ON MINIMIZING AND MAXIMIZING A CERTAIN INTEGRAL WITH STATISTICAL APPLICATIONS[1, 2]

By Jagdish Sharan Rustagi

Carnegie Institute of Technology

1. Summary. We consider here the problem of minimizing and maximizing $\int_{-X}^{X}\varphi(x, F(x))\,dx$ under the assumptions that $F(x)$ is a cumulative distribution function (cdf) on $[-X, X]$ with the first two moments given and that φ is a certain known function having certain properties. The existence of the solution has been proved and a characterization of the maximizing and minimizing cdf's given. The minimizing cdf is unique when $\varphi(x, y)$ is strictly convex in y and is completely characterized for some special forms of φ. The maximizing cdf is a discrete distribution and in the above case turns out to be a three-point distribution. Several statistical applications are discussed.

2. Introduction. Let $x_1 \leqq x_2 \leqq \cdots \leqq x_n$ be n ordered independent observations from a population with cdf $F(x)$ having standard deviation σ. Let $w_n = x_n - x_1$ denote the sample range. Then it is well-known that

$$E(w_n) = \int_{-\infty}^{\infty} x\, d\{F^n(x) + (1 - F(x))^n\} \tag{2.1}$$

and

$$E(x_n) = \int_{-\infty}^{\infty} x\, d\{F^n(x)\}. \tag{2.2}$$

Plackett [9] considered the problem of establishing universal upper and lower bounds for $[E(w_n)]/\sigma$ on the lines of Chebycheff inequalities for moments. Moriguti [14] considered an equivalent case of establishing bounds for $E(x_n)$, but he assumed that the underlying distribution is symmetrical.

Gumbel [10] uses a variational method to derive the solution of the problem of maximizing $E(w_n)$ and $E(x_n)$ over the class of continuous cdf's with given mean and variance and gives a sort of sufficiency condition. Hartley and David [1] consider the same problem of maximizing $E(x_n)$ as in [10], and obtain the solution of the problem of maximizing and minimizing $E(w_n)$ but they assume, in addition, that $F(x)$ is a cdf on the bounded range $[-X, X]$.

Integrating (2.1) and (2.2) by parts, we find that the problems of maximizing (minimizing) $E(w_n)$ and $E(x_n)$ are the same as those of minimizing (maximizing)

Received April 6, 1956; revised July 17, 1956.

[1] This paper is based on the author's doctoral dissertation, accepted by Stanford University.

[2] This research was done under the sponsorship of the Office of Naval Research.

$$\int_{-X}^{X} [F(x)^n + (1 - F(x))^n] \, dx \qquad \text{and} \qquad \int_{-X}^{X} F(x)^n \, dx,$$

respectively, with appropriate restrictions on $F(x)$. We see here that $\varphi(x, y) = y^n + (1 - y)^n$ or $\varphi(x, y) = y^n$ is strictly convex in y for $0 \leqq y \leqq 1$.

There are many other situations in statistics where problems of maximizing and minimizing an integral of a strictly convex function of $F(x)$ occur. In the evaluation of efficiencies of various nonparametric tests of hypotheses, we are faced with integrals of the above type. For example, Birnbaum and Klose [7] have derived a lower bound for the variance of the Mann-Whitney Statistic—an improvement on the lower bound due to V. Dantzig, which is based on minimizing $\int_0^1 [F(x) - x]^2 \, dx$, where $F(x)$ is a cdf on $0 \leqq x \leqq 1$ and $\int_0^1 F(x) \, dx = 1 - p \geqq \frac{1}{2}$. Here again $[F(x) - x]^2$ is strictly convex in $F(x)$.

The above problems suggest a generalization.

We consider the problem of maximizing and minimizing an integral of $\varphi(x, F(x))$, where $\varphi(x, y)$ is strictly convex in y. A very special case of this is the one where $\varphi(x, y)$ is a function of y alone and includes the important applications of minimizing and maximizing $E(w_n)$, $E(x_n)$, etc. Many other authors, to name a few such as Chernoff and Reiter [2], Rubin and Isaacson [13], Karlin [3, 4], Hoeffding [11], Hoeffding and Shrikhande [15], Brunk, Ewing and Utz [6], have also considered related problems.

We have used Karlin's [3, 4] technique in the solution of the minimum problem. We compare our technique with that of [2] and [3] in Section 6 and obtain results similar to those in [2] and [3]. The maximum problem is discussed in Section 8. We find that the maximizing cdf is a discrete distribution and in our case a three-point distribution. David and Hartley [1] have shown further that the minimizing cdf for $E(w_n)$ with given restrictions on mean and variance is a two-point distribution. This does not seem to be true in general. Similar results were obtained in [2], [11], and [15].

The results and techniques of our paper have many other applications besides those discussed above. Many classical inequalities of the Chebycheff type can be obtained with the help of our results. In Section 7 we discuss an example where the techniques of this paper yield the solution to a problem of a different type.

3. Statement of the problem and existence of its solution. Let $S = \{(x, y): -X \leqq x \leqq X, 0 \leqq y \leqq 1\}$ where X is already specified.

Let φ be a function defined on the closed and bounded region S such that

(1) φ is bounded and continuous in S,

(2) φ is strictly convex and twice differentiable in y.

We shall minimize (maximize)

$$I(F) = \int_{-X}^{X} \varphi(x, F(x)) \, dx \tag{3.1}$$

over all $F \varepsilon \mathfrak{A}$ where $\mathfrak{A}$ is the class of all admissible cdf's, i.e., cdf's satisfying the following constraints:

$$\int_{-X}^{X} x\, dF(x) = \mu_1, \qquad \int_{-X}^{X} x^2\, dF(x) = \mu_2, \tag{3.2}$$

and

$$F(x) = \begin{cases} 0, & x < -X, \\ 1, & x > \ \ X. \end{cases}$$

Here μ_1 and μ_2 are such that $\mu_2 > \mu_1^2$ and $\mu_2 < X^2$. In this case there exist cdf's satisfying (3.2) and hence class $\mathfrak{A}$ is non-null.

Integrating by parts the integrals in (3.2), the restrictions become

$$\begin{aligned} \int_{-X}^{X} F(x)\, dx &= X - \mu_1, \\ \int_{-X}^{X} xF(x)\, dx &= \frac{X^2 - \mu_2}{2}. \end{aligned} \tag{3.3}$$

We shall first show that an admissible minimizing (maximizing) cdf exists. Let $\mathfrak{F}$ be the class of all cdf's defined on $[-X, X]$. Then the following is well-known [8]:

LEMMA 3.1. *$\mathfrak{F}$ is convex and compact in the topology of convergence in distribution.* (The compactness of Lemma 3.1 is a restatement of the Helly-Bray lemma.)

Define a transformation

$$T: \mathfrak{F} \to R \qquad \text{such that}$$

$$T \circ F = \left(\int_{-X}^{X} \varphi(x, F(x))\, dx, \int_{-X}^{X} F(x)\, dx, \int_{-X}^{X} xF(x)\, dx \right).$$

It is easy to see that T is a continuous transformation as $\varphi(x, y)$ is continuous in y. But a continuous transformation maps a compact set into a closed and bounded set [8]. Hence we have the following lemma.

