Duration Curves of Logarithmic Normal Distribution Type and Their Applications

By

Shigehisa IWAI

(Received October, 1949)

Synopsis

A statistical study was made on duration curves mainly of hydrological data, such as rainfalls and river discharges, which are the basis of waterworks planning. New proposals were made covering various fields such as application and reliability as well as fundamental theory and method of estimation, and many theoretical and practical results were obtained.

1. Analysis of Logarithmic Normal Distribution

In Germany, France and America, Grassberger, Gibrat and Slade devised and developed many theories as the statistical asymmetric distribution to which duration curve belongs, almost at the same time about 20 years ago, but with no mutual connection. Upon examining these theories, it was found that they could all be reduced to the Logarithmic Normal Distribution, namely "distribution obtained by regarding the logarithmically transformed variable of the original variable as the probability variable of Gauss' normal distribution". As the various distributions were represented by complicated forms in various countries, analogy between them could not be recognized and it was difficult to compare them. These distributions, however, were systematized as follows according to original contrivance.

These Logarithmic Normal Distributions can be divided into two types, namely, partly bounded type and totally bounded type. The former is divided again into a distribution in which origin zero of variable is the lower limit, (represented by suffix E), and a distribution with lower limit $(-b_0)$, (represented by suffix I), while the latter is a distribution in which the upper limit is g and the lower limit $(-b)$, (represented by suffix S). In each case function W which represents the duration curve can be shown by the following expression as an integral function of function V which represents the frequency curve.

$$W_B(\chi)= V_0(\log\chi)\frac{d(\log\chi)}{d\chi}=\frac{k\log e}{\chi\sqrt{\pi}}e^{-k^2(\log(\chi/\chi_0))^2}$$
$$=\frac{1}{2}\varphi_0(\chi)\frac{d\xi}{d\chi},\qquad \xi=k\log(\chi/\chi_0) \qquad \cdots\cdots\cdots (1)$$

$$W_B(\chi)=\frac{1}{2}\{1-\Phi_0(\xi)\} \qquad \cdots\cdots\cdots (1')$$

$$V_I(\chi)=V_0\left(\log(\chi+b_0)\right)\frac{d\left(\log(\chi+b_0)\right)}{d\chi}=\frac{c_0'\log e}{(\chi+b_0)\sqrt{\pi}}e^{-c_0'^2(\log\{(\chi+b_0)/(\chi_0+b_0)\})^2}$$
$$=\frac{1}{2}\varphi_0(\xi')\frac{d\xi'}{d\chi},\qquad \xi'=c_0'\log\{(\chi+b_0)/(\chi_0+b_0)\} \qquad \cdots\cdots\cdots (2)$$

$$W_I(\chi)=\frac{1}{2}\{1-\Phi_0(\xi')\} \qquad \cdots\cdots\cdots (2')$$

$$V_S(\chi)= V_0\left(\log\{(\chi+b)/(g-\chi)\}\right)\frac{d\left(\log\{(\chi+b)/(g-\chi)\}\right)}{d\chi}$$
$$=\frac{(g+b)c_0\log e}{(\chi+b)(g-\chi)\sqrt{\pi}}e^{-c_0^2[\log(\{(\chi+b)/(g-\chi)\}/X_0)]^2}$$
$$=\frac{1}{2}\varphi_0(\xi'')\frac{d\xi''}{d\chi},\qquad \xi''=c_0\left[\log\{(\chi+b)/(g-\chi)\}-\log X_0\right] \qquad \cdots (3)$$
$$X_0=(\chi_0+b)/(g-\chi_0),\quad \text{where}\quad g>\chi>-b,\quad g>b$$

$$W_S(\chi)=\frac{1}{2}\{1-\Phi_0(\xi'')\} \qquad \cdots\cdots\cdots (3')$$

In these equations, φ_0 and Φ_0 are symbols indicating Gauss' error function and error integral, respectively, which are given as follows.*

$$\varphi_0(\theta)=\frac{2}{\sqrt{\pi}}e^{-\theta^2} \qquad \cdots\cdots\cdots (4)$$

$$\Phi_0(\theta)=\frac{2}{\sqrt{\pi}}\int_0^\theta e^{-u^2}du=\int_0^\theta\varphi_0(u)\,du \qquad \cdots\cdots\cdots (5)$$

In each of the above cases, limits and median, χ_0 were expressed as constants in the frequency function V and, therefore, in the duration function W. Furthermore, the following theoretical equations were successfully deduced for the mode, hst (χ), deflection point, wnd (χ) and mean, m (χ), of the frequency curve and for the moments of the t th. degree about the mean, ${}_m\mu_t$.

*These values are easily obtained from adequates tables. For example, from Table V, pp. 456, 457 of "Probability and its Engineering Uses", Thornton C. Fry, D. Van Nostrand, 1928, giving the relation between y and $\varphi(y)$, we can get the relation between θ and $\varphi_0(\theta)$ by substituting $y=\sqrt{2}\,\theta$ and $\varphi_0(\theta)=2\sqrt{2}\,\varphi(y)$. Also from Table IV, pp. 454-455, giving the relation between y and $\Phi(y)$, we can get the relation between θ and $\Phi_0(\theta)$ or $W(\theta)$ by putting $y=\sqrt{2}\,\theta$, $\Phi_0(\theta)=1-\Phi(y)$ or $W(\theta)=\Phi(y)/2$.

For partly bounded type, V_I and W_I;

$$\mathrm{hst}(\chi)=\delta(\chi_0+b_0)-b_0, \quad \log\delta=-K'/(2c_0'), \quad \delta=e^{-K'^2/2} \qquad \cdots\cdots\cdots (6)$$

$$\left.\begin{aligned} &\mathrm{wnd}_1(\chi)=\delta_1(\chi_0+b_0)-b_0, \quad \mathrm{wnd}_2(\chi)=\delta_2(\chi_0+b_0)-b_0, \\ \text{where} \quad &\left.\begin{matrix}\log\delta_1 \\ \log\delta_2\end{matrix}\right\}=\frac{1}{2}\left(3\pm\sqrt{1+(8/K'^2)}\right)\log\delta, \\ &\left.\begin{matrix}\delta_1 \\ \delta_2\end{matrix}\right\}=e^{-\frac{K'^2}{4}\left(3\pm\sqrt{1+(8/K'^2)}\right)}=\delta^{\left(3\pm\sqrt{1+(8/K'^2)}\right)/2}\end{aligned}\right\} \cdots (7)$$

$$K'=1/(c_0'\log e)=2.30259/c_0' \qquad \cdots\cdots\cdots (8)$$

$$\left.\begin{aligned} &\mathrm{m}(\chi)=\mu_1=\gamma_1(\chi_0+b_0)-b_0, \\ \text{where} \quad &\log\gamma_1=-(\log\delta)/2, \quad \gamma_1=\delta^{-0.5}=10^{K'/(4c_0')}\end{aligned}\right\} \cdots\cdots (9)$$

$$\left.\begin{aligned} {}_m\mu_2&=\sigma^2=(\gamma_1^2-1)(\mu_1+b_0)^2, \\ {}_m\mu_3&=(\gamma_1^6-3\gamma_1^2+2)(\mu_1+b_0)^3, \\ {}_m\mu_4&=(\gamma_1^{12}-4\gamma_1^6+6\gamma_1^2-3)(\mu_1+b_0)^4\end{aligned}\right\} \cdots\cdots\cdots (10)$$

If the lower limit $(-b_0)$ in equations frcm (6) to (10) above is put zero, they will naturally become equations for V_E and W_E.

