EXISTENCE OF BOUNDED LENGTH CONFIDENCE INTERVALS[1]

By Rajinder Singh

University of Illinois and Punjab University, Chandigarh

0. Introduction. Let X be a random variable whose distribution P belongs to a family of distributions $\mathcal{P}$. Let h be a real valued function defined on $\mathcal{P}$. This paper is concerned with the existence of estimates of prescribed accuracy for such functions h based on observations on X under different types of sampling plans. By an estimate of prescribed accuracy we mean a confidence interval of prescribed length and confidence coefficient, or a point estimate with prescribed expected loss. The sampling plans considered here are the m-stage, $m \geqq 1$, and sequential sampling plans. It may be pointed out that estimates with prescribed accuracy have been defined in the literature in various ways, [2], [3], [4], [6], [7], [9].

In most problems of estimation, estimates based on samples of fixed sizes have precisions which depend on unknown parameters. Consequently estimates with prescribed accuracy are not available without resort to multistage and sequential sampling plans. In fact, [1] in many non-parametric problems, even sequential sampling plans fail to give estimates with prescribed accuracy. It therefore becomes desirable to know whether, in a given problem of estimation, estimates of prescribed accuracy for the functions h exist under a given type of sampling plan.

It is shown that if h has a bounded length confidence interval based on one-stage sampling plans then h is uniformly continuous on $(\mathcal{P}, d^1)$, and if h has a bounded length confidence interval based on m-stage or sequential sampling plans, then h is continuous on $(\mathcal{P}, d^1)$, where d^1 is the familiar absolute variational distance on $\mathcal{P}$.

Further, if g is a uniformly continuous function of a real variable and h has a bounded length confidence interval based on m-stage sampling plans, then the composite function $g(h)$ has also a bounded length confidence interval based on m-stage sampling plans. If g is simply continuous (but not uniformly so), $g(h)$ has a bounded length confidence interval based on $2m$-stage sampling plans.

1. Definition, notation and statement of the problem. Fixed throughout are Ω, $\mathcal{A}$, $\mathcal{P}$ where $\mathcal{P}$ is a family of probability measures on $\mathcal{A}$, a σ-field of subsets of the set Ω; $X_1, X_2, \cdots$, are random variables such that, for each P in $\mathcal{P}$, $X_1, X_2, \cdots$ are independent and identically distributed. Point sets $\{\omega : Y(\omega) \varepsilon B\}$, where Y is a random variable and B is a Borel set, will be denoted by $(Y \varepsilon B)$

Received September 22, 1960; revised April 1, 1963.

[1] This paper is part of the thesis submitted in partial fulfillment of the requirements for the Ph.D. degree, University of Illinois, June, 1960. This research was partly sponsored by the National Science Foundation.

and $P(\{\omega: Y(\omega) \varepsilon B\})$ will be written as $P(Y \varepsilon B)$. $I(A)$ will denote the characteristic function of A.

By a stopping variable N, sequential or m-stage, we mean a sample size function as in [8]. Note that the stopping variable considered in this paper are closed, that is, $P(N < \infty) = 1$ for each P in $\mathcal{P}$. For each positive integer n, let $\mathcal{A}^n$ be the subfield of $\mathcal{A}$ generated by all events of the form $((X_1, \cdots, X_n) \varepsilon B)$ where B is a Borel set of the n-dimensional Euclidean space. For each stopping variable N, let $\mathcal{A}^N$ be the subfield of $\mathcal{A}$ generated by all events of the form $(N = 0)$ and $(N = n, (X_1, \cdots, X_n) \varepsilon B)$ where B is a Borel set of the appropriate space and $n \geqq 1$.

We may now describe the main problem of this paper in the decision theoretic setting. Let $\mathfrak{D}$ be a Borel subset of a Euclidean space. The set $\mathfrak{D}$ is the decision space. Let W be a non-negative function defined on $\mathcal{P} \times \mathfrak{D}$ and such that $W(P, \cdot)$ is Borel measurable, $P \varepsilon \mathcal{P}$. The function W is the loss function, that is $W(P, d)$ is the loss if decision d is made (or action d is taken) and P obtains. Let C be a non-negative function on the non-negative integers. $C(n)$ is the cost of observing $X_1, \cdots, X_n$. Let N be a stopping variable and Y be based on $\mathcal{A}^N$, the range of Y being a subset of $\mathfrak{D}$. Such a pair may be called a decision procedure. For any such procedure consider the expected loss function, $E_P W(P, Y)$ and the expected cost function $E_P C(N)$. These are function on $\mathcal{P}$.

Ideally, both functions should be uniformly small. However, there is no procedure which uniformly minimizes both except in trivial cases. A possible and often used approach is to put an upper bound on one of them and seek a procedure which minimizes the other in some sense, say the supremum. For example, let $\alpha > 0$ be given and among all procedures (or among all procedures in a certain class, say m-stage procedures) satisfying,

$$E_P W(P, Y) \leqq \alpha, \qquad P \varepsilon \mathcal{P}, \tag{1.1}$$

choose the one minimizing $\sup_P E_P C(N)$.

One question immediately arises. Are there any procedures satisfying (1.1)? For the class $\mathcal{W}_1$ of all loss functions W_{hL} of the form $W_{hL}(P, d) = 1$ if $|d - h(P)| > L/2$, $= 0$ otherwise, where h is a function on $\mathcal{P}$ into the reals and $L > 0$, the above question amounts to investigating the existence of bounded length confidence intervals for h.

There is another class $\mathcal{W}_2$ of loss functions, to which we may occasionally refer. It consists of all loss functions W satisfying $W(P, d) = V(|d - h(P)|)$, where the function V is strictly increasing in its argument. In particular, if $V(|a - b|) = |a - b|^r$, $r > 0$, denote $\mathcal{W}_2$ by $\mathcal{W}_r$.

