ASYMPTOTICALLY SHORTEST CONFIDENCE INTERVALS[1]

BY ABRAHAM WALD[2]

Columbia University

The theory of confidence intervals, based on the classical theory of probability, has been treated by J. Neyman.[3] While Neyman considers the case of small samples, we shall deal here with the limit properties of the confidence intervals if the number of observations approaches infinity.

1. Definitions. We will start with some of Neyman's definitions. Let $f(x, \theta)$ be the probability density function of a variate x involving an unknown parameter θ. Denote by E_n a point of the n-dimensional sample space of n independent observations on x. If $\rho(E_n)$ denotes for each E_n a subset of the real axis, the symbol $P[\rho(E_n)c\theta' \mid \theta'']$ will denote the probability that $\rho(E_n)$ contains θ' under the hypothesis that θ'' is the true value of the parameter. Let $\underline{\theta}(E_n)$ and $\bar{\theta}(E_n)$ be two real functions defined over the whole sample space such that $\underline{\theta}(E_n) \leq \bar{\theta}(E_n)$. The interval $\delta(E_n) = [\underline{\theta}(E_n), \bar{\theta}(E_n)]$ is called a confidence interval of θ corresponding to the confidence coefficient α $(0 < \alpha < 1)$ if $P[\delta(E_n)c\theta \mid \theta] = \alpha$ for all values of θ.

The interval function $\delta(E_n)$ is called a shortest confidence interval of θ corresponding to the confidence coefficient α if

(a) $P[\delta(E_n)c\theta \mid \theta] = \alpha$ for all values of θ, and

(b) for any interval function $\delta'(E_n)$ which satisfies the condition (a) we have

$$P[\delta(E_n)c\theta' \mid \theta''] \leq P[\delta'(E_n)c\theta' \mid \theta''],$$

for arbitrary values θ' and θ''.

The interval function $\delta(E_n)$ is called a shortest unbiased confidence interval of θ if the following three conditions are fulfilled:

(a) $P[\delta(E_n)c\theta \mid \theta] = \alpha$ for all values of θ.

(b) $P[\delta(E_n)c\theta' \mid \theta''] \leq \alpha$ for all values of θ' and θ''.

(c) For any interval function $\delta'(E_n)$ for which the conditions (a) and (b) are satisfied, we have

$$P[\delta(E_n)c\theta' \mid \theta''] \leq P[\delta'(E_n)c\theta' \mid \theta''],$$

for all values of θ' and θ''.

For any relation R we shall denote by $P(R \mid \theta)$ the probability that R holds under the hypothesis that θ is the true value of the parameter. Similarly for

[1] Presented at a joint meeting of the Institute of Mathematical Statistics and the American Mathematical Society in Hanover, September, 1940.

[2] Research under a grant-in-aid from the Carnegie Corporation of New York.

[3] J. NEYMAN, "Outline of a theory of statistical estimation based on the classical theory of probability," *Phil. Trans. Roy. Soc. London*, Vol. 236 (1937), pp. 333–380.

any region Q_n of the n-dimensional sample space the symbol $P(Q_n \mid \theta)$ will denote the probability that the sample point E_n falls in Q_n under the hypothesis that θ is the true value of the parameter.

In all that follows we shall denote a region of the n-dimensional sample space by a capital letter with the subscript n.

A real function $\bar{\theta}(E_n)$ is called a best upper estimate of θ if the following two conditions are fulfilled:

(a) $P[\theta \leq \bar{\theta}(E_n) \mid \theta] = \alpha$ for all values of θ.

(b) For any function $\bar{\theta}'(E_n)$ which satisfies the condition (a) we have

$$P[\theta' \leq \bar{\theta}(E_n) \mid \theta''] \leq P[\theta' \leq \bar{\theta}'(E_n) \mid \theta'']$$

for all values θ' and θ'' for which $\theta' \geq \theta''$.

A real function $\underline{\theta}(E_n)$ is called a best lower estimate of θ if the following two conditions are fulfilled:

(a) $P[\theta \geq \underline{\theta}(E_n) \mid \theta] = \alpha$ for all values of θ.

(b) For any function $\underline{\theta}'(E_n)$ which satisfies the condition (a) we have

$$P[\theta' \geq \underline{\theta}(E_n) \mid \theta''] \leq P[\theta' \geq \underline{\theta}'(E_n) \mid \theta'']$$

for all values of θ' and θ'' for which $\theta' \leq \theta''$.

