SUFFICIENT STATISTICAL ESTIMATION FUNCTIONS FOR THE PARAMETERS OF THE DISTRIBUTION OF MAXIMUM VALUES

By Bradford F. Kimball

New York State Department of Public Service

1. Summary. The problem of estimating from a sample a confidence region for the parameters of the distribution of maximum values is treated by setting up what are called "statistical estimation functions" suggested by the functional form of the probability distribution of the sample, and finding the moment generating function of the probability distribution of these estimation "functions. Such an estimate by the method of maximum likelihood is also treated.

A definition of "sufficiency" is proposed for "statistical estimation functions" analogous to that which applies to "statistics." Also the concept of "stable statistical estimation functions" is introduced.

By means of a numerical illustration, four methods are discussed for setting up an approximate confidence interval for the estimated value of x of the universe of maximum values which corresponds to a given cumulative frequency .99, for confidence level .95. Two procedures for solving this problem are recommended as practicable.

2. Introduction. If the universe comprises a set of maximum values of a large number of quantities, it has been shown that in many cases the probability density function of such a set of values of x is given approximately by

$$f(x) = \alpha e^{-t} e^{-e^{-t}}, \qquad t = \alpha(x - u), \qquad -\infty < x < +\infty, \tag{2.1}$$

where α and u denote parameters, usually unknown [1].

This paper is concerned with the problem of estimation of the parameters α and u on the basis of sample data.

The notion of "sufficiency" is fundamental in the problem of estimation, since it means that the necessary elements of the sample have been used which will result in complete determination of that part of the sample probability distribution function involving the unknown parameters to be estimated. Unfortunately it does not seem to be possible to set up "sufficient statistics" within the usual definition of "statistic" for the above distribution. In this investigation the writer was struck by the fact that certain functions of the data involving one of the parameters could be used to play a very similar role to a set of *sufficient statistics* for determining α and u, in spite of the fact that one function involved the value of α, and hence was not directly determined by the data,—and hence not a "statistic."

Various statistics have been used in the past to estimate the parameters α and u, such as the sample mean, variance, mean deviation and an adjusted modal value (see [2] and [3]). For the reason noted above, sufficient statistics have not been developed. In order to bridge this impasse and meet the *essentials* of the condition of sufficiency, the writer believes that a broader definition of sufficiency is needed. Such a definition is developed in the following section.

3. A broader definition of sufficiency. If the reader reexamines the process of estimating the two parameters of the normal distribution, and the determination of the two parameter confidence region for them from the statistics consisting of the sample mean, and the mean square deviation of the sample values from their mean, he will find that the separate determination of $\bar{x}$ and s^2 is not inherently necessary. The mean a and the variance σ^2 of the universe, are usually estimated from the *pair* of equations

$$E(\bar{x}) = a, \qquad E(s^2) = (n - 1)\sigma^2/n$$

and the boundary of the confidence region is determined from knowledge of the bivariate distribution of $\bar{x}$ and s, which involves the four variables $\bar{x}$, s, a, and σ. The equation of the bounding curve is most easily set up in terms of transformed variables such as

$$U = \sqrt{n}\,(\bar{x} - a)/\sigma, \qquad V = \sqrt{n}\,s/\sigma. \tag{3.1}$$

Then the probability density of U and V is given by

$$f(U,V) = (\text{const.})\,V^{n-2}e^{-(U^2+V^2)/2}$$

and with confidence coefficient β; a bounding curve may be defined implicitly by the two equations

$$\iint f(U, V)\,dU dV = \beta,$$

$$f(U_1, V_1) = \text{constant}$$

where the above integral is taken over the region of the $V \geqq 0$ half of U,V plane bounded by the curve $f(U_1, V_1) = $ constant.

A range of estimate of the parameters a and σ is offered by this confidence region by virtue of the fact that each point of the region corresponds to a *unique* pair of values of a and σ for a given set of sample values $0_n(x_i)$, and the fact that the equation of the bounding curve does not involve the parameters a and σ. Thus one arrives at a determinate range of estimate of a and σ, after the sample values have been observed. In this paper such functions will be referred to as *statistical estimation functions* (see [4]).

