THE EXTENDED PROBABILITY THEORY FOR THE CONTINUOUS VARIABLE WITH PARTICULAR APPLICATION TO THE LINEAR DISTRIBUTION

By
H. P. LAWTHER, JR.

The engineering worker is often confronted with the necessity of utilizing a group of quantities concerning whose numerical values it is known only that they lie between definite upper and lower limits. If a number n of specimens is selected from such a group and the sum of the n values taken, intuition rules that there is negligible probability that this sum will be as great as n times the upper limit or as small as n times the lower limit, and that the most probable value must be intermediate between these two extremes. Some assurance is desired regarding the practical limits within which such a sum may be expected to fall. While the distribution of the individual values within their limits may be unknown directly, yet workable inferences frequently may be made from the nature of the quantities. For example, in many manufacturing operations it is economical to turn out items (such as bearing balls, paper condensers, or spacing washers) in large quantities with rather coarse precision. By means of gauges set to limits narrow as compared with the total spread, the product is then selected into bins, and in the operation of assembly a completed article utilizes the material from a single bin. The contents of any such bin clearly may be expected to follow a linear distribution very closely, and if the relative proportions of the product finding their ways into this bin and its immediate neighbors can be learned, the distribution may be specified with practical accuracy. The linear distribution is thus fundamental to a large class of problems.

On several occasions the writer's speculations have led to problems involving the linear distribution, and the reference literature has been searched for assistance. The special case of the rectangular distribution seems first to have been formulated by Laplace[1] more than a century ago. Rietz,[2] Irwin,[3] Hall,[4] and Craig,[5] in recent years have presented analyses applicable to the study of the linear distribution, each from a somewhat different viewpoint. In the attempt to follow the logical processes of specialists in this field the writer was driven to put forth considerable independent effort in order to arrive at a satisfactory understanding. As the result of this effort still another angle of approach was developed. This method employs steps and terminology familiar to one whose mathematical education may have been limited to that commonly encountered in a college engineering course, and should be readily understandable to a wide field of workers. Encouragement was thus given to treating the case of the generalized linear distribution, a treatment which appears to be new. In the application to practical cases it was necessary to carry out certain tedious computations yielding interesting values and curves. The results of this work are presented with the thought that they may stimulate the understanding and use of a law of considerable application to engineering practice.

It will be understood that when a selection, or the sum of n selections, is spoken of there is meant the dimension of that selection, or the sum of the dimensions of the n selections. Following

---

1 Laplace: *Théorie Analytique des Probabilités,* Troisième Edition (1820), pp. 257-263.

2 Rietz: *On a Certain Law of Probability of Laplace,* Proceedings of the International Mathematical Congress, Toronto (1924), vol. 2, pp. 795-799.

3 Irwin: *On the Frequency Distribution of Means of Samples from a Population having any Law of Frequency with Finite Moments, with special reference to Pearson's type II.* Biometrika, vol. 19 (1927), pp. 226-239.

4 Hall: *The Distribution of Means of Samples of Size N drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable.* Biometrika, vol. 19 (1927), pp. 240-244.

5 Craig: *On the Distributions of Certain Statistics.* American Journal of Mathematics, vol. 54, No. 2 (1932), pp. 353-366.

the usual notation, the symbol $f_n(x)$ will be defined to be such that the integral $\int_a^b f_n(x)\cdot dx$ is equal to the probability that the sum of n selections lies between the values a and b. Neglecting higher orders of infinitesimals, the probability that the sum of n selections lies between x and $x+\Delta x$ would then be equal to the product $f_n(x)\cdot\Delta x$ The sum of n selections clearly is the sum of $n-1$ selections plus the value of an additional selection. The probability that the sum of n selections lies within the interval x to $x+\Delta x$ must then be equal to the summation of the probabilities associated with all possible pairs of values for the sum of the first $n-1$ selections and for the last selection, respectively, that can yield a final sum lying between x and $x+\Delta x$. The values $x-m\cdot\Delta x$ and $m\cdot\Delta x$, where m is an integer, are such a pair, and the totality of these pairs is obtained by extending m to all possible values. Recalling that the probability of the simultaneous occurrence of two independent events is equal to the product of the probabilities associated with their individual occurrences, there may be written in the conventional symbols

$$f_n(x)\cdot\Delta x = \sum_{m=-\infty}^{m=\infty} f_{n-1}(x-m\cdot\Delta x)\cdot\Delta x\cdot f_1(m\cdot\Delta x)\cdot\Delta x\,.$$

Setting $m\cdot\Delta x = \lambda$, and passing to the limiting form, there is obtained

$$f_n(x) = \int_{-\infty}^{\infty} f_{n-1}(x-\lambda)\cdot f_1(\lambda)\cdot d\lambda\,,$$

as the general formula for determining all subsequent f's from $f_1(x)$. The form of the function $f_1(x)$ is, of course, determined for any particular case from the best available physical data. The expression for $f_n(x)$ is then obtained by $n-1$ successive applications of the operation of integration indicated above. For the sake of subsequent brevity, the operator P will be defined to be such that

$$P \cdot \phi(x) = \int_{-\infty}^{\infty} \phi(x-\lambda) \cdot f_1(\lambda) \cdot d\lambda .$$

Using this notation it would be written

$$f_n(x) = P^{n-1} \cdot f_1(x).$$

For the general linear distribution $f_1(x)$ is given as follows:

$$f_1(x) = 0 \,, \quad \text{for} \quad -\infty < x < 0$$

$$f_1(x) = 0 \,, \quad \text{for} \quad a < x < \infty \,.$$

$f_1(x)$ is the equation of a straight line for $0 \leqq x \leqq a$, subject to the conditions: (1) the area under the line from $x=0$ to $x=a$ is unity, (2) no ordinate is negative for any value of x in this interval. By imposing these conditions upon the general equation to a straight line there is obtained

$$f_1(x) = \frac{1}{a}\left[1 + k - \frac{2K}{a}(a-x)\right] \,, \quad \text{for} \quad 0 \leqq x \leqq a$$

where the parameter K is restricted to the values $-1 \leqq K \leqq 1$. With $f_1(x)$ so defined it can be inferred immediately that $f_n(x)$ must be identically zero for all negative values of x, will have some positive finite value everywhere in the interval $0 < x < na$, and must be identically zero for all values of x greater than na. Also, since $f_1(x)$ is discontinuous for $x \to 0$ and for $x \to a$ the application of the operator P must be effected through proper choice of limits of integration. In this connection three possible cases arise:

Case 1: Where x, the sum of n selections, lies in the interval $0 \leqq x \leqq a$ it could have resulted only from a value for the sum of n-1 selections lying in the interval x to 0, coupled with a suitable value for the n-th selection lying in the interval 0 to x. For this case the operator P will be distinguished as follows:

$$f_n(x) = \overset{x}{\underset{0}{P}} \cdot f_{n-1}(x) = \int_0^x f_{n-1}(x-\lambda) \cdot f_1(\lambda) \cdot d\lambda \,, \quad \text{for} \quad 0 \leqq x \leqq a.$$

Case 2: Where x, the sum of n selections, lies in the interval $a \leqq x \leqq (n-1)a$ it could have resulted only from a value for the sum of n-1 selections lying in the interval x to $x-a$, coupled with a suitable value for the n-th selection lying in the interval 0 to a. For this case the operator P will be distinguished as follows:

$$f_n(x) = \overset{a}{\underset{0}{P}} \cdot f_{n-1}(x) = \int_0^a f_{n-1}(x-\lambda) \cdot f_1(\lambda) \cdot d\lambda \text{, for } a \leqq x \leqq (n-1)a.$$

Case 3: Where x, the sum of n selections, lies in the interval $(n-1)a \leqq x \leqq na$ it could have resulted only from a value for the sum of $n-1$ selections lying in the interval $(n-1)a$ to $x-a$, coupled with a suitable value for the n-th selection lying in the interval $x-(n-1)a$ to a. For this case the operator P will be distinguished as follows:

$$f_n(x) = \overset{a}{\underset{x-(n-1)a}{P}} \cdot f_{n-1}(x) = \int_{x-(n-1)a}^{a} f_{n-1}(x-\lambda) \cdot f_1(\lambda) \cdot d\lambda \text{ for, } (n-1)a \leqq x \leqq na.$$

The procedure now is analogous to that employed in establishing the binomial theorem. The first few f's are obtained by hand-power methods, until the sequences can be discerned and the expression for $f_n(x)$ can be inferred. The expression for $f_n(x)$ is then established, first by applying to $f_n(x)$ the operator P and showing that this yields an expression for $f_{n+1}(x)$ wholly consistent with that for $f_n(x)$ when $n+1$ is substituted for n, and finally by showing that it degenerates into $f_1(x)$ when n is taken as 1.

The preliminary steps, while very necessary, are quite tedious, and there would be no value in repeating them here. Suffice it to state that by such means it can be inferred that $f_n(x)$ is of the form

$$\begin{aligned} f_n(x) = \frac{1}{a^n} \cdot \left(-\frac{1}{p}\right)^{n-2} \Bigg\{ & \left(1+K+\frac{2K}{ap}\right)^n [na-x] - \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)[na-a-x] \\ & + \binom{n}{2}\left(1+K+\frac{2K}{ap}\right)^{n-2}\left(1-K+\frac{2K}{ap}\right)^2 [na-2a-x] \\ & - \binom{n}{3}\left(1+K+\frac{2K}{ap}\right)^{n-3}\left(1-K+\frac{2K}{ap}\right)^3 [na-3a-x] \text{-----} \\ & \text{-------------------} \\ & + (-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n-1}[a-x] \Bigg\} \end{aligned}$$

where it is understood that each term including a bracket member of the form $[na-ba-x]$ is to be assigned the value zero for values of x which render this negative. The use of brackets [* *] distinguishes the operand in each term. The symbol $\frac{1}{p}$ denotes the operation of integration with respect to x between the upper limit x and that lower limit for which the integrand vanishes. Thus

$$\frac{1}{p}\cdot[na-ba-x]^m = -\frac{1}{m+1}[na-ba-x]^{m+1}$$

. Where $\frac{1}{p}$ occurs with a negative exponent it signifies the inverse operation of differentiation with respect to x. The symbol $\binom{n}{b}$ means $\frac{n!}{b!(n-b)!}$ and is one of the familiar binomial coefficients.

Preparatory to establishing the validity of the inferred expression for $f_n(x)$, it is convenient to assemble certain working material. First $\varphi_n(x)$ will be defined as follows:

$$\begin{aligned}\varphi_n(x) = \frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-2}\Big\{&\left(1+K+\frac{2K}{ap}\right)^n[na-x] - \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)[na-a-x]\\ &+\binom{n}{2}\left(1+K+\frac{2K}{ap}\right)^{n-2}\left(1-K+\frac{2K}{ap}\right)^2[na-2a-x]\\ &-\binom{n}{3}\left(1+K+\frac{2K}{ap}\right)^{n-3}\left(1-K+\frac{2K}{ap}\right)^3[na-3a-x]----\\ &-------------------\\ &+(-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n-1}[a-x]\\ &+(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x]\Big\},\end{aligned}$$

where the symbols all have the same meaning as before, but here there is no special understanding regarding the bracket members of the form $[na-ba-x]$ and they are to exist for all values of x. Especial note should be made of the inclusion of a final term in $[-x]$. Otherwise the expression is identical in appearance with that for $f_n(x)$ for the interval $0 \leqq x \leqq a$. Next, the typical operation

$${}_0^aP\cdot[B-x] = \int_0^a [B-x+\lambda]\cdot\frac{1}{a}\cdot\left[1+K-\frac{2K}{a}(a-\lambda)\right]\cdot d\lambda\,,$$

is readily evaluated and yields

$$\frac{1}{a}\cdot\left(-\frac{1}{p}\right)\left\{\left(1+K+\frac{2K}{ap}\right)[B+a-x]-\left(1-K+\frac{2K}{ap}\right)[B-x]\right\}.$$

Now it can be written immediately that

$$\begin{aligned}\overset{a}{\underset{o}{P}}\cdot\varphi_n(x)=\frac{1}{a^{n+1}}\cdot\left(-\frac{1}{p}\right)^{n-1}\Big\{&\left(1+K+\frac{2K}{ap}\right)^{n+1}[na+a-x]-\left(1+K+\frac{2K}{ap}\right)^{n}\left(1-K+\frac{2K}{ap}\right)[na-x]\\ &-\binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n}\left(1-K+\frac{2K}{ap}\right)[na-x]\\ &+\binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)^{2}[na-a-x]---\\ &-------------------\\ ---&+(-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)^{2}\left(1-K+\frac{2K}{ap}\right)^{n-1}[2a-x]\\ &-(-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n}[a-x]\\ &+(-1)^{n}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n}[a-x]\\ &-(-1)^{n}\left(1-K+\frac{2K}{ap}\right)^{n+1}[-x]\Big\}\end{aligned}$$

