Double Cumulative and Lorenz Curves in Weather Modification

EDWIN L. CROW

National Center for Atmospheric Research,[1] *Boulder, CO 80307*

(Manuscript received 11 November 1981, in final form 30 April 1982)

ABSTRACT

A graphical presentation of precipitation data has been used for some years in which the cumulative percentage of the total mass falling on various days of a sample of precipitation days, ordered from the largest to the smallest mass, is plotted against the cumulative percentage of days. This type of graph is called a "double cumulative curve" (DCC), but it is essentially the same as the Lorenz curve in economics. The paper reviews the literature, summarizes the properties, shows the DCC's for uniform, log-normal, exponential, gamma and degenerate distributions, studies the average effect of sample size, and presents a formula for testing the significance of the difference between two DCC's. This formula is applied to compare hail data of various types. It is concluded that a sample DCC is a biased estimate of the population DCC, but the bias becomes negligible as the sample size increases beyond ~30.

1. Introduction

A graphical presentation of precipitation data has been used for some years in which the cumulative percentage of the total mass falling on various days of a sample of precipitation days, ordered from the largest to the smallest mass, is plotted against the cumulative percentage of days. An example, from Knight *et al.* (1979), is shown in Fig. 1. A primary advantage of this type of presentation is the clear indication of the wide variation in a distribution. For example, the curve for northeast Colorado in Fig. 1 shows that the largest hail day of the 33 actual hail days in the National Hail Research Experiment (NHRE) accounted for nearly 25% of the total hail mass recorded during the three-year experiment, although it represents only 3% of the days.

The second principal advantage of such "double cumulative curves" (DCC) is also illustrated in Fig. 1: the presentation on the same dimensionless graph of distributions with differing scales of measurement. The hail mass distribution of northeast Colorado can be compared with those of loss-to-risk insurance ratios used in Alberta, tobacco crop damage in South Africa and mass of made tea lost due to hail damage in Kenya. This advantage follows simply from standardizing distributions to the same scale; any differences in central tendency and in dispersion are automatically eliminated. All that remains to be illustrated is the *shape* of a distribution.

Perhaps the first use of such curves in weather modification research (or even in meteorology generally) is a comparison by Huff (1969) of precipitation in storms classified by duration. Morgan *et al.* (1980) and Morgan (1982) have used DCC's extensively to summarize a large amount of diverse precipitation data.

However, essentially the same type of curve has been used in other disciplines, especially economics, for many years. Lorenz (1905) introduced it to study whether wealth was becoming more or less concentrated with time. Several pages of Samuelson's *Economics* text (1976) are devoted to the *Lorenz curves* of income distribution. In economics and most other applications, the cumulation of values is begun with the *smallest* rather than the *largest*, as in weather modification. Partly for that reason and partly because of the greater descriptiveness, the term "double cumulative curve" is used in the present paper, referring primarily to curves begun with the largest value.

Other applications have been made: the distribution of sizes of firms (Hart, 1975), racial segregation in city districts (Jahn *et al.*, 1947), inequality of representation in state legislatures and racial imbalance in school systems (Alker, 1965, 1972), the distribution of fish among fishermen (Thompson, 1976), the distribution of scientific grants (House of Commons, 1975), and the distribution of scientific articles in a particular discipline among scientists and among journals (Leimkuhler, 1967; Allison *et al.*, 1976; Garfield, 1980).

The purpose of the present paper is fivefold: 1) to review briefly the literature of DCC's in disciplines other than the atmospheric sciences in order to apply any useful results; 2) to summarize the properties of DCC's; 3) to show the DCC's for several common

[1] The National Center for Atmospheric Research is sponsored by the National Science Foundation.