The classical idea of sufficiency implies (a) that the estimate be adequate for unique determination of the parameters, and (b) that *all* the sample information pertinent to such estimation be used. In the case of "statistics" the second requirement has been simply and elegantly formulated by the requirement that the probability density function of the sample distribution factor in such a way that one factor be *completely* determined by the statistical estimates and the parameters of the distribution, and that the remaining factor be *independent* of the parameters to be estimated (see [7], or [5] p. 135).

It seems to be possible to carry over this formulation to statistical estimation functions (denoted by T_i). Since one or more of the parameters to be estimated, denoted by $(a_1, a_2, \cdots, a_r)$, are involved in these functions, a requirement that they be adequate for unique determination of these parameters is obviously that there be a *one-to-one correspondence* between the parameter set $(a_1, a_2, \cdots, a_r)$ and the set of estimation functions $(T_1, T_2, \cdots, T_r)$ in the region of estimate. This requirement will be referred to as Requirement (1).

It has been pointed out by a referee that some further requirement as to the independence of the probability density function of $(T_1, T_2, \cdots, T_r)$ relative to the parameters to be estimated is needed.

If one requires that the p. d. f. of $(T_1, T_2, \cdots, T_r)$ be entirely independent of the parameters $(a_1, a_2, \cdots, a_r)$ the estimation functions will furnish "confidence regions" for estimates of the parameters;—see example noted above for the normal distribution.

However, in some cases the mean values $E(T_i)$ may be independent of the parameters, while the p. d. f. may not be; for example, —estimation functions for the two parameters of the Pearson Type III distribution formed from the maximum likelihood functions of that distribution. In such cases, a *point estimation* of the parameters is still possible, and would seem to satisfy the classical requirements of sufficiency.

The author accordingly makes the following proposals:

(a) Statistical estimation functions that satisfy the first two requirements—that of one-to-one correspondence with the parameters to be estimated, and the factoribility condition—be termed *sufficient* for estimation of the parameters. The reasonableness of such a definition is strengthened by the observation that given a set of "sufficient statistics" in the classical sense, statistical estimation functions that satisfy the factoribility condition can always be formed from them, and hence they are subject further only to Requirement (1) to make them *sufficient statistical estimation functions* under the proposed definition.

(b) Statistical estimation functions that satisfy Requirement (1) and also have a p. d. f. which is independent of the parameters to be estimated shall be called *stable*—a term suggested to the author by a referee.

(c) Statistical estimation functions T_i that satisfy Requirement (1) and are such that $E(T_i)$, $(i = 1, 2, \cdots, r)$, be independent of the parameters to be estimated, be called *stable in mean*, and that similarly, if the modal or median values of T_i be independent of these parameters, they be called *stable in mode*, *stable in median*, etc.

Thus a definition of sufficiency applicable to statistical estimation functions is formulated as follows:

The term "statistical estimation function" will be used to denote a function of the sample values and one or more population parameters, used for purposes of statistical estimation.

Given a universe with probability density function involving m parameters $a_1, a_2, \cdots, a_m$ in an admissible region R, and a set of r statistical estimation functions $T_i(0_n; a_1, a_2, \cdots, a_m)$ to be used for estimating the r parameters $a_1, a_2, \cdots, a_r$ relative to the information in a given sample 0_n. Consider the conditions:

(1) The functional form T_i insures a one-to-one correspondence between the points of the r-parameter space $(a_1, a_2, \cdots, a_r)$ contained in R and the points of the r-space defined by $(T_1, T_2, \cdots, T_r)$ for fixed $0_n(x_i)$ and fixed parameter values $a_{r+1}, a_{r+2}, \cdots, a_m$.

(2a) It shall be possible to express the probability density function of the sample 0_n as

$$P(0_n) = g_1(T_1, T_2, \cdots, T_r; a_1, a_2, \cdots, a_m)\cdot g_2(0_n; a_{r+1}, a_{r+2}, \cdots, a_m),$$

where the first factor is uniquely determinable for fixed $(a_1, a_2, \cdots, a_m)$ from the corresponding values of the functions T_i, and the second factor is independent of the parameters to be estimated.

(2b) It shall be possible to express the probability density function of the sample 0_n as

$$P(0_n) = G(T_1, T_2, \cdots, T_r; a_1, a_2, \cdots, a_m)g_2(0_n; a_{r+1}, a_{r+2}, \cdots, a_m),$$

where $G(, , \cdots, ; a_1, a_2, \cdots, a_m)$ is a functional, depending on $a_1, a_2, \cdots, a_m$, which in general involves the values of the T_i for values of $a_1, a_2, \cdots, a_m$ different from those appearing in the rest of the identity. (For example,

$$G(T, a) = \exp \int_0^a T(0_n; a')da'.)$$

(3) The r-variate probability density function of T_i based on $P(0_n; a_1, a_2, \cdots, a_m)$ shall exist.

