Journal of Hydroscience and Hydraulic Engineering
Vol. 4, No. 1, April, 1986, pp.79-97.

STUDIES ON MULTIVARIATE CONDITIONAL MAXIMUM ENTROPY DISTRIBUTION AND ITS CHARACTERISTICS

BY

Noriaki Sogawa, Masao Araki, and Toshihiko Imai

Department of Civil Engineering
Shinshu University, Nagano 380, Japan

SYNOPSIS

The multivariate conditional maximum entropy (MCME) distribution is useful when we estimate the occurrence probability of any variates given the occurrence of the other variates, particularly in the area of scarce hydrological data. The theoretical equation of the MCME distribution with the arbitrary function $g_r(\cdot)$ and the moments as the constraint conditions is derived from the multivariate maximum entropy distribution. It is discussed that the maximum entropy distribution is based on the same level as the Pearson's system of frequency-curves and the Gram-Charlier's series because some distributions are derived by this distribution with the concrete $g_r(\cdot)$ as the constraint condition. Finally, the MCME distribution is applied to the annual rainfall and annual maximum daily rainfall, and the applicability of this distribution to the hydrological data is investigated in detail.

INTRODUCTION

A conditional probability density function is necessary when we estimate the occurrence probability of any variates given the occurrence of the other variates. The conditional probability density functions are available for two-variate normal distribution (Kanda and Fujita (5)) and two-variate gamma distribution (Nagao and Kadoya (7,8)). They are one-variate conditional probability density function given the occurrence of the other one variate. A multivariate probability density function is necessary in order to derive a multivariate conditional probability density function. But only a multivariate normal distribution (Takeuchi (18)) and a multivariate gamma distribution (Krishnamoorthy and Partharathy (6)) have been studied.

Though the multivariate gamma distribution corrects the defect of the multivariate normal distribution, which can express the only symmetrical shape, and applies widely to the population of the non-symmetrical shape, it is difficult to introduce it into the hydrologic frequency analysis because the identification method of the parameters from data is not proved. Besides, it is difficult to select the distribution we should adopt because in many cases hydrological data are scarce.

This paper describes the MCME distribution and its characteristics in order to overcome the problems of the application to the population of the non-symmetrical shape and the selection of distribution in the area of scarce hydrological data.

Former Researches and Investigations

Sonuga (15) in 1972 introduced the concept of entropy into the hydrologic frequency analysis, and proposed the method to estimate the probability density function from the principle of maximum entropy. He derived the probability density function p(x) of a variate x by maximizing the entropy H(x) to be defined in eq. 1 preserving the first two moments of the given data.

$$H(x)=-\int p(x)\ln p(x)dx \quad (1)$$

The maximum entropy distribution itself doesn't have the proper form and the form of this distribution depends on the constraint conditions (how to get the information from data) under which eq. 1 is maximized. The introduction of this distribution in hydrologic frequency analysis should, therefore, be highly evaluated because it gave the theoretical validity to the adoption of distribution. He (16) then derived one-variate maximum entropy distribution given the occurrence of the other one variate by maximizing the mutural entropy $H(x,y)$ to be defined in eq. 2 preserving the first two moments of the given data.

$$H(x,y)=-\int\int p(x,y)\ln p(x,y)dxdy \quad (2)$$

where $p(x,y)$ is the joint probability density function of two variates x and y. But it is very difficult to adopt more than three moments in his models.

We introduced Wragg and Dowson's technique (19) which was developed in the field of information theory, and studied the characteristics of one-variates maximum entropy distribution based on the arbitrary number of moments with the applicability to hydrological data (Sogawa and Araki (11), Sogawa, Araki and Kobayashi (14)). We then extended them to two-variate (Sogawa, Araki and Terashima (12)) and multivariate (Sogawa, Araki and Sato (13)) maximum entropy distributions.

The Purpose and the Outline of This Paper

This paper describes the MCME distribution using the multivariate maximum entropy distribution.

We show the theoretical equation given the arbitrary function $g_r(\cdot)$ as the constraint condition and develop the theory replacing them with moments. Besides, some conditional distributions are derived from the MCME distributions, and the level of the maximum entropy distribution is discussed in the field of the probability distribution. Finally, the applicability to hydrological data is investigated in detail.

THEORY

The Case of the Constraint of the Arbitrary Function $g_r(\cdot)$

Consider firstly a multivariate continuous probability density function in which the variates $x_1, x_2, \ldots, x_n$ assume the values $x_1, x_2, \ldots, x_n$ with probability $p(x_1,x_2,\ldots,x_n)$. The entropy $H(x_1,x_2,\ldots,x_n)$ of this distribution is defined by

$$H(x_1,x_2,\ldots,x_n)=-\int\cdots\int\int p(x_1,x_2,\ldots,x_n)\ln p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n \quad (3)$$

The expectations of this distribution of the arbitrary function $g_r(x_1,x_2,\ldots,x_n)$ for r=1, 2, ..., M are defined by

$$\int\cdots\int\int g_r(x_1,x_2,\ldots,x_n)p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n=E[g_r(x_1,x_2,\ldots,x_n)],\quad r=1,2,\ldots,M \quad (4)$$

where $E[\cdot]$ is the mathematical expectation operater. Eq. 1 contains the normalization condition $g_r(x_1,x_2,\ldots,x_n)=1$. By maximizing eq. 3 subject to eqs. 4 (the principle of maximum entropy), the multivariate maximum entropy distribution is given by Sogawa, Araki and Sato (13) as

$$p(x_1,x_2,\ldots,x_n)=\exp\{-1-\sum_{r=1}^{M}\lambda_r g_r(x_1,x_2,\ldots,x_n)\} \quad (5)$$

where λ_r is the Lagrangian multiplier associated with the normalization and main constraint conditions. Eq. 5 is the general multivariate maximum entropy distribution. We can give $g_r(\cdot)$ the various concrete functions whose expectations exist.

From eq. 5 we see that the MCME distribution is simply of the form

$$p_{x_1x_2\ldots x_m}(x_{m+1},x_{m+2},\ldots,x_n)=\frac{p(x_1,x_2,\ldots,x_n)}{\int\cdots\int p(x_1,x_2,\ldots,x_n)dx_{m+1}dx_{m+2}\ldots dx_n}$$
$$=\frac{\exp\{-1-\sum_{r=1}^{M}\lambda_r g_r(x_1,x_2,\ldots,x_n)\}}{\int\cdots\int\exp\{-1-\sum_{r=1}^{M}\lambda_r g_r(x_1,x_2,\ldots,x_n)\}dx_{m+1}dx_{m+2}\ldots dx_n}$$
$$=\frac{\exp\{-\sum_{r=1}^{M}\lambda_r g_r(x_1,x_2,\ldots,x_n)\}}{\int\cdots\int\exp\{-\sum_{r=1}^{M}\lambda_r g_r(x_1,x_2,\ldots,x_n)\}dx_{m+1}dx_{m+2}\ldots dx_n} \quad (6)$$

which is refered as the general MCME distribution.

The Case of the Constraint of Moments

Given variate x_i $(0<x_i<\infty)$, we calculate the moments about the origin ${}_1\mu_{a_1}$, ${}_2\mu_{a_2}$, $\ldots$, ${}_n\mu_{a_n}$, $\mu_{b_1b_2\ldots b_n}$, and put them into eq. 4, namely,

$$\int_0^\infty\cdots\int_0^\infty\int_0^\infty x_1^{a_1}p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n={}_1\mu_{a_1},\quad a_1=1,2,\ldots,N_{a_1} \quad (7.1)$$
$$\int_0^\infty\cdots\int_0^\infty\int_0^\infty x_2^{a_2}p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n={}_2\mu_{a_2},\quad a_2=1,2,\ldots,N_{a_2} \quad (7.2)$$
$$\cdots$$
$$\int_0^\infty\cdots\int_0^\infty\int_0^\infty x_n^{a_2}p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n={}_n\mu_{a_n},\quad a_n=1,2,\ldots,N_{a_n} \quad (7.n)$$
$$\int_0^\infty\cdots\int_0^\infty\int_0^\infty x_1^{b_1}x_2^{b_2}\ldots x_n^{b_n}p(x_1,x_2,\ldots,x_n)dx_1dx_2\ldots dx_n=\mu_{b_1b_2\ldots b_n},$$
$$b_1=0,1,\ldots,N_{b_1},\ b_2=0,1,\ldots,N_{b_2},\ \ldots,\ b_n=0,1,\ldots,N_{b_n} \quad (7.n+1)$$

where $\mu_{00\ldots0}=1$ expresses the normalization condition. Besides, we exclude the case that b_i $(i=1,2,\ldots,n)$ takes integer more than 1 and the other b_j $(j=1,2,\ldots,i-1, i+1,\ldots,n)$ takes 0 in eqs. 7.n+1 because this information (constraint condition) is contained in eqs. 7.1 - 7.n.