For notational convenience we make the following definition. For any class $\mathcal{Y}$ of random variables (on Ω) and for any class $\mathcal{W}$ of loss functions, let $H(\mathcal{Y}, \mathcal{W})$ be the class of all real-valued functions h on $\mathcal{P}$ such that, for every $\alpha > 0$ and $W \varepsilon \mathcal{W}$, there exists a random variable Y in $\mathcal{Y}$ satisfying $E_P W(P, Y) \leqq \alpha$, $P \varepsilon \mathcal{P}$. Note that by Tchebycheff's inequality $H(\mathcal{Y}, \mathcal{W}_2) \subset H(\mathcal{Y}, \mathcal{W}_1)$. Furthermore $[Y - L/2, Y + L/2]$, restricting Y to $\mathcal{Y}$, is a $1 - \alpha$ confidence interval for h iff h is in $H(\mathcal{Y}, \mathcal{W}_1)$. The study of the class $H(\mathcal{Y}, \mathcal{W}_1)$ is therefore important for our problem. Let $H_m = H(\mathcal{Y}_m, \mathcal{W}_1)$, $m = 1, 2, \cdots, \infty$, where $\mathcal{Y}_m[\mathcal{Y}_\infty]$ is the class of all random variables Y based on m-stage [sequential] sampling plans.

2. Some preliminary results. Let $\mathcal{A}_0$ be a subfield of $\mathcal{A}$ and let $d^{\mathcal{A}_0}(P, Q) = \sup |P(A) - Q(A)|$, where the supremum is taken over all A in $\mathcal{A}_0$ and P, Q are in $\mathcal{P}$. We may note that, as shown in [8], $d^{\mathcal{A}_0}(P, Q)$ is equal to each of the following expressions:

(i) $\frac{1}{2}\int_\Omega |p - q|\, d\nu$, where ν is any σ-finite measure with respect to which P and Q are absolutely continuous, for example, $\nu = P + Q$ and p, q are $\mathcal{A}_0$-measurable and satisfy $P(A) = \int_A p\, d\nu$, $Q(A) = \int_A q\, d\nu$, $A \varepsilon \mathcal{A}_0$.

(ii) $\sup |E_P Y - E_Q Y|$, where the supremum is taken over all $\mathcal{A}_0$-measurable Y into $[0, 1]$.

(iii) $\sup |E_P Y - E_Q Y|$, where the supremum is taken over all Y into $[0, 1]$ that are based on $\mathcal{A}_0$.

We shall write $d^N(P, Q)$ for $d^{\mathcal{A}^N}(P, Q)$ and $d^n(P, Q)$ for $d^{\mathcal{A}^n}(P, Q)$. For our purposes there is no loss of generality in assuming that $d^1(P, Q) = 0$ implies that $P = Q$. It can be easily seen that this assumption implies that $(\mathcal{P}, d^N)$ is a metric space provided $P(N > 0) > 0$ for each P in $\mathcal{P}$.

We also note that for our problem it is enough to consider terminal decision functions Y that are $\mathcal{A}^N$-measurable. That is, if N is a stopping variable and Y is a random variable that is based on $\mathcal{A}^N$ and satisfies

$$P(|Y - h(P)| \leqq L/2) \geqq 1 - \alpha, \qquad P \varepsilon \mathcal{P}, \tag{2.1}$$

then there is a random variable Z that is $\mathcal{A}^N$-measurable and satisfies,

$$P(|Z - h(P)| \leqq L) \geqq 1 - 2\alpha, \qquad P \varepsilon \mathcal{P}. \tag{2.2}$$

To prove this, let $F(y, \omega) = P^{\mathcal{A}^N}(Y \leqq y, \omega)$ be the conditional distribution function of Y given $\mathcal{A}^N$. $F(y, \cdot)$ is an $\mathcal{A}^N$-measurable function for each y and $F(\cdot, \omega)$ is a distribution function for each ω. Let $\cdots, y_{-1}, y_0, y_1 \cdots$ be the sequence $\cdots, -L/2, 0, L/2, \cdots$. Let $K(\omega) = \min\{k \mid \frac{1}{2} \leqq F(y_k, \omega)\}$. Then K is an $\mathcal{A}^N$-measurable random variable. For, for every integer k, $(K(\omega) = k) = (F(y_{k-1}, \omega) < \frac{1}{2}) \cap (F(y_k, \omega) \geqq \frac{1}{2})$ and the last two sets are in $\mathcal{A}^N$. Let $Z(\omega) = y_{k(\omega)}$. This Z satisfies (2.2). For, from (2.1) we have $E_P(P^{\mathcal{A}^N}(|Y - h(P)| \leqq L/2, \omega)) \geqq 1 - \alpha$, $P \varepsilon \mathcal{P}$, which, as shown below, implies that

$$\begin{aligned} P(P^{\mathcal{A}^N}(|Y - h(P)| \leqq L/2, \omega) > \epsilon) &\geqq (1 - \alpha - \epsilon)/1 - \epsilon \\ &= 1 - (\alpha/1 - \epsilon). \end{aligned} \tag{2.3}$$

Therefore, if $\epsilon = \frac{1}{2}$, then the left side of (2.3) $\geqq 1 - 2\alpha$. But $P^{\mathcal{A}^N}(|Y - h(P)| \leqq L/2, \omega) > \frac{1}{2}$ implies that $|Z(\omega) - h(P)| \leqq L$, which proves that Z satisfies (2.2). To justify (2.3), we note that if a random variable V is such that $P(0 \leqq V \leqq 1) = 1$, then $P(V > \epsilon) \geqq (EV - \epsilon)/(1 - \epsilon)$, $0 < \epsilon < 1$.