Definition A. A set of statistical estimation functions T_i which satisfies conditions (1) and (2a) will be said to be a *sufficient* set of estimation functions for estimating the parameters a_i, $(i = 1, 2, \cdots, r)$, relative to the sample 0_n.

Definition B. A set of statistical estimation functions T_i which satisfies conditions (1) and (2b) will be said to be a *functionally sufficient* set of estimation functions for estimating the parameters a_i $(i = 1, 2, \cdots, r)$, relative to the sample O_n.

Definition C. If the conditions (1) and (3) are satisfied, and the p.d.f. of $(T_1, T_2, \cdots, T_r)$ is independent of the parameters a_i, $(i = 1, 2, \cdots, r)$, the functions T_i will be said to be *stable* relative to estimation of these parameters.

Definition D. If the conditions (1) and (3) are met, and $E(T_i)$, $(i = 1, 2, \cdots, r)$ are independent of the parameters to be estimated, the functions T_i will be said to be *stable-in-mean*; and similarly if modal or median values of T_i are independent of these parameters, the estimation functions will be said to be *stable-in-mode, stable-in-median,* etc.

It is not difficult to prove that a set of maximum likelihood functions

$$L_\alpha = \partial[\log P(0_n\,; \alpha, \beta)]/\partial\alpha, \qquad L_\beta = \partial[\log P(0_n\,; \alpha, \beta)]/\partial\beta$$

under the condition that the second order determinant

$$\begin{vmatrix} L_{\alpha\alpha} & L_{\alpha\beta} \\ L_{\beta\alpha} & L_{\beta\beta} \end{vmatrix}$$

exists and does not vanish over the admissible range of α and β, constitutes a set of estimation functions for α and β that are *functionally sufficient* and *stable-in-mean* under the definition given above. The meeting of Condition (2b) is demonstrated by the relation

$$\log P(0_n\,; \alpha, \beta) = \int_{\alpha_0}^{\alpha} L_\alpha(\alpha, \beta_0)\,d\alpha + \int_{\beta_0}^{\beta} L_\beta(\alpha, \beta)\,d\beta + \log P(0_n\,; \alpha_0\,, \beta_0)$$

since the first two terms on the right depend entirely upon the functions L_α and L_β, and the third term on the right becomes independent of α and β, if α_0 and β_0 are arbitrarily chosen, once for all, in the admissable region R.

In general the maximum likelihood functions are not *stable* estimation functions, but in many cases by the introduction of suitable factors which appear in the variance-covariance matrix (see (5.3) and (5.4)) estimation functions may be formed which satisfy Definition C.

4. Sufficient statistical estimation functions for the distribution of maximum values. The probability density function for the sample $0_n(x_i)$ drawn from a universe of maximum values is

$$P(0_n) = \alpha^n e^{-\Sigma e^{-\alpha(x_i-u)}} e^{-\alpha\Sigma(x_i-u)} \tag{4.1}$$

where the summation sign used here and hereinafter refers to summation over all indices from 1 to n. Let $\bar{x}$ denote the sample mean, and define a new set of variables z_i by

$$z_i = e^{-\alpha x_i}, \qquad (i = 1, 2, \cdots, n), \tag{4.2}$$

with mean $\bar{z}$. Also set

$$z_0 = e^{-\alpha u}.$$

Recognizing that the variables $2z_i/z_0$ are independently distributed like χ^2 on two degrees of freedom, the probability density function of $\bar{z}$ is given by

$$P(\bar{z})\,d\bar{z} = [1/\Gamma(n)]e^{-n\bar{z}/z_0}(n\bar{z}/z_0)^{n-1}n\,d\bar{z}/z_0 \tag{4.3}$$

with mean equal to z_0 and variance equal to z_0^2/n.