The multivariate maximum entropy distribution, subject to constraint conditions 7.1 - 7.n+1 can be determined by the principle of maximum entropy as

$$p(x_1,x_2,\ldots,x_n)=\exp\{-1-\sum_{a_1=1}^{N_{a_1}}{}_1\gamma_{a_1}x_1^{a_1}-\sum_{a_2=1}^{N_{a_2}}{}_2\gamma_{a_2}x_2^{a_2}-\ldots-\sum_{a_n=1}^{N_{a_n}}{}_n\gamma_{a_n}x_n^{a_n}-\sum_{b_n=0}^{N_{b_n}}\cdots\sum_{b_2=0}^{N_{b_2}}\sum_{b_1=0}^{N_{b_1}}\delta_{b_1b_2\ldots b_n}x_1^{b_1}x_2^{b_2}\ldots x_n^{b_n}\} \quad (8)$$

where ${}_1\gamma_{a_1}$, ${}_2\gamma_{a_2}$, $\ldots$, ${}_n\gamma_{a_n}$ and $\delta_{b_1b_2\ldots b_n}$ are the Lagrangian mulipliers. They are determined by the use of iteration method (Sogawa, Araki and Sato (13)).

In the case with constraint of moments, eq. 6 becomes

$$p_{x_1x_2\ldots x_m}(x_{m+1},x_{m+2},\ldots,x_n)$$
$$\exp(-\sum_{a_{m+1}=1}^{N_{a_{m+1}}}{}_{m+1}\gamma_{a_{m+1}}x_{m+1}^{a_{m+1}}-\sum_{a_{m+2}=1}^{N_{a_{m+2}}}{}_{m+2}\gamma_{a_{m+2}}x_{m+2}^{a_{m+2}}-\ldots-\sum_{a_n=1}^{N_{a_n}}{}_n\gamma_{a_n}x_n^{a_n}$$
$$-\sum_{b_n=0}^{N_{b_n}}\cdots\sum_{b_2=0}^{N_{b_2}}\sum_{b_1=0}^{N_{b_1}}\delta_{b_1b_2\ldots b_n}x_1^{b_1}x_2^{b_2}\ldots x_n^{b_n})\Big/$$
$$(b_1+b_2+\ldots+b_n\neq0)$$

$$\int_0^\infty \cdots \int_0^\infty \int_0^\infty \exp\Big(-\sum_{a_{m+1}=1}^{Na_{m+1}} {}_{m+1}\gamma_{a_{m+1}} x_{m+1}^{a_{m+1}} - \sum_{a_{m+2}=1}^{Na_{m+2}} {}_{m+2}\gamma_{a_{m+2}} x_{m+2}^{a_{m+2}} - \cdots - \sum_{a_n=1}^{Na_n} {}_n\gamma_{a_n} x_n^{a_n}$$
$$- \sum_{\substack{b_n=0 \\ (b_1+b_2+\ldots+b_n\neq 0)}}^{Nb_n} \cdots \sum_{b_2=0}^{Nb_2} \sum_{b_1=0}^{Nb_1} \delta_{b_1 b_2 \ldots b_n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}\Big) dx_{m+1} dx_{m+2} \cdots dx_n \quad (9)$$

Note that $\delta_{00\ldots0}$ is reduced here both in the numerator and the denominator because it is not related to integral.

We express the distribution given by eq. 9 as ${}_m^nM(Na_1,Na_2,\ldots,Na_n;Nb_1,Nb_2,\ldots,Nb_n)$.

Derivation of Some Concrete Distribution

Various MCME distributions can be derived by eq. 6 assuming concrete function for $g_r(\cdot)$. The following are some conditional distributions derived from eq. 9 with constraint of moments for different number of variates and moments.

1) ${}_1^2M(Na_1,1;0,0)$

$$p_{x_1}(x_2)=\frac{\exp(-{}_2\gamma_1 x_2)}{\int_0^\infty \exp(-{}_2\gamma_1 x_2)dx_2}={}_2\gamma_1\exp(-{}_2\gamma_1 x_2) \quad (10)$$

where ${}_2\gamma_1>0$.

2) ${}_1^2M(Na_1,1;1,1)$

$$p_{x_1}(x_2)=\frac{\exp(-{}_2\gamma_1 x_2-\delta_{11}x_1x_2)}{\int_0^\infty \exp(-{}_2\gamma_1 x_2-\delta_{11}x_1x_2)dx_2}=({}_2\gamma_1+\delta_{11}x_1)\exp(-{}_2\gamma_1 x_2-\delta_{11}x_1x_2) \quad (11)$$

where ${}_2\gamma_1+\delta_{11}x_1>0$ in all x_1.

3) ${}_1^2M(Na_1,2;1,1)$

$$p_{x_1}(x_2)=\frac{\exp(-{}_2\gamma_1 x_2-{}_2\gamma_2 x_2^2-\delta_{11}x_1x_2)}{\int_0^\infty \exp(-{}_2\gamma_1 x_2-{}_2\gamma_2 x_2^2-\delta_{11}x_1x_2)dx_2}$$
$$=2\sqrt{\frac{{}_2\gamma_2}{\pi}}\;\frac{\exp\{-({}_2\gamma_1+\delta_{11}x_1)^2/(4\,{}_2\gamma_2)-{}_2\gamma_1x_2-{}_2\gamma_2x_2^2-\delta_{11}x_1x_2\}}{\operatorname{erfc}\{({}_2\gamma_1+\delta_{11}x_1)/(2\sqrt{{}_2\gamma_2})\}} \quad (12)$$

where $\operatorname{erfc}(z)=\frac{2}{\sqrt{\pi}}\int_z^\infty \exp(-t^2)dt$ and ${}_2\gamma_2>0$.

4) ${}_1^2M(Na_1,2;2,2)$

$$p_{x_1}(x_2)=\frac{\exp(-{}_2\gamma_1x_2-{}_2\gamma_2x_2^2-\delta_{11}x_1x_2-\delta_{12}x_1x_2^2-\delta_{21}x_1^2x_2-\delta_{22}x_1^2x_2^2)}{\int_0^\infty \exp(-{}_2\gamma_1x_2-{}_2\gamma_2x_2^2-\delta_{11}x_1x_2-\delta_{12}x_1x_2^2-\delta_{21}x_1^2x_2-\delta_{22}x_1^2x_2^2)dx_2}$$
$$=2\sqrt{({}_2\gamma_2+\delta_{12}x_1+\delta_{22}x_1^2)/\pi}\;\exp[-({}_2\gamma_1+\delta_{11}x_1+\delta_{21}x_1^2)^2/\{4({}_2\gamma_2+\delta_{12}x_1+\delta_{22}x_1^2)\}-{}_2\gamma_1x_2-{}_2\gamma_2x_2^2-$$
$$\delta_{11}x_1x_2-\delta_{12}x_1x_2^2-\delta_{21}x_1^2x_2-\delta_{22}x_1^2x_2^2]\big/\operatorname{erfc}\{({}_2\gamma_1+\delta_{11}x_1+\delta_{12}x_1+\delta_{21}x_1^2)/(2\sqrt{{}_2\gamma_2+\delta_{12}x_1+\delta_{22}x_1^2})\} \quad (13)$$

where ${}_2\gamma_2+\delta_{12}x_1+\delta_{22}x_1^2>0$ in all x_1.

5) ${}_{n-1}^{\ n}M(Na_1,Na_2,\ldots,Na_{n-1},2;\underbrace{1,1,\ldots,1}_{\longleftarrow n\longrightarrow})$

$$p_{x_1x_2\ldots x_{n-1}}(x_n)=\exp(-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_{n-1}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1}x_1^{b_1}x_2^{b_2}\cdots x_{n-1}^{b_{n-1}}x_n^{b_n})\Big/\int_0^\infty \exp(-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_{n-1}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1}x_1^{b_1}x_2^{b_2}\cdots x_{n-1}^{b_{n-1}}x_n^{b_n})dx_n$$
$$=2\sqrt{{}_n\gamma_2/\pi}\,\exp\{-({}_n\gamma_1+\sum_{b_{n-1}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1}x_1^{b_1}x_2^{b_2}\cdots x_{n-1}^{b_{n-1}})^2/(4\,{}_n\gamma_2)-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_{n-1}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1}x_1^{b_1}x_2^{b_2}\cdots x_{n-1}^{b_{n-1}}x_n\}\Big/\mathrm{erfc}\{({}_n\gamma_1+\sum_{b_{n-1}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1}x_1^{b_1}x_2^{b_2}\cdots x_{n-1}^{b_{n-1}})/(2\sqrt{{}_n\gamma_2})\} \qquad (14)$$

where ${}_n\gamma_2>0$.

