satisfy condition (1) with probability 1, provided Y has positive variance and a finite absolute moment of order 3. Thus condition (1) constitutes a considerable improvement over condition W.

REFERENCES

[1] A. WALD AND J. WOLFOWITZ, "Statistical tests based on permutations of the observations," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 358–372.
[2] H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton, 1946.

ON SUMS OF SYMMETRICALLY TRUNCATED NORMAL RANDOM VARIABLES

BY Z. W. BIRNBAUM AND F. C. ANDREWS[1]

University of Washington, Seattle

1. Introduction. Let X_a be the random variable with the probability density

$$f_a(x) = \begin{cases} Ce^{-x^2/2} & \text{for } |x| \le a \\ 0 & \text{for } |x| > a, \end{cases} \tag{1.1}$$

obtained from the normal probability density $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ by symmetrical truncation at the "terminus" $|x| = a$, and let $S_a^{(m)}$ be the sum of m independent sample-values of X_a. We consider the following *problem*: An integer $m \ge 2$ and the real numbers $A > 0$, $\epsilon > 0$ are given; how does one have to choose the terminus a so that the probability of $|S_a^{(m)}| \ge A$ is equal to ϵ,

$$P(|S_a^{(m)}| \ge A) = \epsilon? \tag{1.2}$$

This problem arises for example when single components of a product are manufactured under statistical quality control, so that each component has the length $Z = k + X$ where X has the probability density $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, and the final product consists of m components so that its total length S is the sum of the lengths of the components. We wish to have probability $1 - \epsilon$ that S differs from mk by not more than a given A. To achieve this we decide to reject each single component for which $|Z - k| = |X| > a$; how do we determine a?

The exact solution of this problem would require laborious computations.[2] In the present paper methods are given for obtaining approximate values of a which are "safe", that is such that

$$P(|S_a^{(m)}| \ge A) \le \epsilon. \tag{1.3}$$

[1] Research done under the sponsorship of the Office of Naval Research.

[2] A similar problem has been studied by V. J. Francis [2] for one-sided truncation; he actually had the exact probabilities for the solution of his problem computed and tabulated for $m = 2, 4$.

In deriving these safe values, use will be made of theorems on random variables with comparable peakedness, for which the reader is referred to a previous paper [1].

2. The safe value a_1. For fixed $a > 0$, we consider the normal random variable Y_a with expectation 0 and with probability density $g_a(Y_a)$ such that $g_a(0) = f_a(0)$. It is easily seen that Y_a has the standard deviation

$$\sigma_a = \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{-t^2/2}\, dt, \tag{2.1}$$

and that $g_a(\xi) < f_a(\xi)$ for $|\xi| \leq a$, $g_a(\xi) > 0 = f_a(\xi)$ for $|\xi| > a$. Hence, applying Theorem 1 in [1], we conclude that

$$P(|S_a^{(m)}| \geq A) \leq \frac{2}{\sqrt{2\pi}} \int_{(A/\sigma_a\sqrt{m})}^{\infty} e^{-t^2/2}\, dt. \tag{2.2}$$

If m, A, and ϵ are given, we determine ξ_ϵ from tables of the normal probability integral so that $\frac{2}{\sqrt{2\pi}} \int_{\xi_\epsilon}^{\infty} e^{-t^2/2}\, dt = \epsilon$, set $\sigma_a = \frac{A}{\xi_\epsilon \sqrt{m}}$ in (2.1), and solve the equation

$$\frac{A}{\xi_\epsilon\sqrt{m}} = \frac{1}{\sqrt{2\pi}} \int_{-a}^{+a} e^{-t^2/2}\, dt \tag{2.3}$$

for a using again tables of the normal probability integral. In view of (2.2) this solution satisfies (1.3) and hence is safe; it will be denoted by a_1.

3. The safe value a_2. A direct application of Theorem 2 in [1] yields the inequality

$$\begin{aligned} &P(|S_a^{(m)}| \geq A) \\ &\leq \frac{1}{2^{m-1}m!} \sum_{\frac{1}{2}(m+A/a)<j\leq m} (-1)^j \binom{m}{j} \left(\frac{A}{a} + m - 2j\right)^m = h_m\left(\frac{A}{a}\right) \end{aligned} \tag{3.1}$$

for $0 \leq A \leq ma$. Hence by equating $h_m(A/a)$ to ϵ and solving for a, we obtain a safe value which will be denoted by a_2. It is of interest to note that (3.1) *is true* not only for $f_a(x)$ defined by (1.1) i.e. truncated normal, but *for any probability density* $f_a(x)$ *which is symmetrical and unimodal,* since these are the only assumptions needed for Theorem 2 in [1].

4. Solution for large m. The random variable X_a has the variance

$$\sigma^2(X_a) = 1 + \frac{2\phi''(a)}{2\phi(a) - 1} \tag{4.1}$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Hence, according to the central limit theorem, we have the approximate equality

$$P(|S_a^{(m)}| \geq A) = \frac{2}{\sqrt{2\pi}} \int_{(A/\sigma(X_a)\sqrt{m})}^{\infty} e^{-t^2/2}\, dt \tag{4.2}$$

for m sufficiently large.