We will extend the above definitions of Neyman to the limit case when n approaches infinity.

Definition I: *A sequence of interval functions* $\{\delta_n(E_n)\}$ $(n = 1, 2, \cdots)$ *is called an asymptotically shortest confidence interval of* θ *if the following two conditions are fulfilled*:

(*a*) $P[\delta_n(E_n) c \theta \mid \theta] = \alpha$ *for all values of* θ.

(*b*) *For any sequence of interval functions* $\{\delta_n'(E_n)\}$ $(n = 1, 2, \cdots, ad\ inf.)$ *which satisfies* (*a*), *the least upper bound of*

$$P[\delta_n(E_n) c \theta' \mid \theta''] - P[\delta_n'(E_n) c \theta' \mid \theta'']$$

with respect to θ' *and* θ'' *converges to zero as* $n \to \infty$.

Definition II: *A sequence of interval functions* $\{\delta_n(E_n)\}$ *is called an asymptotically shortest unbiased confidence interval of* θ *if the following three conditions are fulfilled*:

(*a*) $P[\delta_n(E_n) c \theta \mid \theta] = \alpha$ *for all values of* θ.

(*b*) *The least upper bound of* $P[\delta_n(E_n) c \theta' \mid \theta'']$ *with respect to* θ' *and* θ'' *converges to* α *with* $n \to \infty$.

(*c*) *For any sequence of interval functions* $\{\delta_n'(E_n)\}$ *which satisfies the conditions* (*a*) *and* (*b*), *the least upper bound of*

$$P[\delta_n(E_n) c \theta' \mid \theta''] - P[\delta_n'(E_n) c \theta' \mid \theta'']$$

with respect to θ' *and* θ'' *converges to zero with* $n \to \infty$.

Definition III: *A sequence of real functions* $\{\bar{\theta}_n(E_n)\}$ $(n = 1, 2, \cdots, ad\ inf.)$ *is called an asymptotically best upper estimate of* θ *if the following two conditions are fulfilled*:

(*a*) $P[\theta \leq \bar{\theta}_n(E_n) \mid \theta] = \alpha$ *for all values of* θ.

(b) *For any sequence of functions* $\{\bar{\theta}'_n(E_n)\}$ *which satisfies* (a) *the least upper bound of*

$$P[\theta' \leq \bar{\theta}_n(E_n) \mid \theta''] - P[\theta' \leq \bar{\theta}'_n(E_n) \mid \theta'']$$

in the domain $\theta' \geq \theta''$ *converges to zero with* $n \to \infty$.

DEFINITION IV: *A sequence of real functions* $\{\underline{\theta}_n(E_n)\}$ *is called an asymptotically best lower estimate of* θ *if the following two conditions are fulfilled:*

(a) $P[\theta \geq \underline{\theta}_n(E_n) \mid \theta] = \alpha$ *for all values of* θ.

(b) *For any sequence of functions* $\{\underline{\theta}'_n(E_n)\}$ *which satisfies* (a) *the least upper bound of*

$$P[\theta' \geq \underline{\theta}_n(E_n) \mid \theta''] - P[\theta' \geq \underline{\theta}'_n(E_n) \mid \theta'']$$

in the domain $\theta' \leq \theta''$ *converges to zero with* $n \to \infty$.

2. Two Propositions. PROPOSITION I: *Let* $\{W_n(\theta)\}$ $(n = 1, 2, \cdots, \textit{ad inf.})$ *be for each* θ *a sequence of regions such that the following two conditions are fulfilled:*

(a) $P[W_n(\theta) \mid \theta] = 1 - \alpha$ *for all values of* θ.

(b) *For any sequence of regions* $\{Z_n(\theta)\}$ *which satisfies* (a) *the least upper bound of*

$$P[Z_n(\theta') \mid \theta''] - P[W_n(\theta') \mid \theta'']$$

in the domain $\theta' \geq \theta''(\theta' \leq \theta'')$ *converges to zero with* $n \to \infty$.

Denote by $\rho_n(E_n)$ *the set of all values of* θ *for which* E_n *does not lie in* $W_n(\theta)$. *Then we have*

(c) $P[\rho_n(E_n) c\theta \mid \theta] = \alpha$ *for all values of* θ.

(d) *For any sequence of set functions* $\{\rho'_n(E_n)\}$ *which satisfies* (c), *the least upper bound of*

$$P[\rho_n(E_n) c\theta' \mid \theta''] - P[\rho'_n(E_n) c\theta' \mid \theta'']$$

in the domain $\theta' \geq \theta''(\theta' \leq \theta'')$ *converges to zero with* $n \to \infty$.