For totally bounded type, V_S and W_S; To make it simple, if ξ'' in equation (3) is represented by ξ, then we get equation (11) which is a general equation for arbitrary variable χ. If suffixes h, o, m, 1 and 2 are attached when the values of χ are $\mathrm{hst}(\chi)$, χ_0, $\mathrm{m}(\chi)$, $\mathrm{wnd}_1(\chi)$ and $\mathrm{wnd}_2(\chi)$, then we get equations from (12) to (17).

$$\left.\begin{aligned} &Y=z-A=(X-1)/(X+1), \quad \text{where} \quad z=2\chi/(g+b),\ A=(g-b)/(g+b), \\ &\quad \text{and for} \quad g>\chi>-b, \quad 1>|Y|, \quad \therefore \quad \tanh^{-1}Y=\frac{1}{2}\ln X, \\ &X=(\chi+b)/(g-\chi)=(1+Y)/(1-Y)>0, \quad \xi=\frac{1}{K}\ln(X/X_0), \\ &K=1/(c_0\log e)=2.30259/c_0, \\ &\chi=(gX-b)/(1+X)=(g+b)z/2=(g+b)(Y+A)/2=\{(g-b)+(g+b)Y\}/2\end{aligned}\right\}$$

$$\cdots\cdots\cdots (11)$$

$$\left.\begin{aligned} \mathrm{hst}(\chi)&=\{(g-b)+(g+b)Y_h\}/2, \\ \tanh^{-1}Y_h&=K^2Y_h/4+1.15129\log X_0\end{aligned}\right\} \cdots\cdots\cdots (12)$$

$$\left.\begin{aligned} &\mathrm{wnd}_1(\chi)=\{(g-b)+(g+b)Y_1\}/2, \quad \mathrm{wnd}_2(\chi)=\{(g-b)+(g+b)Y_2\}/2, \\ &\tanh^{-1}Y_1=K^2\left[3Y_1-\left(3Y_1^2+\{2(4-K^2)\}/K^2\right)^{0.5}\right]/8+1.15129\log X_0, \\ &\tanh^{-1}Y_2=K^2\left[3Y_2+\left(3Y_2^2+\{2(4-K^2)\}/K^2\right)^{0.5}\right]/8+1.15129\log X_0\end{aligned}\right\} (13)$$

If we put,

$$\lambda = c_0 \log(X_0)^{-1}, \qquad X_0 = e^{-\lambda K} \qquad \cdots\cdots\cdots\cdots\cdots \text{(14)}$$

$$\left.\begin{aligned} & m(\chi) = \mu_1 = \left\{(g-b)+(g+b)Y_m\right\} \Big/ 2, \\ \text{where} \quad & Y_m = Y_{m\cdot 1} - \Phi_0(\lambda), \quad Y_{m\cdot 1} = \varphi_0(\lambda)\sum_{r=1}^{+\infty}(-1)^{r-1}(A_r - B_r), \\ & A_r = \left\{\varphi_0(\lambda - rK/2)\right\}^{-1}\left\{\Phi_0(\lambda - rK/2)+1\right\}, \\ & B_r = \left\{\varphi_0(\lambda + rK/2)\right\}^{-1}\left\{1-\Phi_0(\lambda + rK/2)\right\} \end{aligned}\right\} \quad \cdots\cdots \text{(15)}$$

Furthermore, using the next general representation,

$$Y_{m\cdot s} = \varphi_0(\lambda)\sum_{r=1}^{+\infty}(-1)^{r-1}(r)^{s-1}\left[A_r + (-1)^s B_r\right], \ (s=1,\ 2,\ 3, \ldots) \qquad \cdots\cdots \text{(16)}$$

$$\left.\begin{aligned} & {}_m\mu_2 = \sigma^2 = (g+b)^2\left\{1 - Y_m^2 - Y_{m\cdot 2}\right\} \Big/ 4, \\ & {}_m\mu_3 = (g+b)^3\left\{-Y_m + Y_m^3 + 3Y_m \cdot Y_{m\cdot 2} + Y_{m\cdot 3}\right\} \Big/ 4, \\ & {}_m\mu_4 = (g+b)^4\left\{3 + 6Y_m^2 - 9Y_m^4 - 8Y_{m\cdot 2} - 36Y_m^2 Y_{m\cdot 2} - 24Y_m Y_{m\cdot 3} - 4Y_{m\cdot 4}\right\} \Big/ 48 \end{aligned}\right\}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{(17)}$$

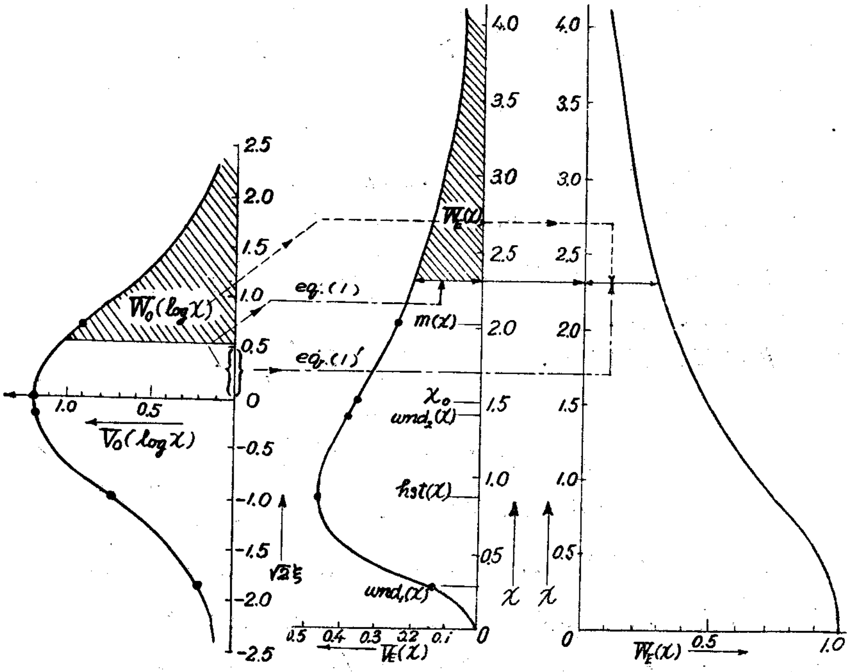

Fig. 1. Frequency Curve and Duration Curve of Partly Bounded Type ($V_B(x)$ and $W_B(x)$ by equations (1), (1′) when $x_0 = -1.5$ and $\sqrt{2}\,k = 3.0$)

Upon investigating the above relations, it was found that the order in which the mode, median and mean were located was always the same in all of the above cases and, furthermore, that these three characteristic values were always located on either the left or right side of the middle point between the upper and lower limits. Four characteristic coefficients, namely, of position, of variation, of skewness and kurtosis were obtained, and upon determining the criterion, it was verified that partly bounded type belonged to Pearson's Type VI and totally bounded type belonged to Pearson's Type I or II.