LEMMA 3.2. *The set Γ of points*

$$\left(\int_{-X}^{X} \varphi(x, F(x))\, dx, \int_{-X}^{X} F(x)\, dx, \int_{-X}^{X} xF(x)\, dx \right), \quad \text{for} \quad F \varepsilon \mathfrak{F},$$

is a closed and bounded set in R.

The restrictions (3.3) define a cross section Γ_1 of a closed and bounded set Γ, and hence Γ_1 is also closed and bounded. Therefore, the minimizing and maximizing points exist and are given by the boundary points of Γ_1 so long as Γ_1 is non-null. But Γ_1 is non-null as $\mathfrak{A}$ is non-null as seen before. Hence the minimizing and maximizing admissible cdf's exist.

4. Reduction of the minimum problem to subsidiary problems and uniqueness of its solution. In this section we first prove the uniqueness of the minimizing cdf, using the property of strict convexity of the function φ in its second argument. In characterizing the solution of the minimizing problem, we use Karlin's method [3] to reduce the main problem of minimizing the integral (3.1) over the class $\mathcal{A}$ of admissible cdf's, to a subsidiary problem of minimizing an integral of a related function over all cdf's $\mathcal{F}$. This reduction together with the uniqueness of the minimizing cdf gives us a characterization of the minimizing cdf which we give in the next section.

LEMMA 4.1. *There is a unique cdf* F_0, *which minimizes* (3.1) *subject to the side conditions* (3.3) *when* $\varphi(x, y)$ *is strictly convex in* y.

PROOF. Suppose the solution is not unique. Let $F_0(x)$ and $F_1(x)$ be two distinct admissible cdf's which minimize (3.1). Let

$$M = \min_{F \varepsilon \mathcal{A}} \int_{-X}^{X} \varphi(x, F(x))\, dx.$$

As φ is strictly convex in y, for $0 < \lambda < 1$,

$$\begin{aligned} &\int_{-X}^{X} \varphi(x, \lambda F_0(x) + (1-\lambda)F_1(x))\, dx \\ &\qquad < \lambda \int_{-X}^{X} \varphi(x, F_0(x)\, dx + (1-\lambda) \int_{-X}^{X} \varphi(x, F_1(x)\, dx \\ &\qquad = \lambda M + (1-\lambda)M \\ &\qquad = M. \end{aligned} \tag{4.0}$$

But M is the minimum, and hence we have a contradiction.

We shall now prove the following lemmas, with the help of which we shall reduce the main problem to a simpler problem.

LEMMN 4.2. $F_0(x)$ *minimizes* (3.1) *if and only if*

$$\int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, y)\,|_{y=F_0(x)} F(x)\, dx \geqq \int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, y)\,|_{y=F_0(x)} F_0(x)\, dx \tag{4.1}$$

for all $F \varepsilon \mathcal{A}$.

PROOF. For any other admissible cdf $F(x)$, define

$$I(\lambda) = \int_{-X}^{X} \varphi(x, \lambda F_0(x) + (1-\lambda)F(x))\, dx, \qquad 0 \leqq \lambda \leqq 1.$$

As φ is twice differentiable in y, $\partial\varphi/\partial y$ exists and is continuous in y, and hence $I(\lambda)$ is differentiable and is given by

$$I'(\lambda) = \int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, \lambda F_0(x) + (1-\lambda)F(x))(F_0(x) - F(x))\, dx.$$

Since φ is strictly convex in y, it follows very easily that $I(\lambda)$ is a strictly convex function of λ. If $F_0(x)$ minimizes (3.1), then $I(\lambda)$ achieves its minimum at $\lambda = 1$; and this is possible if and only if

$$I'(\lambda)\,|_{\lambda=1} \leqq 0,$$

i.e.,

$$\int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, y)\,|_{y=F_0(x)}[F_0(x) - F(x)]\, dx \leqq 0,$$

or,

$$\int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, y)\,|_{y=F_0(x)} F_0(x)\, dx \leqq \int_{-X}^{X} \frac{\partial}{\partial y}(x, y)\,|_{y=F_0(x)} F(x)\, dx.$$

Conversely let (4.1) hold true. Then we have $I'(\lambda)\,|_{\lambda=1} \leqq 0$, and hence by the strict convexity of $I(\lambda)$ we have $I(1) < I(0)$ or,

$$\int_{-X}^{X} \varphi(x, F_0(x))\, dx < \int_{-X}^{X} \varphi(x, F(x))\, dx,$$

i.e., $F_0(x)$ minimizes (3.1). This proves the lemma.

We use the following notation:

$$I_{F_0}(F) = \int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, F_0(x))F(x)\, dx.$$

With the help of the above lemma, we find that the problem P_1 of minimizing $I(F)$ over all admissible cdf's, is related to the problem P_{2F_0} of finding an admissible $F(x)$ which minimizes $I_{F_0}(F)$. In fact, we are interested in finding an F_0 such that F_0 is a solution of P_{2F_0}. This looks like a complicated problem, but we now have a problem linear in F which is relatively easy to deal with.

Because P_1 has a unique solution, Lemma 4.2 implies that there is one and only one F_0 such that F_0 solves P_{2F_0}. This, however, does not mean that P_{2F_0} has a unique solution.

Let $T:\mathfrak{F} \rightarrow \Gamma_2$ be a transformation given by

$$T \circ F = \left(\int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, F_0(x))F(x)\, dx, \int_{-X}^{X} F(x)\, dx, \int_{-X}^{X} xF(x)\, dx\right).$$

Obviously T is bounded and is linear and hence continuous in F. But as $\mathfrak{F}$ is convex and compact in the topology of convergence in distribution by Lemma 3.1, the transformation T maps the convex and compact set into a convex and compact set Γ_2, and hence we have the following result.

Lemma 4.3. *Γ_2 is a convex, closed and bounded set in three dimensions.*

Solving P_{2F_0} corresponds to finding the minimum among all points of Γ_2 for which

$$\int_{-X}^{X} F(x)\, dx = X - \mu_1,$$

$$\int_{-X}^{X} xF(x)\, dx = \frac{X^2 - \mu_2}{2},$$

and this will be a boundary point of the set Γ_2.