We now prove

THEOREM 2.1. *Let N be a stopping variable. Then for every $\epsilon > 0$ and every P in $\mathcal{P}$ there is a $\delta > 0$ such that $d^N(P, Q) < \epsilon$ if $d^1(P, Q) < \delta$. δ does not depend on P if $\lim_{k\to\infty} P(N > k) = 0$ uniformly over $\mathcal{P}$.*

PROOF. Suppose $\epsilon > 0$ and $P \varepsilon \mathcal{P}$. Let k be a positive integer such that

$$P(N > k) < \epsilon/2. \tag{2.4}$$

This is possible because we are considering closed stopping variables. For this k choose a $\delta > 0$ such that

$$d^N(P, Q) \leqq P(N > k) + \epsilon/2, \tag{2.5}$$

where $Q \varepsilon \mathcal{P}$ and satisfies $d^1(P, Q) < \delta$. The existence of such a δ is shown below. Combining (2.4) and (2.5) we get the first part of the theorem. For the second part note that if $\lim_{k\to\infty} P(N > k) = 0$ uniformly over $\mathcal{P}$, then the choice of k in (2.4) does not depend upon P and consequently δ depends only on ϵ.

We now prove (2.5). We can choose a $\delta > 0$ (see [8]) such that $d^1(P, Q) < \delta$ implies that $d^k(P, Q) < \epsilon/2(k + 1)$. Let $A \varepsilon \mathcal{A}^N$. Then $A = A_0 \cup \bigcup_{n=1}^{\infty}(N = n, (X_1, \cdots, X_n) \varepsilon B_n)$, where A_0 is either empty or the set $(N = 0)$ and B_1, $B_2, \cdots$ are Borel subsets of the appropriate spaces. Thus, if $Y = I(A)$ then $Y = \sum_{n=0}^{\infty} Y_n Z_n$, where $Z_n = I(N = n)$ and Y_n is the characteristic function of a set in $\mathcal{A}^n$, $n = 0, 1, \cdots$. If $n \leqq k$, then $(N = n)$ and the events in $\mathcal{A}^k$ are independent of $X_{k+1}, X_{k+2}, \cdots$, and $E_P(Z_n | \mathcal{A}^k) = E_P(Z_n \mid \mathbf{X})$a.e.$[P]$, $P \varepsilon \mathcal{P}$, where $\mathbf{X} = (X_1, X_2, \cdots)$. But the latter conditional expectation may be taken free of P and so also can the former. Therefore, if P and Q satisfy $d^1(P, Q) < \delta$, then

$$\begin{aligned} Q(A) &= E_Q Y > \sum_{n=0}^{k} E_Q Y_n Z_n = \sum_{n=0}^{k} E_Q(E_Q(Y_n Z_n \mid \mathcal{A}^k)) \\ &= \sum_{n=0}^{k} E_Q(Y_n E_Q(Z_n \mid \mathcal{A}^k)) = \sum_{n=0}^{k} E_Q(Y_n E_P(Z_n \mid \mathcal{A}^k)) \\ &> \sum_{n=0}^{k} [E_P(Y_n E_P(Z_n \mid \mathcal{A}^k)) - \epsilon/2(k + 1)] \\ &\geqq P(A) - P(N > k) - \epsilon/2. \end{aligned}$$

This implies that $P(A) - Q(A) < P(N > k) + \epsilon/2$. A similar inequality holds if A is replaced by $\Omega - A$. Combining the two inequalities we get (2.5).

We now prove a theorem which shows that if h has a bounded length confidence interval based on $\mathcal{A}_0$ then h is uniformly continuous on $(\mathcal{P}, d^{\mathcal{A}_0})$.

THEOREM 2.2. *Suppose Y is a random variable based on $\mathcal{A}_0$ and satisfies*

$$P(|Y - h(P)| \leqq L/2) \geqq 1 - \alpha, \qquad P \varepsilon \mathcal{P}. \tag{2.6}$$

Then $|h(P) - h(Q)| \leqq L$ whenever $d^{\mathcal{A}_0}(P, Q) < 1 - 2\alpha$, for P, Q in $\mathcal{P}$.

PROOF. Let $Z = 1$, if $|Y - h(P)| \leqq L/2$; 0, otherwise. Then $|E_P Z - E_Q Z| \leqq d^{\mathcal{A}_0}(P, Q)$ and therefore

$$E_Q Z \geqq E_P Z - d^{\mathcal{A}_0}(P, Q). \tag{2.7}$$

Now

$$Q(|Y - h(P)| \leqq L/2, |Y - h(Q)| \leqq L/2) \geqq Q(|Y - h(P)| \leqq L/2) + Q(|Y - h(Q)| \leqq L/2) - 1 \geqq [1 - \alpha - d^{\alpha_0}(P, Q)] + (1 - \alpha) - 1$$
$$= 1 - 2\alpha - d^{\alpha_0}(P, Q) \qquad \text{by (2.6) and (2.7).}$$

Obviously if $d^{\alpha_0}(P, Q) < 1 - 2\alpha$ then the above inequality is positive. Consequently two events $(|Y - h(P)| \leqq L/2)$ and $(|Y - h(Q)| \leqq L/2)$ occur together with positive probability provided $d^{\alpha_0}(P, Q) < 1 - 2\alpha$. This proves the theorem.

3. Necessary conditions. In this section we derive necessary conditions for h to be in $H_1, H_2, \cdots, H_\infty$. Note that a necessary condition for h to be in H_∞ is also a necessary condition for h to be in H_m since $H_m \subset H_\infty$, $m = 1, 2, \cdots$. In Section 5 we shall discuss some other properties of these classes.

Let (Θ, D) be a metric space with $1 - 1$ mapping $\theta \rightarrow P_\theta$ from Θ into $\mathcal{P}$ and $\hat{h}$ denote a real-valued function on Θ. For functions $\hat{h}$ we denote the class $H(\mathcal{Y}, \mathcal{W})$, defined in Section 1 for the functions h, by $\hat{H}(\mathcal{Y}, \mathcal{W})$.