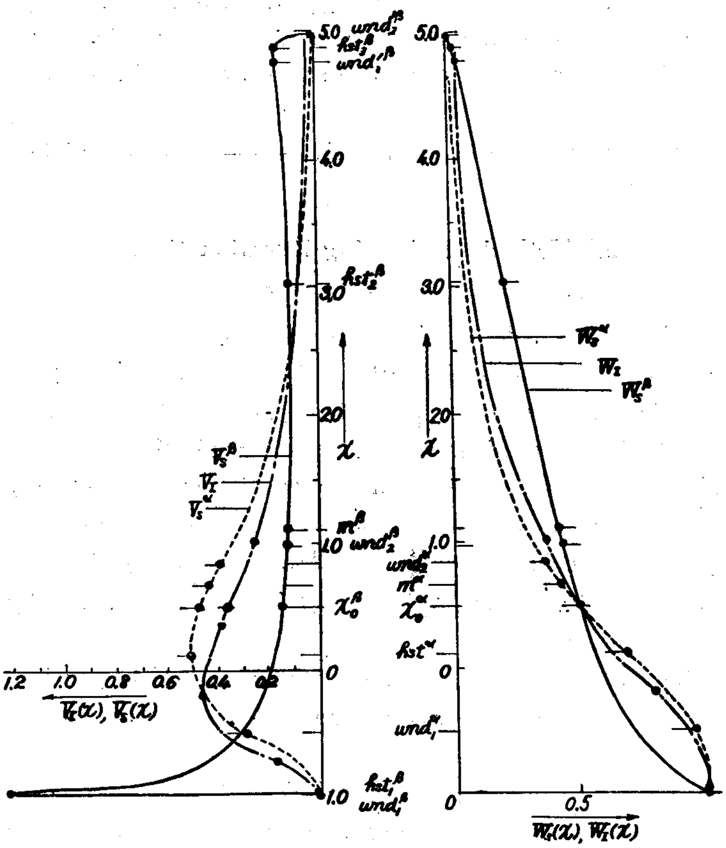

Fig. 2. Frequency Curves and Duration Curves of Partly- and Totally-Bounded Types

($V_I(x)$ and $W_I(x)$ by equations (2), (2′) when $x_0=0.5$, $\sqrt{2}c_0'=3.0$ and $b_0=1.0$, $V_S{}^{\alpha}(x)$ and $W_S{}^{\alpha}(x)$ by equations (3), (3′) when $x_0=0.5$, $\sqrt{2}c_0=3.0$, $b=1.0$, $g=5.0$ and $\sqrt{2}\lambda=1.43136$ and also $V_S{}^{\beta}(x)$ and $W_S{}^{\beta}(x)$ by equations (3), (3′) when $x_0=0.5$, $\sqrt{2}c_0=1.0$, $b=1.0$, $g=5.0$ and $\sqrt{2}\lambda=0.47712$)

In **Fig. 1**, V_B and W_B curves are shown when $\chi_0=1.5$ and $\sqrt{2}\,k=3.0$, the procedure of calculation being shown by the broken lines in the figure. In **Fig. 2**, V_I and W_I curves are shown when $\chi_0=0.5$, $\sqrt{2}\,c_0'=3.0$ and $b_0=1.0$. These curves are the same as the two curves in **Fig. 1**, only the origin being shifted to (-1.0). $V_S{}^{\alpha}$ and $W_S{}^{\alpha}$ curves represent the case when $\chi_0=0.5$, $\sqrt{2}\,c_0=3.0$, $b=1.0$, $g=5.0$ and $\sqrt{2}\,\lambda=1.43136$, while $V_S{}^{\beta}$ and $W_S{}^{\beta}$ curves, when $\chi_0=0.5$, $\sqrt{2}\,c_0=1$, $b=1.0$, $g=5.0$ and $\sqrt{2}\,\lambda=0.47712$.

It is an interesting new fact that even in the case of totally bounded type, if particular values are given to the various constants, then even though it may be similar to $V_S{}^{\alpha}$ in that it produces a frequency curve with one peak, contraverse skewness appears on the oposite side (not illustrated), and also that it sometimes produces a frequency curve with two peaks as in the case of $V_S{}^{\beta}$ (as illustrated).

Each characteristic value in each figure is the value calculated by each of the equations above.

2. Methods Estimating Duration Curves

Besides the two semi-empirical methods of Fuller[(1)] and Hazen[(2)] used in America to estimate flood discharges, the method by Goodrich's Skew Frequency Paper[(3)-(5)] and Foster's method[(6), (7)] based upon Pearson's Type Distribution, many methods based upon this distribution type have been proposed in many countries. First, Gibrat's semi-graphical trial method[(8)-(11)] based upon equation (2) and (2′) was examined, then Grassberger's methods[(12)-(15)]*, in which asymptotic Bruns' Φ-Series was adopted with no modification, were concisely corrected by applying Charlier's Type A Development to V_B of equation (1) and, furthermore, it was possible to propose the following new method by using the above-mentioned analytical results.

If the values of samples for variable χ_t are represented by x_t, then the equations proposed for the totally bounded type and the order of calculation are as follows.

As median χ_0, the value of x_0 located in the middle of the serial numbers of the values of samples arranged in order of their magnitudes, was taken, then the values of g and b were assumed. Then X_0 in equation (3) was calculated using these values.

$$\chi_0=x_0,\quad g \text{ and } b \text{ are assumed},\quad X_0=(x_0+b)/(g-x_0) \qquad \text{.........}\ (18)_1$$

* Furthermore, Grassberger proposed a new method, which was proposed in the Reference No. 16. Some other References are No. 17-19.

On the other hand, if symbol $l(x_i)$ is determined as the following equation and, first, its mean value l_0 and next, the standard deviation $\sigma(l)$ are calculated by all of the N sampled values, we get c_0 and K simultaneously, and also λ is calculated by X_0 obtained above.

$$\left.\begin{array}{ll} l(x_i)=\log\left\{(x_i+b)/(g-x_i)\right\}, & l_0=\sum_{i=1}^{N} l(x_i)/N, \\ \sigma(l)=\sqrt{\sum_{i=1}^{N}\left\{\left(l(x_i)\right)^2-l_0^2\right\}/N}, & c_0=1/\left(\sqrt{2}\,\sigma(l)\right), \\ K=2.30259/c_0, & \lambda=c_0\log X_0 \end{array}\right\} \quad \cdots\cdots (18)_2$$

Thus it was tested whether the mean, $m(\chi)$ obtained by calculating equation (15) contrarily from bottom upward using the above determined values of λ, K, g and b, is the same as the mean value $m(x)$ of the values of samples. If it is not, the trial calculation is repeated until this condition will be sufficed by changing the initially assumed values of g and b.

$$m(\chi) \text{ of equation (15)} = m(x) = \sum_{i=1}^{N} x/N \quad \cdots\cdots\cdots\cdots\cdots (18)_3$$

If satisfactory constants are finally found, the duration function can be estimated by the following equation with these constants.

$$\left.\begin{array}{ll} W_S(\chi)=\frac{1}{2}\left\{1-\Phi_0(\xi)\right\}, & \xi=c_0\, l(\chi)+\lambda, \\ l(\chi)=\log\left\{(\chi+b)/(g-\chi)\right\} & \end{array}\right\} \quad \cdots\cdots (18)_4$$

In the totally bounded distribution, suffice it to determine four constants, hence several different kinds of solution are produced in accordance with the various combination of $\chi_0=x_0$ and equation (12), (14), (15), (16) and (17). Among these, however, the above method seems to be the best one. This method should be appreciated as a precise method which is superior to Slade's method [20], [21] using the analogeous equation of standard deviation and Kimball's method [22] by Ratio Test.