0021-8952/82/081063-08$06.00

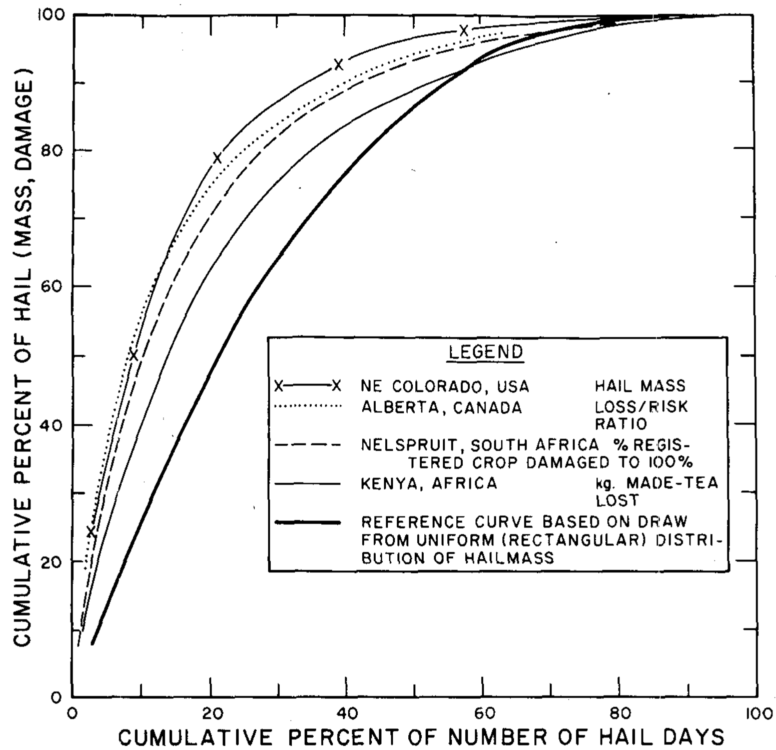

FIG. 1. Cumulative percent of hail mass or hail damage versus the percent of the number of hail days for four different regions, compared to that obtained by a random draw from uniform (rectangular) distribution of hail mass (from Knight *et al.*, 1979).

families of probability distributions (uniform, log-normal, and gamma, including exponential); 4) to study the average effect of sample size on DCC's; and 5) to determine the significance of the difference between two sample DCC's. The log-normal and gamma distributions have been used as models for both rainfall and hailfall data. The average effect of sample size is of interest in weather modification research because the data sets compared with DCC's rarely have the same sample sizes. Thus, any differences observed could conceivably be due to sample size differences rather than intrinsic physical differences such as seeding of the clouds that produced one of the data sets.

The general mathematical properties of DCC's have been studied. Many recent papers have been concerned with the interpolation necessary because income data are usually available only in rather broad categories of income (e.g., Gastwirth and Glauberman, 1976). Gastwirth (1971) gave a general definition of the Lorenz curve. Goldie (1977) and Gail and Gastwirth (1978) derived theorems on the convergence of Lorenz curves as the sample size becomes infinite. Wilk and Gnanadesikan (1968) defined a generalization called "P–P plots," in which the abscissa can be *any* cumulative distribution and proposed them as a means for comparing distributions. Grosh and Morgan (1975) used a graph similar in appearance to a P–P plot to compare predictors of hail days, but their graph was determined quite differently. The present paper is restricted to DCC's, for which the abscissa is the cumulative percentage of items in the sample (or in the theoretical distribution in the limit).

2. Definitions and properties of double cumulative curves

The definitions for a finite population or sample and for a continuous distribution are taken from Gastwirth (1971, 1972), adapted to non-negative variables and to cumulating large values first, rather than small values. The properties are summarized from previous publications, especially Gastwirth (1972), Goldie (1977), and Gail and Gastwirth (1978):

1) Given a set of n ordered numbers, $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n \geqslant 0$, the DCC generated by them is defined at the points i/n, $i = 0, \ldots, n$, by $M(0) = 0$ and $M(i/n) = s_i/s_n$, where $s_i = x_1 + \cdots + x_i$. Thus, $M(i/n)$ represents the fraction of the total s_n (e.g., hail mass over a given area) contributed by the ith largest numbers (e.g., hail days), i.e., by the fraction i/n of largest numbers. The DCC, $M(q)$, is de-

fined for all other q in the interval (0, 1) by linear interpolation between adjacent points $[i/n, M(i/n)]$, thus producing a polygonal or broken-line graph in rectangular coordinates from the origin to the point (1, 1). Both abscissa and ordinate are often expressed as percentages.

2) Let X be a random variable with cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$ (the pdf being assumed to exist) defined over the non-negative real axis:

$$p = \text{Prob}[X \leq x] = F(x) = \int_0^x f(t)dt, \quad x \geq 0. \qquad (1)$$

Thus $F(0) = 0$, $F(\infty) = 1$. The inverse, $x = F^{-1}(p)$, of $F(x)$ is the solution of (1) for x. The mean of X (or first moment of the distribution),

$$\mu = \int_0^\infty tf(t)dt, \qquad (2)$$

is assumed to be finite. The incomplete first moment is defined as

$$\Phi(x) = \frac{1}{\mu}\int_0^x tf(t)dt. \qquad (3)$$

It accumulates the value of the variable from the origin but is normalized so that $\Phi(\infty) = 1$. Kendall and Stuart (1969, p. 48) defined the *Lorenz curve* (their "curve of concentration") as the plot of $\Phi(x)$ as ordinate against $F(x) = p$ as abscissa, the curve being defined parametrically with x as parameter. Gastwirth (1971) defined the Lorenz curve of X by the single equation

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(v)dv, \quad 0 \leq p \leq 1, \qquad (4)$$

by changing the variable of integration, using (1).