Suppose F_0 solves P_{2F_0}. Then F_0 corresponds to a boundary point of Γ_2, and there is a supporting hyperplane of Γ_2 at the minimum point $\gamma = (u_0, v_0, w_0)$, i.e., for some η_0, η_1, η_2 and $\eta_3(\eta_0, \eta_1, \eta_2$ not all zero),

$$\eta_0 u_0 + \eta_1 v_0 + \eta_2 w_0 + \eta_3 = 0, \tag{4.2}$$

and

$$\eta_0 u + \eta_1 v + \eta_2 w + \eta_3 \geqq 0 \tag{4.3}$$

for all other points (u, v, w) belonging to Γ_2, where

$$u = \int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, F_0(x))F(x)\,dx,$$

$$v = \int_{-X}^{X} F(x)\,dx, \qquad w = \int_{-X}^{X} xF(x)\,dx,$$

therefore,

$$\eta_0(u - u_0) + \eta_1(v - v_0) + \eta_2(w - w_0) \geqq 0. \tag{4.5}$$

We shall see below that η_0 can be taken positive and hence can be normalized so as to be equal to one. Therefore, by taking $\eta_0 = 1$ in (4.5) we have

$$\int_{-X}^{X}\left[\frac{\partial \varphi}{\partial y}(x, F_0(x)) + \eta_1 + \eta_2 x\right]F(x)\,dx \geqq \int_{-X}^{X}\left[\frac{\partial \varphi}{\partial y}(x, F_0(x)) + \eta_1 + \eta_2 x\right]F_0(x)\,dx.$$

Hence $F_0(x)$ minimizes

$$\int_{-X}^{X}\left[\frac{\partial \varphi}{\partial y}(x, F_0(x)) + \eta_1 + \eta_2 x\right]F(x)\,dx \tag{4.6}$$

among the class $\mathfrak{F}$ of all cdf's on $[-X, X]$.

Conversely, if $F_0(x) \varepsilon\, \mathfrak{A}$ minimizes (4.6), we have, retracing the steps, that $(u - u_0) + \eta_1(v - v_0) + \eta_2(w - w_0) \geqq 0$. Suppose $F(x)$ is admissible, and hence $v = v_0$ and $w = w_0$, and hence $u - u_0 \geqq 0$, i.e.,

$$\int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, F_0(x))F(x)\,dx \geqq \int_{-X}^{X} \frac{\partial \varphi}{\partial y}(x, F_0(x))F_0(x)\,dx.$$

In other words, F_0 minimizes $I_{F_0}(F)$ over all admissible cdf's $\mathfrak{A}$.

We shall now show that η_0 can be taken positive. Let

$$\Gamma_2^* = \{(u^*, v, w): u^* \geqq u, (u, v, w)\, \varepsilon\, \Gamma_2\}.$$

Then Γ_2^* is obviously convex and $\Gamma_2 \subseteq \Gamma_2^*$. u_0 is the minimum of u subject to the conditions that $v = v_0$ and $w = w_0$. This implies that (u_0, v_0, w_0) is also a minimum point of Γ_2^* and hence is its boundary point. Hence there is an $(\eta_0, \eta_1, \eta_2) \neq (0, 0, 0)$ such that

$$\eta_0(u^* - u_0) + \eta_1(v - v_0) + \eta_2(w - w_c) \geqq 0 \tag{4.7}$$

for points (u^*, v, w) belonging to Γ_2^*. Hence

$$\eta_0(u - u_0) + \eta_1(v - v_0) + \eta_2(w - w_0) \geqq 0$$

for $(u, v, w) \varepsilon \Gamma_2$. Suppose $\eta_0 = 0$. Then we have

$$\eta_1(v - v_0) + \eta_2(w - w_0) \geqq 0,$$

or

$$\int_{-X}^{X} (\eta_1 + \eta_2 x)F(x)\, dx \geqq \int_{-X}^{X} (\eta_1 + \eta_2 x)F_0(x)\, dx,$$

i.e., F_0 minimizes $\int_{-X}^{X} (\eta_1 + \eta_2 x)F(x)\, dx$ over all $F \varepsilon \mathfrak{F}$. Now $\eta_1 + \eta_2 x$ is either nondecreasing or nonincreasing according as $\eta_2 \geqq 0$ or $\eta_2 \leqq 0$, and hence the unique minimizing cdf of the above integral is a two-point distribution with its total mass concentrated at $-X$ and X so that $\mu_2 = X^2$. But such a cdf is not admissible and hence there is a contradiction.

It is easily seen now that η_0 is not negative. Suppose η_0 is negative. Consider then a point $(u_0 + h, v_0, w_0) \varepsilon \Gamma_2^*$ for some $h > 0$, so that from (4.7) we obtain

$$\eta_0 h \geqq 0,$$

which is again a contradiction. Hence η_0 is positive.

REMARK. Another way to show that $\eta_0 \neq 0$ would be as follows: $\eta_0 = 0$ corresponds to boundary points of the set Γ_2 where the supporting hyperplanes are parallel to the u-axis, and hence (v_0, w_0) corresponds to the boundary of the projection Γ_2 on the $(v. w)$ plane. But the conditions on the first two moments are such that the given point (v_0, w_0) will be interior to the projection set, and hence $\eta_0 \neq 0$.

The previous argument applies only in the special case of the first two moments of $F(x)$ being given. In general when more moments are specified, the latter argument will apply if we impose conditions on the given moments such that the given point is interior to the moment space which is analogous to the projection of the set Γ_2. It easily follows then that $\eta_0 > 0$, in general. Let

$$I_{F_0\eta_1\eta_2}(F) = \int_{-X}^{X} \left[\frac{\partial\varphi}{\partial y}(x, F_0(x)) + \eta_1 + \eta_2 x\right] F(x)\, dx,$$

and let the problem $P_{3F_0\eta_1\eta_2}$ be that of finding the minimum of $I_{F_0\eta_1\eta_2}(F)$ over all $F(x) \varepsilon \mathfrak{F}$.

The above results are summarized in the following lemma.

LEMMA 4.4. *F_0 solves P_{2F_0} if F_0 is an admissible cdf which solves $P_{3F_0\eta_1\eta_2}$, and any F_0 which solves P_{2F_0} solves $P_{3F_0\eta_1\eta_2}$ for some η_1 and η_2.*

5. Characterization of the solution. In this section we characterize the solution of the minimum problem in terms of $f_{\eta_1\eta_2}(x)$ which is that value of y for which

$$\frac{\partial\varphi}{\partial y}(x, y) + \eta_1 + \eta_2 x = 0.$$

Let

$$A(x) = B_{\eta_1\eta_2}(x, F_0(x)) = \frac{\partial\varphi}{\partial y}(x, F_0(x)) + \eta_1 + \eta_2 x.$$

Since $\frac{\partial\varphi}{\partial y}(x, y)$ is continuous in y, $A(x)$ can have a discontinuity only if $F_0(x)$ has a jump. But as $\frac{\partial\varphi}{\partial y}(x, y)$ is increasing in y, the discontinuities of $A(x)$ are upward jumps. Also since φ is continuous in the region

$$S = \{(x, y): -X \leqq x \leqq X, 0 \leqq y \leqq 1\},$$

$A(x)$ is bounded in S. We then have the following theorems which will characterize the solution of our problems.

THEOREM 5.1. *If F_0 solves $P_{3F_0\eta_1\eta_2}$, the set $\{x: A(x) \neq 0, -X < x < X\}$ has F_0-measure zero.*

PROOF. Suppose that F_0 is continuous on the right. Consider the set $S_p = \{x: A(x) > 0, -X < x < X\}$. It is a denumerable union of intervals $[x_1, x_2)$. We shall show that

$$F_0(x_2) = F_0(x_1 -)$$

and therefore, the interval $[x_1, x_2)$ has F_0-measure zero. Suppose this were not the case. Then as $A(x) > 0$ and is increasing in y, so that

$$\int_{x_1}^{x_2} A(x)F_0(x_1 -)\, dx < \int_{x_1}^{x_2} A(x)F_0(x)\, dx,$$

there is a contradiction. Consequently, S_p has F_0-measure zero.