The mean value of t of the original distribution (2.1) is known to be Euler's constant, which will be denoted by C. Thus

$$E[\alpha(\bar{x} - u)] = C = .5772157. \tag{4.4}$$

The above considerations point to a set of statistical estimation functions defined as follows

$$X = \sqrt{n}\,[\alpha(\bar{x} - u) - C],$$
$$Y = \sqrt{n}\,[\bar{z}/z_0 - 1]. \tag{4.5}$$

The author was not able to determine the explicit bivariate probability density function of X and Y, but the moment generating function G may be found with some degree of facility if the variables z_i are used in (4.1). Using simplified functions $n\alpha(\bar{x} - u)$ and $n\bar{z}/z_0$,

$$G(\theta_1, \theta_2) = E[e^{\theta_1 n\alpha(\bar{x}-u)} e^{\theta_2 n\bar{z}/z_0}] = (1 - \theta_2)^{n(\theta_1-1)}\Gamma^n(1 - \theta_1). \tag{4.6}$$

Clearly $\bar{x}$ and $\bar{z}$ are not statistically independent. The first and second partial derivatives give

$$G_{\theta_1}(0, 0) = nC, \qquad G_{\theta_2}(0, 0) = n, \qquad G_{\theta_1\theta_1}(0, 0) = n\pi^2/6 + n^2C^2,$$
$$G_{\theta_2\theta_2}(0, 0) = n^2 + n, \qquad G_{\theta_1\theta_2}(0, 0) = n^2C - n. \tag{4.7}$$

Hence the variances of the marginal distributions are

$$\sigma^2[n\alpha(\bar{x} - u)] = n\pi^2/6, \qquad \sigma^2(n\bar{z}/z_0) = n, \tag{4.8}$$

and the covariance is equal to $-n$.

Now the marginal distributions rapidly approach normality with increasing n. The question arises whether the bivariate distribution approaches normality. One way to prove this is as follows: Consider the moment-generating function G_2 of the statistical functions X and Y defined by (4.5). Following methods outlined above, with $\theta_3 = \sqrt{n}\theta_1$, $\theta_4 = \sqrt{n}\,\theta_2$, it is not difficult to show that the logarithm of the moment generating function $G_2(\theta_3, \theta_4)$ is given by

$$\log G_2 = (\sqrt{n}\,\theta_3 - n)\log(1 - \theta_4/\sqrt{n}) - \sqrt{n}\,\theta_4 + n\log\Gamma(1 - \theta_3/\sqrt{n}) - \sqrt{n}\,C.$$

As $n \to \infty$, one notes the relations

$$-n\log(1 - \theta_4/\sqrt{n}) - \sqrt{n}\,\theta_4 = \theta_4^2/2 + o_1(\sqrt{n}),$$
$$n\log\Gamma(1 - \theta_3/\sqrt{n}) - \sqrt{n}\,C\theta_3 = (\theta_3^2/2)(\pi^2/6) + o_2(\sqrt{n}), \tag{4.9}$$
$$\sqrt{n}\,\theta_3\log(1 - \theta_4/\sqrt{n}) = -\theta_3\theta_4 + o_3(\sqrt{n}),$$

where $o_i(\sqrt{n})$ denote functions that approach zero as $\sqrt{n} \to \infty$, uniformly for θ_3 and θ_4 in the neighborhood of zero. The limit

$$\lim_{n\to\infty} \log G_2 = \tfrac{1}{2}[\theta_4^2 - 2\theta_3\theta_4 + (\pi^2/6)\theta_3^2]$$

is recognized as the logarithm of the moment generating function of a normal bivariate distribution.

Thus *the bivariate probability distribution function of the estimation functions X and Y approaches the normal bivariate distribution with zero means and variance-covariance matrix*

$$\begin{Vmatrix} \pi^2/6 & -1 \\ -1 & 1 \end{Vmatrix} \tag{4.10}$$

as n increases without limit, and the means and second order moments thus indicated, hold precisely for all values of n.

The functions X and Y satisfy Condition (1) for sufficiency relative to estimation of the parameters α and u provided α and u can be expressed as single valued functions of X and Y. A condition for this is that the Jacobian of the transformation shall not vanish. This Jacobian may be reduced to

$$[(n\alpha\bar{z})/z_0][\bar{x} - (\Sigma x_i e^{-\alpha x_i})/(\Sigma e^{-\alpha x_i})].$$

Let x_i be ordered so that $x_i \leqq x_{i+1}$. Then for $\alpha > 0$, the second term constitutes a weighted mean with positive weights which monotonically decrease as i increases, when the inequality $x_i < x_{i+1}$ holds. Hence unless all x_i are equal, this weighted mean is less algebraically than $\bar{x}$. Condition (2a) for sufficiency is clearly met by these functions. Thus one concludes that *for $\alpha > 0$, and the case that not all x_i are equal, the estimation functions X and Y constitute a sufficient set of estimation functions for the parameters α and u of distribution* (2.1). Since the moment generating function (see (4.6)) is independent of α and u, these functions are also *stable* estimation functions.