Collecting terms, this becomes

$$\begin{aligned}\frac{1}{a^{n+1}}\cdot\left(-\frac{1}{p}\right)^{n+1-2}\Big\{&\left(1+K+\frac{2K}{ap}\right)^{n+1}[(n+1)a-x]\\ &-\binom{n+1}{1}\left(1+K+\frac{2K}{ap}\right)^{n+1-1}\left(1-K+\frac{2K}{ap}\right)[(n+1)a-a-x]\\ &+\binom{n+1}{2}\left(1+K+\frac{2K}{ap}\right)^{n+1-2}\left(1-K+\frac{2K}{ap}\right)^{2}[(n+1)a-2a-x]---\\ &-------------------\\ &+(-1)^{n}\binom{n+1}{n}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n+1-1}[a-x]\\ &+(-1)^{n+1}\left(1-K+\frac{2K}{ap}\right)^{n+1}[-x]\Big\},\end{aligned}$$

and this is seen to correspond exactly with the expression for $\varphi_n(x)$ if $n+1$ is substituted for n. Consequently, $\varphi_n(x)$ must be the result of $n-1$ successive applications of the operator ${}_{0}^{a}P$ to a certain $\phi_1(x)$ given by

$$\phi_1(x) = \frac{1}{a}\cdot\left(-\frac{1}{p}\right)^{-1}\left\{\left(1+K+\frac{2K}{ap}\right)[a-x]-\left(1-K+\frac{2K}{ap}\right)[-x]\right\}$$
$$= \frac{1}{a}\cdot\left\{\left[1+K-\frac{2K}{a}(a-x)\right]-\left[1-K-\frac{2K}{a}(-x)\right]\right\} \equiv 0,$$

which is seen to be identically zero for all values of x. Now the application of the operator ${}_{0}^{a}P$ to zero yields zero. Therefore $\phi_n(x)$ must be identically zero for all values of x. Finally, it is convenient to evaluate the typical operation

$${}_{x-B}^{a}P\cdot[B-x] = \int_{x-B}^{a}[B-x+\lambda]\cdot\frac{1}{a}\left[1+K-\frac{2K}{a}(a-\lambda)\right]\cdot d\lambda\ .$$

This yields

$$\frac{1}{a}\left(-\frac{1}{p}\right)\left(1+K+\frac{2K}{ap}\right)[B+a-x]\ .$$

The expression for $f_n(x)$ now may be established in straight-forward fashion.

In the interval $0 \leqq x \leqq a$ the expression for $f_n(x)$ will conclude with the term involving $[a-x]$. As has been shown before, the expression for $f_{n+1}(x)$ should then be given by

$$f_{n+1}(x) = {}_{0}^{x}P\cdot f_n(x),\quad \text{for}\quad 0 \leqq x \leqq a\ .$$

This operation may be evaluated readily, and will yield a result that is correct. The form of the result, however, is such that it does not display the desired correspondence with the expression for $f_n(x)$. The expression for $\phi_n(x)$ is introduced here to advantage. Let it be written

$$f_{n+1}(x) = {}_{0}^{x}P\cdot f_n(x) + {}_{x}^{a}P\cdot\phi_n(x), \qquad \text{for } 0 \leqq x \leqq a\ .$$

Since the last term is identically zero its introduction is permissible. Remembering that $\phi_n(x)$ consists of all the terms of $f_n(x)$ plus a final term in $[-x]$, a rearrangement may be made giving

$$f_{n+1}(x) = {}^a_0P \cdot f_n(x) + {}^a_0P \cdot \frac{1}{a^n} \cdot \left(-\frac{1}{p}\right)^{n-2} (-1)\left(1-K+\frac{2K}{ap}\right)^n [-x], \text{ for } 0 \leqq x \leqq a.$$

Using the operations that have been evaluated above, there is obtained immediately

$$\begin{aligned} f_{n+1}(x) = \frac{1}{a^{n+1}}\left(-\frac{1}{p}\right)^{n-1} \Bigg\{ & \left(1+K+\frac{2K}{ap}\right)^{n+1}[na+a-x] - \left(1+K+\frac{2K}{ap}\right)^n \left(1-K+\frac{2K}{ap}\right)[na-x] \\ & - \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^n \left(1-K+\frac{2K}{ap}\right)[na-x] - - - - - - \\ & + \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)^2 [na-a-x] \\ & - \\ & - - - - + (-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)^2\left(1-K+\frac{2K}{ap}\right)^{n-1}[2a-x] \\ & - (-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^n [a-x] \Bigg\} \\ & + \frac{1}{a^{n+1}}\left(-\frac{1}{p}\right)^{n-1}\left\{(-1)^n\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^n[a-x]\right\}, \text{ for } 0 \leqq x \leqq a. \end{aligned}$$

Collecting terms, this becomes

$$\begin{aligned} f_{n+1}(x) = \frac{1}{a^{n+1}} \cdot \left(-\frac{1}{p}\right)^{n+1-2} \Bigg\{ & \left(1+K+\frac{2K}{ap}\right)^{n+1}[(n+1)a-x] \\ & - \binom{n+1}{1}\left(1+K+\frac{2K}{ap}\right)^{n+1-1}\left(1-K+\frac{2K}{ap}\right)[(n+1)a-a-x] \\ & + \binom{n+1}{2}\left(1+K+\frac{2K}{ap}\right)^{n+1-2}\left(1-K+\frac{2K}{ap}\right)^2[(n+1)a-2a-x] - - \\ & - - - - - - - - - - - - - - - - - - \\ & - - - - + (-1)^{n-1}\binom{n+1}{n-1}\left(1+K+\frac{2K}{ap}\right)^2\left(1-K+\frac{2K}{ap}\right)^{n-1}[2a-x] \\ & + (-1)^n\binom{n+1}{n}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^n[a-x] \Bigg\}, \end{aligned}$$

for $0 \leqq x \leqq a$

and this is seen to be wholly consistent with the formula for $f_n(x)$ for the same interval with the substitution of $n+1$ for n.