THEOREM 3.1.

(a) *If the mapping $\theta \rightarrow P_\theta$ from (Θ, D) into $(\mathcal{P}, d^1)$ is continuous, then $\hat{h} \varepsilon \hat{H}_\infty$ implies that $\hat{h}$ is continuous on (Θ, D).*

(b) *If the mapping $\theta \rightarrow P_\theta$ is uniformly continuous, then $\hat{h} \varepsilon \hat{H}_1$ implies that $\hat{h}$ is uniformly continuous on (Θ, D).*

PROOF. (a) For every θ in Θ define h on $\mathcal{P}$ by taking $h(P_\theta) = \hat{h}(\theta)$. It suffices to show that h is continuous on $(\mathcal{P}, d^1)$. For this let $\epsilon > 0$. Then by the definition of H_∞ it follows that there exists a stopping variable N and an $\mathcal{A}^N$-measurable random variable Y satisfying $P(|Y - h(P)| \leqq \epsilon/2) \geqq \frac{2}{3}$, $P \varepsilon \mathcal{P}$. Again, by Theorem 2.1, there is a $\delta > 0$ such that for $Q \varepsilon \mathcal{P}$, $d^1(P, Q) < \delta$ implies $d^N(P, Q) < \frac{1}{3}$. Consequently by Theorem 2.2, $d^1(P, Q) < \delta$ implies that $|h(P) - h(Q)| \leqq \epsilon$. Thus h is continuous at Q. Proof for (b) is similar and is therefore omitted.

Suppose $\mathcal{P}$ is dominated by a σ-finite measure μ, then for each θ in Θ there is a density (Radon-Nikodym derivative) p_θ, $\mathcal{A}^1$-measurable, corresponding to P_θ such that $P_\theta(A) = \int_A p_\theta \, d\mu$, $A \varepsilon \mathcal{A}^1$. We can now state the following

COROLLARY 3.1. *Suppose for each ω, $p_{(\cdot)}(\omega)$ is a continuous function from (Θ, D) into the real numbers, then $\hat{h} \epsilon \hat{H}_\infty$ implies that $\hat{h}$ is continuous on (Θ, D).*

REMARK.

(1) If (Θ, D) is, in particular, $(\mathcal{P}, d^1)$ then Theorem 3.1 reads as follows: (a) $h \varepsilon H_\infty$ implies that h is continuous on $(\mathcal{P}, d^1)$, (b) $h \varepsilon H_1$ implies that h is uniformly continuous on $(\mathcal{P}, d^1)$.

(2) It may be noted that continuity of h in (a) cannot be replaced by its uniform continuity as in (b). That is, if h has a bounded length confidence interval in m-stage ($m \geqq 2$) or sequential sampling plans then h need not be uniformly continuous on $(\mathcal{P}, d^1)$. For this see Examples 3.1 and 3.2 below.

EXAMPLE 3.1. Let $\mathcal{P}$ be the class of normal distributions with mean μ and variance σ^2. Let $\theta = (\mu, \sigma)$, $h(P_\theta) = \mu$, and $\theta_i = (\mu_i, \sigma)$, $i = 1, 2$. Then $d^1(P_{\theta_1}, P_{\theta_2}) = 2G(|\mu_1 - \mu_2|/2\sigma) - 1$, where G is the distribution function of a random variable normal with mean 0 and variance 1. It is clear that h is not uniformly continuous on $(\mathcal{P}, d^1)$. However μ has a bounded length confidence interval based on two-stage sampling plans as shown in [10].

EXAMPLE 3.2. Let $\mathcal{P}$ be the class of exponential distributions having density function $f(x) = 1/\theta e^{-x/\theta}$ if $x > 0$, $= 0$ otherwise. Let $h(P_\theta) = \theta^2$. It will be shown that h has a $(1 - \alpha)$ confidence interval of length L based on a two-stage sampling plan although h is not uniformly continuous on $(\mathcal{P}, d^1)$. For this let $\bar{X} = (\sum_{i=1}^4 X_i)/4$ and $b = L^2\alpha/5120$, where $L > 0$, $\alpha > 0$ are preassigned. Let N be the smallest integer $\geqq 4 + \exp(\bar{X}/b)$ and $Y = (\sum_{i=5}^N X_i^2)/2(N - 4)$. Then $[Y - L/2, Y + L/2]$ is a $1 - \alpha$ confidence interval for θ^2. For, by Tchebycheff's inequality, $P(|Y - \theta^2| \leqq L/2) \geqq 1 - 4L^{-2}E_\theta(Y - \theta^2)^2 = 1 - 4L^{-2} \operatorname{var}_\theta Y = 1 - 4L^{-2}\{E_\theta \operatorname{var}_\theta (Y \mid N) + \operatorname{var}_\theta E_\theta(Y \mid N)\} = 1 - 4L^{-2}E_\theta \operatorname{var}_\theta (Y \mid N) = 1 - 4L^{-2}E_\theta(5\theta^4/(N - 4)) \geqq 1 - 4L^{-2}E_\theta(5\theta^4 b e^{-\bar{X}}) = 1 - 20\theta^4 bL^{-2} (E_\theta(e^{-X_1/4}))^4 = 1 - \alpha$. Now

$$d^1(P_{\theta_1}, P_{\theta_2}) = \frac{1}{2}\int_0^\infty |(1/\theta_1)e^{-x/\theta_1} - (1/\theta_2)e^{-x/\theta_2}|\, dx$$
$$= |e^{-x_0/\theta_1} - e^{-x_0/\theta_2}| \leqq |\theta_1 - \theta_2|/\theta_2,$$

where $x_0 = (\theta_1\theta_2 \log(\theta_2/\theta_1))/(\theta_2 - \theta_1)$ is the point of intersection of two density curves. Hence, for $|\theta_1 - \theta_2|/\theta_2 < \delta$, we have $d^1(P_{\theta_1}, P_{\theta_2}) < \delta$, but $|\theta_1^2 - \theta_2^2| = (|\theta_1 - \theta_2|/\theta_2)\theta_2|\theta_1 + \theta_2|$ is not bounded, proving thereby that θ^2 is not uniformly continuous on $(\mathcal{P}, d^1)$.