Consider now the set $S_q = \{x: A(x) < 0, -X < x < X\}$. Because all discontinuities are upward jumps and $A(x)$ is continuous on the right, S_q is an open set. Hence S_q is denumerable union of intervals (x_1, x_2). Then we shall see that $F_0(x_2) = F_0(x_1) = 0$ and that the F_0-measure of the interval (x_1, x_2) is zero. We also prove this by contradiction, as otherwise,

$$\int_{x_1}^{x_2} A(x)F_0(x_2)\, dx < \int_{x_1}^{x_2} A(x)F_0(x)\, dx.$$

Since S_q is a denumerable union of such intervals, S_q has also F_0-measure zero.

Hence the above arguments show that the set $\{x: A(x) \neq 0, -X < x < X\}$ has F_0-measure zero.

REMARKS. 1. It is easy to see that if $A(-X) > 0$, then $F_0(-X) = 0$ and if $A(X) < 0$, F_0 is continuous at X.

2. The following corollary shows that the integral of $A(x)$ is zero over intervals on which F is constant.

COROLLARY. *If $F_0(x)$ be such that $F_0(x) = c$, $0 < c < 1$ for $a \leqq x < b$ and $F_0(x) < c$ for $x < a$, $F_0(x) > c$ for $x > b$, then*

$$\int_a^b A(x)\, dx = 0.$$

PROOF. Suppose $\int_a^b A(x)\,dx < 0$ and $b < X$. Replace $F_0(x)$ on the interval $[a, b + \delta)$ for any small number $\delta > 0$, by the constant quantity $F_0(b + \delta)$. Let ν be the increase in $I_{F_0\eta_1\eta_2}(F)$ due to this replacement. Then

$$\nu = \int_a^{b+\delta} A(x)F_0(b_i + \delta)\,dx - \int_a^{b+\delta} A(x)F_0(x)\,dx$$

$$= (F_0(b + \delta) - c)\int_a^{b+\delta} A(x)\,dx - \int_a^{b+\delta} A(x)[F_0(x) - c]\,dx$$

$$\leqq [F_0(b + \delta) - c]\left(\int_a^b A(x)\,dx + 2M\delta\right), \quad \text{where} \quad |A(x)| < M.$$

Letting $\delta \to 0$, we find that ν becomes negative and hence there is a contradiction. The case where $b = X$ is trivial.

If we suppose that $\int_a^b A(x)\,dx > 0$ and a is a point of continuity of F_0, then by an argument similar to that above, we get the contradiction when $a > -X$, by replacing F_0 on $(a - \delta, b)$ by $F_0(a - \delta)$ and letting $\delta \to 0$. In case $a = -X$ or there is a jump in F_0 at a, the proof is trivial.

REMARK. If $K(x)$ is a function satisfying the properties of the function $A(x)$, then the problem P_K (corresponding to $P_{3F_0\eta_1\eta_2}$) of finding a cdf F_0 such that

$$I_K(F_0) = \min_{F\epsilon\mathfrak{F}} I_K(F) = \min_{F\epsilon\mathfrak{F}} \int_{-X}^{X} K(x)F(x)\,dx,$$

gives the same results as stated in Theorem 5.1 and its corollary, i.e., if F_0 is the solution of P_K,

(a) the set $\{x: K(x) \neq 0, -X < x < X\}$ is of F_0-measure zero.

(b) $\int K(x)F_0(x)\,dx$ is zero over intervals where F_0 is constant.

THEOREM 5.2. *If F_0 solves $P_{3F_0\eta_1\eta_2}$, then F_0 has no jumps on the open interval $(-X, X)$ and hence $A(x)$ is continuous on $(-X, X)$.*

PROOF. Let F_0 have a jump at x_0, $-X < x_0 < X$. Then by Theorem 5.1, $A(x_0) = 0$. But since $\partial\varphi/\partial y$ is strictly increasing in y, x_0 is the right-hand end-point of an interval on which $A(x) < 0$. By the same arguments as in the proof of the Theorem 5.1, we see that on this interval $F_0(x) = F_0(x_0)$, and hence F_0 has no discontinuity. But because discontinuities of $A(x)$ arise on account of jumps of F_0, there are no discontinuities in $A(x)$, or $A(x)$ is continuous on the interval $(-X, X)$.

Let $f_{\eta_1\eta_2}(x)$ be defined with $0 \leqq f_{\eta_1\eta_2}(x) \leqq 1$ such that $B_{\eta_1\eta_2}(x, f_{\eta_1\eta_2}(x)) = 0$. (The function $f_{\eta_1\eta_2}$ is defined on that subset of $[-X, X]$ for which there exists a y between 0 and 1 such that $B_{\eta_1\eta_2}(x, y) = 0$.)

As $\partial\varphi(x, y)/\partial y$ is continuous and strictly increasing in y, $f_{\eta_1\eta_2}(x)$ is continuous wherever it is defined. If $0 < f_{\eta_1\eta_2}(x_0) < 1$, then $f_{\eta_1\eta_2}$ is defined in some interval about x_0 (the interval is one-sided if $x_0 = \pm X$). Graphically $f_{\eta_1\eta_2}$ represents a number of curve segments which terminate when $f_{\eta_1\eta_2}(x)$ is zero or one.

More specifically, $f_{\eta_1\eta_2}(x)$ is defined on the union of closed intervals at the end-points of such of which it is either zero or one. Let $[a_i, b_i]$ and $[a_j, b_j]$ be two such intervals, not separated by any others such that $b_i \leqq a_j$, then $f_{\eta_1\eta_2}(b_i) = f_{\eta_1\eta_2}(a_j)$. If there are an infinite number of intervals $[a_j, b_j]$ in the neighborhood of b_i, it follows that

$$f_{\eta_1\eta_2}(b_i) = f_{\eta_1\eta_2}(a_j) = f_{\eta_1\eta_2}(b_j)$$

for b_j sufficiently close to b_i. Hence the following definition of a function $g_{\eta_1\eta_2}$ has a meaning.

DEFINITION. Define $g_{\eta_1\eta_2}$ to be that unique function on $[-X, X)$ which is continuous on $[-X, X)$ such that

$$g_{\eta_1\eta_2}(x) = \begin{cases} f_{\eta_1\eta_2}(x), & \text{where } f_{\eta_1\eta_2} \text{ is defined,} \\ 0 \text{ or } 1, & \text{elsewhere,} \end{cases}$$

and

$$g_{\eta_1\eta_2}(X) = 1,$$

provided that the subset of $[-X, X)$ for which $f_{\eta_1\eta_2}$ is defined, is non-null.

THEOREM 5.3. *If F_0 solves $P_{3F_0\eta_1\eta_2}$, then for $-X \leqq x < X$, F_0 coincides with $g_{\eta_1\eta_2}$ except on intervals on which F_0 is constant.*

PROOF. From Theorems 5.1 and 5.2, we know that F_0 has no jumps on $(-X, X)$ and F_0 cannot increase when $A(x) \neq 0$. Therefore, F_0 remains constant until it intersects with $f_{\eta_1\eta_2}$.