Also for partly bounded type, various methods of solution were studied using various combinations of $\chi_0=x_0$ and equation (6), (8), (9) and (10) as was done above, and it was found that the solution by using equation (9), and ${}_m\mu_2$, ${}_m\mu_3$ in equation (10) was the same as that of Slade's method of moments [23].

Now the following method based upon an absolutely different principle from the above is proposed for partly bounded type.

In W_I it is sufficient to determine the three constants χ_0, b_0 and c_0'. First, an approximate value of χ_0 is obtained from the following equation.

$$\log \chi_0 = \sum_{t=1}^{N} x_t / N \quad \text{.............................} \quad (19)_1$$

In this case it is more theoretical to put $\chi_0 = x_0$ as in the case of equation (18), but better results are obtained by the above equation. If x_s and x_t are two sampled values which come in the same order when counted from the beginning and end of the serial numbers, respectively, hence we get the following relation from equation (2) and (2').

$$\xi_s' = c_0' \log\left\{(x_s + b_0)/(\chi_0 + b_0)\right\} = -c_0' \log\left\{(x_t + b_0)/(\chi_0 + b_0)\right\} = -\xi_t'$$

By solving this, b_0 is obtained from the following equation.

$$b_0 = (x_s x_t - \chi_0^2)/\left\{2\chi_0 - (x_s + x_t)\right\}, \quad \text{where} \quad s < N/2, \; t = N - (s-1)$$

(suffixes s and t are both the orders in the serial numbers when counted from the beginning)

.... $(19)_2$

Otherwise, b_0 may be determined as the arithmetic mean value of the values calculated by the above equation, by taking several pairs of x_s and x_t within the parts of the duration curve in which a better fittness is wanted, for example, according to Kimball's principle, the upper and lower ten percent may be adopted.

Thus c_0' is determined as follows by the formula of ordinary standard deviation.

$$c_0' = 1 \Big/ \sqrt{2 \sum_{t=1}^{N} \left[\log\left\{(x_t + b_0)/(\chi_0 + b_0)\right\}\right]^2 \Big/ (N-1)} \quad \text{......} \quad (19)_3$$

Otherwise c_0' may be obtained from,

$$c_0' = 2\xi_s' / \left\{\log(x_s + b_0) - \log(x_t + b_0)\right\}, \quad \text{.........} \quad (19')_3$$

wherein ξ_s' is given by $W_I(\xi_s') = (2s-1)/2N$

Lastly, duration curves can be deduced from the following equations.

$$W_I(\chi) = \frac{1}{2}\left\{1 - \Phi_0(\xi')\right\}, \quad \xi' = c_0' \log\left\{(x + b_0)/(\chi_0 + b_0)\right\} \quad \text{...} \quad (19)_4$$

This principle can be effectively applied also to totally bounded type.

In all of the above calculations, ratios of x and χ to the mean of x, $\mathrm{m}(x)$, were used instead of x and χ respectively, and by utilizing adequate tables, the calculations becomes very easy.

3. Several Examples of Application

First, the results of the application of the above methods to the records of annual flood discharges of Tone River at Kurihashi, from 1917 to 1941, will be explained as an example. After picking out the yearly maximum discharges of these 25 years and arranging them from the maximum to the minimum in the order of their magnitudes, we get Q_i given in **Table 1.** The arithmetic mean value of these Q_i's, $\mathrm{m}(Q)$, is 4 000.8 m³/sec and the ratio of Q_i to this $\mathrm{m}(Q)$ gives the values of $Q_i/\mathrm{m}(Q)$ in **Table 1.** We may proceed with the estimation taking the original sampled value Q_i as x_i, but in the following, $Q_i/\mathrm{m}(Q)$ is taken as x_i in orter to simplify the calculation. Hence, in this case, the variable χ should always be measured also in the unit of $1/\mathrm{m}(Q)$.

Table 1. The Record of Maximum Annual Flood Discharge of Tone River at Kurihashi, 1917-1941

Serial Number, i	Annual Flood Discharge, Q_i	Ratio to the Mean, $Q_i/\mathrm{m}(Q)=x_i$	Serial Number, i	Annual Flood Discharge, Q_i	Ratio to the Mean, $Q_i/\mathrm{m}(Q)=x_i$
	m³/sec			m³/sec	
1	10692	2.67246	14	3341	0.83508
2	9433	2.35777	15	3203	0.80058
3	6866	1.71615	16	3076	0.76884
4	5569	1.39197	17	2981	0.74510
5	5569	1.39197	18	2530	0.63237
6	5309	1.32698	19	2057	0.51414
7	5289	1.32198	20	1993	0.49815
8	4792	1.19776	21	1923	0.48065
9	4358	1.08929	22	1665	0.41616
10	4333	1.08303	23	1177	0.29419
11	4210	1.05228	24	960	0.23995
12	4209	1.05203	25	917	0.22920
13	3568	0.89182	Arithmetic Mean, $\mathrm{m}(Q)=4000.8$ m³/sec		

Totally bounded distribution;

Since $N=25$, the median in the sampled values of x_i becomes $x_{13}=0.89182$ and according to equation $(18)_1$, this value is immediately adopted as the value of χ_0, the median in this population. Next, the values of g and $(-b)$ are assumed with reference to the values of x_1 and x_{25}, the maximum and minimum values in **Table 1** respectively. Upon putting $g=6.85$ and $(-b)=-0.05$, we get

$X_0=0.15807$. Then calculating each value of the 25 $l(x_i)$'s by equation $(18)_2$, we get $l_0=-0.81202$ as the arithmetic mean value of them and also $\sigma(l)=0.3000$, $c_0=2.35702$ and $K=0.97691$. Finally $\lambda=1.88832$ is obtained from $\log X_0=-0.80115$, the logarithmic value of X_0, and the value of c_0 calculated above. In this case it should be $l_0=\log X_0$ theoretically, and this relation is tolerably satisfied by these values of l_0 and $\log X_0$. Next, we check whether the value of $\mathrm{m}(\chi)$, obtained by putting the above calculated values of K and λ into equation (15), will become 1 or not, since the value of $\mathrm{m}(x)$ in equrtion $(18)_3$ is naturally equal to 1. In this procedure, values of A_r and B_r in equation (15) can be calculated with the help of proper general mathematical tables, but we should calculate the asymytotic values of them separately because the rates of convergence of these alternative serieses are not the same. In this example, we get $\mathrm{m}(\chi)=1.038\sim1.021$ after obtaining $\varphi_0(\lambda)=0.03202$ and calculating the value of A_r up to its 18th term and B_r up to its 10th term.