(3) As observed earlier, the necessary condition stated in the theorem is also a necessary condition for h to be in H_m, for every positive integer m. Example 3.3 given below shows that for h to be in H_m this condition is not sufficient.

EXAMPLE 3.3. Let $\mathfrak{F}$ be the class of all distribution functions on the real line having a unique median. Let, for every real x, $P_F(X < x) = F(x)$. Let $\mathcal{P} = \{P_F \mid F \varepsilon \mathfrak{F}\}$ and $h(P_F) =$ median of the distribution function F. It is easily seen that h is continuous on $(\mathcal{P}, d^1)$. However it is shown in [5] that $h \not\varepsilon H_m$ for any finite integer $m \geqq 1$.

(4) Whether the continuity of h on $(\mathcal{P}, d^1)$ is sufficient for h to be in H_∞ is still an open question. That the condition in this case is sufficient appears to be a reasonable conjecture.

(5) The following example shows that the necessary condition in Theorem 3.1 (b) is not sufficient. That is, h uniformly continuous on $(\mathcal{P}, d^1)$ need not imply that $h \varepsilon H_1$.

EXAMPLE 3.4. Let $\mathcal{P}$ be the class of Poisson distributions with parameter λ such that $P_\lambda(X = x) = e^{-\lambda}\lambda^x/x!$, $x = 0, 1, 2, \cdots$. Let $h(P_\lambda) = \lambda$. It is easy to verify that

$$d^1(P_{\lambda_1}, P_{\lambda_2}) = \frac{1}{2}\sum_0^\infty \left|\frac{e^{-\lambda_1}\lambda_1^x}{x!} - \frac{e^{-\lambda_2}\lambda_2^x}{x!}\right| = \sum_{x=0}^{x_0}\left|\frac{e^{-\lambda_1}\lambda_1^x}{x!} - \frac{e^{-\lambda_2}\lambda_2^x}{x!}\right| < e^{|\lambda_1 - \lambda_2|} - 1,$$

where x_0 is the largest integer $\leqq (\lambda_2 - \lambda_1)/\ln (\lambda_2/\lambda_1)$. This shows that h is uniformly continuous on $(\mathcal{P}, d^1)$. However, as is well-known, λ does not have estimates of prescribed accuracy based on one-stage sampling plans.

4. Applications. Theorems proved in Sections 2 and 3 are useful in proving the non-existence of m-stage, $m \geqq 1$, and sequential plans for finding estimates of prescribed accuracy for h. If, for a given h, bounded length confidence intervals do not exist under some sampling plans then, as observed earlier at the end of Section 1, point estimates for h with expected loss (loss $W \varepsilon \mathcal{W}_2$) bounded by a preassigned number α cannot exist under the same sampling plans. In Examples 4.1 through 4.3 we shall show that h is not uniformly continuous on $(\mathcal{P}, d^1)$. This will, by Theorem 3.1, imply that h can have neither bounded length confidence intervals nor point estimates with bounded expected loss (loss $W \varepsilon \mathcal{W}_2$) under one-stage sampling plans.

Example 4.1. Let $\mathcal{P}$ be the class of distributions with density functions $(1/\sigma)f((x - \mu)/\sigma)$, where μ is real, $\sigma > 0$ and f is a given density function. Here the density function f is known but the location and scale parameters are unknown. Let $\theta = (\mu, \sigma)$, $h(P_\theta) = \mu$ and $\theta_i = (\mu_i, \sigma)$, $i = 1, 2$. Then

$$d^1(P_{\theta_1}, P_{\theta_2}) = \frac{1}{2}\int (1/\sigma)\,|f((x - \mu_1)/\sigma) - f((x - \mu_2)/\sigma)|\,dx$$

$$= \frac{1}{2}\int |f(y) - f(y + (\mu_1 - \mu_2)/\sigma)|\,dy,$$

which tends to zero as $\sigma \to \infty$. This implies that by choosing σ sufficiently large we can make $d^1(P_{\theta_1}, P_{\theta_2}) < \delta$. But, on the other hand, $|h(P_{\theta_1}) - h(P_{\theta_2})| = |\mu_1 - \mu_2|$ is unbounded. This proves that h is not uniformly continuous on $(\mathcal{P}, d^1)$.

Example 4.2. Let $\mathcal{P}$ be the class of gamma distributions with density functions $f(x) = (1/\Gamma(r))\theta^r x^{r-1}e^{-\theta x}$, $x > 0$ and, $= 0$ otherwise. Here θ and r are arbitrary positive real numbers. Let $h(P_\theta) = \theta^s$, where $s > 0$ is a real number. It is easy to see that $d^1(P_{\theta_1}, P_{\theta_2}) \leqq |\theta_1^r - \theta_2^r|/\min(\theta_1^r, \theta_2^r)$. Hence for $|\theta_1^r - \theta_2^r|/\min(\theta_1^r, \theta_2^r) < \delta$, we have $d^1(P_{\theta_1}, P_{\theta_2}) < \delta$ but $|\theta_1^r - \theta_2^r|$ is unbounded. Thus h is not uniformly continuous on $(\mathcal{P}, d^1)$.

