A FLUCTUATION THEOREM AND DISTRIBUTION FREE TEST

by
Charles H. Kraft
and
Constance van Eeden

Techncial Report 37

TR 37

A FLUCTUATION THEOREM AND A DISTRIBUTION-FREE TEST

Charles H. Kraft[1] and Constance van Eeden[2]
University of Minnesota

Let $(X_1,\ldots,X_n)$ be completely independent random variables with continuous and symmetric (about zero) distributions. If

$N^* =$ number of positive sums in $\{X_1, X_1 + X_2, X_1+ X_2+ X_3,\ldots,X_1+\ldots+X_n\}$

it has been shown by Anderson [1], [2] that N^* has, for each n, a distribution which is the same for the given class of sequences $(X_1,\ldots,X_n)$. See Baxter [3] for this and a related result.

We show here that the random variable, N,

$N =$ number of positive sums in $\{\sum_{j\in T} X_j;\quad T\subset(1, 2,\ldots,n)\}$

also has a distribution which is constant for the same class of sequences $(X_1,\ldots,X_n)$. In particular, the distribution of N is given by

$$P(N = k) = \frac{1}{2^n}, \quad k = 0, 1,\ldots, 2^n-1 .$$

Proof: The conditional, given $(|X_1|,|X_2|,\ldots,|X_n|)$, distribution of $(X_1,\ldots,X_n)$ is uniform on the 2^n points $\{(\epsilon_1|X_1|,\epsilon_2|X_2|,\ldots, \epsilon_n|X_n|);\ \epsilon_i = -1, \text{ or } 1\}$. Let $0 = S_o < S_1 <\ldots< S_{2^n-1}$ be the ordered partial sums from $\{|X_1|,\ldots,|X_n|\}$. Let $S_k = \sum_{i\in T} |X_i|$, and, for $j = 1,\ldots, n$, $\delta_j = \begin{cases} -1 & j\in T \\ 1 & j\in' T \end{cases}$.

Then, (see lemma),

$$N = N(\delta_1|X_1|,\delta_2|X_2|,\ldots,\delta_n|X_n|)$$

$$= \text{number of positive sums in } \{S_i - S_k,\ i = 0,\ldots,2^n-1,\ i\neq k\} .$$

---

1) This research was supported in part by the National Science Foundation under Grant Number G-19126.

2) This research was supported in part by the National Institutes of Health under Grant Number 2G-43(C7+8).

Since this last expression is clearly equal to 2^n-1-k and

$$P\{(\delta_1|X_1|,\dots,\delta_n|X_n|)\Big|(|X_1|,\dots,|X_n|)\} = \frac{1}{2^n}$$ the result follows.

Lemma:

$N(\delta_1|X_1|,\dots,\delta_n|X_n|)$ = number of positive sums in $\{S_i-S_k;\ i=0,1,\dots,2^n-1,\ i\neq k\}$.

Proof: A one-to-one correspondence between the partial sums from $(\delta_1|X_1|,\dots,\delta_n|X_n|)$ and the elements of $\{S_i-S_k,\ i=0,1,\dots,2^n-1,\ i\neq k\}$ is given below where $T = T(S_k)$ is defined above.

To $\sum_{j\in A}\delta_j|X_j|$, make correspond $S_{i(A)}-S_k$ where

$$\begin{cases} S_{i(A)} = \sum_{j\in(A-T)\cup(T-A)}|X_j| & A \neq T \\ i(A) = 0 & A = T . \end{cases}$$

To S_i-S_k, make correspond $\sum_{j\in(B-T)\cup(T-B)}\delta_j|X_j|$ where B is defined by $S_i = \sum_{j\in B}|X_j|$.

The number of positive sums in the two sets is the same since the correspondents of the first map are equal, i.e., $\sum_{j\in A}\delta_j|X_j| = S_{i(A)}-S_k$.

<u>Consistency of tests based on N (or N^*).</u>

Let T_n denote one of N^*/n or $N/(2^n-1)$. A test of the null hypothesis, $X_1,\dots,X_n$ are independent and symmetric about zero is to reject H_o if $T_n > k_n$ where k_n is, for large n, approximately the solution to $B_t(\frac{1}{2},\frac{1}{2}) = 1-\alpha$ (B_t is the incomplete beta-function) if $T_n = N^*/n$ and $k_n = 1-\alpha$ if $T_n = N/(2^n-1)$.

Since, for any random variable Z with $P(0 \leq Z \leq 1) = 1$,

$$P(Z \geq 1-\epsilon) \geq 1 - \frac{1-EZ}{\epsilon},$$

a lower bound for the power of these tests is $1 - \frac{1-ET_n}{1-k_n}$. Therefore the tests will be consistent if $ET_n \to 1$. Further,

$$1-ET_n = \sum_{k=0}^{n} P\{\frac{Y_k-k\theta}{\sqrt{k}\ \sigma} < -\frac{\sqrt{k}\ \theta}{\sigma}\}\cdot p_k \tag{1}$$

where

$$p_k = \begin{cases} \frac{1}{n}, & k=1,\ldots,n \qquad \text{for } T_n = N^*/n \\ \binom{n}{k}(\frac{1}{2})^n, & k=0,\ldots,n \qquad \text{for } T_n = N/2^n-1 , \end{cases}$$

and where it is assumed that $(X_1,\ldots,X_n)$ are independently and identically distributed with a distribution $F(\frac{x-\theta}{\sigma})$ for some F, and for $\theta>0$, $\sigma > 0$, and Y_k is a random variable distributed as $\sum_{i=1}^{k} X_i$. From (1) and the Helly-Bray theorem it follows that the test will be consistent if

$$(2) \qquad P\{\frac{Y_k-k\theta}{\sqrt{k}\,\sigma} < -\frac{\sqrt{k}\,\theta}{\sigma}\} \to 0 .$$

If, for some alternative, ET_n does not converge to one there will be a size for which T_n will not be consistent. For example (2) is constant for the Cauchy distribution and T_n is not consistent.

References.

[1] Anderson, E. Sparre (1949). On the number of positive sums of random variables. Skand. Aktuarietidskr. 36 27-36.

[2] Anderson, E. Sparre (1953). On the fluctuations of sums of random variables. Math. Scand. 1 263-285.

[3] Baxter, Glen (1962). On a generalization of the finite arcsine law. Ann. Math. Stat. 33 909-915.