JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 9, 468-476 (1964)

Fluctuation and Periodicity*

S. SHERMAN

Wayne State University, Detroit 2, Michigan

Submitted by Richard Bellman

I. INTRODUCTION

In the theory of fluctuations of sums of independent, identically distributed, random variables X_1, X_2, X_3, $\cdots$ a central result is Spitzer's formula [1, p. 330] which says

$$\sum_{n=0}^{\infty} \varphi_n(\lambda) \; {}^{n} = \exp \sum_{k=1}^{\infty} \psi_k(\lambda)\, t^k/k \tag{1}$$

for λ real and $|t| < 1$, where

$$S_0 = 0, \qquad S_k = \sum_{j=1}^{k} X_j, \qquad 1 \leqslant k,$$

$$\varphi_n(\lambda) = \mathscr{E}\,(\exp\,(i\lambda \max\,(S_0, S_1, \cdots, S_n)),$$

$$\psi_k(\lambda) = \mathscr{E}\,(\exp\,(i\lambda \max\,(S_0, S_k)).$$

E. S. Andersen [2-4] has studied a measure of the fluctuation of the partial sums other than max $(S_0, S_1, \cdots, S_n)$, his measure being H_n, the number of sides of the least concave majorant (this being a trivial change from [4]) of the graph of $(0, S_0)$, $(1, S_1)$, $\cdots$, (n, S_n), where a vertex of a side is a point (k, S_k), $0 \leqslant k \leqslant n$, which is on the least concave majorant of the aforementioned graph. When the common distribution function is continuous he finds that for $|s| < 1$, $|t| < 1$,

$$H(t, s) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P(H_n = m)\, t^n s^m = (1 - t)^{-s} \tag{2}$$

and so distribution free.

There is another natural definition of the "number of sides" used by Andersen [4, pp. 2-3]: "*Remark on the definition of H_n*.

If we plot in a coordinateplane the points (k, S_k) for $k = 0, 1, \cdots, n$ and

* Research partially supported by National Science Foundation, Grant No. G24334.

connect the consecutive points with straight segments, then we obtain an open polygon from (0, 0) to (n, S_n). To this polygon there exists a unique, *least*, *upper concave* polygon from (0, 0) to (n, S_n). The point on this *concave* polygon with abscissa k has ordinate $T_k^{(n)}$. The definition of H_n given in 2 defines H_n as the number of points (k, S_k) for $k = 1, \cdots, n$ on the *concave* polygon (see Fig. 1). If the common distribution of the random variables

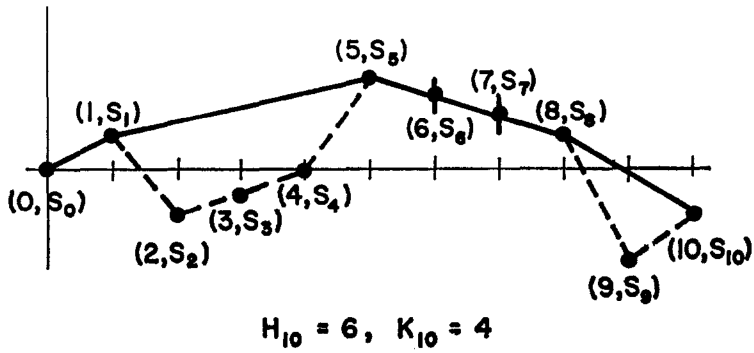


FIG. 1. "Number of sides"

$X_1, X_2, \cdots$ is continuous then H_n is equal to the number of straight segments which form the concave polygon. This indicates that a natural alternative definition of H_n would have been to define H_n to be the number of straight segments in the *concave* polygon, we may denote this number by K_n. Evidently $H_n \geqslant K_n$."

(The italicized words are changed from the Andersen text in order to be consistent with the departure from the Andersen definition. There is a corresponding change in the figure.) As indicated, when the X's have the same continuous distribution, $H_n = K_n$ with probability 1, so one has a corresponding formula for the generating functions. When the X's have the same discrete distribution it is no longer true that $H_n = K_n$ with probability 1 and Andersen [4] has developed somewhat complicated analogues to (2) for both the H_n and K_n cases.

Here another alternative is explored. A random variable J_n, the "number of sides," is defined in Section II so that $H_n \geqslant J_n \geqslant K_n$ and a random variable π_n, the number of periods, is defined so that for a particular sample point

$$
\begin{aligned}
&(X_1, X_2, \cdots, X_n)\\
&= (\underbrace{X_1, X_2, \cdots X_{n/\pi_n}}_{1\ \text{period}}, X_1, X_2, \cdots, X_{n/\pi_n}, \cdots, X_1, \cdots X_{n/\pi_n})\\
&\longleftarrow\qquad\qquad \pi_n \text{ periods} \qquad\qquad\longrightarrow
\end{aligned}
$$

and $X_1, X_2, \cdots, X_{n/\pi_n}$ is aperiodic.