Example 4.3. Let $\mathcal{P}$ be the class of distributions with density functions $f(x) = 1/\theta$, $0 < x < \theta$, and $= 0$ otherwise. Here $d^1(P_{\theta_1}, P_{\theta_2}) \leqq |\theta_1 - \theta_2|/\min(\theta_1, \theta_2)$. Hence for $|\theta_1 - \theta_2|/\min(\theta_1, \theta_2) < \delta$, we have $d^1(P_{\theta_1}, P_{\theta_2}) < \delta$ but $|\theta_1 - \theta_2|$ is unbounded.

Example 4.4. Let $\mathcal{P}$ be the class of all distributions on the real line with finite rth moment. Let h denote moment of some order $i \leqq r$. It is well known that h is not continuous on $(\mathcal{P}, d^1)$. Hence it follows from Theorem 3.1 that h has neither bounded length confidence intervals nor point estimates with bounded expected loss (loss $W \epsilon \mathcal{W}_2$) even under sequential sampling plans. The non-existence of statistical procedures for this problem was obtained in [1] by a different approach.

EXAMPLE 4.5. Let $\mathcal{P} = \{P_\theta : 0 < \theta < 1\}$ such that $P_\theta(X_1 = 1) = \theta$, $P_\theta(X_1 = 0) = 1 - \theta$. Let $h(P_\theta) = \ln \theta$. It is easy to see that $d^1(P_{\theta_1}, P_{\theta_2}) = |\theta_1 - \theta_2|$ so that $\ln \theta$ is not uniformly continuous on $(\mathcal{P}, d^1)$. Further if N is an m-stage non-randomized stopping variable then $\lim_{n\to\infty} P(N > n) = 0$ uniformly for all θ. Hence by Theorems 2.1 and 2.2, $\ln \theta$ can have neither bounded length confidence intervals nor point estimates with bounded expected loss (loss $W \varepsilon \mathcal{W}_2$) under m-stage non-randomized sampling plans.

5. Properties of $H(\mathcal{Y}, \mathcal{W})$. For proving various properties of $H(\mathcal{Y}, \mathcal{W})$ we shall need the following assumption about $\mathcal{Y}$. It will be stated explicitly whenever this assumption is used.

ASSUMPTION 5.1. If Y_1, Y_2 are in $\mathcal{Y}$, then $\psi(Y_1, Y_2)$ is in $\mathcal{Y}$ for all Borel measurable functions ψ into the real line.

Let $\mathcal{Y}_m^0[\mathcal{Y}_\infty^0]$ be the class of all random variables Y that are $\mathcal{A}^N$-measurable, N being a non-randomized m-stage [sequential] stopping variable. Further we denote the class $H(\mathcal{Y}_m^0, \mathcal{W}_1)$ by H_m^0, $m = 1, 2, \cdots, \infty$.

LEMMA 5.1.

(i) *If N_1, N_2 are non-randomized stopping variables then $\hat{N} = \max(N_1, N_2)$ is a non-randomized stopping variable.* (ii) *If N_1, N_2 are m-stage non-randomized stopping variables, then $\hat{N} = \max(N_1, N_2)$ is an m-stage non-randomized stopping variable.*

PROOF.

(i) For every integer $n \geqq 1$,

$$(\hat{N} \leqq n) = (\max(N_1, N_2) \leqq n) = (N_1 \leqq n) \cap (N_2 \leqq n).$$

The last two sets are in $\mathcal{A}^n$. Consequently $(\hat{N} \leqq n)$ is in $\mathcal{A}^n$. Hence, for every integer $n \geqq 1$, the event $(\hat{N} \leqq n)$ is independent of $X_{n+1}, X_{n+2}, \cdots$ and that the conditional distribution of $\hat{N}$ given $(X_1, X_2, \cdots)$ is free of P. This proves (i).

(ii) By induction on m. It is easy to see that the proposition is true for $m = 1$. Let $m > 1$ and suppose that the proposition is true for $1, 2, \cdots, m - 1$. Let N_1, N_2 be m-stage non-randomized stopping variables. Then, by definition, there exist $(m - 1)$-stage non-randomized stopping variables M_1, M_2 such that N_i is $\mathcal{A}^{M_i}$-measurable, $i = 1, 2$. Let $\hat{M} = \max(M_1, M_2)$. Then by the induction hypothesis $\hat{M}$ is $(m - 1)$-stage non-randomized stopping variable. Also since $\mathcal{A}^{M_i} \subset \mathcal{A}^{\hat{M}}$, N_i is $\mathcal{A}^{\hat{M}}$-measurable, $i = 1, 2$. This implies that $\hat{N}$ is $\mathcal{A}^{\hat{M}}$-measurable and the lemma is proved.

LEMMA 5.2. *For $m = 1, 2, \cdots, \infty$, $\mathcal{Y}_m^0$ satisfies Assumption* 5.1.

PROOF. Let m be a positive integer. Let $Y_i \varepsilon \mathcal{Y}_m^0$, $i = 1, 2$. Then there is an m-stage non-randomized stopping variable N_i such that Y_i is $\mathcal{A}^{N_i}$-measurable. Let $\hat{N} = \max(N_1, N_2)$. Then, by Lemma 5.1, $\hat{N}$ is an m-stage non-randomized stopping variable and Y_i is $\mathcal{A}^{\hat{N}}$-measurable. This implies the desired result. The proof for $\mathcal{Y}_\infty^0$ is similar and is omitted.

THEOREM 5.1. *Under Assumption 5.1,*

(a) *if, for every integer* $n > 1$, f *is a uniformly continuous function of* n *real variables then* $h_i \varepsilon H(\mathcal{Y}, \mathcal{W}_1)$, $i = 1, \cdots, n$, *implies that* $f(h_1, \cdots, h_n) \varepsilon H(\mathcal{Y}, \mathcal{W}_1)$.

(b) $H(\mathcal{Y}, \mathcal{W}_1)$ *is closed under passages to limits with respect to uniform convergence.*

Proof.

(a) The proof is elementary and is, therefore, omitted.

