1. Notation and some lemmas

Let $m(2j) = \binom{2j}{j}$, $m(2j-1) = 0$ for $j = 1, 2, \ldots$ and $m(0) = 1$. Let $D(0) = S(0) = 1$ and for $n = 1, 2, \ldots$ define the matrices

$$D(n) = ||m(i+j)||, \qquad i, j = 0, 1, 2, \ldots, n$$

$$T(n) = ||m(2i+2j-2)||, \qquad i, j = 1, 2, \ldots, n$$

$$S(n) = ||m(2i+2j)||, \qquad i, j = 0, 1, 2, \ldots, n.$$

Further, for $n = 0, 1, 2, \ldots,$ consider the partitioned matrix

$$D(n+1) = \begin{Vmatrix} D(n) & B(n) \\ B'(n) & m(2n+2) \end{Vmatrix}$$

where $B'(n)$ is the transpose of the $(n+1)$ by 1 column vector $B(n)$. Note that the last entry in $B(n)$ is always zero and write $B(n) = \begin{Vmatrix} B^*(n) \\ 0 \end{Vmatrix}$ where $B^*(n)$ is n by 1.

We now prove several lemmas needed in Section 2.

Lemma 1.1 : *If $j \geqslant i \geqslant 1$ and $j = i+r$ then*

(a) $$\sum_{k=0}^{2i-1} \binom{2i-1}{k} \binom{2j-1}{k+r} = m(2i+2j-2)$$

(b) $$\sum_{k=0}^{2i} \binom{2i}{k} \binom{2j}{k+r} = m(2i+2j).$$

Proof: The easy proof is omitted.

Lemma 1.2 : *det $T(n) = 2^n =$ det $S(n)$.*

Proof: Define the $2n$-dimensional row vectors $x^1, x^2, \ldots, x^n$ as follows:

$$x^i = (0_{n-i} \vdots A_{2i-1} \vdots 0_{n-i})$$

where 0_{n-i} is a $(n-i)$-dimensional row vector of zeroes and A_{k-1} is a k-dimensional vector with components $\binom{k-1}{j}$, $j = 0, 1, 2, \ldots, k-1$, in that order. Using Lemma 1.1 it can be seen that $T(n)$ is a matrix whose (i, j) element is the inner product of x^i and x^j. The determinant of such a matrix can be evaluated as follows (Bellman, 1975, p. 49)

$$\det T(n) = \sum_{(i)} \begin{vmatrix} x^1_{i_1} & x^1_{i_2} & \cdots & x^1_{i_n} \\ x^2_{i_1} & x^2_{i_2} & \cdots & x^2_{i_n} \\ \cdots & \cdots & \cdots & \cdots \\ x^n_{i_1} & x^n_{i_2} & \cdots & x^n_{i_n} \end{vmatrix}^2 \quad \ldots \quad (1.1)$$

$$= \sum_{(i)} \Delta(i)^2, \text{ say}$$

where the sum is over all indices $1 \leqslant i_1 \leqslant i_2 \leqslant \ldots \leqslant i_n \leqslant 2n$ and x^i_j is the j-th coordinate in x^i.

Before we evaluate the right side in (1.1) we note that if $M = \left\| \begin{matrix} x^1 \\ x^2 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{matrix} \right\|$ is the n by $2n$ matrix, its j-th and $(2n+1-j)$-th columns are identical, $j = 1, 2, \ldots, n$. Hence $\Delta(i) = 0$ whenever two indices are either equal or add up to $2n+1$. Further the right side in (1.1) depends only on the absolute values of $\Delta(i)$. Consider only the case $\Delta(i) \neq 0$.