5. Maximum likelihood estimation functions. General theory points to the use of the method of maximum likelihood as giving the most efficient solution (see [5]). With

$$f(x) = \alpha e^{-e^{-\alpha(x-u)}} e^{-\alpha(x-u)} \tag{5.1}$$

the maximum likelihood estimation functions are

$$\begin{aligned} L_u &= -n\alpha(\bar{z}/z_0 - 1) \\ L_\alpha &= n[1/\alpha - (\bar{x} - u) + \partial(\bar{z}/z_0)/\partial\alpha] \end{aligned} \tag{5.2}$$

with variance-covariance matrix

$$\begin{Vmatrix} n\alpha^2 & n(1 - C) \\ n(1 - C) & (n/\alpha^2)[\pi^2/6 + (1 - C)^2] \end{Vmatrix}. \tag{5.3}$$

Thus with

$$\begin{aligned} &X = \sqrt{n}\,(\bar{z}/z_0 - 1], \; Y = \sqrt{n}\,[\alpha(u - \bar{z}e^{-\alpha u}(u + \bar{z}_\alpha/\bar{z})) - (\alpha\bar{x} - 1)]/B \\ &B = \sqrt{\pi^2/6 + (1 - C)^2}, \end{aligned} \tag{5.4}$$

where

$$z_\alpha = \partial[\Sigma e^{-\alpha x_i}/n]/\partial\alpha,$$

the bivariate distribution of X and Y rapidly approaches normality with increasing n, with zero means, unit variances, and correlation coefficient given by (negative, since sign of L_u has been reversed)

$$r = -(1 - C)/(\sqrt{\pi^2/6 + (1 - C)^2}). \tag{5.5}$$

With non-vanishing Jacobian, X and Y constitute a sufficient set of estimation functions for the parameters α and u (see (3.2) above). Furthermore the unit variances and correlation value given above are exact for all values of n. By setting up the moment generating function it is not difficult to show that these functions are also *stable* estimation functions for all values of n.

The theory of maximum likelihood further shows that if $\hat{u}$ and $\hat{\alpha}$ are defined as the u and α solutions of the equations

$$L_u = 0, \qquad L_\alpha = 0 \tag{5.6}$$

the distribution of $\sqrt{n}\,(\hat{u} - u)$ and $\sqrt{n}\,(\hat{\alpha} - \alpha)$ will approach normality asymptotically with zero means and variance-covariance matrix which is the reciprocal of the above matrix (multiplied by n); namely,

$$\begin{Vmatrix} (1/\alpha^2)[1 + (1 - C)^2/(\pi^2/6)] & -(1 - C)/(\pi^2/6) \\ -(1 - C)/(\pi^2/6) & \alpha^2/(\pi^2/6) \end{Vmatrix} \tag{5.7}$$

6. Numerical applications. As an illustration of the application of the methods outlined above for determining the parameters of the distribution of maximum values from an observed sample, data is taken from the 57 year record of annual maximum flood flows previously used as an illustration by the author ([6] p. 324). There is some evidence to indicate that such a series follows approximately the distribution of maximum values. At any rate the series serves pretty well as a numerical illustration.

Confidence regions for u and α can be determined by four methods based upon the preceding theory. In order to make the numerical illustration more cogent, we shall answer the following question by each of the methods. What is the confidence interval (with confidence level .95) for annual flood x corresponding to a cumulated frequency of .99 (often referred to as a 100 yr. flood) based upon our observed 57 yr. sample, under the assumption that the distribution of maximum values (2.1) applies to this data?