3) We define a *population double cumulative curve* (or DCC of a random variable X) as

$$M(q) = 1 - L(1 - q), \quad q = 1 - p. \qquad (5)$$

The DCC is thus the Lorenz curve with origin moved to (1, 1) and axes rotated 180° in the plane of the paper. It can alternatively be defined as the plot of

$$\Psi(x) = 1 - \Phi(x) = \frac{1}{\mu}\int_x^\infty tf(t)dt \qquad (6)$$

against

$$G(x) = q = 1 - p = \int_x^\infty f(t)dt. \qquad (7)$$

In fact this definition is often more convenient because the integrals may be tabulated at convenient values of x, as for gamma and log-normal distributions.

The following properties hold for both finite set (sample) and population DCC's, except for (f).

4) The DCC $M(q)$ is a continuous and increasing function of q.

5) $M(q)$ is a concave function of q; i.e., line segments connecting any two points on it lie below it or on it. Its slope $M'(q)$ thus never increases as q increases.

6) The slope $M'(q)$ of a population DCC is equal to one at $q = G(\mu)$.

7) The degenerate distribution obtained when X equals a constant has the 45° line $\Psi = G$ as its DCC. The area between the DCC and the 45° line is called the *area of concentration*. Its ratio to the area above the 45° line (which is ½) is the Gini coefficient of concentration, or Gini index of inequality.

8) Two cumulative distribution functions have the same DCC if and only if they differ (at most) by a scale transformation (Thompson, 1976).

9) Given a concave, non-decreasing function $M(q)$ that satisfies $M(0) = 0$ and $M(1) = 1$, then there is a cumulative distribution function (or random variable) that has $M(q)$ for its DCC (Thompson, 1976).

10) If a population has a finite mean and a continuous cumulative distribution function, then the DCC of a sample from the population converges with a probability of one to the population DCC as the sample size becomes infinite (Goldie, 1977; Gail and Gastwirth, 1978).

11) If a population has a finite variance and a continuous cumulative distribution function, then the DCC of a sample from the population is, for any given abscissa, asymptotically normally distributed as the sample size becomes infinite (Goldie, 1977; Gail and Gastwirth, 1978).

3. Double cumulative curves for common distributions

a. Uniform distribution

The pdf is

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise.} \end{cases} \qquad (8)$$

The mean is $\mu = (a + b)/2$. From (6)–(7)

$$G(x) = q = \begin{cases} 1, & x < a, \\ \dfrac{b-x}{b-a}, & a \leq x \leq b, \\ 0, & b < x. \end{cases} \qquad (9)$$

$$\Psi(x) = \begin{cases} 1, & x < a, \\ \dfrac{b^2-x^2}{b^2-a^2}, & a \leq x \leq b, \\ 0, & b < x. \end{cases} \qquad (10)$$

Solving (9) for x and substituting in (10) gives the DCC,

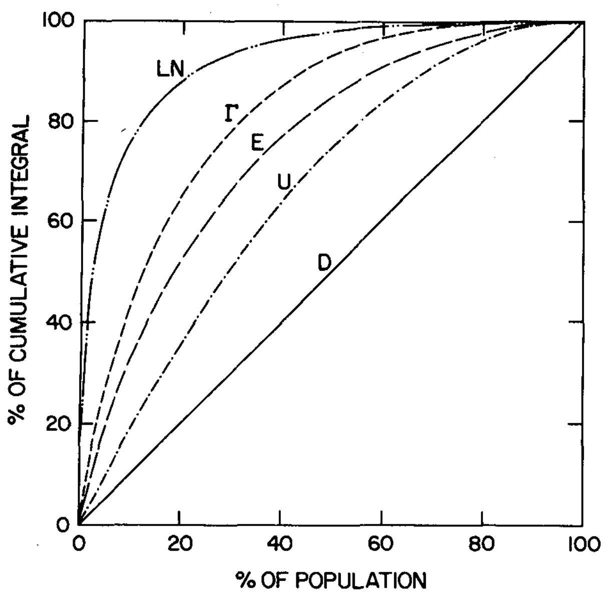

FIG. 2. Population double cumulative curves for degenerate D, uniform U, exponential E, gamma Γ ($\alpha = ½$) and log-normal LN ($\sigma = 2$) distributions.

$$M(q) = \frac{2bq - (b-a)q^2}{a+b}, \quad 0 \leqslant q \leqslant 1. \tag{11}$$

We note that its shape depends on the ratio a/b. However, if the initial value a is zero, then $M(q) = 2q - q^2$, which is independent of the final value b.