Now for some interval between $x=a$ and $x=na$, say $(na-ba-a) \leqq x \leqq (na-ba)$ where b is an integer having any value from zero to $n-2$, the expression for $f_n(x)$ will conclude with the term involving $[na-ba-x]$. For the interval immediately preceding, namely $(na-ba-2a) \leqq x \leqq (na-ba-a)$, the expression will include the additional term involving $[na-ba-a-x]$. Therefore in evaluating the operation ${}_0^aP \cdot f_n(x)$, which as has been shown before should yield the expression for $f_{n+1}(x)$, the integration of all but the $[na-ba-a-x]$ term will be carried over the complete range of λ from 0 to a. The term involving $[na-ba-a-x]$ will not enter into $f_n(x-\lambda)$ until λ reaches the value $(x-na+ba+a)$, and so the integration of it will be between the limits $(x-na+ba+a)$ and a. Using the operations that have been evaluated previously there is obtained immediately

$$
\begin{aligned}
f_{n+1}(x) = \frac{1}{a^{n+1}}\left(-\frac{1}{p}\right)^{n-1} \Bigg\{ & \left(1+K+\frac{2K}{ap}\right)^{n+1}[na+a-x] - \left(1+K+\frac{2K}{ap}\right)^{n}\left(1-K+\frac{2K}{ap}\right)[na-x] \\
& - \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n}\left(1-K+\frac{2K}{ap}\right)[na-x] \\
& + \binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)^{2}[na-a-x] --- \\
& ------------------------ \\
& --- + (-1)^{b}\binom{n}{b}\left(1+K+\frac{2K}{ap}\right)^{n-b+1}\left(1-K+\frac{2K}{ap}\right)^{b}[na-ba+a-x] \\
& - (-1)^{b}\binom{n}{b}\left(1+K+\frac{2K}{ap}\right)^{n-b}\left(1-K+\frac{2K}{ap}\right)^{b+1}[na-ba-x] \\
& + (-1)^{b+1}\binom{n}{b+1}\left(1+K+\frac{2K}{ap}\right)^{n-b}\left(1-K+\frac{2K}{ap}\right)^{b+1}[na-ba-x] \Bigg\},
\end{aligned}
$$

for $(na-ba-a) \leqq x \leqq (na-ba)$.

Collecting terms, this becomes

$$f_{n+1}(x)=\frac{1}{a^{n+1}}\cdot\left(-\frac{1}{p}\right)^{n+1-2}\Big\{\left(1+K+\frac{2K}{ap}\right)^{n+1}[(n+1)a-x]$$

$$-\binom{n+1}{1}\left(1+K+\frac{2K}{ap}\right)^{n+1-1}\left(1-K+\frac{2K}{ap}\right)[(n+1)a-a-x]--$$

$$--+(-1)^{b+1}\binom{n+1}{b+1}\left(1+K+\frac{2K}{ap}\right)^{n+1-b-1}\left(1-K+\frac{2K}{ap}\right)^{b+1}[(n+1)a-(b+1)a-x]\Big\},$$

$$\text{for } (na-ba-a)\leqq x\leqq(na-ba).$$

and this is seen to be wholly consistent with the expression for $f_n(x)$ for the same interval with the substitution of $n+1$ for n.

Finally, for the interval $na \leqq x \leqq (na+a)$ the expression for $f_n(x)$ is identically zero. For the interval immediately preceding, namely $(na-a) \leqq x \leqq na$, it consists of the single term involving $[na-x]$. Therefore

$$f_{n+1}(x)={}^{a}_{x-na}P\; f_n(x)={}^{a}_{x-na}P\cdot\frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-2}\left(1+K+\frac{2K}{ap}\right)^n[na-x]$$

$$=\frac{1}{a^{n+1}}\left(-\frac{1}{p}\right)^{n+1-2}\left(1+K+\frac{2K}{ap}\right)^{n+1}[(n+1)a-x] \text{ for } na\leqq x\leqq(n+1)a$$

and this is seen to be wholly consistent with the expression for $f_n(x)$ for the corresponding interval.

Setting $n=1$ in the general expression for $f_n(x)$ for the interval $0 \leqq x \leqq a$ there is obtained

$$f_1(x)=\frac{1}{a}\cdot\left(-\frac{1}{p}\right)^{-1}\Big\{\left(1+K+\frac{2K}{ap}\right)[a-x]\Big\}.$$

Carrying out the indicated operations, this becomes

$$f_1(x)=\frac{1}{a}\cdot\left[1+K-\frac{2K}{a}(a-x)\right], \text{ for } 0\leqq x\leqq a$$

and this is the $f_1(x)$ chosen at the start.

In the derivation of $f_{n+1}(x)$ from $f_n(x)$ it was assumed that the two forms of the operator P, namely ${}^{a}_{0}P$ and ${}^{a}_{x-B}P$, are commutative with the operator $\frac{1}{p}$ when applied to an operand of the form $[B-x]$. It is very easy to show that ${}_0^aP\cdot\frac{1}{p}\cdot[B-x]$ yields a result identical with that of $\frac{1}{p}\cdot{}_0^aP\cdot[B-x]$ and that ${}_{x-B}^{\;\;a}P\cdot\frac{1}{p}\cdot[B-x]$ yields a result identical with that of $\frac{1}{p}\cdot{}_{x-B}^{\;\;a}P\cdot[B-x]$, and the space will not be taken here to give this demonstration. The formula for $f_n(x)$ may thus be regarded as firmly established.

It has been shown that the complete expression $\phi_n(x)$ is identically zero for all values of x Therefore, if in the interval $(na-ba-a) \leqq x \leqq (na-ba)$ the desired function $f_n(x)$ can be represented by the partial expression

$$f_n(x)=\frac{1}{a^n}\cdot\left(-\frac{1}{p}\right)^{n-2}\left\{\left(1+K+\frac{2K}{ap}\right)^n[na-x]-\binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)[na-a-x]\right.$$
$$+\binom{n}{2}\left(1+K+\frac{2K}{ap}\right)^{n-2}\left(1-K+\frac{2K}{ap}\right)^2[na-2a-x]---$$
$$\left.----+(-1)^b\binom{n}{b}\left(1+K+\frac{2K}{ap}\right)^{n-b}\left(1-K+\frac{2K}{ap}\right)^b[na-ba-x]\right\},$$

it follows that it may equally well be represented in the same interval by the negative of the remainder of the complete expression, or

$$f_n(x)=\frac{1}{a^n}\cdot\left(-\frac{1}{p}\right)^{n-2}\left\{-(-1)^{b+1}\binom{n}{b+1}\left(1+K+\frac{2K}{ap}\right)^{n-b-1}\left(1-K+\frac{2K}{ap}\right)^{b+1}[na-ba-a-x]\right.$$
$$-(-1)^{b+2}\binom{n}{b+2}\left(1+K+\frac{2K}{ap}\right)^{n-b-2}\left(1-K+\frac{2K}{ap}\right)^{b+2}[na-ba-2a-x]---$$
$$-----\;-(-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n-1}[a-x]$$
$$\left.-(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x]\right\}.$$