Let

$$\Phi_n(\lambda) = \mathscr{E}(\exp(i\lambda J_n)) \quad \text{and} \quad \Psi_n(\lambda) = \mathscr{E}(\exp(i\lambda\pi_n).$$

Then for λ real and $|t| < 1$

$$\sum_{n=0}^{\infty} \Phi_n(\lambda)\, t^n = \exp \sum_{k=1}^{\infty} (\Psi_k(\lambda)\, t^k)/k, \tag{3}$$

i.e., the "Spitzer formula" holds this time relating the characteristic function of the "number of sides" and the characteristic function of the "number of periods." A similar method could have been used to deduce the original Spitzer formula. Some connections are deduced between the distribution of J_n in the case of a classical random walk on the integers and the dimension of certain subspaces of a free Lie algebra on two generators. The author wishes to acknowledge helpful discussions with Glen Baxter and Marcel Schützenberger.

II. The Noncommutative Witt Identity

Let $a_1 < a_2 < \cdots$ constitute a countable totally ordered set A which generates a free semigroup (= free monoid [5, p. 18]) F so that elements $f \in F$ are words generated from the a_i's and the binary associative operation is given by juxtaposition. Suppose F is lexicographically ordered so that for example $a_1^2 a_2 < a_1 a_2 < a_2 a_1$. A word is a standard word if it is aperiodic and lexicographically less than each of its cyclic permutations. The collection of all standard words [6, p. 83] is denoted by H. Note $a_1^2 a_2 \in H$ since $a_1^2 a_2 < a_1 a_2 a_1$ and $a_1^2 a_2 < a_2 a_1^2$. Further $(a_1 a_2)^2 \notin H$ since $(a_1 a_2)^2$ is periodic. Of course $H \subset F$ implies H is totally ordered. By $\Pi\{1 - h : h \in H; <\}$ is meant a formal infinite product of binomial terms in the real algebra generated by F, the binomial terms taken in increasing lexicographic order in the h's. The noncommutative Witt identity [7, Lemma 3] is

$$\Pi\{1 - h : h \in H; <\} = 1 - \sum a_i\,. \tag{4}$$

If one considers the obvious homomorphism which sends the generators of the free semigroup F into the generators of a free abelian semigroup he gets the classical Witt identity [8, pp. 155-156; 9, pp. 169-170]. If one takes inverses of both sides, [7, just prior to Lemma 3]

$$\Pi\{(1 - h)^{-1} : h \in H; >\} = \sum \{f : f \in F\},$$

where

$$(1 - h)^{-1} = 1 + h + h^2 + \cdots.$$

This says that each nonunit $f \in F$ has a unique monotone decreasing factorization into elements of H,

$$f = h_1 h_2 \cdots h_k ,$$
$$h_1 \geqslant h_2 \geqslant \cdots \geqslant h_k ,$$
$$h_i \in H, \qquad 1 \leqslant i \leqslant k.$$

Whenever one has a free semigroup F and a subset of F such that each nonunit element of F admits of a unique monotone decreasing factorization into elements of that subset, then one says one has a (generalized) Witt identity. Related considerations are given in [10].

III. Application to Fluctuation of Sums of Independent Random Variables

Let X_1, X_2, $\cdots$ be independent random variables with the same discrete distribution:

$$P(X_j = a_i) \text{ independent of } j,$$

$$\sum_{i=1}^{\infty} P(X_j = a_i) = 1.$$

Corresponding to the sample point ω, the first n components are

$$X_1(\omega) = a_{i_1}, \cdots, X_n(\omega) = a_{i_n} .$$

Let

$$S_0(\omega) = 0, \qquad S_k(\omega) = \sum_{m=1}^{k} a_{i_m}, \qquad 1 \leqslant k.$$

Plot $\{(k, S_k), 0 \leqslant k \leqslant n\}$ and take the least concave majorant (see Fig. 2). Note that

$$f(\omega) = a_{i_1}, \qquad a_{i_2} \cdots a_{i_n} = (-1, 2, -2, 3, -1, 2, -1, 1, 0) \in F$$

in the case of Fig. 2.