This parabola is plotted for $a = 0$ (along with DCC's for other distributions) in Fig. 2.

b. Exponential distribution

The pdf is

$$f(x) = \beta^{-1}e^{-x/\beta}, \quad x \geqslant 0\ (\beta > 0). \tag{12}$$

The mean is $\mu = \beta$. From (6) and (7),

$$M(q) = q(1 - \ln q), \tag{13}$$

which is independent of β. It is plotted in Fig. 2. It is more distant from the diagonal than the DCC for the uniform distribution, reflecting greater skewness. Its slope is $-\ln q$, which is infinite at the origin, whereas that for the uniform distribution with $a = 0$ is 2 at the origin.

c. Gamma distribution

The pdf is

$$f(x) = \frac{1}{\beta\Gamma(\alpha)}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}, \quad x > 0 \tag{14}$$

($\alpha > 0$, $\beta > 0$). The mean is $\alpha\beta$. From (1), with a change of variable of integration,

$$F(x) = I(\alpha^{-1/2}x/\beta, \alpha - 1), \tag{15}$$

where $I(\cdot, \cdot)$ is the incomplete Γ-function ratio tabulated by Pearson (1965). Similarly from (3),

$$\Phi(x) = I\{(\alpha + 1)^{-1/2}x/\beta, \alpha\}. \tag{16}$$

From (6) and (7) the DCC for any given α can be plotted by simply looking up (15) and (16) in Pearson's tables for conveniently varying values of the dimensionless parameter $\alpha^{-1/2}x/\beta$. Thus the DCC for the family of gamma distributions does not depend on the scale parameter β, but it does depend on the shape parameter α. One example has already been plotted in Fig. 2, the special case of the exponential distribution, for which $\alpha = 1$. Although a closed, analytical form (13) was obtained for the DCC in the exponential case, that is not generally possible. The DCC for the more skewed distribution with $\alpha = 1/2$ is also plotted in Fig. 2. That value of α is approximately the value found in fitting a gamma distribution to both rain and hail area total masses in the National Hail Research Experiment (Crow *et al.*, 1979).

d. Log-normal distribution

The pdf of the two-parameter log-normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\{-(\ln x - \nu)^2/(2\sigma^2)\},$$
$$x > 0. \tag{17}$$

The mean is $\mu = \exp(\nu + \sigma^2/2)$. From (1)

$$F(x) = N\{(\ln x - \nu)/\sigma\} = N(y), \tag{18}$$

where $N(y)$ is the standard (i.e., mean 0, variance 1) normal cdf. Similarly from (3)

$$\Phi(x) = N\{(\ln x - \nu - \sigma^2)/\sigma\} = N(y - \sigma). \tag{19}$$

From (6) and (7) the DCC for any given σ can be plotted by simply looking up (18) and (19) in a standard normal table for conveniently varying values of the dimensionless parameter $y = (\ln x - \nu)/\sigma$. Thus the DCC for the family of two-parameter log-normal distributions does not depend on the scale parameter ν, but it does depend on the shape parameter σ. It is plotted in Fig. 2 for $\sigma = 2$, which is the approximate value found in fitting a log-normal distribution to both rain and hail area total masses in the National Hail Research Experiment (NHRE) (Crow *et al.*, 1979). Although both the log-normal distribution with $\sigma = 2$ and the gamma distribution with $\alpha = ½$ are consistent with the NHRE data, Fig. 2 shows that their DCC's are quite different.

e. Degenerate distribution

Since consideration has been limited in this paper to populations with pdf's, the degenerate distribution (i.e., the random variable is a constant) is not strictly included but is a limiting case. It is intuitively evident from the finite sample case in Section 2a that each value of the random variable contributes the same proportion of the total. Hence the DCC is the diagonal straight line from the origin to (1, 1), as shown in Fig. 2.

It is interesting to note that the DCC of the family of uniform distributions in Section 3a approaches the DCC of the degenerate distributions when the parameter a becomes infinite, $b - a$ being held fixed. Likewise the DCC's of the family of gamma and log-normal distributions approach the diagonal line when α and σ respectively become infinite.