(b) Suppose h_n is in $H(\mathcal{Y}, \mathcal{W}_1)$, $n = 1, 2, \cdots$ and $\lim_{n\to\infty} h_n(P) = h(P)$, uniformly for all P in $\mathcal{P}$. Therefore there exist a positive integer m such that $|h_m(P) - h(P)| < L/4$, for all P in $\mathcal{P}$. Also there exists a Y in $\mathcal{Y}$ such that

$$P(|Y - h_m(P)| \leqq L/4) \geqq 1 - \alpha, \qquad P \text{ in } \mathcal{P}.$$

Then $[Y - L/2, Y + L/2]$ is a $1 - \alpha$ confidence interval for h. For,

$$P(|Y - h(P)| \leqq L/2) \geqq P(|Y - h_m(P)| \leqq L/4, |h_m(P) - h(P)| \leqq L/4)$$
$$\geqq 1 - \alpha, \qquad P \text{ in } \mathcal{P}.$$

It may be noted that this theorem is true of H_m^0, $m = 1, 2, \cdots, \infty$. This is an immediate consequence of Lemma 5.2.

Theorem 5.2. *Let* g *be a continuous function of one variable. Then* (i) *the composite function* $g(h) \varepsilon H_{2m}^0$ *if* $h \varepsilon H_m^0$, $m = 1, 2, \cdots$; (ii) *the composite function* $g(h) \varepsilon H_\infty^0$ *if* $h \varepsilon H_\infty^0$.

Remark. (a) Compare with Theorem 5.1 (a). There uniform continuity of g was assumed. Here only continuity is assumed. (b) The change from m-stage to $2m$-stage sampling plan in (i) is brought about as follows: we first use an m-stage sampling plan to get a provisional bounded length confidence interval for h. From this we obtain a closed real number interval which brackets h with a certain probability. On this interval g is uniformly continuous. As in Theorem 5.1 (a) we now use another m-stage sampling plan to get a bounded length confidence interval for $g(h)$ as desired. Lastly these two m-stage sampling plans are combined to give a $2m$-stage sampling plan.

Proof. (i) Suppose $h \varepsilon H_m^0$. Let $L > 0$, $\alpha > 0$. Since h is in H_m^0, it is clear that there is a non-randomized m-stage stopping variable N and an $\mathcal{A}^N$-measurable random variable Y such that Y is integer valued and satisfies

$$P(|Y - h(P)| \leqq 2) \geqq 1 - \alpha/2, \qquad P \varepsilon \mathcal{P}.$$

Let $J_k = [k - 4, k + 4]$ if k is an integer. Since g is continuous on J_k it is uniformly continuous on J_k. Thus there is a positive $L_k < 1$ such that if a and b are in J_k and satisfy $|a - b| < 2L_k$ then $|g(a) - g(b)| < L/2$. If k is an integer, let Y^k be an $\mathcal{A}^{N^k}$-measurable random variable where N^k is a non-randomized m-stage stopping variable and such that $P(|Y^k - h(P)| \leqq L_k) \geqq 1 - \alpha/2$, $P \varepsilon \mathcal{P}$. Such N^k and Y^k exist since $h \varepsilon H_m^0$. This will imply, as shown below, the existence of a non-randomized $2m$-stage stopping variable $\hat{N}$ and a random variable $\hat{Y}$, $\mathcal{A}^{\hat{N}}$-measurable such that

$$P(\hat{Y} \varepsilon B \mid Y = k) = P(Y^k \varepsilon B) \tag{5.1}$$

for all $P \varepsilon \mathcal{P}$, all Borel sets B and all k. This implies that

$$[g(\hat{Y}) - L/2, g(\hat{Y}) + L/2]$$

is a $1 - \alpha$ confidence interval for $g(h)$. For

$$P(|\hat{Y} - h(P)| \leqq L_Y) = \sum_{k=-\infty}^{\infty} P(|\hat{Y} - h(P)| \leqq L_k \mid Y = k) P(Y = k)$$

$$= \sum_{k=-\infty}^{\infty} P(|\hat{Y} - h(P)| \leqq L_k) P(Y = k) \geqq 1 - \alpha/2,$$

$$P \varepsilon \mathcal{P}.$$

Consequently,

$$P(|g(\hat{Y}) - g(h(P))| \leqq L/2)$$

$$\geqq P(|\hat{Y} - h(P)| < L_Y, |Y - h(P)| \leqq 2) \geqq 1 - \alpha, \qquad P \varepsilon \mathcal{P}.$$

Thus $g(h)$ is in H_{2m}^0.

We now prove (5.1). For each k we can express N^k, Y^k as follows: $N^k = \sum_{n=1}^{\infty} n s_n^k(X_1, \cdots, X_n)$, $Y^k = \sum_{n=1}^{\infty} s_n^k(X_1, \cdots, X_n) t_n^k(X_1, \cdots, X_n)$. Here $s_n^k(X_1, \cdots, X_n) = I(N^k = n)$ and the functions s_n^k and t_n^k are Borel measurable on the appropriate spaces. If, for positive integers p, k, n,

$$N = p, \qquad Y = k, \qquad N^k = n - p, \tag{5.2}$$

take $\hat{N} = n$. That is, $\hat{N}$ is defined by

$$(\hat{N} = n) = \bigcup_{p=1}^{n-1} \bigcup_{k=-\infty}^{\infty} (N = p, Y = k, s_{n-p}^k(X_{p+1}, \cdots, X_n) = 1).$$

Furthermore if, in conjunction with (5.2), we have $Y^k = y$, define $\hat{Y} = y$. That is, for every real y, we define $\hat{Y}$ by

$$(\hat{Y} = y) = \bigcup_{n=2}^{\infty} \bigcup_{p=1}^{n-1} \bigcup_{k=-\infty}^{\infty} (N = p, Y = k, s_{n-p}^k(X_{p+1}, \cdots, X_n) = 1,$$

$$t_{n-p}^k(X_{p+1}, \cdots, X_n) = y).$$

Clearly $\hat{N}$ is a non-randomized $2m$-stage stopping variable and $\hat{Y}$ is $\mathfrak{A}^N$-measurable.