Without loss of generality let

$$1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n,\ n+1 \leqslant i_{k+1} < \ldots < i_n \leqslant 2n \ (k = 0, 1, 2, \ldots, n).$$

In view of the above observations, none of the i_j's $(j \geqslant k+1)$ can be any of the k numbers $2n+1-i_l (l \leqslant k)$. Thus, given the first k indices among $1, 2, \ldots, n$ the remaining $n-k$ indices among $n+1, n+2, \ldots, 2n$ are *uniquely* determined and these latter indices may then be replaced by their differences from $2n+1$ since the corresponding columns in M are identical. Hence, whatever k, $(k = 0, 1, 2, \ldots, n)$ and whatever $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$, $\Delta(i)$ is the determinant of the first n columns of M, with (perhaps) the columns permuted. Since we need only $\Delta(i)^2$ and for $i_l = l \ (l = 1, 2, \ldots, n)$ its value is 1 we have from (1.1)

$$\det T(n) = \Sigma\, 1$$

where the sum is over all $k = 0, 1, 2, \ldots, n$ and all choices $1 \leqslant i_1 < i_2 \ldots < i_k \leqslant n$. That is

$$\det T(n) = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

We only sketch the proof of the other determinantal value. Define $(n+1)$ row vectors $x^0, x^1, \ldots, x^n$ each $(2n+1)$ dimensional as follows

$$x^i = (0_{n-i} \vdots A_{2i} \vdots 0_{n-i})$$

In this case the only non-zero entry $(=1)$ in x^0 is in the $(n+1)$st column and hence only those indices with some $i_k = n+1$, $(k = 0, 1, \ldots, n)$ and no two indices adding to $2n+2$ contribute a non-zero $(=1)$ value to the sum which gives $\det S(n)$. Hence

$$\det S(n) = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Lemma 1.3 : *The cofactors of the zero entries in $D(n)$ are all zero.*

Proof: We will only consider the case n odd $(= 2k-1)$ since the other case is essentially the same. Further we will only consider the cofactors of the zero entries (a) in row 1 and (b) in row 2 since the case of the other rows is seen to fall in one of these cases.

Case (a) : Let $a_1, a_2, \ldots, a_k$ be the cofactors of the zero entries in row 1 read from left to right. Since the inner product of the vector of cofactors of row 1 with the vector of elements of row $2j$, $j \geqslant 1$, are all zero, we get the following system of equations satisfied by the a's

$$a_1 m(2) + a_2 m(4) + \ldots = 0$$

$$a_1 m(4) + a_2 m(6) + \ldots = 0$$

and so on. That is $T(k)\ a = 0$ where $a = \left\| \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{matrix} \right\|$ and $T(k)$ is as defined in Section 1. Since $T(k)$ is non-singular the proof in case (a) is completed.

Case (b) : In this case the role of $T(k)$ is taken by $S(k)$ which is also non-singular. The lemma is completely proved.

Lemma 1.4 : $\quad det\ D(n) = 2^n.$

Proof : It suffices to consider the case n even since the other case is similar. Let $n = 2k$. It is easy to see that a sequence of interchanges of columns and then a sequence (the same number—by symmetry of $D(n)$) of interchanges of rows transforms $D(n)$ into the form

$$\begin{Vmatrix} S(k) & 0 \\ 0 & T(k) \end{Vmatrix}$$

whose determinant is det $S(k)$. det $T(k) = 2^{2k} = 2^n$ by Lemma 1.2.

Lemma 1.5 : $B^{*\prime}(n)D^{-1}(n-1)\,B\,(n-1) = 0, \quad n = 1, 2, \ldots$

Proof : Set $B^{*\prime}(n) = (b_0^*, b_1^*, \ldots, b_{n-1}^*)$ and $B'(n-1) = (b_0, b_1, \ldots, b_{n-1})$ and $D^{-1}(n-1) = \| d^{ij} \|$, $i, j = 0, 1, \ldots, n-1$. Note that either $b_0 = b_2 = \ldots = b_1^* = b_3^* = \ldots = 0$ or $b_0^* = b_2^* = \ldots = b_1 = b_3 = \ldots = 0$. Further, by Lemma 1.3, $d^{ij} = 0$ for $i+j =$ odd. Denoting the left side of the equality in the statement of the lemma by c, we have

$$c = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} b_i^* b_j\, d^{ij}.$$

The typical term in this sum is zero if $i+j =$ odd since then $d^{ij} = 0$. If $i+j =$ even, then either i and j are both even or they are both odd and in either case $b_i^* b_j = 0$ and thus $c = 0$.