COROLLARY. *If $g_{\eta_1\eta_2}$ is a cdf, then $F_0(x) = g_{\eta_1\eta_2}(x)$.*

REMARKS. 1. We can represent a conceivable situation by Fig. 1.

2. It must be noted that the corollary to Theorem 5.1 puts a strong restriction on the intervals on which F_0 is constant.

3. The solution in the general case may not be completely specified, but we shall consider in the following some special cases where the minimizing cdf is completely characterized.

Special Cases.

I. When $\partial\varphi(x, y)/\partial y$ is nonincreasing in x.

THEOREM 5.4. *If $\partial\varphi(x, y)/\partial y$ is nonincreasing in x and $\eta_2 < 0$, then $F_0(x) = g_{\eta_1\eta_2}(x)$ for $-X < x < X$.*

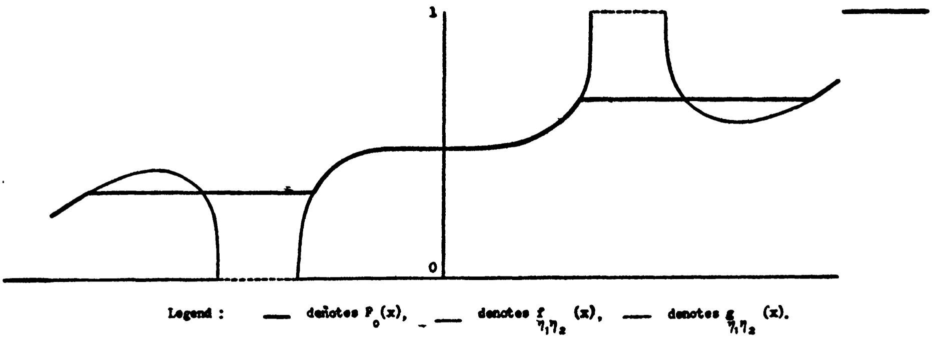


FIG. 1

PROOF. If the conditions of the theorem hold, $A(x)$ is a decreasing function of x and hence $f_{\eta_1\eta_2}$ is increasing in x so that $g_{\eta_1\eta_2}$ is a cdf. We get the result of the theorem, then, by the corollary to Theorem 5.3.

II. When φ is a function of y alone, i.e., $\varphi(x, y) = \psi(y)$.

LEMMA 5.1. *If* $\varphi(x, y) = \psi(y)$, *then corresponding to* F_0, *which is the solution of* P_1, η_2 *is negative.*

PROOF. If $\eta_2 \geqq 0$, $\eta_1 + \eta_2 x$ is nondecreasing. Also as $\psi'(y)$ is nondecreasing in x, the function $A(x) = \psi'(F_0) + \eta_1 + \eta_2 x$ is nondecreasing in x. Therefore, $f_{\eta_1\eta_2}(x)$ is nonincreasing, and hence from Theorem 5.3, it follows that F_0 is constant on $[-X, X)$. But for such F_0, $\mu_2 = X^2$, and hence F_0 is not admissible. Therefore, $\eta_2 < 0$.

Theorem 5.4 and Lemma 5.1 imply, then, the following theorem.

THEOREM 5.5. *If* $\varphi(x, y) = \psi(y)$, *the solution* F_0 *of* P_1 *is given by* $g_{\eta_1\eta_2}$ *for some* η_1, η_2.

REMARK. Unfortunately it is not always true that $\eta_2 < 0$ as assumed in Theorem 5.4. In fact for side conditions corresponding to small variance, one has $\eta_2 > 0$. It might still happen that $f_{\eta_1\eta_2}$ is nondecreasing, and then the result of Theorem 5.4 still holds. In any case Theorem 5.3 with the corollary to Theorem 5.1 gives a useful characterization of the solution of our problem.

6. Comparison of our Technique with that of Gumbel [10] and Chernoff and Reiter [2].

(i) *Gumbel's Method.* The problem considered by Gumbel is that of maximizing

$$\int_0^1 x(F)nF^{n-1}\,dF, \tag{6.1}$$

with restrictions

$$\int_0^1 x(F)\,dF = 0, \qquad \int_0^1 x^2(F)\,dF = 1.$$

A variational technique has been used to derive the form obtained for the maximizing cdf is given by equating to zero, the first variation of

$$\int \{nx(F)F^{n-1} + \eta_1 x(F) + \eta_2 x^2(F)\}\,dF, \tag{6.2}$$

i.e.,

$$nF^{n-1} + \eta_1 + 2\eta_2 x(F) = 0, \qquad 0 \leqq F \leqq 1. \tag{6.3}$$

The above equation gives a sort of sufficiency condition as any admissible F given by (6.3) does maximize the integral (6.2) and hence maximizes (6.1). David and Hartley [1] have given an ingenious argument to prove the sufficiency of the solution, but that seems unnecessary. However, the above equation does not give the necessity of the solution, since this approach does not provide an argument for proving that the constants η_1 and η_2, to make the cdf admissible, always exist.

This method also extends to the case of a bounded random variable as treated in this paper.

We shall use the above approach for our problem. Integrating by parts, we have

$$\int_{-X}^{X} \varphi(x, F(x))\, dx = x\varphi(x, F(x)) \Big]_{-X}^{X} - \int_{-X}^{X} x \frac{d\varphi}{dx}(x, F)\, dx.$$

Now

$$d\varphi(x, F(x)) = \frac{\partial\varphi}{\partial x}(x, F(x))\, dx + \frac{\partial\varphi}{\partial y}(x, F(x))\, dF.$$

When $\varphi(x, F(x))$ is a function of F alone, say, $\psi(F(x))$, $d\psi = \psi'(F)\, dF$, and hence the problem of minimizing

$$\int_{-X}^{X} \psi(F)\, dx$$

is the same as that of maximizing

$$\int_0^1 x(F)\psi'(F)\, dF.$$

Hence using Gumbel's approach, we maximize

$$\int_0^1 [x(F)\psi'(F) + \eta_1 x(F) + \eta_2 x^2(F)]\, dF$$

and get the following equation satisfied by the admissible maximizing cdf,

$$\psi'(F) + \eta_1 + 2\eta_2 x(F) = 0.$$

In the above case, our technique would also lead to a similar equation. But in the general case when we consider $\varphi(x, F(x))$, Gumbel's approach does not seem to apply.

(ii) *Chernoff and Reiter Method.* Chernoff and Reiter [2] consider the problem of minimizing and maximizing

$$\int g(x)\, dF(x),$$

with side conditions

$$\int x\, dF(x) = c_1, \qquad \int x^2\, dF(x) = c_2$$

such that $c_2 > c_1^2$ and $g(x)$ is a continuous function of x.