6) ${}_{n-2}^{\ n}M(Na_1,Na_2,\ldots,Na_{n-2},2,2;\underbrace{1,1,\ldots,1}_{\longleftarrow n\longrightarrow})$

$$p_{x_1x_2\ldots x_{n-2}}(x_{n-1},x_n)=\exp(-{}_{n-1}\gamma_1x_{n-1}-{}_{n-1}\gamma_2x_{n-2}^2-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_n=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_n}x_1^{b_1}x_2^{b_2}\cdots x_n^{b_n})\Big/\int_0^\infty\int_0^\infty \exp(-{}_{n-1}\gamma_1x_{n-1}-{}_{n-1}\gamma_2x_{n-1}^2-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_n=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_n}x_1^{b_1}x_2^{b_2}\cdots x_n^{b_n})dx_{n-1}dx_n$$
$$=2\sqrt{{}_{n-1}\gamma_2/\pi}\,\exp(-{}_{n-1}\gamma_1x_{n-1}-{}_{n-}\gamma_2x_{n-1}^2-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_n=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_n}x_1^{b_1}x_2^{b_2}\cdots x_n^{b_n}\Big/[\int_0^\infty \exp\{-{}_n\gamma_1x_n-{}_n\gamma_2x_n^2-\sum_{b_{n-2}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-2}01}x_1^{b_1}x_2^{b_2}\cdots x_{n-2}^{b_{n-2}}x_n+({}_{n-1}\gamma_1+\sum_{b_n=0}^{1}\sum_{b_{n-2}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1b_n}x_1^{b_1}x_2^{b_2}\cdots x_{n-2}^{b_{n-2}}x_n^{b_n})^2/(4\,{}_{n-1}\gamma_2)\}\mathrm{erfc}\{({}_{n-1}\gamma_1+\sum_{b_n=0}^{1}\sum_{b_{n-2}=0}^{1}\cdots\sum_{b_2=0}^{1}\sum_{b_1=0}^{1}\delta_{b_1b_2\ldots b_{n-1}1b_n}x_1^{b_1}x_2^{b_2}\cdots x_{n-2}^{b_{n-2}}x_n^{b_n})/(2\sqrt{{}_{n-1}\gamma_2})\}]dx_n \qquad (15)$$

where ${}_{n-1}\gamma_2>0$ and we exclude the case that $b_{n-1}=b_n=0$.

The Level of Maximum Entropy Distribution

The maximum entropy distribution can express many types of distributions which are well-known in statistics. That is, the uniform distribution (Amari (1)), the normal distribution, the logarithmic normal distribution, the exponential distribution (Takasao and Ikebuchi (17)), the Gumbel distribution (Jowitt (4)), and the Cauchy distribution (Campenhout and Cover (2)) are derived by using the one-variate maximum entropy distribution. And the uniform distribution, the normal distribution, and the exponential distribution with no correlation between variates are derived by using the two-variate maximum entropy distribution (Sogawa, Araki and Terashima (12)), and moreover the uniform distribution and the exponential distribution with no correlation among variates are derived by using the multivariate maximum entropy distribution (Sogawa, Araki and Sato (13)). In the same manner as the above maximum entropy distribution, we can derive many distributions which contain six distributions described in former paragraph by using the MCME distribution. The kind of distribution which we can derive depends on only the function $g_r(\cdot)$.

On the other hand, Pearson showed that many probability density functions which are well-known in statistics were the solutions for the following differential equation.

$$\frac{1}{p(x)}\frac{dp(x)}{dx}=\frac{k_0+x}{k_1+k_2x+k_3x^2} \tag{16}$$

where k_0, k_1, k_2 and k_3 are the constants. The probability density function p(x) is divided into three main types and ten transition types, and they are called the Pearson's system of frequency-curves (Elderton and Johnson (3)). They contain J type distribution, U type distribution, symmetric distribution, non-symmetric distribution and so on. There are the both side finite distributions, the one side finite - the other side infinite distributions, and the both side infinite distributions in them.

There are the Gram-Charlier's series which are more general than the Pearson's system of frequency-curves in a sense (Sato (10)). This system of the distribution does not compose the set of many different functions but infinite series whose terms are the particular functions. They are given by

$$p(x)=\sum_{i=0}^{\infty} A_i\Phi_i(x) \tag{17.1}$$

or

$$p(x)=\exp\{\sum_{i=0}^{\infty} A_i\Phi_i(x)\} \tag{17.2}$$

where A_i is the coefficient and $\Phi_i(x)$ is the function of x. The Charlier's A type series has the normal distribution and its differentiated as $\Phi_i(x)$ with continuous variate x in eq. 14.1. The Charlier's C type series has the Hermite polinomials as $\Phi_i(x)$ with continuous variate x eq. 14.2. The Charlier's B type series has the Poisson distribution and its differences with discrete variate x in eq. 14.1.

Considering now comparison among above three systems, the constraint conditions of the maximum entropy distribution, that is, eqs. 4 are equivalent to the relation among each constant in Pearson's system of frequency-curves and to the type of $\Phi_i(\cdot)$ in the Gram-Charlier's series. The maximum entropy distribution with the constraints of the arbitrary functions $g_r(\cdot)$, that is, eq. 5 and eq. 6 are equivalent to the differntial equation eq. 16 in the Pearson's system of frequency -curves and to the infinite series eqs. 17.1 and 17.2 in the Gram-Charlier's series.

This shows that the maximum entropy distribution is based on the same level as the Pearson's system of frequency-curves and the Gram-Charlier's series. The maximum entropy distribution with the concrete function $g_r(\cdot)$ is equivalent to the one of the distributions of the Pearson's system of frequency-curves and to the one of the Gram-Charlier's series.

THE APPLICATION TO HYDROLOGICAL DATA AND DISCUSSION OF THEIR RESULTS

The Application to Annual Rainfall

We applied the MCME distribution to the annual rainfall at the four rainfall gaging stations (Nagano, Matumoto, Ueda and Karuizawa in Nagano prefecture) in Fig. 1.

We subtracted 500mm from the real annual rainfall so as to be identified the Lagrangian multipliers and normalized it by using the expectation (Wragg and Dowson (19)). When we illustrate the curve of the MCME distribution, we returned it to the basical value. We identified the MCME distribution by changing the objective points in which the occurrence probabilities are estimated, the conditional points in which the occurrence values are given as the condition, the conditional number which is the number of the conditional points, and how of adoption of the moments.

The whole term is divided into three terms as follows.

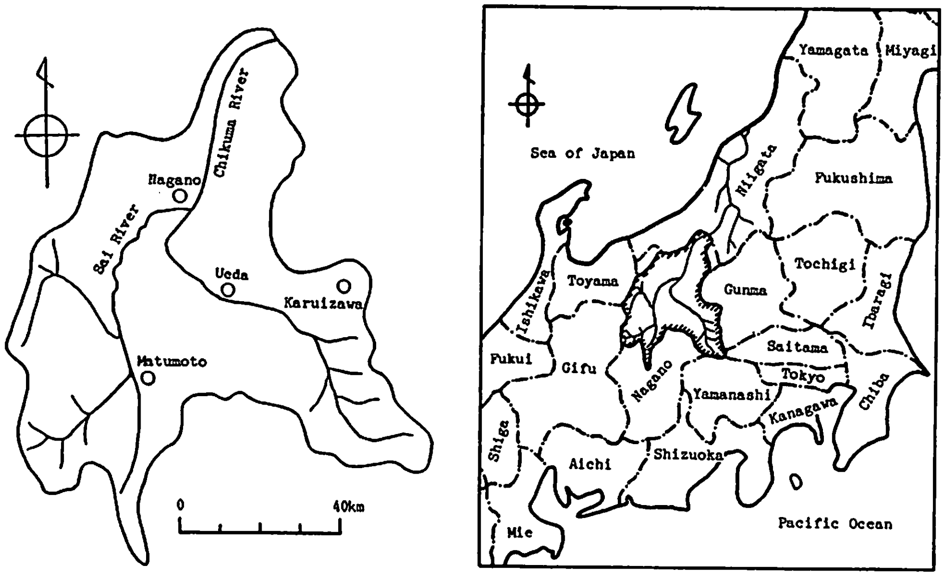


Fig. 1 Four rainfall gaging stations in the Chikuma river

1) Term of identification : Identification means to identify the parameters of the MCME distribution. This is the term in which there are both the objective points data and the conditional points data, and they are utilized for the identification of the MCME distribution.