Remembering that $\binom{n}{1}=\binom{n}{n-1}$, $\binom{n}{2}=\binom{n}{n-2}$, etc., this last may be rewritten

$$f_n(x)=\frac{1}{a^n}\cdot\left(\frac{1}{p}\right)^{n-2}\left\{\left(1-K+\frac{2K}{ap}\right)^n[x]-\binom{n}{1}\left(1-K+\frac{2K}{ap}\right)^{n-1}\left(1+K+\frac{2K}{ap}\right)[x-a]\right.$$
$$+\binom{n}{2}\left(1-K+\frac{2K}{ap}\right)^{n-2}\left(1+K+\frac{2K}{ap}\right)^2[x-2a]-----$$
$$\left.--+(-1)^{n-b-1}\binom{n}{n-b-1}\left(1-K+\frac{2K}{ap}\right)^{b+1}\left(1+K+\frac{2K}{ap}\right)^{n-b-1}[x-na+ba+a]\right\}.$$

where again it is understood that each term including a bracket member of the form $[x-ba]$ is to be assigned the value zero for values of x which render this negative. The having of these two forms of expression for $f_n(x)$ is very valuable in computation work, since it limits the number of terms that have to be handled to $\frac{n+2}{2}$ at most.

Setting K equal to zero gives the special case of the rectangular distribution, and the expressions for $f_n(x)$ reduce to the forms

$$f_n(x)=\frac{1}{a^n}\left(\frac{1}{p}\right)^{n-2}\left\{[na-x]-\binom{n}{1}[na-a-x]+\binom{n}{2}[na-2a-x]---\right\},$$

and

$$f_n(x)=\frac{1}{a^n}\left(\frac{1}{p}\right)^{n-2}\left\{[x]-\binom{n}{1}[x-a]+\binom{n}{2}[x-2a]-----\right\}.$$

Carrying out the indicated operations, these become

$$f_n(x)=\frac{1}{a^n(n-1)!}\left\{[na-x]^{n-1}-\binom{n}{1}[na-a-x]^{n-1}+\binom{n}{2}[na-2a-x]^{n-1}----\right\}$$

and

$$f_n(x)=\frac{1}{a^n(n-1)!}\left\{[x]^{n-1}-\binom{n}{1}[x-a]^{n-1}+\binom{n}{2}[x-2a]^{n-1}---\right\}.$$

This last expression is the one usually found in the literature, and it was originally developed by Laplace as the limiting form of an urn problem.

Setting K equal to plus one or to minus one gives either of the extreme cases of the "right triangular distribution" For K equal to plus one the expressions for $f_n(x)$ reduce to

$$f_n(x)=\frac{2^n}{a^n(n-1)!}\cdot\left\{\left(1+\frac{1}{ap}\right)^n[na-x]^{n-1}-\binom{n}{1}\left(1+\frac{1}{ap}\right)^{n-1}\left(\frac{1}{ap}\right)[na-a-x]^{n-1}\right.$$
$$\left.+\binom{n}{2}\left(1+\frac{1}{ap}\right)^{n-2}\left(\frac{1}{ap}\right)^2[na-2a-x]^{n-1}\cdots\cdots\right\},$$

and

$$f_n(x)=\frac{2^n}{a^n(n-1)!}\cdot\left\{\left(\frac{1}{ap}\right)^n[x]^{n-1}-\binom{n}{1}\left(\frac{1}{ap}\right)^{n-1}\left(1+\frac{1}{ap}\right)[x-a]^{n-1}\right.$$
$$\left.+\binom{n}{2}\left(\frac{1}{ap}\right)^{n-2}\left(1+\frac{1}{ap}\right)^2[x-2a]^{n-1}---\right\}.$$

The function $f_n(x)$ normally has no direct practical application, but it is of interest to see its trend with increasing values of n. There are shown on Figure 1 several members of the family of curves originating with the rectangular distribution, and on Figure 2 corresponding members of the family originating with the right triangular distribution. In both figures the interval a has been taken as unity, and there have actually been plotted the curves $y = n \cdot f_n(n \cdot x)$. This change in variable places all curves to a common base, and at the same time preserves the property of the total area under each being unity.

For the sum of n selections the practical worker wants to know the minimum value x'; or the maximum value x''; or, most often of all, the shortest interval x' to x'' associated with a certain probability value. The probability that the sum of n selections will be less than x' is given by

$$F_n(x') = \int_0^{x'} f_n(x) \cdot dx,$$

and the probability that the sum will exceed x'' is given by

$$F_n(x'') = \int_{x''}^{na} f_n(x) \cdot dx.$$

Noting that $\int_0^{x'} [x - ba] \cdot dx = \int_{ba}^{x'} [x - ba] \cdot dx = \frac{1}{p} \cdot [x' - ba]$ and that

$\int_{x''}^{na} [na - ba - x]\, dx = \int_{x''}^{na-ba} [na - ba - x]\, dx = -\frac{1}{p} \cdot [na - ba - x'']$, since the bracket members are assigned the value zero for values of x which render them negative, there may be written immediately

$$F_n(x') = \frac{1}{a^n} \cdot \left(\frac{1}{p}\right)^{n-1} \left\{ \left(1 - K + \frac{2K}{ap}\right)^n [x'] - \binom{n}{1} \left(1 - K + \frac{2K}{ap}\right)^{n-1} \left(1 + K + \frac{2K}{ap}\right) [x' - a] \right.$$
$$\left. + \binom{n}{2} \left(1 - K + \frac{2K}{ap}\right)^{n-2} \left(1 + K + \frac{2K}{ap}\right)^2 [x' - 2a] - - - \right\},$$

and

$$F_n(x'') = \frac{1}{a^n} \cdot \left(-\frac{1}{p}\right)^{n-1} \left\{ \left(1 + K + \frac{2K}{ap}\right)^n [na - x''] \right.$$
$$\left. - \binom{n}{1} \left(1 + K + \frac{2K}{ap}\right)^{n-1} \left(1 - K + \frac{2K}{ap}\right) [na - a - x''] + \binom{n}{2} \left(1 + K + \frac{2K}{ap}\right)^{n-2} \left(1 - K + \frac{2K}{ap}\right)^2 [na - 2a - x''] - \right\}.$$