In the process of reduction of our main problem, we have an intermediary problem P_{2F_0} of finding the minimum of

$$I_{F_0}(F) = \int_{-X}^{X} \frac{\partial\varphi}{\partial y}(x, F_0(x))F(x)\, dx$$

over all admissible cdf's $\mathfrak{A}$. Now as

$$f(x) = \int_{-X}^{x} \frac{\partial \varphi}{\partial y}(x, F_0(x))\, dx$$

is continuous in x, integrating by parts $I_{F_0}(F)$, we have

$$I_{F_0}(F) = \text{constant} - \int_{-X}^{X} f(x)\, dF(x),$$

or we maximize

$$\int_{-X}^{X} f(x)\, dF(x)$$

over all admissible cdf's $\mathfrak{A}$. Now as $f(x)$ is continuous, by the methods of Chernoff and Reiter, the necessary condition for the maximum is given by the following.

(a) There is an η_1 and η_2 such that when x is a point of continuity of $F(x)$,

$$B_{\eta_1\eta_2}(x, F_0(x)) = \frac{\partial \varphi}{\partial y}(x, F_0(x)) + \eta_1 + 2\eta_2 x = 0,$$

except on a set of F_0-measure zero

(b) F_0 has no jump in $-X < x < X$, otherwise either $B_{\eta_1\eta_2}(x, F_0(x)) > 0$ or $B_{\eta_1\eta_2}(x, F_0(x-)) < 0$. Hence we get a result similar to our result obtained in Section 5, i.e., the set

$$\{x : B_{\eta_1\eta_2}(x, F_0(x)) \neq 0, \quad -X < x < X\}$$

has F_0-measure zero.

7. Examples. In this section we discuss some examples to illustrate the method of obtaining the minimizing cdf for our problem for some specified function φ. We also discuss an example of the special case $\varphi(x, y) = \psi(y)$. We have included an example where the methods of the paper have been used to solve a problem of a different type.

EXAMPLE 1. Consider the problem of finding

$$\min_{F \varepsilon \mathfrak{A}} \frac{1}{2} \int_{-X}^{X} [F(x) - x]^2\, dx,$$

when $\mathfrak{A}$ denotes the admissible class of cdf's as in Section 3.

This is the special case of a more general problem where we minimize $\int_{-X}^{X} \psi(F(x) - x)\, dx$ for $F \varepsilon \mathfrak{A}$. Here $\psi(y - x) = \frac{1}{2}(y - x)^2$, ψ being a strictly convex, bounded, and continuous function of its argument. This problem has also been discussed by Birnbaum and Klose [7], as a lemma to derive a lower bound for the variance of the Mann-Whiteny Statistic.

If $\varphi(x, y) = \psi(y - x)$, φ is strictly convex in y and it is easy to verify that $(\partial^2\varphi)/(\partial x \partial y) < 0$. We know that the solution exists and is unique, and the problem is reduced to that of finding an η_1 and η_2 so that F_0 solves $P_{3F_0\eta_1\eta_2}$ where $P_{3F_0\eta_1\eta_2}$ is the problem of finding $F \varepsilon \mathfrak{F}$ which minimizes

$$\int_{-X}^{X} [\psi'(F_0(x) - x) + \eta_1 + \eta_2 x]F(x)\, dx.$$

Then by the theorems of Section 5, we know that $F_0(x)$ is given in terms of $g_{\eta_1\eta_2}(x)$, where $g_{\eta_1\eta_2}(x)$ is uniquely expressed in terms of the function

$$f_{\eta_1\eta_2}(x) = x + \psi'^{-1}(-\eta_1 - \eta_2 x).$$

Returning to our example, we have the function

$$f_{\eta_1\eta_2}(x) = (1 - \eta_2)x - \eta_1,$$

$$f_{\eta_1\eta_2}(x) = 0 \quad \text{for} \quad x_1 = \frac{\eta_1}{1 - \eta_2}, \qquad f_{\eta_1\eta_2}(x) = 1 \quad \text{for} \quad x_2 = \frac{1 + \eta_1}{1 - \eta_2}.$$

Case 1. $\eta_2 < 1$. Then $f_{\eta_1\eta_2}$ is increasing. Define $g_{\eta_1\eta_2}(x)$ by the following.

$$g_{\eta_1\eta_2}(x) = \begin{cases} 0, & x < \max(-X, x_1), \\ 1, & x \geqq \min(x_2, X), \\ (1 - \eta_2)x - \eta_1, & \text{elsewhere}. \end{cases}$$

As $g_{\eta_1\eta_2}$ is a cdf, $F_0(x) = g_{\eta_1\eta_2}(x)$.

Case 2. $\eta_2 \geqq 1$, $f_{\eta_1\eta_2}(x)$ is nonincreasing, and hence the solution is either a one-point distribution or a two-point distribution with total probability concentrated at $-X$ and X. In both cases, then, the solution is not admissible.

EXAMPLE 2. Consider the same problem as in Example 1, but with an additional restriction on the cdf $F(x)$, i.e., $F(x) \geqq x$. Now let $F(x)$ be a cdf on [0, 1].

Under this additional restriction, the class $\mathfrak{A}^*$ of admissible cdf's is also compact and convex. Then the solution to this problem exists. It is unique since $\varphi(x, y) = \frac{1}{2}(y - x)^2$ is strictly convex in y.

Applying the methods used to prove Lemma 4.2, we see that the problem is the same as that of minimizing

$$\int_0^1 [F_0(x) - x]F(x)\, dx$$

over all $F \varepsilon \mathfrak{A}^*$. It is easy to see that the set analogous to Γ_2 of Lemma 4.3 here is also convex, closed and bounded, and hence, applying the method of Lemma 4.4, we reduce the problem to that of finding the cdf's corresponding to

$$\min_{F \varepsilon \mathfrak{F}} \int_0^1 [F_0(x) - x + \eta_1 + \eta_2 x]F(x)\, dx$$

or

$$\min_{F \varepsilon \mathfrak{F}} \int_0^1 [F_0(x) + \eta_1 + \eta_3 x]F(x)\, dx, \qquad \eta_3 = \eta_2 - 1$$

where $\mathfrak{F}$ is the class of all cdf's F on [0, 1] such that $F(x) \geqq x$. We can now apply the methods of Section 5. Define the function $f_{\eta_1\eta_3}$ with $x \leqq f_{\eta_1\eta_3}(x) \leqq 1$ such that

$$f_{\eta_1\eta_3}(x) = -\eta_1 - \eta_3 x.$$

Define the function $g_{\eta_1\eta_3}$ such that

$$g_{\eta_1\eta_3}(x) = \begin{cases} f_{\eta_1\eta_3}(x), & \text{where } f_{\eta_1\eta_3} \text{ is defined,} \\ x \text{ or } 1, & \text{elsewhere on } [0, 1], \end{cases}$$

$$g_{\eta_1\eta_3}(x) \text{ is continuous on } [0, 1], \text{ and } g_{\eta_1\eta_3}(1) = 1.$$

Then $g_{\eta_1\eta_3}$ gives the solution F_0 of the problem if $g_{\eta_1\eta_3}(x)$ is a cdf. We shall give an explicit characterization of $g_{\eta_1\eta_3}$ in the various possible cases. Let the point of intersection of $y = -\eta_1 - \eta_3 x$ and $y = x$ be denoted by $x^* = -[\eta_1/(1 + \eta_3)]$. Let x^{**} be such that $-\eta_1 - \eta_3 x^{**} = 1$.