Lemma 1.6 : *Let A and D be square symmetric matrices not necessarily of the same order. Then*

$$\text{(i)} \qquad det \begin{Vmatrix} A & B \\ B' & D \end{Vmatrix}$$
$$= det\ A.\ \ det \,\| D - B'A^{-1}B \|$$
$$= det\ D.\ \ det \,\| A - BD^{-1}B' \|$$

provided the indicated inverses exist.

$$\text{(ii)} \qquad \begin{Vmatrix} A & B \\ B' & D \end{Vmatrix}^{-1} = \begin{Vmatrix} A^{-1} + FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{Vmatrix}$$

where $E = D - B'A^{-1}B$, $F = A^{-1}B$ and the indicated inverses exist.

Proof: The reader is referred to Rao (1973, pp. 32-33).

2. Proof of the conjecture

Before stating Theorem 2.1 we need one more matrix. Using the notation introduced in Section 1, let

$$A(n+1) = \begin{Vmatrix} D(n) & G(n) \\ G'(n) & m(2n+2) \end{Vmatrix}$$

for $n = 0, 1, 2, \ldots$ where $G(n) = \begin{Vmatrix} B^*(n) \\ 2 \end{Vmatrix}$. Note that $G(n) - B(n) = 2e(n+1)$ where $e(k)$ is a k-dimensional column vector whose k-th coordinate is 1 and the others are zero.

Theorem 2.1: *$\det A(n) = 0$ for every $n = 2, 3, \ldots$*

Proof: As a start, clearly

$$\det A(2) = \det \begin{Vmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 0 \end{Vmatrix} = 0.$$

By Lemma 1.6, part (i), we have, for $n \geqslant 2$

$$\det A(n) = \det \begin{Vmatrix} D(n-1) & G(n-1) \\ G'(n-1) & m(2n) \end{Vmatrix}$$

$$= \det D(n-1) . \det \| m(2n) - G'(n-1) D^{-1}(n-1) G(n-1) \|$$

$$= 2^{n-1}(m(2n) - G'(n-1) D^{-1}(n-1) G(n-1)).$$

Thus the theorem will be proved if we show that, for $n \geqslant 1$,

$$m(2n+2) = G'(n) D^{-1}(n) G(n). \qquad \ldots \quad (2.1)$$

Consider the right side in (2.1), replace $G(n)$ by $B(n) + 2e(n+1)$ and, using Lemma 1.6, part (ii), replace

$$D^{-1}(n) \text{ by } a. \begin{Vmatrix} P & Q \\ Q' & 1 \end{Vmatrix}$$

where

$$P = a^{-1}D^{-1}(n-1)+D^{-1}(n-1)B(n-1)B'(n-1)D^{-1}(n-1)$$

$$Q = -D^{-1}(n-1)B(n-1)$$

$$a^{-1} = m(2n)-B'(n-1)D^{-1}(n-1)B(n-1) \qquad \ldots \quad (2.2)$$

assuming, for the moment, that the right side is non-zero and thus defines a non-zero a. Multiplying out the resulting quantity we see that (2.1) is equivalent to

$$a^{-1}m(2n+2) = B^{*\prime}(n)PB^{*}(n)+4+4Q'B^{*}(n). \qquad \ldots \quad (2.3)$$

Here we have used the fact that $G(n) = \begin{Vmatrix} B^{*}(n) \\ 2 \end{Vmatrix}$. Substituting for P from (2.2) we see that

$$B^{*\prime}(n)P\,B^{*}(n) = a^{-1}B^{*\prime}(n)D^{-1}(n-1)B^{*}(n)+\{B^{*\prime}(n)D^{-1}(n-1)B(n-1)\}^{2}$$
$$= a^{-1}B^{*\prime}(n)D^{-1}(n-1)B^{*}(n)$$

in view of Lemma 1.5. Substituting for Q we see that $Q'B^{*}(n) = 0$, again by Lemma 1.5. Thus (2.3) is equivalent to

$$m(2n+2) = B^{*\prime}(n)D^{-1}(n-1)B^{*}(n)+4a \qquad \ldots \quad (2.4)$$

We proceed to show (2.4).