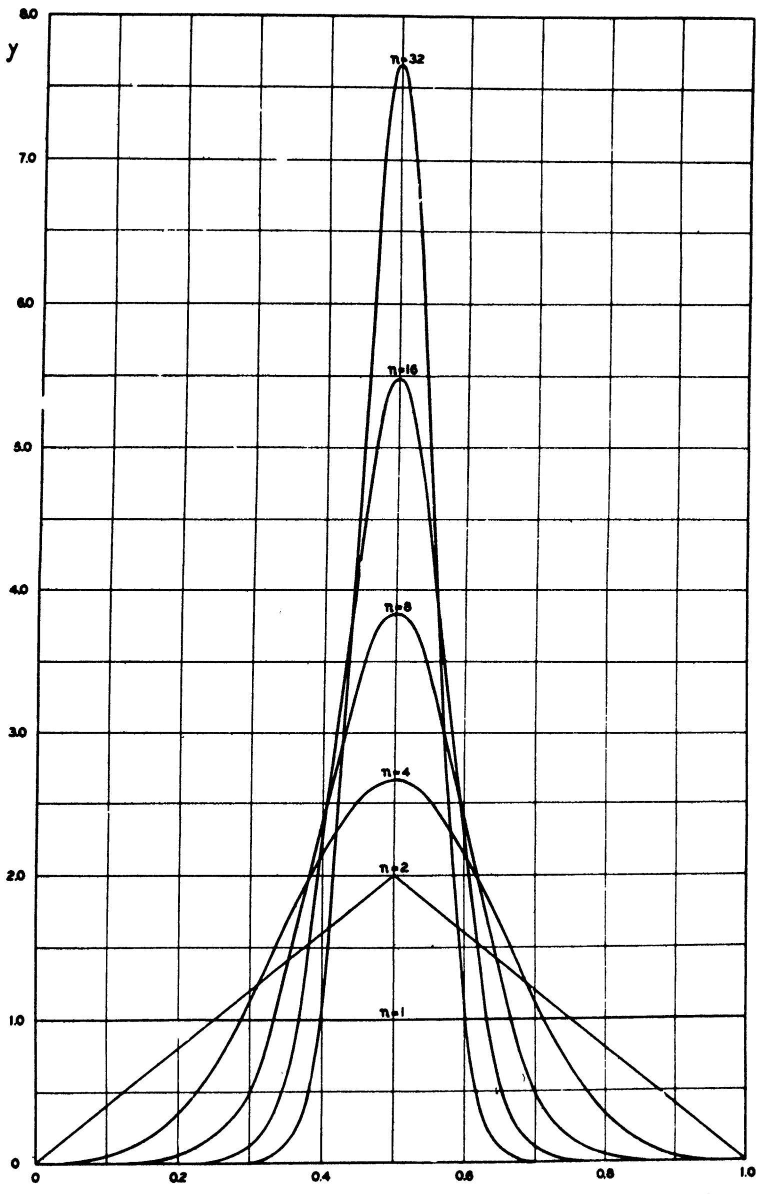

FIGURE -# 1

GRAPHS OF CURVES $y = n \cdot f_n(n \cdot x)$ FOR RECTANGULAR DISTRIBUTION

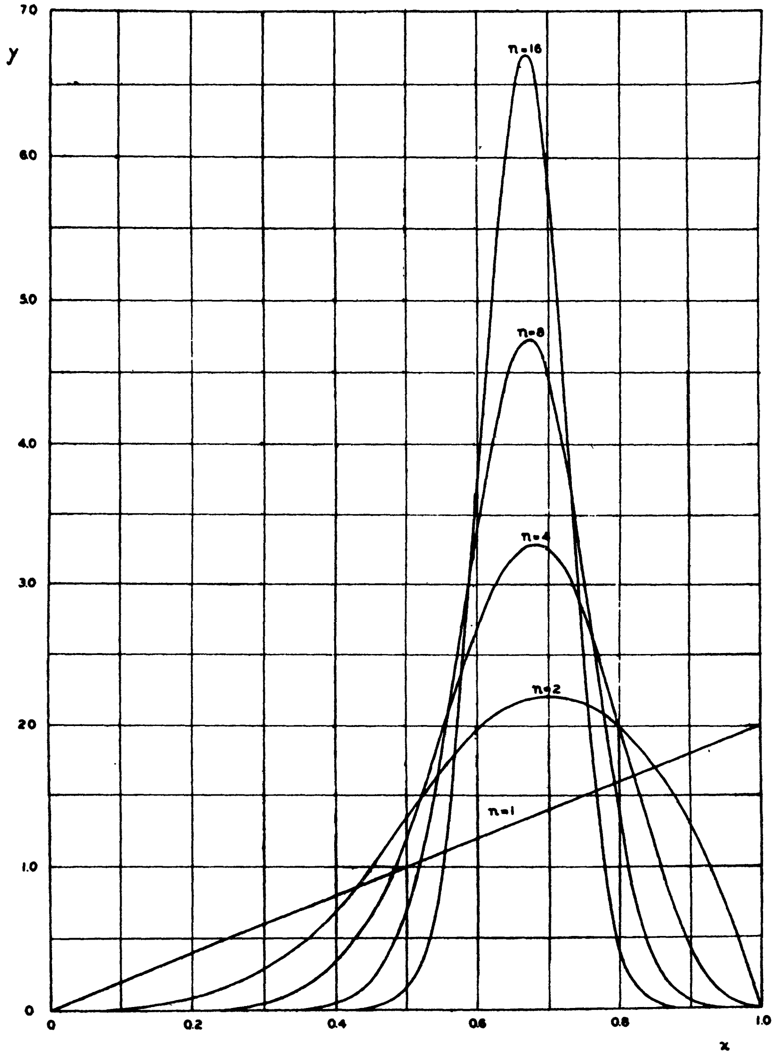

FIGURE - #2

GRAPHS OF CURVES $y = n \cdot f_n(n \cdot x)$ FOR RIGHT TRIANGULAR DISTRIBUTION

For $x' = na$ or for $x'' = 0$ these last expressions respectively should equate to unity, the total area under any probability curve. It is simple to verify that they do so, and the demonstration will be given for one of them. For the interval $0 \leqq x'' \leqq a$ the expression for $F_n(x'')$ concludes with the term containing $[a-x'']$. Let there be added and subtracted a term involving $[-x'']$ to give