Case I. $\eta_3 > -1$.

(a) $x^* \leqq 0$

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leqq x < 1. \\ 1, & x \geqq 1. \end{cases}$$

This $g_{\eta_1\eta_3}$ is a cdf, but it is not admissible.

(b) $x^* > 1$

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0, \\ -\eta_1 - \eta_3 x, & 0 < x < x^{**}, \\ 1, & x > x^{**}. \end{cases}$$

If $x^{**} > 0$, $g_{\eta_1\eta_3}$ is a cdf and hence $F_0(x) = g_{\eta_1\eta_3}(x)$.

If $x^{**} < 0$, the solution is a one-point distribution with mass concentrated at $x = 0$, which is not admissible.

(c) $0 < x^* < 1$. Consider (i) $-1 < \eta_3 < 0$

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0 \\ -\eta_1 - \eta_3 x, & 0 \leqq x < x^* \\ x, & x^* \leqq x \leqq 1 \\ 1, & x \geqq 1. \end{cases}$$

This is again a cdf, and hence $F_0(x) = g_{\eta_1\eta_3}(x)$.

(ii) $\eta_3 > 0$

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0, \\ -\eta_1 - \eta_3 x, & 0 \leqq x \leqq x^*, \\ x, & x^* \leqq x \leqq 1, \\ 1, & x \geqq 1. \end{cases}$$

This $g_{\eta_1\eta_3}$ is not a cdf. Here $A(x) = F_0(x) + \eta_1 + \eta_3 x$. At $x = 0$, $A(x) > 0$ if $F_0(x) > -\eta_1$ and $A(x) < 0$ if $F_0(x) < -\eta_1$. Suppose there is a jump at $x = 0$ such that $F_0(0) > -\eta_1$, then $A(0) > 0$, and $F_0(x)$ is then continuous at $x = 0$. Hence there is a contradiction. Let there be a jump at $x = 0$ such that $c = F_0(0) < -\eta_1$, $A(0) < 0$ and hence we take as a possible minimizing cdf

$$G(x) = \begin{cases} 0, & x < 0, \\ c, & 0 \leqq x < c, \\ x, & c \leqq x < 1, \\ 1, & x > 1. \end{cases}$$

The remark after Theorem 5.1 puts the following restriction on c,

$$\int_0^c (c + \eta_1 + \eta_3 x)\, dx = 0,$$

i.e.,

$$(c + \eta_1)c + \frac{\eta_3}{2} c^2 = 0, \qquad \text{or} \qquad c = -[\eta_1/(\eta_3/2 + 1)],$$

since $c = 0$ gives an inadmissible cdf. Incidentally this shows also, as is evident from the value of c itself that

$$x^* < c < -\eta_1 .$$

The unique value of c which is obtained from the constraints exists if and only if

$$(7.1) \qquad (1 - 2\mu_1)^3 = (1 - 3\mu_2)^2.$$

This condition is obtained by eliminating c between the equations

$$\int_0^1 F(x)\, dx = 1 - \mu_1 = c^2 + \frac{1 - c^2}{2} = \frac{c^2 + 1}{2}$$

and

$$2 \int_0^1 xF(x)\, dx = 1 - \mu_2 = c^3 + \frac{2}{3}(1 - c^3) = \frac{c^3 + 2}{3}.$$

Hence $G(x)$ is admissible and $F_0(x) = G(x)$ if and only if μ_1 and μ_2 are such that (7.1) is satisfied.

Case II. $\eta_3 < -1$.

(a) $x^* < 0$. We then have

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0, \\ -\eta_1 - \eta_3 x, & 0 \leqq x < x^{**}, \\ 1, & x \geqq x^{**}. \end{cases}$$

If $x^{**} > 0$, $g_{\eta_1\eta_3}$ is a cdf and hence $F_0(x) = g_{\eta_1\eta_3}(x)$. If $x^{**} < 0$, the minimizing cdf is a one-point distribution and is not admissible.

(b) $x^* > 1$. The minimizing cdf is the same as in Case I (a) and is not admissible.

(c) $0 < x^* < 1$. Then we have

$$g_{\eta_1\eta_3}(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leqq x \leqq x^*, \\ -\eta_1 - \eta_3 x, & x^* < x < x^{**}, \\ x, & x > x^{**}. \end{cases}$$

$g_{\eta_1\eta_2}$ is a cdf, and hence $F_0(x) = g_{\eta_1\eta_2}(x)$.

EXAMPLE 3. Let $x_1 \leqq x_2 \leqq \cdots \leqq x_n$ be n ordered, independent observations from a cdf $F(x)$. Consider the problem of maximizing $E(x_n)$ with restrictions (3.2). The same problem for cdf's defined over the whole real line with restrictions on mean and variance has been discussed by Gumbel [10] and David and Hartley [1].

$$E(x_n) = \int_{-X}^{X} x \, d\{F(x)\}^n.$$

Integrating by parts, the above problem reduces to that of finding

$$\min_{F \varepsilon \mathfrak{A}} \int_{-X}^{X} F^n(x) \, dx.$$

Now as $\varphi(x, y) = y^n$ is strictly convex in y and is a function of y alone, the solution $F_0(x)$ is given by the function $g_{\eta_1\eta_2}(x)$, where

$$g_{\eta_1\eta_2}(x) = \begin{cases} 0, & x < \max\left(-\dfrac{\eta_1}{\eta_2}, -X\right), \\ 1, & x \geqq \min\left(X, -\dfrac{n+\eta_1}{\eta_2}\right); \\ \left(-\dfrac{\eta_1 + \eta_2 x}{n}\right)^{1/n-1}, & \text{elsewhere.} \end{cases}$$

Here η_1 and η_2 are determined by the following four cases:

	$\max(x_1, -X)$	$\min(x_2, X)$
Case 1.	x_1	x_2
Case 2.	$-X$	x_2
Case 3.	$-X$	X
Case 4.	x_1	X

We give below the equations determining η_1, η_2 in the above cases.

Case 1. $\mu_1 = -\dfrac{1}{\eta_2}(1 + \eta_1)$,

$$\mu_2 = \frac{1}{\eta_2^2}\left(\eta_1^2 + 2\eta_1 + \frac{n^2}{2n-1}\right).$$

Case 2. $\mu_1 = -\dfrac{n-1}{\eta_2}\xi^{n/(n-1)} - \dfrac{1+\eta_1}{\eta_2}, \qquad \xi = \dfrac{1}{n}(-\eta_1 + \eta_2 X)$,

$$\mu_2 = \left(\frac{n+\eta_1}{\eta_2}\right)^2 - \frac{2}{\eta_2^2}\left[\frac{n^2(n-1)}{2n-1} + (n-1)\eta_1 - (n-1)\xi^{n/(n-1)} \times \left(\eta_1 + \frac{n^2}{n-1}\xi\right)\right].$$

Case 3. $\mu_1 = X + \frac{n-1}{\eta_2}(\zeta^{n/n-1} - \xi^{n/n-1}), \qquad \zeta = \frac{1}{n}(-\eta_1 - \eta_2 X),$

$$\mu_2 = X^2 - \frac{2n}{\eta_2}\left[\frac{n-1}{n}\eta_1(\zeta^{n/n-1} - \xi^{n/n-1}) + \frac{n(n-1)}{2n-1} \cdot (\zeta^{(2n-1)/(n-1)} - \xi^{(2n-1)/(n-1)}).\right.$$