4. Effect of sample size on DCC's

Observational data provide an estimate of the long-term (i.e., population) DCC of a given phenomenon, presumably giving a more accurate estimate the larger the sample size. Before showing the average effect of sample size for various distributions, it is desirable to illustrate the random variation in DCC's from one sample to another of the same size. This is done in Fig. 3 with six samples of size five from the uniform distribution. The samples were drawn from a table of three-digit random numbers. One of the samples, re-ordered, was 859, 611, 545, 154, 052. These sum to 2221, so the largest observation contributes 38.7% of the total, compared with 36.0% for the population [from (11) and Fig. 2]. However, the average contribution of the largest observation from all six samples is 34.0%, and the standard deviation of the six contributions is 5.3%. On the other hand, if all six samples are combined to form one sample of size 30, then the six largest observations (20% of the sample in number) contribute 34.6% of the total.

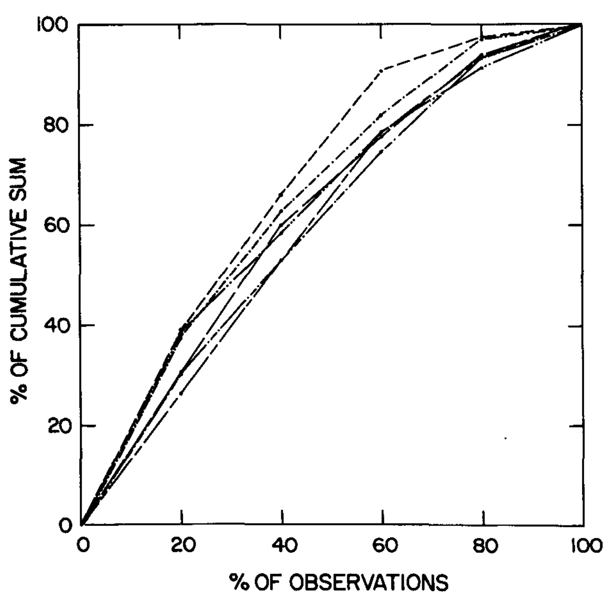

FIG. 3. Double cumulative curves for six random samples of size 5 from the uniform distribution with lower boundary zero.

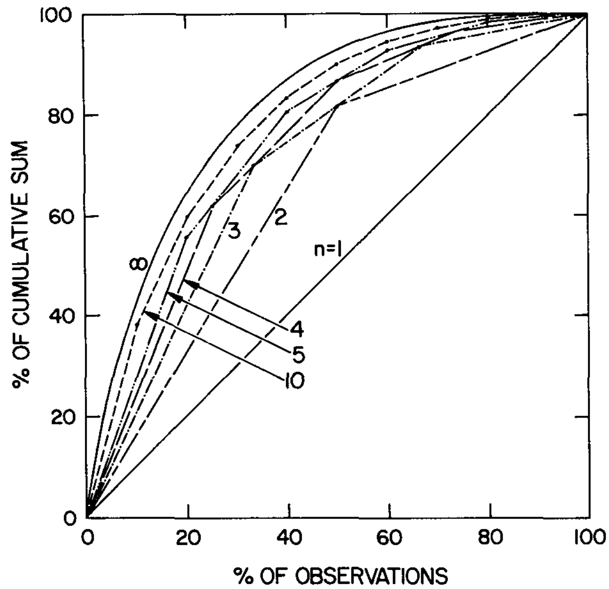

FIG. 4. "Average" double cumulative curves for samples of sizes 1, 2, 3, 4, 5, 10 and ∞ from the gamma distribution with shape parameter $\alpha = 1/2$.

To assess the average effect of sample size it is unnecessary to take repeated samples of random numbers, since the expected values of the ordered observations from common distributions are either known from exact closed formulas (for the uniform distributions, Cramér, 1946, p. 372) or have been tabulated by numerical integration. The tables for the exponential and gamma distributions are given by Harter (1964) for various values of α, and those for the log-normal distribution are given by Gupta *et al.* (1974) for $\sigma = 1$.

Thus, we are easily able to present in Fig. 4 the "average" DCC's for samples from the gamma distribution with $\alpha = 1/2$ for sample sizes $n = 1, 2, 3, 4, 5, 10$ and ∞. These are simply the partial sums of the expected values of the ordered observations divided by the total sums. The curve for $n = \infty$ is the population DCC for the gamma distribution already plotted in Fig. 2.