This value of $\mathrm{m}(\chi)$ is very close to 1. But in many cases it is necessary to repeat such trial calculation procedure, changing the assumed initial values of g and $(-b)$ in equation $(18)_1$, in order to make the value of $\mathrm{m}(\chi)$ approach 1 to this extent. But if we get accustomed to this procedure, the calculation would be accomplished merely upon a couple of trials. The above obtained values of the constants were found as the best values after such trial calculation and with these values, we can represent the duration curve as follows in the form of equation $(18)_4$.

$$\left.\begin{aligned} W_S(\chi)=\frac{1}{2}\left\{1-\Phi_0(\xi)\right\},\quad & \xi=2.35702\ l(\chi)+1.88832, \\ l(\chi)=\log\left\{(\chi+0.05)/(6.85-\chi)\right\} & \end{aligned}\right\}\quad \cdots (20)$$

On the other hand upon obtaining the values of g and $(-b)$ by Kimball's Ratio Test [22], we also got the same values of $g=6.85$ and $(-b)=-0.05$ as the most suitable values. Hence the duration curve in this case will be obtained by substituting $\lambda=-c_0 l_0=1.91393$ into the above equation instead of $\lambda=c_0\log(X)^{-1}=1.88832$. Besides, upon following Slade's Method [20], we obtained $g=7.10$, $(-b)=-0.01$, $c_0=2.24950$ and $\lambda=1.94002$.

Fig. 3 shows the duration curves obtained by the above three methods drawn on Hazen's Logarithmic-Probability Paper. In this figure the black points illustrate the relation between x_i of **Fig. 1** and its probability on the samples, $(2i-1)/(2N)$, and from this we can investigate the goodness of fit of these presumed duration curves. Of the three curves, the curve obtained by the newly proposed method coincides with the curve obtained by Kimball's Method. However, a slight difference between the two results is recognized when the values of 10, 20, 50, 100, 500, 1 000 and 10 000 years flood discharges are calculated accurately by each method as shown in **Table 2.**

In this table the column marked with ※, on the contrary, represents flood years corresponding to a discharge of 10 000 m³/sec and all the values of the discharges are calculated multiplying each χ by m(Q)=4 000.8 m³/sec.

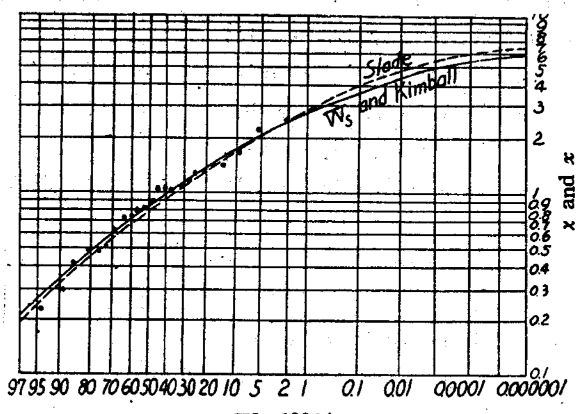

Fig. 3. Duration Curves estimated by Slade's, Kimball's and Iwai's Methods (Totally Bounded Type)

Partly bounded distribution;

The same data of River Tone is adopted to explain the case of the distribution of this type. Since $\log \chi_0 = -0.07950$ from equation $(19)_1$, the value of χ_0 becomes $\chi_0 = 0.83272$ which is pretty close to $x_{13} = 0.89182$. Then this value of χ_0 and two values of x_i, one taken from the maximum and the other from the minimum of **Table 1,** are put into equation $(19)_2$. In this example we can take 12 pairs of x_i's since $N=25$. But here, according to Kimball's Principle, attaching great importance to the goodness of fit within the domains of 10% from both the upper and lower limits, integer 3 which is the closest to the ratio of 25/10 is adopted and 3 values of b_0 are obtained from equation $(19)_2$ by using x_1 and x_{25}, x_2 and x_{24}, x_3 and x_{23} as the values of the 3 pairs of x_s and x_t. Then the arithmetic mean of the above obtained three values of b_0 is regarded as the final value of b_0. In this way the value of b_0 is obtained as $b_0 = 0.2496$. Finally the value of c_0' is obtained as $c_0' = 3.415$ from eqation $(19)_3$ and putting these three constants χ_0, b_0 and c_0' into equation $(19)_4$, we may represent the duration function as follows.

$$W_I(x) = \frac{1}{2}\left\{1 - \Phi_0(\xi')\right\}, \qquad \xi' = 3.415 \log(\chi + 0.2496) - 0.1173 \; \dots \; (21)$$

On the other hand, if we do the calculation according to the above principle adopting $\chi_0 = x_0 = x_{13} = 0.89182$ which is the median of the samples instead of χ_0

from equation $(19)_1$, we get $b_0=0.5756$ and $c_0=4.516$. If the result of this estimation is illustrated as W_I'-curve in comparison with W_I-curve, it becomes as shown in **Fig. 4** and the values in detail are as given in **Table 2.**

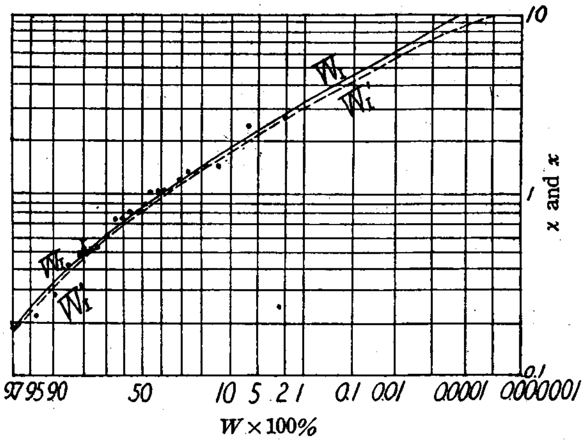

Fig. 4. Duration Curves estimated by Iwai's two Methods (Partly Bounded Type)

Comparing these two curves, W_I-curve is to be appreciated as more suitable than W_I'-curve and this fact tells us that we must hesitate to consider x_0, the median on the samples, as the value of χ_0, because this assumption is not always safe, especially when N is small.