PROPOSITION II: *Let* $\{W_n(\theta)\}$ *be for each* θ *a sequence of regions such that the following three conditions are fulfilled:*

(a) $P(W_n(\theta) \mid \theta] = 1 - \alpha$ *for all values of* θ.

(b) *The greatest lower bound of* $P[W_n(\theta') \mid \theta'']$ *converges to* $1 - \alpha$ *with* $n \to \infty$.

(c) *For any sequence* $\{W'_n(\theta)\}$ *which satisfies* (a) *and* (b), *the least upper bound of*

$$P[W'_n(\theta') \mid \theta''] - P[W_n(\theta') \mid \theta'']$$

with respect to θ' *and* θ'' *converges to* 0 *with* $n \to \infty$.

Denote by $\rho_n(E_n)$ *the set of all values of* θ *for which* E_n *does not lie in* $W_n(\theta)$. *Then we have*

(d) $P[\rho_n(E_n) c\theta \mid \theta] = \alpha$ *for all values of* θ.

(e) *The least upper bound of* $P[\rho_n(E_n) c\theta' \mid \theta'']$ *converges to* α *with* $n \to \infty$.

(f) *For any sequence of setfunctions* $\{\rho'_n(E_n)\}$ *which satisfies* (d) *and* (e), *the least upper bound of*

$$P[\rho_n(E_n) c\theta' \mid \theta''] - P[\rho'_n(E_n) c\theta' \mid \theta'']$$

with respect to θ' *and* θ'' *converges to* 0 *with* $n \to \infty$.

The validity of the above propositions follows easily from the identity

$$P[\rho_n(E_n)c\theta' \mid \theta''] = 1 - P[W_n(\theta') \mid \theta''].$$

3. Assumptions on the probability density function. For any function $\psi(x)$ denote by $E_\theta\psi(x)$ the expected value of $\psi(x)$ under the assumption that θ is the true value of the parameter, i.e.

$$E_\theta\psi(x) = \int_{-\infty}^{+\infty} \psi(x)f(x, \theta)\, dx.$$

For any x, for any positive δ, and for any real value θ' denote by $\varphi_1(x, \theta', \delta)$ the greatest lower bound, and by $\varphi_2(x, \theta', \delta)$ the least upper bound of $\frac{\partial^2}{\partial\theta^2} \log f(x, \theta)$ in the interval $\theta' - \delta \leq \theta \leq \theta' + \delta$.

Throughout this paper the following assumptions on $f(x, \theta)$ will be made:

Assumption I: *The expectation* $E_{\theta'} \frac{\partial}{\partial\theta} \log f(x, \theta'')$ *is a continuous function of* θ' *and* θ'', *and for any pair of sequences* $\{\theta'_n\}$ *and* $\{\theta''_n\}$ $(n = 1, 2, \cdots, ad\ inf.)$ *for which*

$$\lim_{n=\infty} E_{\theta'_n} \frac{\partial}{\partial\theta} \log f(x, \theta''_n) = 0$$

also

$$\lim_{n=\infty} (\theta'_n - \theta''_n) = 0.$$

Furthermore

$$E_{\theta'}\left[\frac{\partial}{\partial\theta} \log f(x, \theta'')\right]^2$$

is a bounded function of θ' *and* θ'', *and* $E_\theta\left[\frac{\partial}{\partial\theta} \log f(x, \theta)\right]^2 = d(\theta)$ *has a positive lower bound.*

Assumption II: *There exists a positive value* k_0 *such that the expectations* $E_{\theta'}\varphi_1(x, \theta'', \delta)$ *and* $E_{\theta'}\varphi_2(x, \theta'', \delta)$ *are uniformly continuous functions of* θ', θ'' *and* δ *where* δ *takes only values for which* $|\delta| \leq k_0$. *Furthermore it is assumed that* $E_{\theta'}[\varphi_i(x, \theta'', \delta)]^2$ $(i = 1, 2)$ *are bounded functions of* θ', θ'' *and* δ $(|\delta| \leq k_0)$.

Assumption III: *The relations*

$$\int_{-\infty}^{+\infty} \frac{\partial}{\partial\theta} f(x, \theta)\, dx = \int_{-\infty}^{+\infty} \frac{\partial^2}{\partial\theta^2} f(x, \theta)\, dx = 0$$

hold.