2) Term of checking : Checking means to compare the observation data with the values obtained from the MCME distribution. This is the term in which there are both the objective points data and the conditional points data, and they are utilized for the checking of the MCME distribution.

•3) Term of estimation : Estimation means to obtain the values from the MCME distribution where there are no observation data of the objective points. This is the term in which there are no objective points data and there are the conditional points data, and they are utilized for the estimation of the MCME distribution.

(a) Relation between the how of the adoption of moments and change of distribution

The explanation of the distributions in this section is shown in Table 1, and the parameter values of these distributions, that is, the values of the Lagrangian multipliers are shown in Table 2. The change of the distributions by how of the adoption of moments is summarized as follows:

1) In the MCME distribution where Matumoto x_1 is the condition and Karuizawa x_2 is the object, the how of the adoption of moments is as follows. Firstly, when the first two moments in Karuizawa x_2 are adopted, no moment is adopted in MK1, the first one moment is adopted in MK2, the first two moments are adopted in MK3 and first three moments are adopted in MK4 in Matumoto x_1. If we express the product moments $\int\int x_1^{b_1} x_2^{b_2} p(x_1,x_2)dx_1dx_2$ as (b_1,b_2), we adopted the two moments, that is, (0,0) and (1,1). The examples of these distributions are shown in Fig. 2. The MK2 is much similar to the MK1 in the shape, the expectation and the standard deviation. The MK3 is sharper than the MK2 and differs from the MK2 on the points of the expectation and the standard deviation. The MK4 is similar to the MK3 not only in the sharpness but also in the expectation and the standard deviation. If we express the MCME distribution as $^{n}_{m}M(Na_1, Na_2, \ldots, Na_n; Nb)$ (where Nb is the number of the product moments), the MK1-MK4 are the distributions which are expressed as $^{2}_{1}M(Na_1,2;2)$. Though the Lagrangian multipliers of the number of Na_1 to the condition x_1 are therefore reduced both in the numerator and the denominator, the other Lagrangian multipliers, which are identified here, are influenced considerably by them. The shapes of the MCME distributions consequently have the differences between themselves. Besides, we have known that the even-order moments were very

Table 1 The explanation of the symbol of the MCME distribution applied to annual rainfall where N_b is the number of product moment in ${}^{n}_{m}M(N_{a_1}, N_{a_2}, \ldots, N_{a_n}; N_b)$

(a) $p_{x_1}(x_2)$

Symbol	Condition	Object	Distribution
MK 1	Matumoto (x_1)	Karuizawa (x_2)	${}^{2}_{1}M(0, 2; 2)$
MK 2	"	"	${}^{2}_{1}M(1, 2; 2)$
MK 3	"	"	${}^{2}_{1}M(2, 2; 2)$
MK 4	"	"	${}^{2}_{1}M(3, 2; 2)$
MK 5	"	"	${}^{2}_{1}M(2, 2; 5)$
NM 1	Nagano (x_1)	Matumoto (x_2)	${}^{2}_{1}M(2, 2; 2)$
NM 2	"	"	${}^{2}_{1}M(4, 4; 2)$

(b) $p_{x_1 x_2}(x_3)$

Symbol	Condition	Object	Distribution
NMK 1	Nagano, Matumoto (x_1) (x_2)	Karuizawa (x_3)	${}^{3}_{2}M(2, 2, 2; 2)$
NMK 2	"	"	${}^{3}_{2}M(2, 2, 2; 3)$
NUK 1	Nagano, Ueda (x_1) (x_2)	"	${}^{3}_{2}M(2, 2, 2; 2)$
NUK 2	"	"	${}^{3}_{2}M(2, 2, 2; 3)$
MUK 1	Matumoto, Ueda (x_1) (x_2)	"	${}^{3}_{2}M(2, 2, 2; 2)$
MUK 2	"	"	${}^{3}_{2}M(2, 2, 2; 3)$

(c) $p_{x_1 x_2 x_3}(x_4)$

Symbol	Condition	Object	Distribution
NMUK 1	Nagano, Matumoto, Ueda (x_1) (x_2) (x_3)	Karuizawa (x_4)	${}^{4}_{3}M(2, 2, 2, 2; 2)$

Table 2 The Lagrangian multipliers of the MEME distribution applied to annual rainfall

(a) $p_{x_1}(x_2)$

	$_2\gamma_1$	$_2\gamma_2$	δ_{11}	δ_{12}	δ_{21}	δ_{22}
MK 1	-0.15828E+02	0.72780E+01	0.94557E+00			
MK 2	-0.13070E+02	0.68602E+01	-0.84158E+00			
MK 3	-0.74173E+01	0.11697E+02	-0.15988E+02			
MK 4	-0.74153E+01	0.11895E+02	-0.15985E+02			
MK 5	-0.29850E+02	0.24760E+02	0.24837E+02	-0.23047E+02	-0.17072E+02	0.94557E+01

x_2 First one moment 0.80685E+03 mm

	$_2\gamma_1$	$_2\gamma_2$	$_2\gamma_3$	$_2\gamma_4$	δ_{11}
NM1	-0.40860E+01	0.76915E+01			-0.11305E+02
NM2	-0.12286E+01	0.45097E+01	0.11882E+01	-0.54487E-01	-0.11381E+02

x_2 First one moment 0.55488E+03 mm

(b) $p_{x_1x_2}(x_3)$

	$_3\gamma_1$	$_3\gamma_2$	δ_{011}	δ_{101}	δ_{111}
NMK 1	-0.14103E+02	0.80461E+01			-0.38869E+01
NMK 2	-0.31574E+01	0.14546E+02	-0.17014E+02	-0.89181E+01	
NUK 1	-0.16082E+02	0.98808E+01			-0.35544E+01
NUK 2	-0.89208E+01	0.20148E+02	-0.22456E+02	-0.89176E+01	
MUK 1	-0.16808E+02	0.10534E+02			-0.40041E+01
MUK 2	-0.72122E+01	0.23341E+02	-0.22456E+02	-0.17013D+02	

x_3 First one moment 0.80685E+03 mm

(c) $p_{x_1x_2x_3}(x_4)$

	$_4\gamma_1$	$_4\gamma_2$	δ_{1111}	x_4 First one moment
NMUK 1	-0.14449E+02	0.72869E+01	-0.72969E+00	0.80685E+03 mm

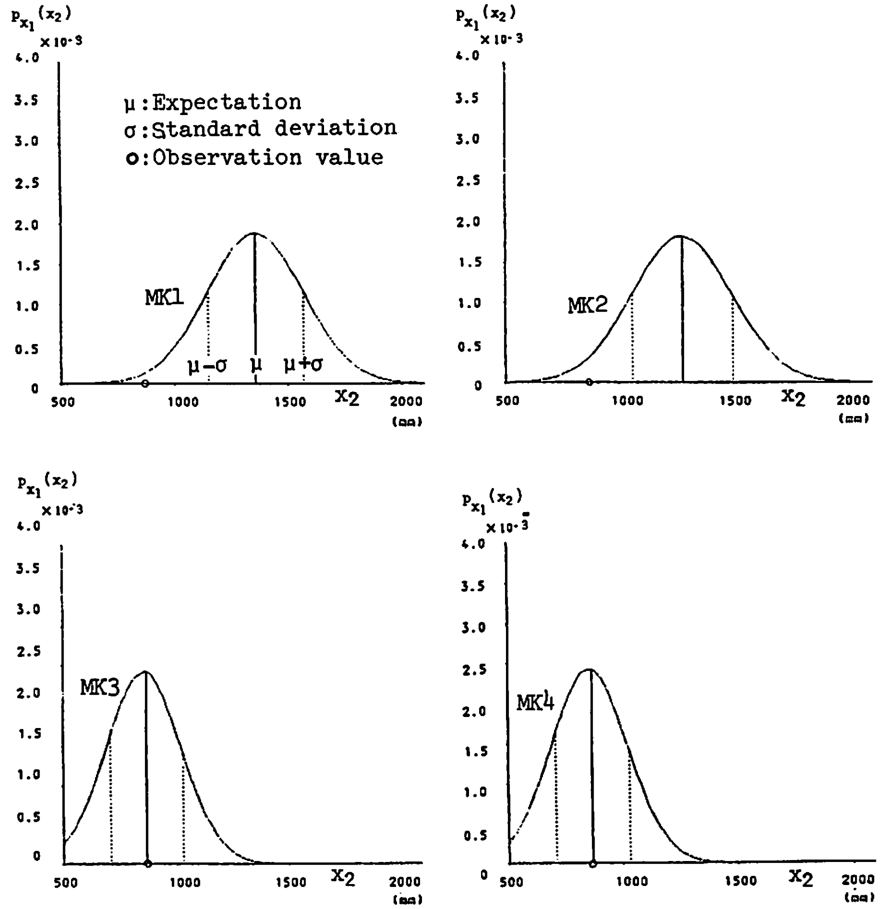


Fig. 2 The change of distribution $p_{x_1}(x_2)$ changing the how of adoption of moments (Product moments : (0,0), (1,1), Condition : Matumoto x_1, Object : Karuizawa x_2)

effective to improve the applicability to histograms in one-variate and two-variate maximum entropy distribution. We could show here too that the second moment varied pretty well the shapes of the MCME distribution.