In the following we will consider a different method apart from the above principle. If $\chi_0=x_0=x_{13}=0.89182$ is assumed, then we have only to estimate the remaining two constants, b_0 and c_0', in some way or other. For this purpose, b_0 is represented by μ_1, χ_0 and γ_1 from equation (9) and substituted into the σ^2 of equation (10), resulting in eliminating b_0 and establishing the following cubic equation of γ_1.

$$\gamma_1^3+\gamma_1^2-U\gamma_1+U=0, \quad \text{where } U=\left\{\sigma/(\mu_1-\chi_0)\right\}^2 \quad \text{.............. (22)}$$

Using the three relations, $\chi_0=x_0$, $\mu_1=\mathrm{m}(x)=\sum_{i=1}^{N}x_i/N$ and $\sigma=\sqrt{{}_m\mu_2}=\left\{\sum_{i=1}^{N}\left(x_i-\mathrm{m}(x)\right)^2/N\right\}^{0.5}$, U becomes a constant and this equation becomes solvable. Then the values of b_0 and c_0' can be easily calculated from equation (8) and (9). Solving the cubic equation by this method, $\gamma_1=1.08046$ was obtained as the most suitable root and hence $b_0=0.45270$ and $c_0'=4.13917$

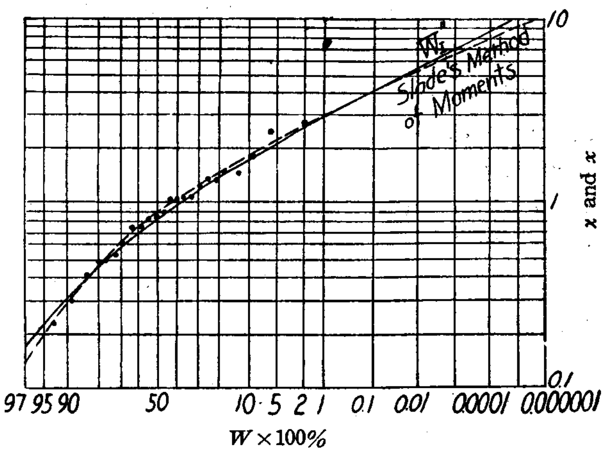

Fig. 5. Duration Curves estimated by Slade's and Iwai's Methods (Partly Bounded Type)

are obtained. Thus we can draw the duration curve shown as W_I'' in **Fig. 5**, of which precise values are tabulated in **Table 2**.

In this case, if we solve using equation (9) and the first two equations in equation (10), adopting $\mu_1=\sum_{i=1}^{N} x_i/N$, ${}_m\mu_2=\sum_{i=1}^{N}\left(x_i-\mathrm{m}(x)\right)^2/(N-1)^*$ and ${}_m\mu_3=\sum_{i=1}^{N}\left(x_i-\mathrm{m}(x)\right)^3/(N-1)^*$, exactly the same results as the results from Slade's Method of Moments [23] are obtained. According to this principle we got $\chi_0=0.903073$, $b_0=0.741201$ and $c_0'=4.810706$ and the duration curve became as shown in **Fig. 5** and the precise values as given in **Table 2**.

In short, when we select and solve any three of the following five equations, $\chi_0=x_0$, equations (6), (9) and (10), the three unknown constants, χ_0, b_0 and c_0' are decided simultaneously.

Thus in this manner, some other new methods of estimation were devised besides the above two. Good results, however, could not be obtained by the other method upon actual application on the same data.

Furthermore, Gibrat's Method by the semigraphical trial calculation [8]-[11] was applied and $\chi_0=0.94$, $b_0=0.7$ and $c_0'=5.19$ were obtained, this result being shown in **Fig. 6** and **Table 2** with the result of Grassberger's New Method [16]. Besides the above methods, Grassberger's old methods [12]-[14] and the combined methods of these and the writer's method were applied to the same example but good results were not acquired by most of them.

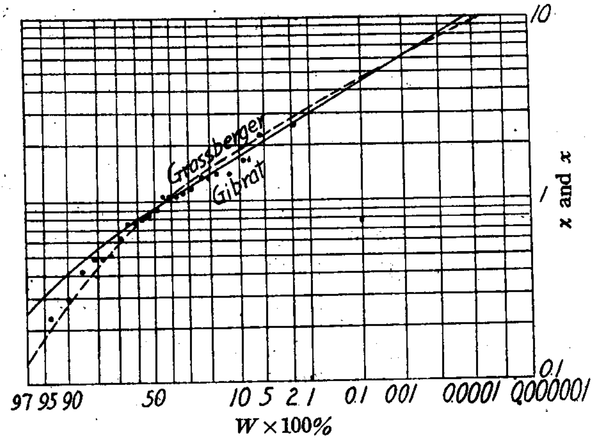

Fig. 6. Duration Curves estimated by Gibrat's Method (Partly Bounded Type) and Grassberger's Method

Besides the above-mentioned methods, all of which depend upon the Logarithmic Normal Distribution, Foster's Method [6] which is based upon Pearson's Distribution and Hazen's Menhod [2] which was found semi-experimentally in

* In these equations, both denominators are not N but (N-1) and this assumption is not theoretical because it allows a hypothesis of an unbiased distribution of the basic variable x. In the above calculation, however, (N-1) was taken in compliance with the principle of the original paper [23] and also in the calculations following explaining the same examples solved by the methods of Foster and Hazen.

America were applied and their duration curves are as shown in **Fig. 7.** Moreover, the results of these two methods are listed in the lower part of **Table 2** in comparison with the above results together with Goodrich's Method [3], [4] utilizing the special skew frequency paper and Fuller's experimental Method [1]. Among them the discharge obtained by Fuller's Formula, in which T' does not mean the statistically correct value of flood year T,

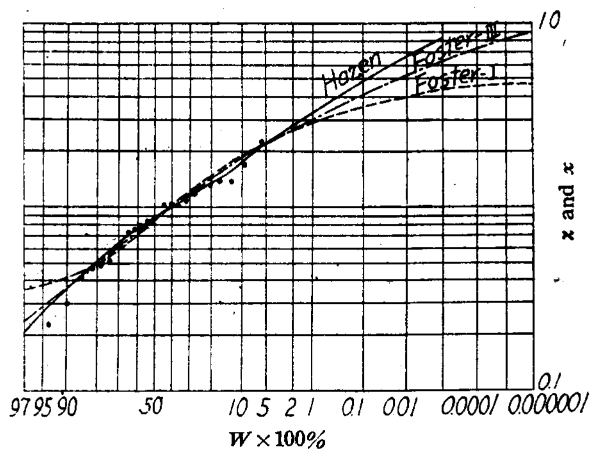

Fig. 7. Duration Curves estimated by Foster's-III Method (Partly Bounded Type), Foster's-I Method (Totally Bounded Type) and Hazen's Method

Table 2. Estimated Flood Discharges of Tone River at Kurihashi by Various Methods.