Method 1. (Based on estimation functions of section 4.) In this case the statistical estimation functions X_1 and Y_1 defined from (4.5) by $X_1 = X\sqrt{6}/\pi$, $Y_1 = Y$, are used. The "best values" of u and α are taken as the solutions of $X_1 = 0$, $Y_1 = 0$, found by trial and error. As a starting point values of u and α may be estimated from $X_1 = 0$ and the standard deviation of x_i (see

[2] or [6]), the mean deviation of x_i, or an adjusted modal value (see [3]). A few trials gives

$$\hat{u} = 179.7, \qquad \hat{\alpha} = .01998.$$

Approximating the distribution function of X_1 and Y_1 by the limiting normal bivariate distribution (4.10), with confidence level of .95 the equation of the bounding constant probability ellipse is found to be

$$X_1^2 + (1.5594)X_1Y_1 + Y_1^2 = 2.3491 \tag{6.1}$$

where the constants are independent of the sample values. This ellipse, by virtue of the one-to-one correspondence between (X_1, Y_1) and (u, α) bounds u and α based upon the observed sample (see [4]).

For cumulated frequency .99, the distribution of maximum values (2.1) yields

$$t = \alpha(x - u) = 4.60015$$

Thus the analytic problem is that of determining the maximum and minimum value of

$$x = g(u, \alpha) = 4.60015/\alpha + u \tag{6.2}$$

which occurs on the ellipse (6.1).[1]

The writer solved this graphically. It was found necessary to compute three values of $\bar{z}$,—at $\alpha = .01$, .015 and .025, in addition to the value of $\bar{z}$ at $\alpha = .01998$ previously found. From these computations the curves $\alpha = .01$, $\alpha = .015$, $\alpha = .01998$ and $\alpha = .025$ were drawn on the chart of the ellipse (6.1). The u = const. curves were quite easily determined by points on the α = const. curves found from their X_1 coordinates which are linear functions of u (see (4.5)). The extreme values of $g(u, \alpha)$ will be found to occur near the extreme values of α on the ellipse. A construction of several u = const. curves near these extremes enables one to determine several successive values of $g(u, \alpha)$ at points where these curves cross the ellipse. The answers were

$$\begin{aligned} &\text{Max. } g(u, \alpha) = 507.4 \text{ at } u = 192, \alpha = .01459, \\ &\text{Min. } g(u, \alpha) = 360.0 \text{ at } u = 172, \alpha = .02447, \\ &\text{and } g(\hat{u}, \hat{\alpha}) = 409.9. \end{aligned} \tag{6.3}$$

Method 2. (Based on maximum likelihood statistical estimation functions (5.4)). For purposes of comparison the writer carried through the solution using the maximum likelihood estimation functions X_2 and Y_2 defined by (5.4).

[1] Since with non-vanishing Jacobian of (X_1, Y_1) relative to (u, α), no singular point of the (u, α) coordinate system can lie within the ellipse, it is clear from the form of the function $g(u, \alpha)$ that its maximum and minimum values will lie on the boundary of the ellipse. A similar remark applies to Methods 2–4.

In this case the equation of the bounding ellipse was

$$X_2^2 + (.62614)X_2Y_2 + Y_2^2 = 5.4042. \tag{6.4}$$

The determination of the network of α = const., u = const. curves was much more complicated in this case. The results were

Solution of $X = 0$, $Y = 0$, gave $\hat{u} = 180.6$, $\hat{\alpha} = .01924$; $g(\hat{u}, \hat{\alpha}) = 419.7$

$$\begin{aligned} &\text{Max. } g(u, \alpha) = 509.5 \text{ at } u = 187, \alpha = .01426 \\ &\text{Min. } g(u, \alpha) = 364.4 \text{ at } u = 172, \alpha = .02391. \end{aligned} \tag{6.5}$$

The slightly smaller range of estimate of $g(u, \alpha)$ resulting from the use of the second method was forecast from the general theory which predicts a narrowing of range of variation of u and α for same confidence level. Both bivariate distributions involve exact moments of the first and second degree for finite n, and both approach normality rapidly with increasing n. Hence comparable results were to be expected. Of course the form of the function $g(u, \alpha)$ in relation to the different types of estimation functions used in the two cases might modify the comparability of the results.

Method 3. (Based on limiting distribution of maximum likelihood statistics $\hat{u}$ and $\hat{\alpha}$, with variances unknown.) The use of the limiting distribution of the estimation functions $\sqrt{n}\,(\hat{u} - u)$, $\sqrt{n}\,(\hat{\alpha} - \alpha)$ led to results which were not entirely expected by the author. Taking

$$\begin{aligned} X_3 &= A\alpha(\hat{u} - u)/B, \qquad Y_3 = A(\hat{\alpha}/\alpha - 1) \\ A &= \pi\sqrt{n}/\sqrt{6}, \qquad B = \sqrt{\pi^2/6 + (1 - C)^2}, \end{aligned} \tag{6.6}$$

with

$$r = -(1 - C)/B,$$

the equation of the bounding ellipse is the same as (6.4), (no reversal of sign of r occurs because sign of r in (6.4) was reversed by reversing sign of L_u in (5.4)).