Fig. 4 shows, at least for the gamma distribution, that the DCC's for very small sample sizes fall on the average considerably short of the population DCC, but the latter is approached rapidly, and the systematic difference may be considered negligible for samples of size 30 or more. The curve for $n = 20$ lies about halfway between those for $n = 10$ and $n = \infty$ in Fig. 4.

Fig. 4 indicates that the ordinate of the "average" DCC for any given abscissa is a monotonically non-

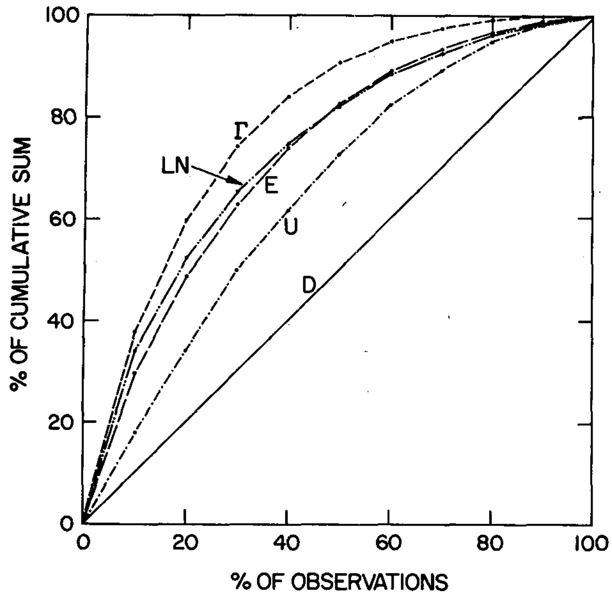

FIG. 5. "Average" double cumulative curves for samples of size 10 from degenerate D, uniform U, exponential E, gamma Γ (α = ½) and log-normal LN (σ = 1) distributions.

decreasing function of the sample size. Furthermore, the straight line segments connecting the discrete values of the DCC for sample size $n = i + 1$ pass exactly through the discrete values of the DCC for sample size $n = i$. These two properties also seem to hold for distributions other than the gamma.

Fig. 5 shows the "average" DCC's for a sample of size ten from the degenerate, uniform, exponential, gamma with α = ½, and log-normal with $\sigma = 1$. These may be compared with the corresponding population DCC's in Fig. 2, except that the DCC for the log-normal distribution in Fig. 2 is for $\sigma = 2$. (The expected values of ordered log-normal observations are available only for $\sigma = 1$.) The log-normal DCC for $\sigma = 1$ and $n = 10$ is close to the exponential DCC for $n = 10$ and crosses it once, whereas the two pdf's are quite different, even when they have the same mean.

We now explain why the sample DCC's in Figs. 4 and 5 are described above as "average" in quotation marks. They are not, in fact, the expected values of random DCC's of the indicated sample size, but rather the expected values of the numerators divided respectively by the expected values of the denominators, which are not the same as the expected values of the fractions because of the nonlinearity of the function. The expected value of the fraction is not readily available, but it will be obtained for small sample sizes from the uniform distribution to illustrate the difference.

The results, obtained by n-dimensional integration, are summarized in Table 1. They show that the true expected sample DCC, at least for the uniform distribution, does not fall as far short of the population DCC as the substitute "average" used in this paper. The substitute "average" does seem to be sufficient to illustrate that the sample DCC will usually deviate systematically (i.e., "on the average") from the population DCC. The *random* deviation for any sample of finite size may well exceed the systematic deviation; see Section 5.

5. Significance of difference between DCC's

It is possible to test for the significance of the difference between two DCC's calculated from large independent samples by using property 11) (asymptotic normal distribution) of Section 2. Gail and Gastwirth (1978) gave the asymptotic variance of the Lorenz curve $L(p)$ at abscissa p, and this is also the asymptotic variance of the DCC $M(q)$ at abscissa $q = 1 - p$:

$$\mathrm{Var}M(q) = \frac{1}{n}\left(\frac{\sigma_1{}^2}{\mu^2} + \frac{\sigma_3{}^2\Phi^2}{\mu^4} - \frac{2\sigma_{13}\Phi}{\mu^3}\right), \quad (20)$$

where μ is the mean (2), $\Phi = \Phi[Q(p)]$ is the incomplete first moment (3), $\sigma_3{}^2$ is the variance of the distribution sampled,

$$\sigma_1{}^2 = 2\int_0^p\left[\int_0^t sQ'(s)ds\right](1-t)Q'(t)dt, \quad (21)$$

and

$$\sigma_{13} = \sigma_1{}^2 + \left[\int_0^p sQ'(s)ds\right] \times \left[\int_p^1 (1-t)Q'(t)dt\right]. \quad (22)$$