Note that $m(2n+2)$ and $D^{-1}(n-1)$ can be related in the following manner. Now

$$D(n+1) = \left\Vert \begin{array}{c:cc} D(n-1) & B(n-1) & B^{*}(n) \\ \hdashline B'(n-1) & m(2n) & 0 \\ B^{*\prime}(n) & 0 & m(2n+2) \end{array} \right\Vert \qquad \ldots \quad (2.5)$$

Using Lemma 1.6 part (i) for the indicated partition, we see that

$$\det D(n+1) = \det D(n-1) \det \left\| \begin{pmatrix} m(2n) & 0 \\ 0 & m(2n+2) \end{pmatrix} - \begin{pmatrix} B'(n-1) \\ B^{*\prime}(n) \end{pmatrix} D^{-1}(n-1)(B(n-1),\ B^{*}(n)) \right\|$$

Recalling that the D matrix, and so its inverse, is symmetric we have by Lemma 1.5, $B'(n-1)D^{-1}(n-1)\, B^{*}(n) = 0$ and the above reduces to

$$\frac{\det D(n+1)}{\det D(n-1)} = \det \left\| \begin{pmatrix} m(2n) & 0 \\ 0 & m(2n+2) \end{pmatrix} - \begin{pmatrix} B'(n-1)D^{-1}(n-1)B(n-1) & 0 \\ 0 & B^{*\prime}(n)D^{-1}(n-1)B^{*}(n) \end{pmatrix} \right\|$$

$$= \det \| m(2n)-B'(n-1)D^{-1}(n-1)B(n-1) \|. \det \| m(2n+2)-B^{*\prime}(n)D^{-1}(n-1)B^{*}(n) \|$$

By Lemma 1.4, $\det D(n) = 2^{n}$ and hence (2.4) is equivalent to showing

$$4 = [m(2n)-B'(n-1)D^{1}(n-1)B(n-1)].(4a) \qquad \ldots \quad (2.6)$$

But this is precisely how a is defined (see (2.2)). The theorem is proved if $a \neq 0$. Since relation (2.6) follows from (2.5), Lemma 1.5 and part (i) of Lemma 1.6 (without appeal to part (ii) of Lemma 1.6) it follows that $a \neq 0$. The theorem is completely proved.

Corollary (Conjecture) : *The arc sine density* (0.1) *is the only non-discrete distribution* (*having moments of all orders*) *that satisfies* (0.2).

Proof : If $c = 1$, the corollary follows from the theorem. Indeed $\det A(n) = 0$ for $n = 2, \ldots$ implies the corollary (see Norton, 1978). If $c \neq 1$, define i.i.d. r.v.'s X', Y' by $X = cX'$, $Y = cY'$. Then (0.2) becomes $X'+Y' \sim X'Y'$ which corresponds to the case $c = 1$.

3. Conclusion

This author has also obtained partial results pertaining to another conjecture of Norton's on the construction of finite random variables satisfying (0.2). These will be published at a later date.

Acknowledgment

The author is grateful to Professor B. Ramachandran for his critical reading of this manuscript and for his interest in the problem. The author wishes to thank the Indian Statistical Institute, New Delhi for the research facility provided and to the University of Michigan, Flint for the sabbatical support.

References

Bellman, R. (1975): *Introduction to Matrix Analysis*, McGraw Hill.

Norton, R. M. (1975): On properties of the arc sine law. *Sankhyā*, Series A, 37, 306-8.

——— (1978): Moment properties and the arc sine law. *Sankhyā*, Series A, 40, 192-198.

Rao, C. R. (1973): *Linear Statistical Inference and its Applications*, 2nd ed., Wiley.

Paper received : January, 1978.