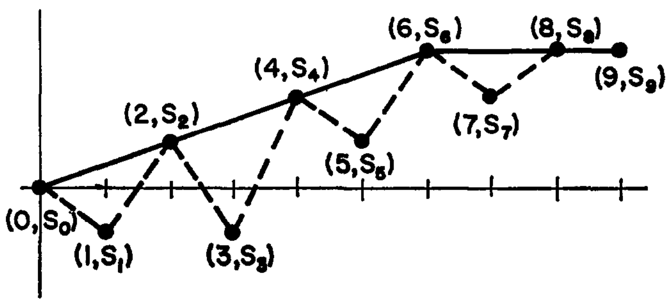
From these words choose in stages a subset Sp to play the role of H. If f is a word let $\mu(f)$, the slope of f, be the sum of the letters divided by $|f|$, the number of letters. A word f is in $A(\mu_1)$, the alphabet associated with slope μ_1, if $\mu(f) = \mu_1$ and $H_{|f|} = 1$. The pairs of successive points of contact of the least concave majorant and the original set $\{(0, S_0), (1, S_1), \cdots, (n, S_n)\}$

give rise to elements of $A(\mu)$ (possibly different μ for different pairs of successive points of contact). Thus in Fig. 2

$$(-1, 2) = a_{i_1}a_{i_2} = a_{i_5}a_{i_6} \in A(\mu_1) = A\left(\tfrac{1}{2}\right);$$

$$(-2, 3) = a_{i_3}a_{i_4} \in A(\mu_1) = A\left(\tfrac{1}{2}\right);$$

$$(-1, 1) = a_{i_7}a_{i_8} \in A(\mu_2) = A(0), \quad \text{and} \quad 0 = a_{i_9} \in A(\mu_2) = A(0).$$



$$(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9) = (-1, 2, -2, 3, -1, 2, -1, 1, 0)$$
$$(-1, 2) = a_{i_1}a_{i_2} = a_{i_5}a_{i_6} < a_{i_3}a_{i_4} = (2, -2),$$
$$0 = a_{i_9} < a_{i_7}a_{i_8} = (-1, 1), \qquad 0 = a_{i_9} < a_{i_1}a_{i_2} = (-1, 2),$$
$$J_9(\omega) = 4, \quad H_9(\omega) = 5, \quad K_9(\omega) = 2, \quad \pi_9(\omega) = 1$$
$$Z_9(\omega) = 1, \quad Z_4(\omega) = 2.$$

FIG. 2. — "Number of sides" in new sense

For each fixed μ totally order $A(\mu)$. Let $Sp(\mu)$ be the set of standard words formed from $A(\mu)$ and let $Sp = \bigcup_\mu Sp(\mu)$. Introduce a total order in Sp as follows: If $\mu(f_1) > \mu(f_2)$, then $f_1 > f_2$. If $f_1, f_2 \in Sp(\mu)$, then order these f lexicographically according to the order in the alphabet $A(\mu)$. For example in Fig. 2

$$(-1, 2) = a_{i_1}a_{i_2} = a_{i_5}a_{i_6} < a_{i_3}a_{i_4} = (-2, 3);$$

$$(-1, 1) = a_{i_7}a_{i_8} > a_{i_9} = 0$$

and of course,

$$(-1, 2) = a_{i_1}a_{i_2} > a_{i_9} = 0.$$

Then

$$(-1, 2, -2, 3) = a_{i_5}a_{i_6}a_{i_3}a_{i_4} \in Sp(\mu_1) = Sp\left(\tfrac{1}{2}\right) \subset Sp;$$

$$(-1, 1) = a_{i_7}a_{i_8} \in A(\mu_2) = A(0) \subset Sp(\mu_2) \subset Sp,$$

and $$(-1, 2, -2, 3) = a_{i_5}a_{i_6}a_{i_3}a_{i_4} > a_{i_7}a_{i_8} = (-1, 1).$$

Now each $f \in F$ has a unique monotone factorization (empty factorization in the case of the unit) in terms of the words in Sp.

$$f = h_{i_1} h_{i_2} \cdots h_{i_k}$$
$$h_{i_j} \in Sp, \qquad 1 \leqslant j \leqslant k$$
$$h_{i_1} \geqslant h_{i_2} \geqslant \cdots \geqslant h_{i_k}.$$

Let $J_n(\omega) = k$, the "number of sides" of the graph of the least concave majorant. For example in Fig. 2

$$f = (a_{i_1}a_{i_2}a_{i_3}a_{i_4})\,(a_{i_5}a_{i_6})\,(a_{i_7}a_{i_8})\,(a_{i_9}) = (-1, 2, -2, 3)\,(-1, 2)\,(-1, 1)\,(0),$$
$$J_9(\omega) = 4, \qquad H_9(\omega) = 5, \qquad K_9(\omega) = 2, \qquad \pi_9(\omega) = 1.$$