2) The MK5 is the MCME distribution that adopts the first two moments of Matumoto x_1 and Karuizawa x_2, respectively, and (0,0), (1,1), (1,2), (2,1), (2,2) moments between Matumoto x_1 and Karuizawa x_2. We show the one example of the MK5 in Fig. 3. As compared the MK5 (=$_1$M(2,2;5)) with the MK3 (=$_1$M(2,2;2)), when Matumoto x_1 is less than about 1000mm, the MK5 is sharper than the MK3, and when Matumoto x_1 is more than about 1000mm, the MK3 is sharper than the MK5. When Matumoto x_1 is considerably small as shown in Fig. 3, the MK5 has the very sharp shape.

3) We show the NM1 which has the first two moments of Nagano x_1 and Matumoto x_2 respectively and the NM2 which has the first four moments of x_1 and x_2 respectively in Fig. 4, where we adopted both (0,0) and (1,1) as the product moments. Though both are similar generally, the NM2 is sharper than the NM1 when x_1 is larger than about 1300mm.

4) We obtained the NMK1 and the NMK2 which have Nagano x_1 and Matumoto x_2 as the conditions, the NUK1 and NUK2 which have Nagano x_1 and Ueda x_2 as the conditions, and the MUK1 and MUK2 which have Matumoto x_1 and Ueda x_2 as the conditions, while Karuizawa x_3 is the object. When we express the product moment $\iiint x_1^{b_1} x_2^{b_2} x_3^{b_3} p(x_1,x_2,x_3)dx_1dx_2dx_3$ as (b_1,b_2,b_3), the NMK1, the NUK1 and the MUK1 have (0,0,0)

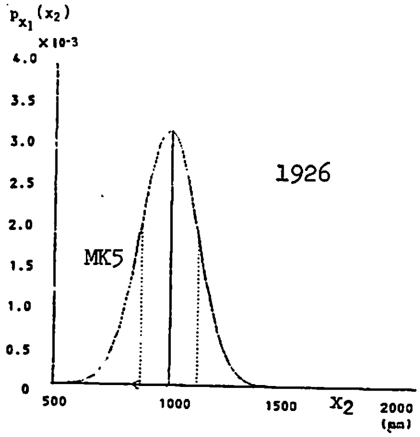


Fig. 3 The change of distribution $p_{x_1}(x_2)$ changing the how of adoption of moments (Product moments : (0,0), (1,1), (1,2), (2,1), (2,2), Condition : Matumoto x_1, Object : Karuizawa x_2)

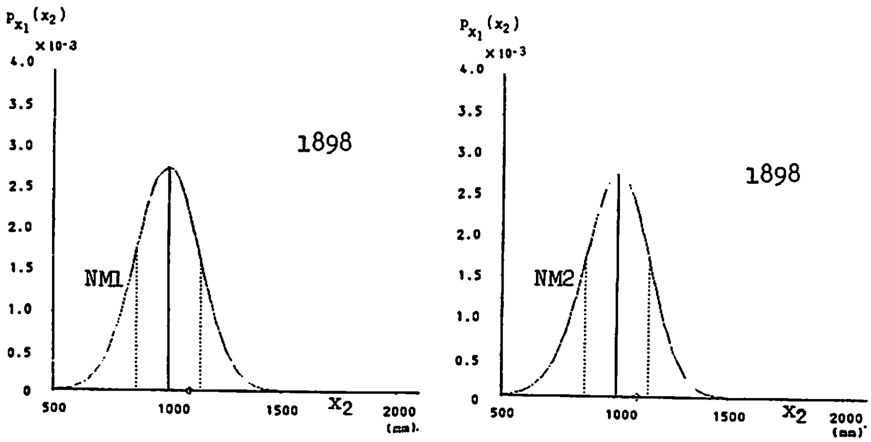


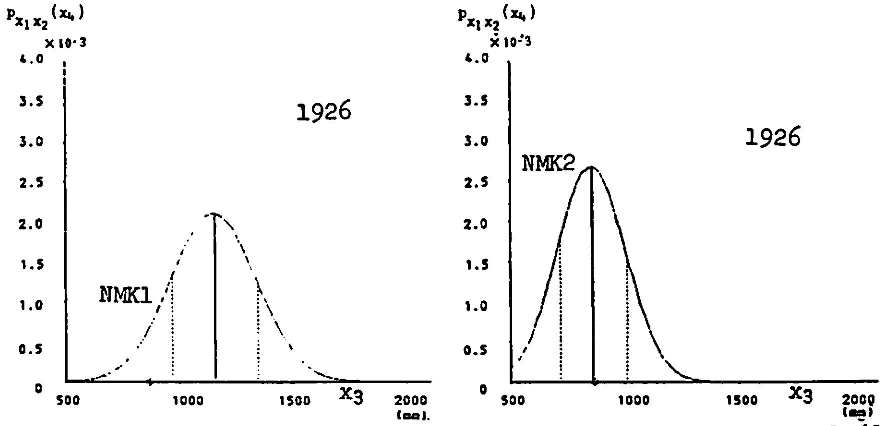
Fig. 4 The change of distribution $p_{x_1}(x_2)$ changing the how of adoption of moments (Product moments : (0,0), (1,1), Condition : Nagano x_1, Object : Matumoto x_2)

and (1,1,1), and the NMK2, the NUK2 and MUK2 have (0,0,0), (1,0,1) and (0,1,1). We show some examples of these distributions in Fig. 5. As compared with each other, we can see that ${}_2^3M(2,2,2;3)$ is much sharper than ${}_2^3M(2,2,2;2)$, and their expectations and standard deviations are different considerably.

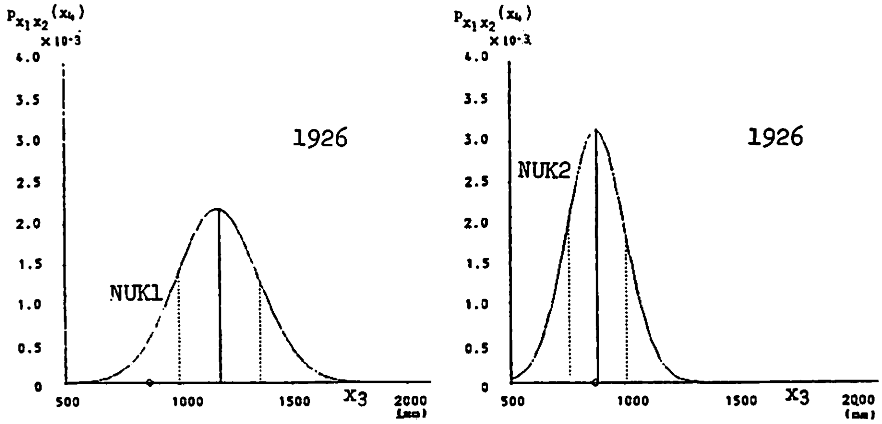
(b) The change of the distribution shape induced by the change of the number of constraint conditions

We added the NMUK1 which has Nagano x_1, Matumoto x_2 and Ueda x_3 as the conditions, and Karuizawa x_4 as the object. This is the ${}_3^4M(2,2,2,2;2)$. When we express the product moments $\iiiint x_1^{b_1}x_2^{b_2}x_3^{b_3}x_4^{b_4}p(x_1,x_2,x_3,x_4)dx_1dx_2dx_3dx_4$ as (b_1,b_2,b_3,b_4), we adopted (0,0,0,0) and (1,1,1,1) in this distribution. We show the one example of the NMUK1 in Fig. 6. We could not identify ${}_3^4M(2,2,2,2;4)$ which has (0,0,0,0), (1,0,0,1), (0,1,0,1) and (0,0,1,1) in this research.