Method \ Flood Years $T=1/W$	10 yrs.	20 yrs.	50 yrs.	100 yrs.	500 yrs.	1 000 yrs.	10 000 yrs.	※10 000 m^3/sec
	m^3/sec	m^3/sec	m^3/sec	m^3/sec	m^3/sec	m^3/sec	m^3/sec	yrs.
Logarithmic Normal Distribution, Totally Bounded: Iwai, W_S (1948)	7446.3	8908.0	10706.2	11969.8	14591.0	15589.9	18393.7	34.6
Logarithmic Normal Distribution, Totally Bounded: Kimball (1938)	7308.9	8755.9	10541.5	11799.8	14419.1	15418.9	18241.2	37.6
Logarithmic Normal Distribution, Totally Bounded: Slade (1934)	7288.6	8807.3	10703.5	12050.8	14874.4	15957.2	19008.9	35.3
Logarithmic Normal Distribution, Partly Bounded: Iwai, W_I (1944)	6961.4	8481.7	10522.1	12122.4	16083.2	17883.6	23605.0	39.5
Logarithmic Normal Distribution, Partly Bounded: Iwai, W_I' (1944)	7015.9	8320.0	10007.5	11278.8	12556.9	15584.7	20136.0	48.0
Logarithmic Normal Distribution, Partly Bounded: Iwai, W_I'' (1948)	7093.9	8462.1	10254.7	11620.4	14496.7	16332.2	21422.2	43.9
Logarithmic Normal Distribution, Partly Bounded: Slade (1934)	7185.6	8514.1	10217.5	11811.5	14460.0	15756.6	20196.3	23.5
Logarithmic Normal Distribution, Partly Bounded: Gibrat (1932)	7222.8	8812.5	10970.8	12676.4	16923.0	18890.4	26103.3	33.3
Logarithmic Normal Distribution: Grassberger (1936)	7377.4	8919.4	11025.0	12974.1	16801.0	18718.7	26275.3	32.1
Pearson Type Distribution: Foster, III, (1924)	7246.7	8703.8	11953.6	12036.8	16125.2	16579.3	20842.6	28.9
Pearson Type Distribution: Foster, I, (1924)	7453.9	8900.2	11701.9	11780.0	14315.3	14596.9	16456.9	28.8
Hazen, (1930)	7778.8	8819.7	11660.3	12608.5	18076.8	18684.5	26911.4	26.6
Goodrich, II, (1926) $(fR)=(R-0.18)$	7088.9	8411.4	10001.8	11188.8	13671.0	14720.9	17928.2	49.0
Goodrich, V, (1926) $(fR)=(R-0.19)/(4.8-R)$	7970.8	9190.2	10433.4	11223.6	12569.8	13037.8	14189.5	34.9
Fuller, (1913) $(1+0.8\log T')$	7201.4	8161.6	9437.9	10420.1	12638.5	13603.0	16803.3	

is so different in nature compared to that obtained by all other methods that we cannot compare it precisely with the others. It was cognizable on not only the example of Tone River but also on many of the other rivers that the value of the coefficient in Fuller's Formula, 0.8, is too small and it must be enlarged to the extent of 1.3 or 1.4 in Japan.

After applying these various methods to many Japanese rivers, we found, as can be understood from **Table 2**, that the more statistical methods produce the better results and, among them, the methods based upon the Logarithmic Normal Distribution are superior to the others and that the above proposed method of W_S is the best one in totally bounded type and W_I in partly bounded type. Furthermore, such studies indicated to us that the conditions under which floods occur at most rivers in Japan are quite different from the conditions in foreign countries and the methods of estimating flood discharges in our river control works are yet in general very rudimentary. Thus it was made clear that researches in this line is of absolute necessity in the future.

Since we obtained $b_0=0.2496>0$ for the above case of W_I of the partly bounded type, the value of the lower limit $(-b_0)$ becomes negative, resulting in giving a negative minimum flood discharge. Therefore, if we persist in obtaining a value of b_0 which will give a positive minimum flood discharge, we must select the values which will satisfy the condition $b_0<0$ from all of the values of b_0 calculated from equation $(19)_2$ with various pairs of x_s and x_t. Then all pairs of values of c_0 which are calculated by putting the above selected values of b_0 into equations $(19)_3$ and $(19')_3$ are compared. Finally the best value of b_0 is adopted by picking up the value of b_0 which will make the two values of c_0 become the closest to each other. According to this procedure we can discriminate and find the most suitable values of b and g in the totally bounded type.

These methods of estimating duration curves proposed above were very useful in various problems besides the problem pertaining to annual flood discharges. For instance, duration curves, the variable of which was taken as the rainfall intensity i which changes with each different rainfall duration time t_r, were estimated according to the principle of equation (19). Then a new method of statistically estimating a curved surface formed by these curves was proposed by equation of the following form,

$$i=K_1 t_r^{-K_2}+K_3 t_r^{-K_4}, \quad \text{.............................. (23)}$$

where K_1, K_2, K_3 and K_4 are constants and two of them, K_1 and K_2 are decided corresponding to the probability W.

Hence we can easily draw a storm rain intensity curve separately corresponding to any arbitrary value of W. These Probability Storm Rain Curves are very useful in estimating the above mentioned probable flood discharges from rainfall, but here, a rational and economical method of planning sewer net was accomplished applying this estimation of probable rainfall into Vicari-Hauff's Method and upon application of this method to some real examples very interesting results were obtained.

Also duration curves of daily discharges for one year of many representative rivers in Japan were estimated according to the principle of equation (19) and the characteristics of the varying conditions of the discharge were comprehended by means of several statistical costants. **Fig. 8** shows these duration curves and their equations for the basin area of 100 km^2 of 11 representative rivers of Japan.

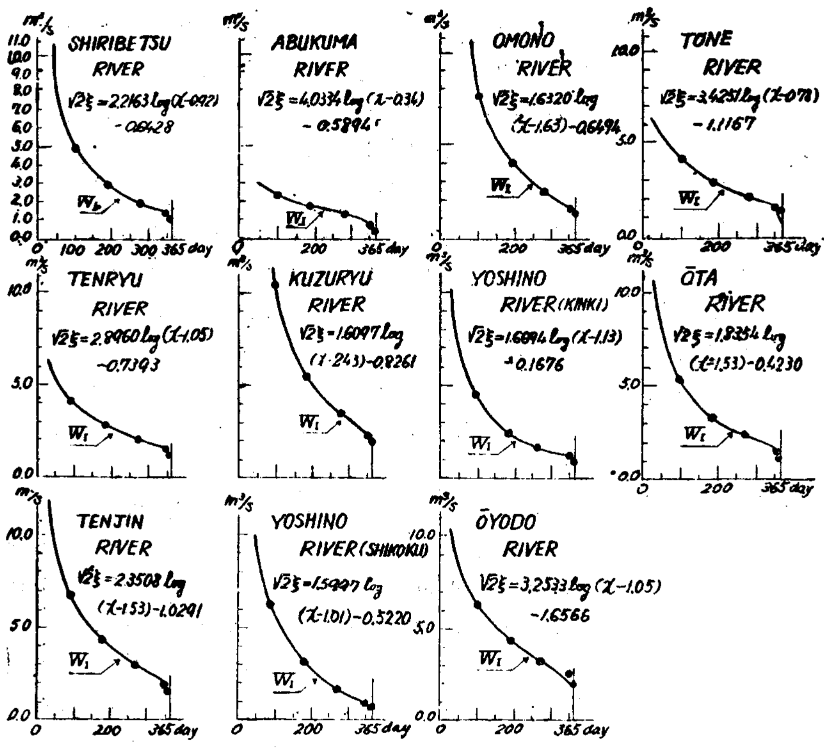

Fig. 8. Duration Curves of the Daily Discharge for one Year of 11 Representative Rivers in Japan

Furthermore, the duration curve based upon the totally bounded distribution type, is regarded recently as being very important not only for the aforementioned hydrological data but also for other problems in which the existence of both the upper and lower limits must be provided in the frequency curve. But the writer applied this to the gradation curve of soil. In this case, assuming that the densities of all soil particles contained in a scoop of soil were uniform, the weight of each particle should be proportional to the cube of the diameter. Hence, it could be readily presumed that the gradation curve pertaining to weights of particles could be theoretically deduced from a duration curve which is the integration curve of a frequency curve drawn corresponding to all sizes of particles and the number of particles represented by the size.