Using the inverse method where the range in u and α, with $\hat{u} = 180.6$, $\hat{\alpha} = .01924$, is determined from the range of (X_3, Y_3) within the ellipse (6.4), the maximum and minimum obtained for $g(u, \alpha)$ was

$$\begin{aligned} &\text{Max. } g(u, \alpha) = 490.2 \text{ at } u = 193.2, \alpha = .01549 \\ &\text{Min. } g(u, \alpha) = 353.8 \text{ at } u = 174.0, \alpha = .02558. \end{aligned} \tag{6.7}$$

This result does not agree closely with the previous results. The reason for this discrepancy may be that since the variances indicated by (5.7) are *not* exact for finite n, a variation of α from the central value predicted by (5.6) tends to exaggerate the departure of the distribution of X and Y from the limiting normal distribution through its effect upon the variances. The plausibility of such an explanation is strengthened by the numerical results of a solution of our problem by Method 4.

Method 4. (Based on limiting distribution of maximum likelihood statistics $\hat{u}$ and $\hat{\alpha}$, with variances estimated by taking $\alpha = \hat{\alpha}$ as observed from the sample.) In this case the unknown variances are estimated by taking $\alpha = \hat{\alpha}$ as observed from the sample studied. In order to avoid confusion let α_0 denote this value of α as used in the variance formulae. Thus the estimating functions X_4 and Y_4 become

$$(6.8) \qquad X_4 = A\alpha_0(\hat{u} - u)/B, \qquad Y_4 = A(\hat{\alpha} - \alpha)/\alpha_0$$

and the approximating distribution of (X_4, Y_4) is taken as the same limiting normal distribution used in Method 3. With

$$u_0 = \hat{u} = 180.6, \qquad \alpha_0 = \hat{\alpha} = .01924$$

the extreme values of $g(u, \alpha)$ on the ellipse were

$$(6.9) \qquad \begin{aligned} &\text{Max. } g(u, \alpha) = 507.4 \text{ at } u = 188.6, \alpha = .01443 \\ &\text{Min. } g(u, \alpha) = 362.8 \text{ at } u = 169.7, \alpha = .02382. \end{aligned}$$

These results agree closely with the results obtained by Methods 1 and 2.

The confidence intervals in $g(u, \alpha)$ obtained were, in summary.

Method 1	360.0 to 507.4
Method 2	364.4 to 509.5
Method 3	353.8 to 490.2
Method 4	362.8 to 507.4.

From the analysis of the four methods presented above, one might recommend the following two procedures for finding the confidence interval for x in a problem of the above description, as practicable:

Procedure 1. Use Method 1.

Procedure 2. Determine the maximum likelihood estimates $\hat{u}$ and $\hat{\alpha}$ from (5.6) by trial and error. Then use Method 4. Presumably the second procedure would be more open to question, especially for small values of n.

REFERENCES

[1] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the smallest and largest member of a sample," *Proc. Camb. Phil. Soc.*, Vol. 24 (1928).

[2] E. J. Gumbel, "The return period of flood flows," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 163–190, formulas (27) and (28).

[3] E. J. Gumbel, "Statistical control-curves for flood discharges," *Trans. Am. Geophysical Union*, Report of Section of Hydrology, 1942, p. 489 ff., formulas (6) and (16).

[4] A. Wald, "The fitting of straight lines if both variables are subject to error," *Annals of Math. Stat.*, Vol. 11 (1940), pp. 284–300, especially footnote 13 page 290 and statement (2) at top of page 293.

[5] S. S. Wilks, *Mathematical Statistics*, Princeton Univ. Press, 1943, p. 139.

[6] B. F. Kimball, "Limited type of primary probability distribution applied to annual maximum flood flows," *Annals of Math. Stat.*, Vol. 13 (1942), p. 318–325.

[7] J. Neyman, "Su un teorema concernente le considdetto statistiche sufficienti," *Giornale dell'Istituto Italiano degli Attuari*, Vol. 6 (1935), p. 320 ff.