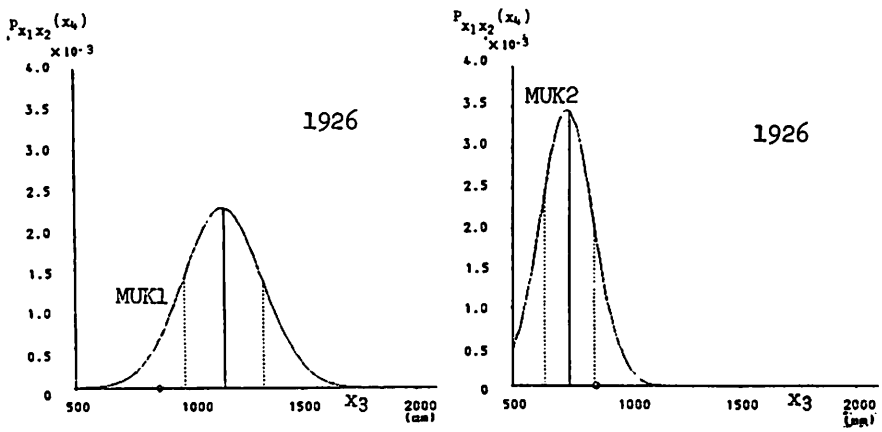
We use the distribution which has Matumoto as one of conditions and Karuizawa as the object in investigating the change of the distribution shape induced by the change of number of conditions. We divide now these distributions into two groups by the how of the adoption of the product moments.



(a) Condition : Nagano x_1, Matumoto x_2, Product moments : NMK1=(0, 0,0), (1,1,1), NMK2=(0,0,0), (1,0,1), (0,1,1)



(b) Condition : Nagano x_1, Ueda x_2, Product moments : NUK1=(0,0, 0), (1,1,1), NUK2=(0,0,0), (1,0,1), (0,1,1)



(c) Condition : Matumoto x_1, Ueda x_2, Product moments : MUK1=(0, 0,0), (1,1,1), MUK2=(0,0,0), (1,0,1), (0,1,1)

Fig. 5 The change of distribution $p_{x_1x_2}(x_3)$ changing the how of adoption of moments (Object : Karuizawa x_3)

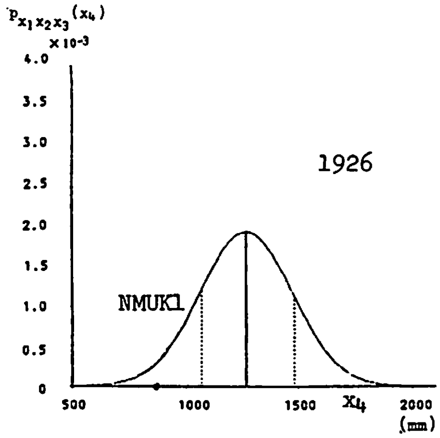


Fig. 6 The change of distribution $p_{x_1x_2x_3}(x_4)$ changing the number of conditions (Product moments : (0,0,0,0), (1,1,1,1), Conditions : Nagano x_1, Matumoto x_2, Ueda x_3, Object : Karuizawa x_4)

1) 1st group : MK3, NMK1, MUK1, NMUK1
This group contains the distributions whose product moments are composed by (0,0,...,0) and (1,1,...,1). The distributions become gradually flat with the increase of the number of conditions, that is, from 1 condition (MK3) to 2 conditions (NMK1 and MUK1) and from 2 conditions to 3 conditions (NMUK1). We think that the shapes of the distributions have the tendency to disperse in increasing of the number of conditions because of the how of above adoption of the product moments.

2) 2nd group : MK3, NMK2, MUK2
This group contains the distributions whose product moments are composed by (0,0,...,0) and the combination of (0,...,0,1,0,...,0,1) as the correlation between the one of conditions and the object. The distributions become gradually sharp by the increase of the number of conditions, that is, from 1 condition (MK3) to two conditions (NMK2 and MUK2). We think that this is why the information of the correlation between variates is adopted well by using the combination of product moment (0,...,0,1,0,...,0,1). We think that the MUK2 is sharper than the NMK2 because the correlation between Karuizawa and Ueda is larger than the correlation between Karuizawa and Nagano.

Finally, we show the expectation μ and $\mu+\sigma$ (σ:standard deviation) given by the MK3 with the real annual rainfall of Karuizawa in Fig. 7. The accuracy of the conformity between the term of identification and the term of checking is almost the same. This tendency is kept in the values which are obtained from the other distributions.

The Application to Annual Maximum Daily Rainfall

It is very impotant to obtain the probable hydrologic variate of annual maximum daily rainfall to make a flood control plan. If the number of existing data are not enough to obtain the probable hydrologic variate, it is necessary to supply the data by using the suitable method. We estimate here the annual maximum daily rainfall in Matumoto given the occurrence of daily rainfall in Nagano.

(a) The estimation of annual maximum daily rainfall under assumed simultaneous occurrence

Assuming the simultaneous occurrence between Nagano and Matumoto on the annual maximum daily rainfall, we estimated the daily rainfall in Matumoto. We obtained the MCME distribution ${}_1M(2,2;2)$ which has the annual maximum daily rainfall in Nagano as the condition x_1 and the daily rainfall of that day in Matumoto as the object x_2. We adopted here (0,0) and (1,1) as the product moments.

The data relates to a 20-yr. annual maximum daily rainfall for the period 1965-1984 in the term of identification and a 9-yr. annual maximum daily rainfall for the period 1956-1964 in the term of checking. We show the Lagrangian multipliers of this distribution in Table 3, and some examples of the shapes of this distribution in term of checking in Fig. 8. We can see that the shapes of this distribution are flat in general, σ becomes large because of it and the width of $\mu \pm \sigma$ becomes large. Though the expectations don't always have good agreement to the observation values, the accuracy of the conformity in the term of checking doesn't decrease comparing with one in the term of identification.

We show the Thomas plot of values of annual maximum daily rainfall in Matumoto given by annual maximum daily rainfall in Nagano as the conditions on log normal curve paper in Fig. 9. Fig. 10 is the Thomas plot of the observation values of annual maximum daily rainfall in Matumoto in the same manner as Fig. 9.

Though we assumed the simultaneous occurrence of annual maximum daily rainfall in Nagano and Matumoto, it didn't occur practically and the values obtained from the MCME distribution are smaller than the observation values. This is why the obtained values shifts to left side to the observation values in Fig. 10.

(b) The estimation of annual maximum daily rainfall by usimg simulation rainfall

We obtained here ${}^2_1M(2,2;2)$ which has the daily rainfall over 25mm in Nagano as the condition x_1 and the simultaneous daily rainfall in Matumoto as the object x_2. We adopted here (0,0) and (1,1) as the product moments.