Applying basic equation (1) and (1′) and following the above principle, Grassberger [14] has succeeded in representing this gradation curve in the form of $1-W=\frac{1}{2}(1+\Phi_0)$. The writer developing this method, also successfully deduced a function of the same form in the case of equations (3) and (3′). Thus a new statistical method of estimating this curve by trial was successfully derived for any given soil sample and new appreciable results were obtained by investigating the correlation between statistical constants and physical properties of many soil examples in Japan. Hereafter, it is expected that these results will prove to be very useful in deciding the proper gradation of concrete aggregate and filter sand in water filtration plants or in understanding their characteristics. Also this method of estimation will be considered very effective in investigating statistically the factor of safety of structures [24]. The writer is now engaged in studying these interesting problems.

4. Conclusions

(1) The so-called Slade's Type Distribution, that is Logarithmic Normal Distribution, was adopted as the statistical asymmetric distribution of hydrological data, and then the distribution was divided into two types, namely, one with a limiting value on one side only and the other with limiting values on both sides. Frequency curves and duration curves were analyzed mathematically in each type and then their asymmetricities were distinctly determined.

(2) Furthermore, upon studying all past methods of estimating duration curves based upon not only this type of distribution but also many other types, it was found that the method based upon this Slade's Type Distribution was the most suitable, and new methods have been proposed by using the above-mentioned analytical results.

(3) Upon practical application of all of the above methods mainly to Japanese data, it was found that, in exactness, equation (18), and, in practice, equation (19) gave the best results. Thus many remarkable results were obtained

concerning flood discharges, river flow conditions and precipitations, and in connection with statistical economics, these results proved to be useful guides in planning various water works, such as river improvement, water supply, water power and sewerage works. Moreover, besides such hydrological problems, these theories and methods will become very effective in other fields, such as soil properties and factors of safety in structures.

(4) An extensive study was theoretically made pertaining to those statistically estimated duration curves, the following being the brief results.

In examining the Reliability of duration curves estimated statistically, Pearson's Chi-Square Method which considered the Number of Degrees of Freedom and v. Mises' Omega-Square Method [25] with which the duration function can be tested directly were applied, giving up the former idea based upon the old theory of errors.

In the case of Small Samples, however, it was found that they all lacked preciseness. As the method effective in this case, a new precise method of testing the Measure of Goodness of Fit based upon polynomial distribution was devised. This result should be appreciated as cultivating a new Stochastical Field in such problems on duration curves.

The writer was very fortunate recently to obtain the new thesis [26] of Prof. Thomas, Harvard University, and admired his new conception and theory and in the near future hopes to make a discussion on it from various points of view.

The above results summerized in this paper had been got under the kind instruction of Prof. Tojiro Ishihara, Kyoto University, and already been published in Japan since 1944. The writer thanks him very much and earnestly expects the world-wide criticism to this paper.

References

(1) W. E. Fuller: Flood Flows, Trans. Amer. Civ. Eng., vol. 77, (1914).

(2) A. Hazen: Flood Flows, New York, 1930.

(3) R. D. Goodrich: Straight-line plotting of skew-frequency data, Trans. A. S. C. E., vol. 91, (1927)

(4) R. M. Harris: Straight-line treatment of hydraulic duration curves, Univ. Washington Eng. Exper. Sta. Bull., No. 65, (1932).

(5) F. Reinhold: Einfache zeichnerische Darstellung von Dauerkurven, Die Wasserwirtsch., Nr. 13, (1933).

(6) H. A. Foster: Theoretical frequency curves and their application to engineering problems, Trans. A. S. C. E., vol. 87, (1924).

(7) D. L. Sokolowski: Anwendung der Verteilungskurven auf dem Gebiete der Hydrologie, Wasserwirtsch. u. Technik, Heft 10-11, 12, 13-14, 18-19, (1936).

(8) R. Gibrat: Sur l'ajustement mathématique des courbes de débits d'un cours d'eau, Revue Générals de l'Electricité, 14, Mai, (1932) and Comptes Rendus, H. d. S. de l'Académie des Scieces, 1 Sem., Tome 194, (1932).

(9) R. Gibrat: Aménagement hydroélectique des cours d'eau, Rev. Gén. de l'Elect., 15 et 22, Oct., (1932) and Rev. Gén. de l'Hydraulique, No. 11, (1936).

• (10) A. Coutagne: Comment caractériser la variabilité intrinsèque des pércipitation, Rev. Gén. de l'Hyd., 18, (1937).

(11) W. Kummer: Die Auswertung unsymmetrischer Verteilungsreihen der Grosszahlforschung, Schweiz. Bauzeit., Nr. 11, (1933).

(12) H. Grassberger: Die Anwendung der Wahrscheinlichkeitsrechnung auf die Wasserführung der Gewässer, die Wasserwirtsch., Nr. 1-6, (1932).

(13) H. Grassberger: Untersuchung über die Hochwässer der Jangtsekiang, Die Wasserwirtsch., Nr. 6, (1933).

(14) H. Grassberger: Die Aufbau der Böden: Die Wasserwirtsch., Nr. 17-19, (1933).

(15) F. Schaffernak: Hydrographie, Wien, 1935, S. 315-316, S. 321-324.

(16) H. Grassberger: Die Anwendung der Wahrscheinlichkeitsrechnung auf Hochwasserfragen, Deutsche Wasserwirtsch., Nr. 9-10, (1936).

(17) C. Reindl: Zur Voraussage grösster Hochwässer, Wasserkr. u. -wirtsch., Heft 13/14, (1938).

(18) V. Felber: Die Anwendung der Wahrscheinlichkeitsrechnung auf die Berechnung von Regenlinie, Wassrkr. u. -wirtsch., Heft 1/2, (1939).

(19) V. Felber: Neues Verfanren zur Berechnung der Hochwassermengen aus den Niederschlägen, Deutsche Wasserwirtsch., Nr. 2-3, (1939).

(20) J. J. Slade: An asymmetric probability function, Trans. A. S. C. E., vol. 62, (1936). p. 58.

(21) T. Saville: Floods in the United States, Geol. Survey Water-Supply Paper, 771, (1936), pp. 398-432.

(22) B. F. Kimball: Probability-distribution-curve for floodcontrol studies, Trans. Amer. Geophys. Union, I, (1938), pp. 460-483.

(23) (20) pp. 47-48.

(24) A. M. Freudenthal: The safety of structures, Proc. A. S. C. E. vol. 71, (1945).

(25) v. Mises: Wahrscheinlichkeitsrechnung, Leipzig und Wien, 1931, S. 316-335.

(26) H. A. Thomas, Jr.: Frequency of minor floods, Journal of the Boston Soc. of Civ. Eng., vol. 35, (1948).