Also

$$P(\hat{Y} \varepsilon B, Y = k) = \sum_{n=2}^{\infty} \sum_{p=1}^{n-1} P(N = p, Y = k, s_{n-p}^k(X_{p+1}, \cdots, X_n) = 1,$$

$$t_{n-p}^k(X_{p+1}, \cdots, X_n) \varepsilon B)$$

$$= \sum_{n=2}^{\infty} \sum_{p=1}^{n-1} P(N = p, Y = k) P(s_{n-p}^k(X_{p+1}, \cdots, X_n) = 1,$$

$$t^k_{n-p}(X_{p+1}, \cdots, X_n) \,\varepsilon\, B)$$

$$= \sum_{p=1}^{\infty} P(N = p, Y = k) \sum_{n=p+1}^{\infty} P(N^k = n - p, Y^k \,\varepsilon\, B)$$

$$= P(Y = k)P(Y^k \,\varepsilon\, B).$$

This completes the proof of (i). Proof of (ii) is similar and is omitted.

This theorem has an important application. If in a problem $\mathcal{P} = \{P_\theta \mid \theta \,\varepsilon\, \Theta$, a subset of the real line$\}$ and it is known that θ is in $\hat{H}^0_m[\hat{H}^0_\infty]$ then the restriction to Θ of any function g continuous from the real line into the real line is in

$$\hat{H}^0_{2m}[\hat{H}^0_\infty].$$

We shall now state a theorem which is an analogue of Theorem 5.1 for the class $H(\mathcal{Y}, \mathcal{W}_r)$. It may be recalled that $H(\mathcal{Y}, \mathcal{W}_r)$ is the class of all real-valued functions h on $\mathcal{P}$ such that, if $\alpha > 0$, then there is a Y in $\mathcal{Y}$ satisfying

$$E_P(|Y - h(P)|^r) \leqq \alpha, \qquad P \,\varepsilon\, \mathcal{P}.$$

THEOREM 5.3. *Under Assumption 5.1,*

(a) *if, for every integer $n \geqq 1$, f is a uniformly continuous function of n real variables, then $h_i \,\varepsilon\, H(\mathcal{Y}, \mathcal{W}_r)$, $i = 1, \cdots, n$, implies that*

$$f(h_1, \cdots, h_n) \,\varepsilon\, H(\mathcal{Y}, \mathcal{W}_r)$$

(b) *$H(\mathcal{Y}, \mathcal{W}_r)$ is closed under passages to a limit with respect to uniform convergence.*

PROOF. (a) Proof is elementary and is omitted. (b) Let $\alpha > 0$ and $\lambda_r = 1$ if $r \leqq 1$, $= 2^{r-1}$ if $r \geqq 1$. Suppose h_n is in $H(\mathcal{Y}, \mathcal{W}_r)$, $n = 1, 2, \cdots$ and

$$\lim_{n\to\infty} h_n(P) = h(P)$$

uniformly for all P in $\mathcal{P}$. Therefore there exists a positive integer m such that $|h_m(P) - h(P)| \leqq \alpha/2\lambda_r$, $P \,\varepsilon\, \mathcal{P}$. Also there exists a Y in $\mathcal{Y}$ such that

$$E_P|Y - h_m(P)|^r \leqq \alpha/2\lambda_r, \qquad P \,\varepsilon\, \mathcal{P}.$$

Then Y is an estimate for h with expected loss $\leqq \alpha$. For $E(|Y - h(P)|^r) < \lambda_r(E|Y - h_m(P)|^r + |h_m(P) - h(P)|^r) < \alpha$, $P \,\varepsilon\, \mathcal{P}$. It may be noted that this theorem is true of $H(\mathcal{Y}^0_m, \mathcal{W}_r)$, $m = 1, 2, \cdots, \infty$. This is an immediate consequence of Lemma 5.2.

6. Acknowledgments. I wish to acknowledge with gratitude the help, guidance and encouragement of Dr. D. L. Burkholder in the preparation of this paper. I am also indebted to the referee for his detailed and careful examination of this paper and for his many useful suggestions.

REFERENCES

[1] BAHADUR, R. R. and SAVAGE, L. J. (1956). The non-existence of certain statistical procedures in non-parametric problems. *Ann. Math. Statist.* **27** 1115–1122.

[2] Birnbaum, A. (1954). Statistical methods for Poisson processes and exponential populations. *J. Amer. Statist. Assoc.* **49** 254–266.

[3] Birnbaum, A. and Healy, W. C. Jr. (1960). Estimates with prescribed variance based on two-stage sampling. *Ann. Math. Statist.* **31** 662–676.

[4] Chapman, D. G. (1950). Some two-sample tests. *Ann. Math. Statist.* **21** 601–606.

[5] Farrell, R. (1959). Sequentially determined bounded length confidence intervals. Ph.D. Thesis, Univ. of Illinois.

[6] Ghurye, S. G. (1958). Note on sufficient statistics and two-stage procedures. *Ann. Math. Statist.* **29** 155–166.

[7] Graybill, F. A. (1958). Determining sample size for a specified width confidence interval. *Ann. Math. Statist.* **29** 282–287.

[8] Hoeffding, W. and Wolfowitz, J. (1958). Distinguishability of sets of distributions. *Ann. Math. Statist.* **29** 700–718.

[9] Moshman, Jack. (1958). A method for selecting the size of the initial sample in Stein's two sample procedure. *Ann. Math. Statist.* **29** 1271–1275.

[10] Stein, C. M. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* **16** 243–258.