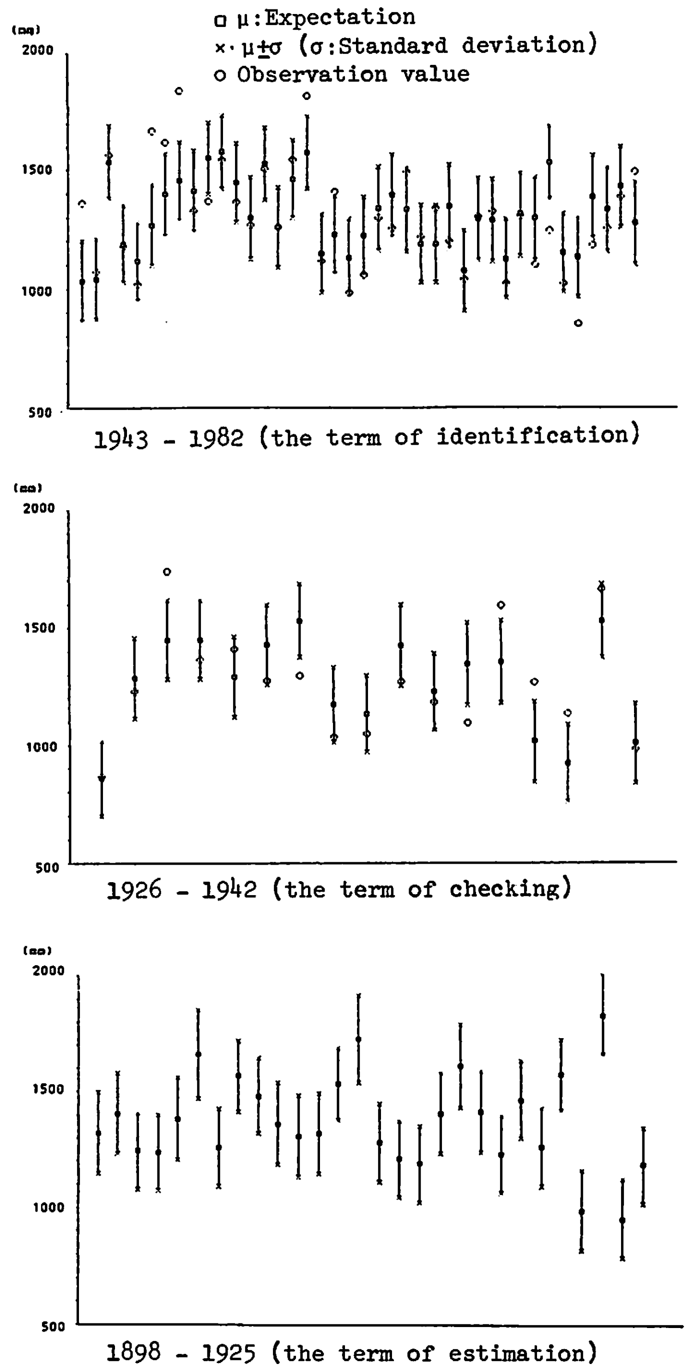


Fig. 7 The annual rainfall obtained by using the MK3 distribution

The data relate to a 20-yr. daily rainfall for the period 1965-1984 (147 rainfalls) in the term of identification and a 9-yr. daily rainfall for the period 1956-1964 (74 rainfalls) in the term of checking. We show the Lagrangian multipliers of this distribution in Table 4, and the shapes of this distribution in 1959 (4 rainfalls) in Fig. 11. These show the shape of the exponential distribution and the shape of the flat

Table 3 The Lagrangian multipliers of the MCME distribution $p_{x_1}(x_2)$ applied to annual maximum daily rainfall asuming simultanious occurrence

$2\gamma 1$	$2\gamma 2$	$\delta 1 1$	x_2 First one moment
0.30977E+01	0.43992E+00	-0.30490E+01	0.46500E+02mm

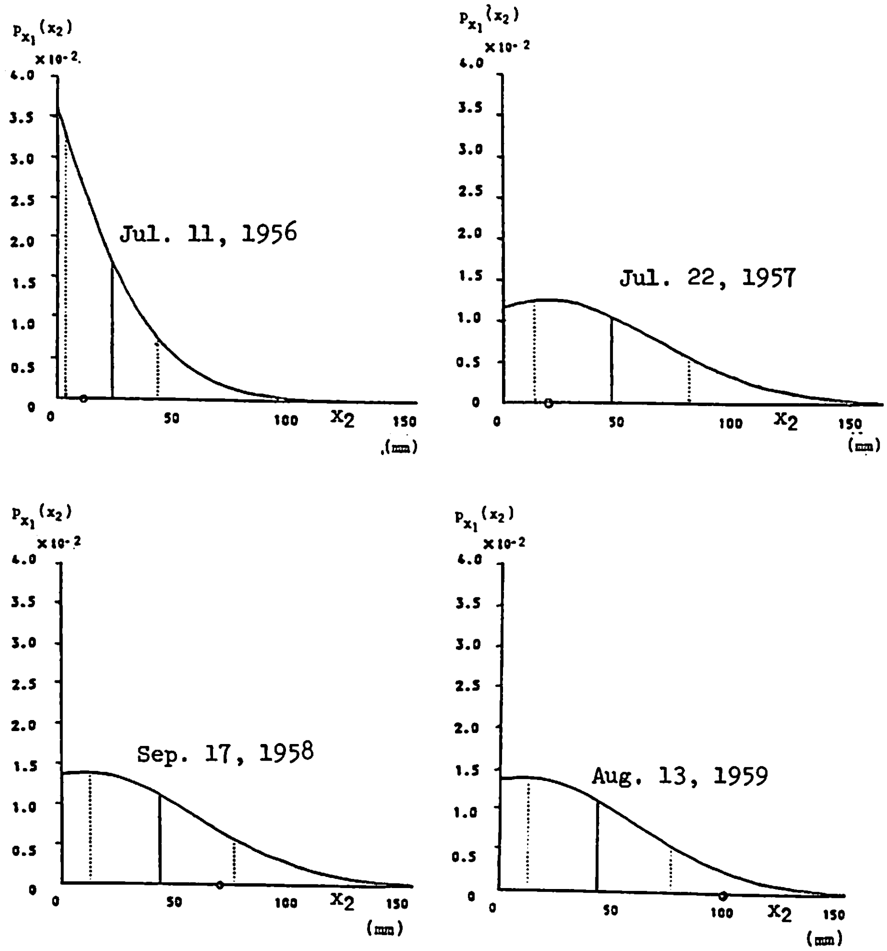
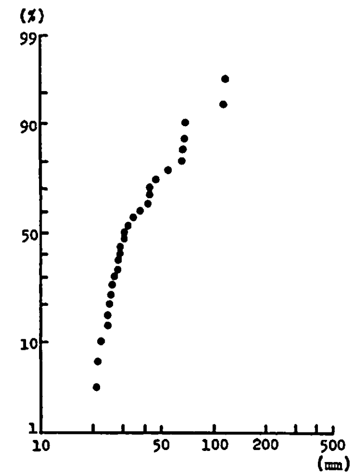


Fig. 8 The MCME distribution $p_x(x_2)$ of annual maximum daily rainfall assumed simultaneous occurrence (Condition ; Nagano x_1, Object : Matumoto x_2)



Fig, 9 The Thomas plot of estimation values of annual maximum daily rainfall in Matumoto given by one in Nagano

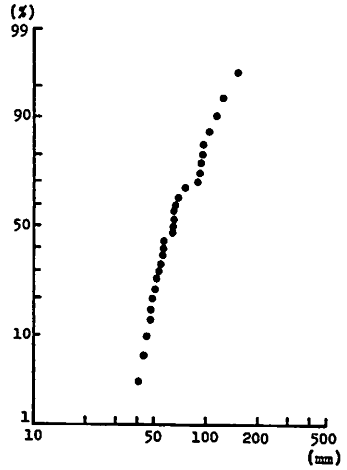


Fig. 10 The Thomas plot of the observation values of annual maximum daily rainfall in Matumoto

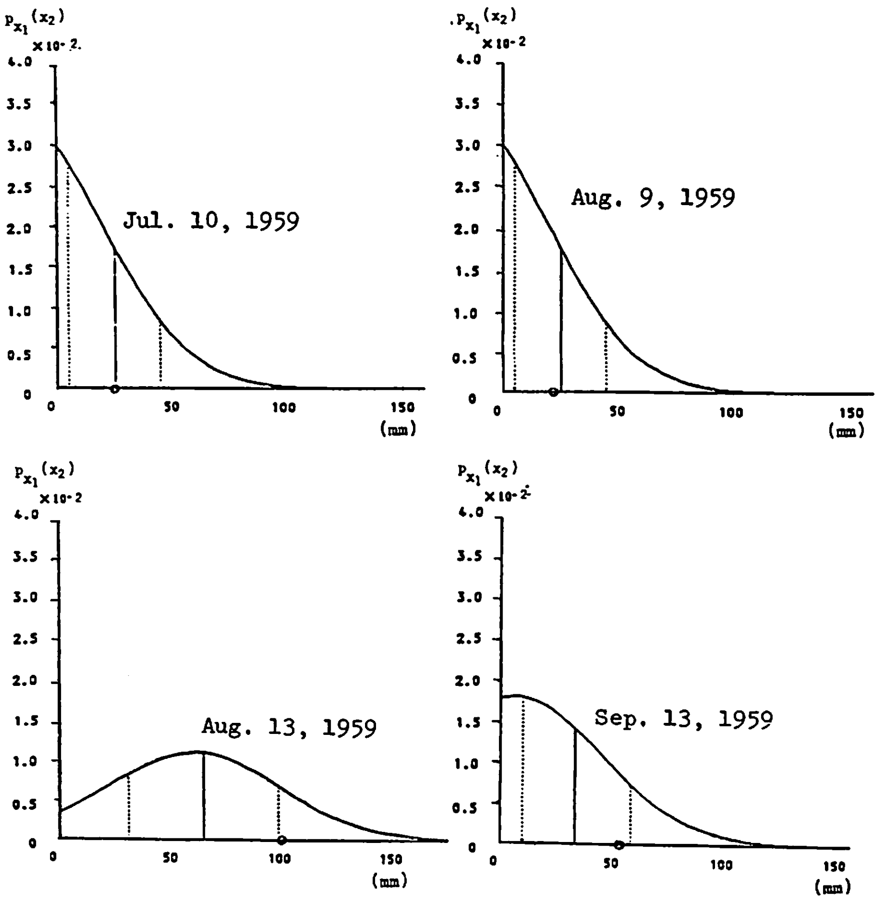


Fig. 11 The MCME distribution $p_{x_1}(x_2)$ of annual maximum daily rainfall by using simulation rainfall (Condition : Nagano x_1, Object : Matumoto x_2)

Table 4 Lagrangian multipliers of the MCME distribution $p_{x_1}(x_2)$ applied to annual maximum daily rainfall by using simulation rainfall

$_2\gamma_1$	$_2\gamma_2$	δ_{11}	x_2 First one moment
-0.19152E+01	0.35428E+00	-0.18308E+01	0.32664E+02mm

normal distribution.