$$F_n(x'') = \frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-1}\left\{\left(1+K+\frac{2K}{ap}\right)^n[na-x'']\right.$$
$$-\binom{n}{1}\left(1+K+\frac{2K}{ap}\right)^{n-1}\left(1-K+\frac{2K}{ap}\right)[na-a-x''] - \cdots + (-1)^{n-1}\binom{n}{n-1}\left(1+K+\frac{2K}{ap}\right)\left(1-K+\frac{2K}{ap}\right)^{n-1}[a-x'']$$
$$\left. +(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x'']\right\} - \frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-1}(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x''].$$

From inspection it is seen that this may be written

$$F_n(x'') = {}_0^aP^{n-1}\cdot\left(-\frac{1}{p}\right)\cdot\Phi_1(x'') - \frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-1}(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x''],$$

where $\Phi_1(x'')$ is the expression introduced previously. Upon carrying out the operations it is found that $-\frac{1}{p}\cdot\Phi_1(x'') = 1$, and ${}_0^aP\cdot[1] = 1$. It follows therefore that

$$F_n(x'') = 1 - \frac{1}{a^n}\left(-\frac{1}{p}\right)^{n-1}(-1)^n\left(1-K+\frac{2K}{ap}\right)^n[-x''] \text{ .for } 0 \leqq x'' \leqq a$$

and it is apparent now that for $x'' = 0$ this expression is equal to unity. Thus it is seen that the function $F_n(x)$ also possesses an end-for-end symmetry similar to that of $f_n(x)$ The complete expression corresponding to $F_n(x)$ is equal to unity instead of zero, however, and where the desired function is represented by a partial expression it can also be equally well represented by one minus the remaining terms of the complete expression.

Referring to Figure 3 it is seen that the sum of area A plus area B, or $F_n(x') + F_n(x'')$, gives the probability that

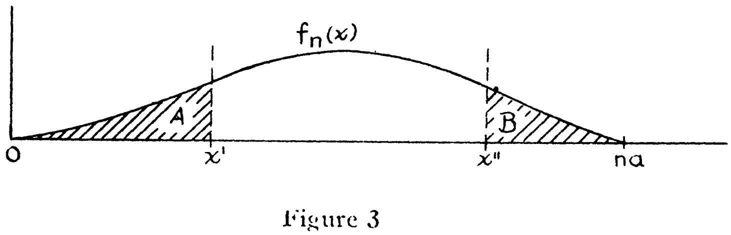

Figure 3

the sum of n selections lies outside the interval x' to x''. In an actual problem usually either the sum of A and B would be assigned and it would be desired to find the interval x' to x'' associated with this expectancy, or the interval x' to x'' would be assigned and it would be desired to find the expectancy associated with these limits. With $f_n(x)$ of the character shown in Figure 3 and with the sum alone of A and B fixed, any number of pairs of values are possible for x' and x''. It is also clear that the length of the interval $x''-x'$ will depend upon the relative magnitudes of A and B. There are two special cases, however, which cover all normal demands. It is seen at a glance that the interval $x''-x'$ will be shortest for a given A plus B when $f_n(x'')=f_n(x')$. Purely from the standpoint of deviations, this shortest interval represents the optimum results of which the group is capable. Where the absolute magnitude is of primary concern it might be specified that $x'=na-x''$.

The function which is of final interest, then, is represented by the sum $F_n(x')+F_n(x'')$, subject either to the restriction that $f_n(x')=f_n(x'')$ or to the restriction that $x'=na-x''$. For the special case of the rectangular distribution $f_n(x)$ is symmetrical about the line $x=\frac{na}{2}$, and the two restrictions imply the same thing. The symmetry of the rectangular distribution permits giving formal expression to the sum $F_n(x')+F_n(x'')$ as a function of $(x''-x')$ when $x'=na-x''$. Under this condition the sum becomes equal to $2F_n(x')$ and x' may be written as $\left[\frac{na-(x''-x')}{2}\right]$ and there is obtained

$$F_n(x')+F_n(x'')=2F_n\left[\frac{na-(x''-x'')}{2}\right], \quad \text{for } x'=na-x''$$

For linear distributions other than the rectangular this simplicity of expression is not possible.

On Figures 4 and 5 are shown graphs of the curves $y = F_n(n \cdot x')$ for several values of n for the rectangular and for the right triangular distribution respectively. Here again the interval a has been taken as unity and change in variable has been made to place all curves to a common base. The values of $F_n(n \cdot x'')$ may be read directly from the same curves, since $F_n(n \cdot x'') = 1 - F_n(n \cdot x')$ for corresponding values of x' and x''. Finally, on Figures 6 and 7 are shown curves for the sum $\left[F_n(x') + F_n(x'')\right]$ plotted as a function of $\left(\frac{x'' - x'}{na}\right)$, subject to the restriction that $f_n(x') = f_n(x'')$, for several values of n for the rectangular and for the right triangular distribution respectively. The values for Figure 6 were computed directly from the formula given in the paragraph above. For Figure 7, however, the values were derived graphically from Figures 2 and 5.

Figures 6 and 7 are applicable immediately to practical problems. As a simple example, suppose there is at hand a group whose individuals are known to lie within the limits of D and $D+a$ and to follow a right triangular distribution with the larger probability associated with the larger limit, and it is desired to know for the sum of eight selections what limits may be expected to be associated with a probability value of 0.01. Referring to Figure 7 it is seen that the curve for $n = 8$ reaches an ordinate value of 0.01 at an abscissa value of approximately 0.45. Referring now to Figure 2, the distance 0.45 is fitted in between the two legs of the curve for $n = 8$, and values for $\frac{x'}{8}$ and $\frac{x''}{8}$ of 0.43 and 0.88 are found. Consequently it can be concluded that for sums of eight selections from this group the probability is 0.01 that the values will lie outside the interval $8D + 3.44a$ to $8D + 7.04a$.