It can be reasonably expected that the cumulative distribution of $S_a^{(m)}$ differs from its limiting normal probability integral by less than the cumulative distribution of the sum $U_a^{(m)}$ of m independent uniform variables in $(-a, +a)$ differs from its limiting normal probability integral. Already for $m = 4$ the cumulative distribution of $U_a^{(m)}$ differs from the corresponding normal cumulative by less than .0075. Equally good or better approximation may, therefore, be expected for the distribution of $S_a^{(m)}$, so that the error in the approximate equality (4.2) between the two-tail probabilities should be less than .015 for $m = 4$, and still less for $m > 4$.

Equating the right-hand term of (4.2) to ϵ and solving for $\sigma^2(X_a)$, we obtain

$$\sigma^2(X_a) = 1 + \frac{2\phi''(a)}{2\phi(a) - 1} = \frac{1}{m}\left(\frac{A}{\xi_\epsilon}\right)^2,$$

an equation which can be solved for a with the aid of tables of $\phi(x)$ and $\phi''(x)$. We denote this value of a by α_1.

5. Use of the different solutions in practice. From the foregoing it appears that the following procedure may be followed in solving our problem in any definite case:

If m is large, α_1 is very close to the exact solution of (1.3) and may be used safely.

If m is not large but $m \geq 5$, it is conjectured that α_1 is such that the left-hand term in (1.3), for $a = \alpha_1$, differs from ϵ by less than 0.015.

If $m \leq 4$, the larger of a_1 and a_2 should be used. Table I contains the A for which a_1 and a_2 have the same value, say a'; a_1 or a_2 should be used if the given A is greater or smaller, respectively, than the tabulated value. The value a_1 is easily computed from a table of the normal probability integral by the procedure of section 2. The value a_2 can be obtained by reading off A/a_2 from Table II.

TABLE I

Values of A for which $a_1 = a_2 = a'$ for given m, ϵ

ϵ \ m	2		3		4	
	A	a'	A	a'	A	a'
.001	4.568	2.357	5.446	2.008	6.152	1.842
.002	4.258	2.228	5.059	1.918	5.717	1.779
.005	3.808	2.047	4.512	1.799	5.111	1.697
.01	3.438	1.910	4.074	1.712	4.632	1.640
.02	3.034	1.765	3.614	1.630	4.131	1.589
.05	2.456	1.581	2.970	1.533	3.425	1.529

TABLE II

Values of A/a_2 for given m, ϵ

ϵ \ m	2	3	4
.001	1.937	2.712	3.339
.002	1.911	2.637	3.213
.005	1.859	2.507	3.011
.01	1.800	2.379	2.824
.02	1.718	2.217	2.600
.05	1.553	1.937	2.240

6. Examples. 1) $A = 3.8$, $m = 4$, $\epsilon = .05$. Since A is greater than the value 3.425 in Table I, we compute $a_1 = 2.162$. From Table II we would obtain $A/a_2 = 2.240$ and thus $a_2 = 1.696 < a_1$. 2) $A = 3, m = 4, \epsilon = .02$. Since $A < 4.131$, we read $A/a_2 = 2.600$ from Table II and obtain $a_2 = 1.153$ which will be greater than a_1. 3) $A = 5$, $m = 30$, $\epsilon = .05$. Using the method of section 4 we obtain $\alpha_1 = 1.62$.

REFERENCES

[1] Z. W. Birnbaum, "On random variables with comparable peakedness," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 76–81.
[2] V. J. Francis, "On the distribution of the sum of n sample values drawn from a truncated normal population," *Roy. Stat. Soc. Jour. Suppl.*, Vol. 8 (1946), pp. 223–232

A CERTAIN CUMULATIVE PROBABILITY FUNCTION

By Sister Mary Agnes Hatke, O.S.F.

St. Francis College, Ft. Wayne, Indiana

Graduations of empirically observed distributions show that the cumulative probability function $F(x) = 1 - (1 + x^{1/c})^{-1/k}$ is a practical tool for fitting a smooth curve to observed data. The graduations are comparable with those obtained by the Pearson system, Charlier, and others and are accomplished with simple calculations. Given distributions are graduated by the method of moments. Theoretical frequencies are obtained by evaluation of consecutive values of $F(x)$ by use of calculating machines and logarithms, and by differencing $NF(x)$. No integration nor heavy interpolation is involved, such as may be required in graduation by a classical frequency function. Burr [1] constructed tables of ν_1, σ, α_3, and α_4 values for the function $F(x)$ for certain combinations of integral values of $1/c$ and $1/k$. In these tables curvilinear interpolation must be used in finding an $F(x)$ with desired moments. The writer constructed more extensive tables for the same cumulative function with c and k a variety of real positive numbers less than or equal to one, such that linear interpolation can be used to determine the parameters c and k for an $F(x)$ that has α_3 and α_4 approximately the same as those of the distribution to be graduated. These tables have been deposited with Brown University. Microfilm or photostat copies may be obtained upon request to the Brown University Library.

The writer used the definitions of cumulative moments and the formulas for the ordinary moments ν_1, σ, α_3, and α_4 in terms of cumulative moments as developed by Burr. These latter moments were tabulated for the function $F(x)$ having various combinations of parameters c and k, c ranging from 0.050 to 0.675 and k from 0.050 to 1.000, each at intervals of 0.025. Within these ranges only those combinations of c and k were used which yielded α_3 of approximately 1 or less and α_4 values of 6 or less, since such moments are most common in practice.

It can be verified that over most of the area of the table α_3 values obtained