TABLE 1. (a) Percent by which the "average" sample DCC, calculated as the partial sum expected value divided by the total sum expected value, falls short of the uniform population DCC. (b) Percent by which the true expected value of the sample DCC falls short of the uniform population DCC. (c) Percent by which the former "average" DCC falls short of the true expected DCC.

n		Largest value (%)	Two largest values (%)	Three largest values (%)
	(a)	11.1	0	—
2	(b)	7.6	0	—
	(c)	3.8	0	—
	(a)	10.0	6.2	0
3	(b)	5.8	4.7	0
	(c)	4.4	1.6	0
	(a)	8.6	6.7	4.0
4	(b)	4.5	4.4	3.2
	(c)	4.3	2.3	0.9
	(a)	7.4	*	*
5	(b)	3.6	*	*
	(c)	4.0	*	*

* Not calculated.

Here $Q(p) = F^{-1}(p)$, the inverse function of (1), and $Q'(p)$ is the derivative of $Q(p)$.

Thus the variance of the ordinate of the DCC depends on the distribution sampled. This fact limits the usefulness of (20), but one may be able to evaluate (20)–(22) for several types of distributions and hence conclude that a given difference in DCC's is either significant or not significant essentially regardless of the distribution. It does appear that (21) and (22) would require numerical integration for gamma and log-normal distributions. It would be highly desirable to develop a distribution-free method for testing the significance of the difference between two DCC's.

One example in which (20)–(22) can be evaluated in closed form is the exponential distribution (Section 3b; Gail and Gastwirth, 1978). The integrations are easy, and the end result is

$$\left.\begin{aligned} &\mathrm{Var}M(q) \\ &\quad = [p + pq + 2q \ln q - (p + q \ln q)^2]/n \\ &\quad = K(q)/n. \end{aligned}\right\} \quad (23)$$

If independent random samples of sizes n_1 and n_2 are taken from two exponential distributions (which need not have the same means), then the variance of the difference between the two resulting DCC's at abscissa q is

$$K(q)(1/n_1 + 1/n_2). \quad (24)$$

The coefficient $K(q)$ is 0 at $q = 0$ and $q = 1$ and is a maximum at $q = 0.166$.

For the sake of concrete illustration, let us apply the above results for the exponential distribution to the NHRE and Kenya DCC's in Fig. 1, which were derived from 33 and 78 hail days, respectively. At $q = 0.2$ the standard deviation of the difference is, from (24), 0.061. The observed difference in Fig. 1 is 0.15, which is 2.4 times the standard deviation and hence significant at the 5% level. On the other hand, the other two hail data curves in Fig. 1 are not significantly different from the NHRE curve at $q = 0.2$. Assuming other types of underlying distribution might result in different conclusions.

6. Concluding remarks

The history and properties of double cumulative curves, used most in economics with slightly different definition and known therein as Lorenz curves, have been summarized for both samples of data and theoretical distributions. Numerical or graphical results for four distributions and samples from them have shown that the sample DCC's are biased estimates of the population DCC's, but the bias becomes negligible as the sample size increases beyond ~30. Hence, it is not important that hailfall or rainfall data from different weather modification experiments compared by means of DCC's have different sample sizes. The larger the sample sizes are, the smaller will be the random fluctuations affecting the DCC's. A method of assessing these random fluctuations is available in Section 5. The advantage of DCC's in putting widely disparate measurements like hail mass and crop damage on the same scale does have a cost; only the shapes, not the means or scales, are compared.

Acknowledgments. The author is indebted to Griffith M. Morgan, Jr., and G. Brant Foote for discussions prior to writing the paper and comments on the paper. Fig. 1 is reproduced from Knight *et al.* (1979) by courtesy of the authors. The paper has benefited further from helpful comments by two reviewers.

REFERENCES

Alker, H. R., Jr., 1965: *Mathematics and Politics.* Macmillan, 152 pp.

——, 1972: Measuring sociopolitical inequality. *Statistics: A Guide to the Unknown*, J. M. Tanur, F. Mosteller, W. H. Kruskal, R. F. Link, R. S. Pieters and G. R. Rising, Eds., Holden-Day, 343–351.

Allison, P. D., D. de Solla Price, B. C. Griffith, M. J. Moravcsik and J. A. Stewart, 1976: Lotka's Law: A problem in its interpretation and application. *Soc. Stud. Sci.*, **6**, 269–276.