If f is not the identity let $\pi[f]$ be the largest integer r such that for some $f_1 \in F$ one has $f = (f_1)^r$. If some sample point ω gives rise to n-letter $f(\omega) \in F$, let $\pi_n = \pi[f(\omega)]$. Let

$$\Phi_n(\lambda) =_{df} \mathscr{E}\,(\exp\,(i\lambda J_n)) \qquad \text{and} \qquad \Psi_n(\lambda) =_{df} \mathscr{E}\,(\exp\,(i\lambda \pi_n)).$$

THEOREM. *If X_1, X_2, $\cdots$ are independent random variables with the same discrete distribution, then for λ real and $|\,t\,| < 1$,*

$$\sum_{n=0}^{\infty} \Phi_n(\lambda)\, t^n = \exp \sum_{k=1}^{\infty} \Psi_k(\lambda)\, t^k/k. \tag{5}$$

PROOF. From the definition of Sp

$$\pi\{(1-h)^{-1} : h \in Sp; >\} = \sum \{f : f \in F\}.$$

Suppose one associates with each $h \in Sp$, $\hat{h} =_{df} P(h)\, t^{|h|} \exp\,(i\lambda)$ with $|\,h\,| =_{df}$ the number of letters in h. If

$$f = h_{i_1} h_{i_2} \cdots h_{i_s}$$

with

$$h_{i_1} \geqslant h_{i_2} \geqslant \cdots \geqslant h_{i_s} \qquad \text{and} \qquad h_{i_j} \in Sp, \qquad 1 \leqslant j \leqslant s,$$

then let

$$\hat{f} = \hat{h}_{i_1} \hat{h}_{i_2} \cdots \hat{h}_{i_s} = P(f) \exp\,(is\lambda)\, t^{|f|},$$

where the independence of the random variables X_1, X_2, $\cdots$ is used. (Only the independence of h_{i_1}, h_{i_2}, $\cdots$ at this stage.)

Then

$$\sum_{k=0}^{\infty} \Phi_k(\lambda)\, t^k = \sum \{\hat{f} : f \in F\}$$

$$= \pi\{(1 - \hat{h})^{-1} : h \in Sp; >\}$$

$$= \exp \sum_{h \in Sp} \sum_{k=1}^{\infty} P(h^k) \exp(i\lambda k)\, t^{k|h|}/k$$

$$= \exp \sum_{h \in Sp} \sum_{k=1}^{\infty} |\, h \,|\, P(h^k) \exp(i\lambda k)\, t^{k|h|}/k \,|\, h \,|$$

$$= \exp \sum_{f \in F} P(f) \exp(i\lambda\pi[f])\, t^{|f|}/|\, f \,|$$

$$= \exp \sum_{k=1}^{\infty} \Psi_k(\lambda)\, t^k/k,$$

where in next to the last equality one used (1) the fact that each aperiodic word has a unique cyclic permutation making it a standard word and (2) if one goes from f to f' by cyclically permuting the letters, then $P(f) = P(f')$. Thus the theorem is proved.

Consider the classical random walk on the integers with only steps of ± 1 permitted, both being equally likely. What can be said about $P(J_n = k)$?

Suppose with [9, p. 169, 171] one lets $M_2(d)$ be the number of circular words composed from two different letters but of length d and period d (= the number of basic commutators of weight d in 2 generators). Then

$$P(\pi_r = k) = 2^{-r} \frac{r}{k} M_2\left(\frac{r}{k}\right) = 2^{-r} \frac{r}{k} \cdot \frac{k}{r} \sum_{dk/r} \mu\left(\frac{r}{dk}\right) 2^d$$

$$\Psi_r(\lambda) = \sum_{k=1}^{r} \exp(i\lambda k)\, 2^{-r} \sum_{dk/r} \mu\left(\frac{r}{dk}\right) 2^d$$

$$\sum_{n=0}^{\infty} \Phi_n(\lambda)\, t^n = \exp \sum_{r=1}^{\infty} \left[\frac{(t/2)^r \sum_{k=1}^{r} \exp(i\lambda k) \sum_{dk/r} \mu\left(\frac{r}{dk}\right) 2^d}{r}\right].$$

The continuous analogue to Eq. (5) is Eq. (2), since for the case of a continuous distribution $J_n = H_n$ with probability one and $\mathscr{E}(\pi_n) = 1$.