Case 4. $\mu_1 = X + \frac{n-1}{\eta_2}\zeta^{n/(n-1)},$

$$\mu_2 = X^2 - \frac{n}{\eta_2}\left[\frac{(n-1)\eta_1}{n}\zeta^{n/(n-1)} + \frac{n(n-1)}{2n-1}\xi^{(2n-1)/(n-1)}\right].$$

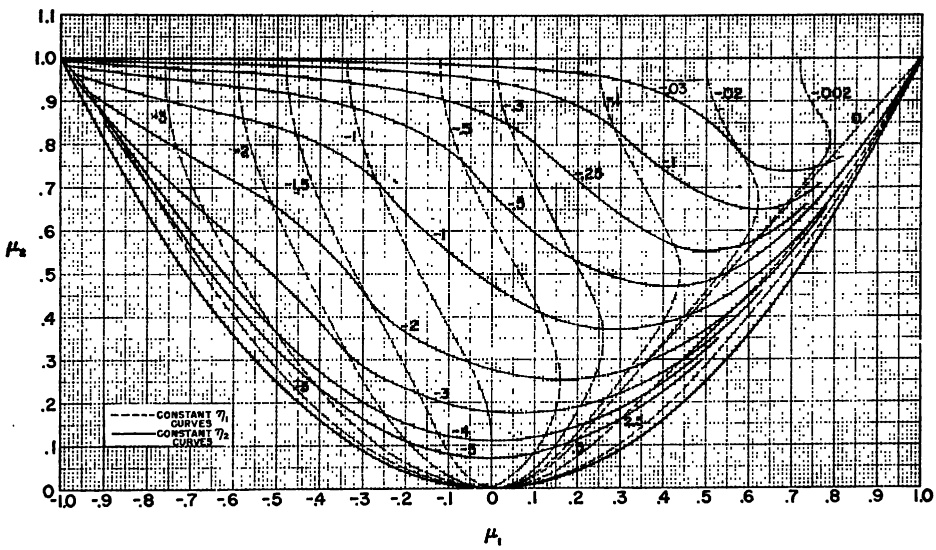
Fig. 2 represents η_1 and η_2 in terms of μ_1 and μ_2 for the case $X = 1$.

Similar results can be easily obtained for maximizing the expectations of the range and the smallest observation with similar restrictions on the underlying cdf's.

8. Characterization of the solution of the maximum problem and some examples. In this section we find the solution to the problem of maximizing

$$I(F) = \int_{-X}^{X} \varphi(x, F(x))\, dx$$

over all admissible cdf's. The existence of the maximizing cdf has already been established by Lemma 3.2. We shall show now that the solution is a discrete distribution and in our case, is at most a three-point distribution, i.e., a distribu



η_1 and η_2 in terms of μ_1 and μ_2

FIG. 2

tion concentrating all its mass at just three or fewer points. Some illustrations have been given at the end of this section.

THEOREM 8.1. *The solution to the problem of maximizing $I(F)$ over the class $\mathcal{A}$ of admissible cdf's, is at most a three-point distribution.*

PROOF. The inequality (4.0) shows that a convex combination of two maximizing admissible cdf's which is itself also admissible, gives a value which is smaller than the maximum. Hence the maximum of $I(F)$ occurs for cdf's which correspond to the extreme points of the convex set $\mathcal{A}$.

By Theorems 21.1 and 21.3 of Karlin and Shapley [5], it is then easy to see that the maximizing cdf is at most a three-point distribution.

REMARK. It is important to note here that in some cases, the maximizing admissible cdf can be further reduced to a two-point distribution [1], [3].

We shall illustrate the above results by a few examples.

EXAMPLE 1. Suppose we want to minimize $E(x_n)$ given in Example 4 of the last section, over all admissible cdf's. The problem is the same as that of maximizing

$$\int_{-X}^{X} F^n(x)\, dx$$

over all admissible cdf's. As $\varphi(x, y) = y^n$, is strictly convex in y, the maximizing admissible cdf of the above integral is at most a three-point distribution.

Similarly the minimizing admissible cdf of $E(w_n)$ is at most a three-point distribution. David and Hartley [1] claim that it can be further reduced to a two-point distribution.

EXAMPLE 2. Suppose we are interested in finding the maximizing cdf of

$$\frac{1}{2}\int_{-X}^{X} [F(x) - x]^2\, dx$$

such that $F(x)$ satisfies side conditions (3.3).

Now $\varphi(x, y) = \frac{1}{2}(y - x)^2$ is strictly convex function of y. Hence the solution of the above problem is at most a three-point distribution.

9. Acknowledgment. The author is deeply grateful to Professor Herman Chernoff who suggested the problem and gave generous help and guidance throughout the entire work.

REFERENCES

[1] H. O. HARTLEY AND H. A. DAVID, "Universal bounds for mean range and extreme observations," *Ann. Math. Stat.*, Vol. 25 (March, 1954), pp. 85–89.

[2] H. CHERNOFF AND S. REITER, "Selection of a Distribution Function to Minimize an Expectation Subject to Side Conditions," Technical Report No. 23 (March 12, 1954), Applied Math. and Statistics Laboratory, Stanford University.

[3] S. KARLIN, "Notes on Theory of Games," California Institute of Technology, mimeo. (1954).

[4] S. KARLIN, "Notes on Theory of Games and Decision Functions," Stanford University, mimeo. (1955).

[5] S. KARLIN AND L. S. SHAPLEY, "Geometry of Moment Spaces," *Memoirs of American Mathematical Society*, No. 12.
[6] H. D. BRUNK, G. M. EWING AND W. R. UTZ, "Minimizing Integrals in Certain Classes of Monotone Functions with Applications," University of Missouri, Department of Mathematics, September, 1954.
[7] Z. W. BIRNBAUM AND O. M. KLOSE, "Bounds for the variance of the U-statistic in terms of probability $Y < X$," Abstract, *Bull. Amer. Math. Soc.*, 1955.
[8] M. LOÈVE, "Probability Theory," D. Van Nostrand, New York, 1955.
[9] R. L. PLACKETT, "Limits of the ratio of mean range to standard deviation," *Biometrika*, Vol. 34 (1947), pp. 120–122.
[10] E. J. GUMBEL, "The maxima of the mean largest value and of the range," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 76–84.
[11] W. HOEFFDING, "The extrema of the expected value of a function of independent random variables," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 268–275.
[12] S. H. KHAMIS, "On the reduced moment problem," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 113–122.
[13] H. RUBIN AND S. ISAACSON, "On Minimizing an Expectation Subject to Certain Side Conditions," Technical Report No. 25 (July, 1954).
[14] S. MORIGUTI, "Extremal properties of extreme value distributions," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 523–536.
[15] W. HOEFFDING AND S. S. SHRIKHANDE, "Bounds for the distribution function of a sum of independent identically distributed random variables," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 439–449.