These computations were made to only two decimal places, so that the final results may easily err by 1, 2, or 3 in the second decimal place.

A more complete discussion of the problem, the origin of the approximations, and tables showing a representative collection of actual values can be found in Memorandum Report 24 of the Statistical Research Group, Princeton University, which bears the same title as this note. Copies may be obtained from its Secretary, Box 708, Princeton, N. J.

REFERENCES

[1] F. J. Anscombe, "The transformation of Poisson, binomial, and negative binomial data", *Biometrika*, Vol. 35 (1948), pp. 246–254.
[2] M. S. Bartlett, "The square root transformation in the analysis of variance", *Jour. Roy. Stat. Soc., Suppl.*, Vol. 3 (1936), pp. 68–78.
[3] M. S. Bartlett, "The use of transformations", *Biometrics*, Vol. 3 (1947), pp. 39–51.
[4] J. H. Curtiss, "On transformations used in the analysis of variance", *Annals of Math. Stat.*, Vol. 14 (1943), pp. 107–122.
[5] Churchill Eisenhart, "The assumptions underlying the analysis of variance", *Biometrics*, Vol. 3 (1947), pp. 1–21.
[6] Frederick Mosteller and John W. Tukey, "The uses and usefulness of binomial probability paper", *Jour. Am. Stat. Assn.*, Vol. 44 (1949), pp. 174–212.

REMARK ON THE ARTICLE "ON A CLASS OF DISTRIBUTIONS THAT APPROACH THE NORMAL DISTRIBUTION FUNCTION" BY GEORGE B. DANTZIG[1]

By T. N. E. Greville

Federal Security Agency

In this interesting and valuable article, Dr. Dantzig showed that, under certain conditions, a sequence of frequency distributions connected by a linear recurrence formula converges to the normal distribution. Among several applications of his results which are discussed, the author mentions their relation to certain types of smoothing formulas, and has shown that if a linear smoothing formula and the data to which it is applied satisfy certain conditions, the iteration of the smoothing process produces a sequence of smoothed distributions which, upon normalization, approaches the normal frequency curve.

In a summary paragraph at the end of the article, it is stated that "successive application of one or many such linear formulas will usually smooth *any* set of values to the normal curve of error." The entire article was concerned with frequency distributions, and a careful reading makes it clear that the author intended the quoted statement to apply only to data in this form. However, its rather general wording seems to have led a number of readers to interpret it as being applicable to other types of data, such as time series, which frequently may not satisfy the conditions assumed. Moreover, it is easy to overlook the

[1] *Annals of Math. Stat.*, Vol. 10 (1939), pp. 247–253.

restrictions imposed on both the original data and the smoothing formula as they are stated only by implication, and not explicitly, even though they have the effect of excluding important classes of smoothing formulas, such as those commonly employed by actuaries.

The approach to the normal distribution is shown to depend on the vanishing of a certain limit denoted as Γ' which is a function of the moments of the original data and of a distribution in which the weights employed in the smoothing formula are interpreted as frequencies. At this point, objection may be taken to Dr. Dantzig's proof, since the smoothing formulas most frequently used contain negative weights. However, it has been shown elsewhere[2] that the occurrence of negative weights will not of itself prevent the sequence of smoothed distributions from approaching the normal curve. A somewhat more serious difficulty arises if, as is commonly the case, the smoothing formula has the property of reproducing polynomials of a specified degree. If the degree reproduced is two or more, this implies the vanishing of the second moment of the weight distribution, in which case the limit Γ' does not vanish. In fact, it has been shown by DeForest[3] and Schoenberg that the iteration of smoothing formulas which reproduce polynomials of higher degree gives rise to a sequence of limiting distributions which have the general appearance of the normal curve in the center portion and of a damped sine curve in the tails. This is, however, at best, a technical exception to Dantzig's statement, as one is still faced with his basic proposition that repeated application of a smoothing formula to a frequency distribution will cause the smoothed distribution to be dominated by the characteristics of the smoothing formula rather than those of the original data.

While he did not intend the statement to refer to data not in the form of a frequency distribution, some readers seem to have interpreted it as being of general application, and, for that reason, I should like to point out a few of the considerations involved in applying iterated smoothing to other types of data, such as, for example, a time series or the values of a mathematical function. The limit Γ', on whose vanishing Dantzig's theorem depends, involves the second and fourth moments of the original data (as well as of the weight distribution) and, therefore, can be computed only if these moments exist. For this it is necessary (but, of course, not sufficient) that the function being smoothed shall tend toward zero as the independent variable approaches positive or negative infinity.

In order to iterate a smoothing formula an infinite number of times, it is obviously necessary to have an infinite set of original values. Therefore, in smoothing, for example, a finite time series, one would have to make some assumption regarding the values of the series outside the range for which they

---

[2] I. J. SCHOENBERG, "Some analytical aspects of the problem of smoothing," *Courant Anniversary Volume*, Interscience Publishers, New York, 1948.

[3] H. H. WOLFENDEN, "On the development of formulae for graduation by linear compounding, with special reference to the work of Erastus L. DeForest," *Trans. Actuarial Soc. Am.*, Vol. 26 (1925), pp. 81–121.

are actually available. Of course, if it were assumed that the values were zero outside this range, Dantzig's theorem would apply. However, under this assumption, infinite iteration of a smoothing formula would not be a rational procedure, as it would smooth each value to zero, and the incidental fact that the sequence of smoothed distributions, while approaching zero, also approach the form of a normal distribution, would not be a very valuable one. In this connection, an important distinction between time series and frequency data is that, in dealing with the former, one is interested in the magnitude of individual values as well as in the general form and shape of the distribution. In practice it might be preferable not to make any assumption about the values outside the given range but rather to employ special devices to obtain smoothed values near the ends of this range. In such a case, the smoothing process would be a function of the range (if not of the actual values) of the original data distribution. Such a process was not considered by Dantzig, and is clearly excluded by his definition of a linear smoothing formula, which requires that the formula be completely independent of the data to which it is applied.

The somewhat academic question of the effect of iteration of a smoothing formula on a function of infinite range for which the moments do not exist, is a difficult one, to which I cannot give a general answer. Schoenberg does not consider this problem, but merely gives the weight distribution to be applied to the original data in order to obtain the limiting smoothed distribution. Two trivial examples may, however, serve to illustrate the nature of the considerations involved. If the original data are values of a polynomial of a specified degree, and if a smoothing formula which reproduces that degree is successively applied, it will of course continue indefinitely to reproduce the original values. On the other hand, if the smoothing formula reproduces only polynomials of lower degree, a bias is introduced. As a simple example, we may consider the case of smoothing the function $y = x^2$ by a formula consisting of three weights each equal to $1/3$ to be applied to the given value and its two immediate neighbors. It is easily shown that the smoothed value is $x^2 + 1/3$, and the effect of successive application of this formula is to add $1/3$ each time. Thus each smoothed value would tend toward infinity as the number of smoothings increases; however, the entire distribution would always remain a parabola of the same form as originally.

Finally, I should like to emphasize that, in common with Dr. Dantzig, I do not regard infinite repetition of the smoothing operation as a practical procedure, but consider it preferable to select, in the first instance, a smoothing formula which is likely to have the desired effect and then to perform the smoothing in a single step. In this way, one is more likely to secure the result desired without losing sight of important characteristics of the original data.