Cramer, H., 1946: *Mathematical Methods of Statistics.* Princeton University Press, 575 pp.

Crow, E. L., A. B. Long, J. E. Dye, A. J. Heymsfield and P. W. Mielke, Jr., 1979: Results of a randomized hail suppression experiment in northeast Colorado, Part II: Surface data base and primary statistical analysis. *J. Appl. Meteor.*, **18**, 1538–1558.

Gail, M. H., and J. L. Gastwirth, 1978: A scale-free goodness-of-fit test for the exponential distribution based on the Lorenz curve. *J. Amer. Stat. Assoc.*, **73**, 787–793.

Garfield, E., 1980: Bradford's Law and related statistical patterns. *Current Contents*, **20**(19), 5–10.

Gastwirth, J. L., 1971: A general definition of the Lorenz curve. *Econometrica*, **39**, 1037–1039.

——, 1972: The estimation of the Lorenz curve and Gini index. *Rev. Econ. Stat.*, **54**, 306–316.

——, and M. Glauberman, 1976: The interpolation of the Lorenz curve and Gini index from grouped data. *Econometrica*, **44**, 479–484.

Goldie, C. M., 1977: Convergence theorems for empirical Lorenz curves and their inverses. *Adv. Appl. Probab.*, **9**, 765–791.

Grosh, R. C., and G. M. Morgan, Jr., 1975: Radar-thermodynamic hail day determination. *Preprints Ninth Conf. Severe Local Storms*, Norman, Amer. Meteor. Soc., 454–459.

Gupta, S. S., G. C. McDonald and D. I. Galarneare, 1974: Moments, product moments and percentage points of the order statistics from the lognormal distribution for samples of size twenty and less. *Sankhyā*, **B36**, 230–260.

Hart, P. E., 1975: Moment distributions in economics: an exposition. *J. Roy. Stat. Soc.*, **A138**, 423–434.

Harter, H. L., 1964: Expected values of exponential, Weibull and gamma order statistics. Aerospace Research Laboratories Rep. 64-31, Wright–Patterson AFB, 106 pp.

House of Commons, 1975: *Second Report on Scientific Research in British Universities* (First Report from the Select Committee on Science and Technology. Session 1975–76). Cmnd. 504, Her Majesty's Stationery Office.

Huff, F. A., 1969: Climatological assessment of natural precipitation characteristics for use in weather modification. *J. Appl. Meteor.*, **8,** 401–410.

Jahn, J., C. F. Schmid and C. Schrag, 1947: The measurement of ecological segregation. *Amer. Sociol. Rev.*, **12,** 293–303.

Kendall, M. G., and A. Stuart, 1969: *The Advanced Theory of Statistics.* Vol. I, *Distribution Theory.* 3rd ed., Hafner, 439 pp.

Knight, C. A., G. B. Foote and P. W. Summers, 1979: Results of a randomized hail suppression experiment in northeast Colorado. Part IX: Overall discussion and summary in the context of physical research. *J. Appl. Meteor.*, **18,** 1629–1639.

Leimkuhler, F. F., 1967: The Bradford distribution. *J. Doc.,* **23,** 197–207.

Lorenz, M. O., 1905: Methods of measuring the concentration of wealth. *J. Amer. Stat. Assoc.*, **9,** 209–219.

Morgan, G. M., Jr., 1982: Precipitation at the ground. Hailstorms of the Central High Plains, Vol. 1: *The National Hail Research Experiment*, C. A. Knight and P. Squires, Eds., NCAR/Colorado Associated University Press, 282 pp.

Morgan, G. M., Jr., D. Vento, J. F. Mezeix, P. Admirat, B. Federer, A. Waldvogel, L. Wojtiw and E. Wirth, 1980: A comparison of hailfall intensity and spatial variability in the United States and Europe based on data from instrumented networks. *Papers Presented at Third WMO Scientific Conf. on Weather Modification*, Clermont-Ferrand, 21–25 July 1980, Groupement National d'Etudes des Fléaux Atmosphériques, Grenoble, 791 pp.

Pearson, K., Ed., 1965: *Tables of the Incomplete* Γ*-function.* Cambridge University Press, 164 pp.

Samuelson, P. A., 1976: *Economics*, 10th ed. McGraw Hill, 917 pp.

Thompson, W. A., Jr., 1976: Fisherman's luck. *Biometrics*, **32,** 265–271.

Wilk, M. B., and R. Gnanadesikan, 1968: Probability plotting methods for the analysis of data. *Biometrika*, **55,** 1–17.