The Theorem gives information about the distribution of J_n. Under the same assumptions what can be said about H_n? Let σ be a cyclic permutation of $(1, 2, \cdots, n)$ such that the chord joining $(0, 0)$ to $(n, X_{\sigma_1} + X_{\sigma_2} + \cdots + X_{\sigma_n})$ is not below any vertex $(k, X_{\sigma_1} + \cdots + X_{\sigma_k})$, $1 \leqslant k \leqslant n$. Let Z_n be number of such vertices on the chord. While σ is not necessarily unique, Z_n is unique. In Fig. 2, $Z_9 = 1$ and $Z_4 = 2$.

Let

$$\bar{\Phi}_n(\lambda) = \mathscr{E}\,(\exp i\lambda H_n)$$

and

$$\bar{\Psi}_n(\lambda) = \mathscr{E}\,(\exp i\lambda Z_n).$$

Then by redefining $\hat{h}$ one can show that for λ real and $|\,t\,| < 1$

$$\sum_{n=0}^{\infty} \bar{\Phi}_n(\lambda)\, t^n = \exp \sum_{k=1}^{\infty} \bar{\Psi}_k(\lambda)\, t^k/k.$$

When applied to distributions which are continuous except for one jump this formula is consistent with [4, p. 3, Eq. (4.2)].

Appropriate redefinition of $\hat{h}$ yields, for discrete distributions of the X's, analogues to Eq. (4.12) of [2] and Eqs. (3.6) and (9.12) of [3] as well as equations involving the joint distribution of random variables of this kind.

A slightly different analysis gives results about the distribution of K_n and can be used to give an independent derivation of the last equation. No derivation will be given here but the result will be stated. Let

$$\tilde{\Phi}_n(\lambda) = \mathscr{E}\,(\exp i\lambda K_n)$$

then

$$\sum_{n=0}^{\infty} \tilde{\Phi}_n(\lambda)\, t^n = \exp \sum_{k=1}^{\infty} \frac{t^k}{k}\, \mathscr{E}[1 - (1 - e^{i\lambda})^{Z_k}]\,.$$

The classical Spitzer formula, Eq. (1), could have been established as was the main theorem by requiring for

$$h = a_{i_1} a_{i_2} \cdots a_{i_k}$$

$$S_0 = 0, \qquad S_k = \sum_{j}^{k} = 1 a_{i_j}$$

that

$$\begin{aligned} \hat{h} &= P(h) \exp\,(i\lambda \max\,(S_0\,, S_1\,, \cdots, S_k))\, t^k \\ &= P(h) \exp\,(i\lambda \max\,(S_0\,, S_{|h|}))\, t^{|h|}. \end{aligned}$$

In establishing the classical Spitzer formula some of the niceties about standard words could be avoided. One could also establish a formula encompassing the current formula and the classical one. If one is to use least concave majorants to get this kind of formula, then the "additivity" of the functional which is the coefficient of λ is one limitation that must be observed. In comparing Spitzer's proof of his formula with the proof suggested just above one notes that the latter does not use the decomposition of a permutation into cycles [1, p. 326 Theorem 2.2].

When one considers the variety of derivations of the classical Spitzer's formula, it becomes of interest to see whether these techniques could be used to derive Eq. (5). In particular a Wiener-Hopf derivation of Eq. (5) would be desirable. This does not seem to be a transparent problem.

References

1. Spitzer, F. A combinatorial lemma and its application to probability theory. *Trans. Am. Math. Soc.* **82**, 323-339 (1956).
2. Andersen, E. Sparre. On the fluctuations of sums of random variables. *Math. Scand.* **1**, 263-285 (1953).
3. Andersen, E. Sparre. On the fluctuations of sums of random variables II. *Math. Scand.* **2**, 195-223 (1954).
4. Andersen, E. Sparre. On the distribution of the random variable H_n. Tech. Sci. Note No. 1, Contract No. AF 61(052)-42, February 27, 1959.
5. Chevalley, C. "Fundamental Concepts of Algebra." Academic Press, New York, 1956.
6. Chen, K. T., Fox, R. H., and Lyndon, R. C. Free differential calculus IV. The quotient groups of the lower central series. *Ann. Math.* **68**, 81-95 (1958).
7. Schützenberger, M. P., and Sherman, S. On a formal product over the conjugate classes of a free group. *Math. Anal. Appl.*, in press.
8. Witt, E. Treue Darstellung Liescher Ringe. *J. Reine Angew. Math.* **177**, 152-160 (1937).
9. Hall, M. "The Theory of Groups." Macmillan, New York, 1959.
10. Schützenberger, M. P. Certain infinite formal products and their combinatorial applications. Colloquium on combinatorial methods in probability theory held Aug. 1-10 at Aarhus, pp. 58-63.