The simulation rainfalls are generated by the obtained MCME distribution as follows:

1) The uniform random number R is generated over the interval [0,1].

2) $\hat{x}_2$, which has the R as the cumulative probability, is obtained by the above distribution with the condition x_1.

3) $\hat{x}_2$ is regarded as the simulation rainfall.

It is not desirable to generate the uniform random number over the interval [0,1] because they contain the abnormally large values and small values. We generated therefore the random numbers over the interval $[\alpha, 1-\alpha]$ being based on the probability of exceedance $\alpha = 1/(n+1)$ of Thomas plot related to the maximum value of sample size n (Nagao and Kadoya (9)). The above interval is [0.005,0.995] to 221 rainfalls for the period 1956-1984.

We show the Thomas plot of annual maximum daily rainfalls in Matumoto, which are the maximum values of these simulation rainfalls in each year, in Fig. 12. The plot points in Fig. 12 expand on both sides comparing with Fig. 10. This is why the shape of the MCME distribution becomes flat because the correlation between the daily rainfalls over 25mm in Nagano and the daily rainfalls generated simultaniously in Matumoto is low, and the number of occurrance rainfalls is small.

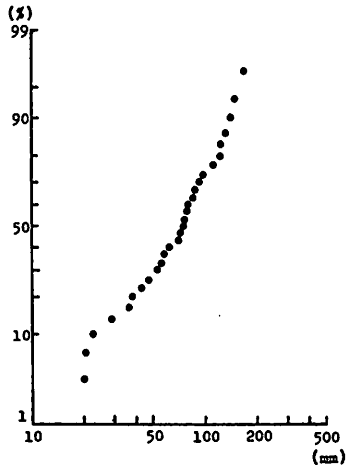


Fig. 12 The Thomas plot of estimation values of annual maximum daily rainfall in Matumoto given by simulation rainfall

CONCLUSION

This paper describes the theory of the MCME distribution, its characteristics and the application of that distribution to hydrological data. We summarize the result of this research as follows.

1) We could derive the theory of the MCME distribution with the arbitrary function $g_r(\cdot)$ and arbitrary order moments as the constraint conditions.

2) We could derive some distributions by giving the number of variates and moments in the MCME distribution with the moments as the constraint conditions.

3) We discussed that the maximum entropy distribution was based on the same level as the Pearson's system of frequency-curves and the Gram-Charlier's series by using 2).

4) Applying the MCME distribution to annual rainfall, we investigated the change of the distribution shape induced by the how of the adoption of moments and the change of the number of conditions.

5) Applying the MCME distribution to annual maximum daily rainfall, we obtained the annual maximum daily rainfall of the objective point in assumption of the simultaneous occurrence between two points and in use of simulation rainfall.

In the near future we wish to increase the number of applications to hydrological data and to investigate the conformity of the MCME distribution to well-known distributions in statistics.

REFERENCES

1. Amari, S. : Continuous information and signal space, Information Theory, No.1, Diamond Pub. Co., pp.141-142, 1970 (in Japanese).
2. Campenhout, J.M. and T.M. Cover : Maximum entropy and conditional probability, IEEE Transactions on Information Theory, Vol.IT-27, No.4, pp.483-489, 1981.
3. Elderton, W.P. and N.L. Johnson : System of frequency curves, Cambrige University Press, p45, 1969.
4. Jowitt, P.W. : The extreme-value type-1 distribution and the principle of maximum entropy, Journal of Hydrology, Vol.42, pp.23-38, 1979.
5. Kanda, T. and M. Fujita : Frequency analysis of hydrogical data, Hydrology, No.1, Gihodo Pub. Co., Tokyo, p.65, 1982 (in japanese).
6. Krishnamoorthy, A.S. and M. Parthasarathy : A multivariate gamma-type distribution, Annual of Mathematical Statistics, Vol.22, pp.549-557, 1951.
7. Nagao, M. and M. Kadoya : Study on two-variate gamma distribution and its engineering application (1) —— Fundamental theory of two-variate exponential distribution ——, Disaster Prevention Research Institute Annuals, Kyoto University, No.13-B, pp.105-133, 1970 (in Japanese).
8. Nagao, M. and M. Kadoya : Study on two-variate gamma distribution and its engineering application (2) —— Estimation of population parameters ——, Disaster Prevention Research Institute Annuals, Kyoto University, No.14-B, pp.43-56, 1971 (in Japanese).
9. Nagao, M. and M. Kadoya : Study on two-variate gamma distribution and its engineering application (4) —— Synthesis of short-time rainfall data applying two-variate exponential distribution theory ——, Disaster Prevention Research Institute Annuals, Kyoto University, No.14-B, pp.77-85, 1971 (in Japanese).
10. Sato, R. : Frequency distribution function and frequency distribution curve, Mathmatical Statistics, No.2, Baifukan Co., Tokyo, pp.46-48, 1952 (in japanese).
11. Sogawa, N. and M. Araki : On introduction of MEP in hydrologic frequency analysis, Proceedings of the Japan Society of Civil Engineering, No.335, pp.89-95, 1983 (in Japanese).
12. Sogawa, N., M. Araki and A. Terashima : Study on bivariate MEP distribution and its characteristics, Proceedings of the 28th Japanese Conference on Hydraulics, pp.397-402, 1984 (in Japanese).
13. Sogawa, N., M. Araki and K. Sato : A study on multivariate maximum entropy distribution and its basic characteristic, Proceedings of the Japan Society of Civil Engineerings, Hydraulic and Sanitary Engineering (in Japanese, unpublished).
14. Sogawa, N., M. Araki and N. Kobayashi : Applicability of the uni-variate MEP distribution to various hydrologic data, Journal of the Faculty of Engineering, Shinshu University, No.58, 1985 (in Japanese).
15. Sonuga, J.O. : Principle of maximum entropy in hydrologic frequency analysis, Journal of Hydrology, Vol.17, pp.177-191, 1972.
16. Sonuga, J.O. : Entropy principle applied to the rainfall-runoff process, Journal of Hydrology, Vol.30, pp.81-94, 1976.
17. Takasao, T. and S. Ikebuchi : Stochastic aspects on the runoff system and model making based on the concept of entropy, Disaster Prevension Research Institute Annuals, Kyoto University, No.23-B-2, pp.193-209, 1980 (in Japanese).
18. Takeuchi, K. : Random variable, Mathematical statistics, Nol, Toyo Economics Co., Tokyo, p,47, 1963 (in Japanese).
19. Wragg, A. and D.C. Dowson : Fitting continuous probability density functions over [0,∞] using information theory, Vol.IT-16, pp.226-230, 1970.

APPENDIX - NOTATION

The following symbols are used in this paper:

A_i = constant of the Gram-Charlier's series;

$g_r(\cdot)$	= arbitrary function;
k_i	= constant of the Pearson's differential equation;
M	= number of arbitrary function $g_r(\cdot)$;
n	= number of variate x_i;
Na_i	= maximum order of moment related with $x_i^{a_i}$;
Nb_i	= maximum order of product moment related with $x_i^{b_i}$;
$p(\cdot)$	= probability density function;
$p_.(\cdot)$	= conditional probability density function;
x_i	= variate;
${}_i\gamma_{a_i}$, $\delta_{b_1 b_2 \ldots b_n}$	= Lagrangian multiplier with the constraint of moments;
λ_r	= Lagrangian multiplier with the constraint of arbitrary function $g_r(\cdot)$; and
$\Phi_i(\cdot)$	= a function or the function derived by a function in the Gram-Charlier's series.