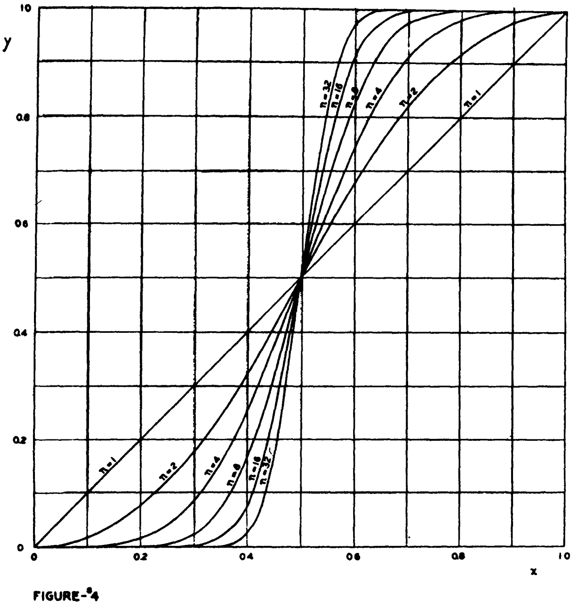

FIGURE-8.4

GRAPHS OF CURVES $y = F_n(n \cdot x)$ FOR RECTANGULAR DISTRIBUTION

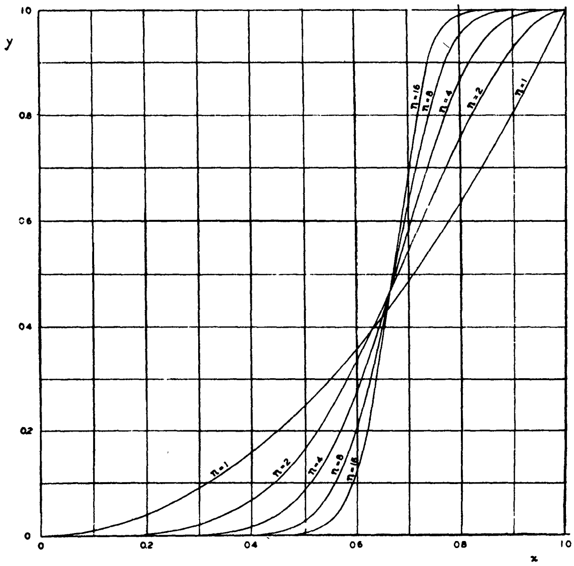

FIGURE-8.5

GRAPHS OF CURVES $y = F_n(n \cdot x)$ FOR RIGHT TRIANGULAR DISTRIBUTION

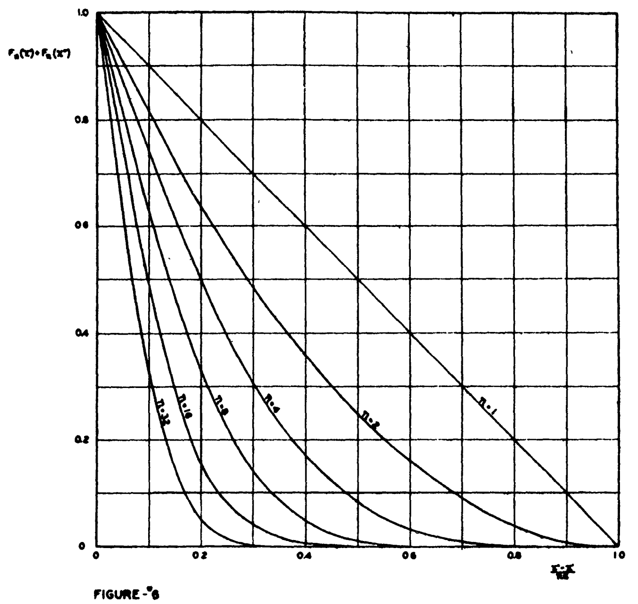

FIGURE-#6

GRAPHS OF CURVES OF $\left[F_n(x')+F_n(x'')\right]$ vs $\left[\frac{x''-x'}{na}\right]$ FOR RECTANGULAR DISTRIBUTION

FOR CONDITION THAT $f_n(x')=f_n(x'')$

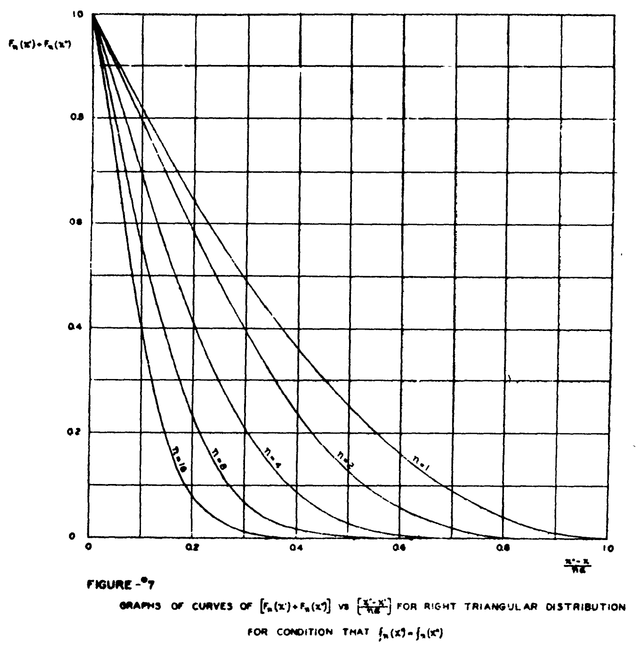

FIGURE-#7

GRAPHS OF CURVES OF $\left[F_n(x')+F_n(x'')\right]$ vs $\left[\frac{x''-x'}{na}\right]$ FOR RIGHT TRIANGULAR DISTRIBUTION

FOR CONDITION THAT $f_n(x')=f_n(x'')$

While this study has been concerned primarily with the linear distribution, it is obvious that the results may find occasional wider application. The curves for $f_2(x)$, $f_4(x)$, etc., are of a character suggestive of distributions that might occur not infrequently in engineering and physics. If any one of them, say $f_n(x)$, is found to fit the group at hand with practical accuracy, then the sequence $f_n(x)$, $f_{2n}(x)$, $f_{3n}(x)$, etc. clearly will give the distribution curves associated with the sums of one, two, three, etc., selections from this group.

H. P. Lawther, Jr