The above assumption means simply that we may differentiate with respect to θ under the integral sign. In fact

$$\int_{-\infty}^{+\infty} f(x, \theta)\, dx = 1$$

identically in θ. Hence

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty} f(x, \theta)\, dx = \frac{\partial^2}{\partial\theta^2}\int_{-\infty}^{+\infty} f(x, \theta)\, dx = 0.$$

Differentiating under the integral sign, we obtain the relations in Assumption III.

ASSUMPTION IV: *There exists a positive* η *such that*

$$E_\theta\left[\frac{\partial}{\partial\theta}\log f(x, \theta)\right]^{2+\eta}$$

is a bounded function of θ.

4. Some theorems. The assumptions on $f(x, \theta)$ made in this paper become identical with the assumptions I–IV formulated in a previous paper[4] if a certain set ω involved in those assumptions is put equal to the whole real axis $(-\infty, +\infty)$. Hence we can make use of all results obtained in that paper putting $\omega = (-\infty, +\infty)$. Among others, the following statements have been proved there:

(A) Denote $\sum_{\alpha=1}^{n} \frac{1}{\sqrt{n}} \frac{\partial}{\partial\theta} \log f(x_\alpha, \theta)$ by $y_n(\theta, E_n)$ and let $R_n(\theta)$ be the region defined by the inequality $y_n(\theta, E_n) \geq A_n(\theta)$ where $A_n(\theta)$ is chosen such that $P[R_n(\theta) \mid \theta] = 1 - \alpha$. Then for any sequence of regions $\{Z_n(\theta)\}$ for which $P[Z_n(\theta) \mid \theta] = 1 - \alpha$, the least upper bound of

$$P[Z_n(\theta') \mid \theta''] - P[R_n(\theta') \mid \theta'']$$

in the set $\theta'' \geq \theta'$ converges to 0 with $n \to \infty$.

(B) Let $S_n(\theta)$ be the region defined by the inequality $y_n(\theta, E_n) \leq B_n(\theta)$ where $B_n(\theta)$ is defined such that $P[S_n(\theta) \mid \theta] = 1 - \alpha$. Then for any sequence of regions $\{Z_n(\theta)\}$ for which $P[Z_n(\theta) \mid \theta] = 1 - \alpha$, the least upper bound of

$$P[Z_n(\theta') \mid \theta''] - P[S_n(\theta') \mid \theta'']$$

in the set $\theta'' \leq \theta'$ converges to 0 with $n \to \infty$.

(C) Denote by $T_n(\theta)$ the region defined by $| y_n(\theta, E_n) | \geq C_n(\theta)$ where $C_n(\theta)$ is chosen such that

(a) $P[T_n(\theta) \mid \theta] = 1 - \alpha$.

Then $T_n(\theta)$ satisfies also the following two conditions:

(b) The greatest lower bound of $P[T_n(\theta')\theta'']$ converges to $1 - \alpha$ with $n \to \infty$.

(c) For any sequence of regions $\{Z_n(\theta)\}$ which satisfies (a) and (b), the least upper bound of

$$P[Z_n(\theta') \mid \theta''] - P[T_n(\theta') \mid \theta'']$$

converges to 0 with $n \to \infty$.

---

[4] A. WALD, "Some examples of asymptotically most powerful tests," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 396–408.

On account of Propositions I and II we easily get the following theorems:

THEOREM I: *Denote by $\xi_n(E_n)$ the set of all values of θ for which $y_n(\theta, E_n) \leq A_n(\theta)$ and $A_n(\theta)$ is defined such that $P[y_n(\theta, E_n) > A_n(\theta) \mid \theta] = 1 - \alpha$. Then $\xi_n(E_n)$ satisfies the following two conditions:*

(a) *$P[\xi_n(E_n) c\theta \mid \theta] = \alpha$ for all values of θ.*

(b) *For any sequence of setfunctions $\{\xi_n'(E_n)\}$ which satisfies the condition (a), the least upper bound of*

$$P[\xi_n(E_n)c\theta' \mid \theta''] - P[\xi_n'(E_n)c\theta' \mid \theta'']$$

in the set $\theta'' \geq \theta'$ converges to 0 with $n \to \infty$.

THEOREM II: *Denote by $\zeta_n(E_n)$ the set of all values of θ for which $y_n(\theta, E_n) \geq B_n(\theta)$ and $B_n(\theta)$ is defined such that $P[y_n(\theta, E_n) < B_n(\theta) \mid \theta] = 1 - \alpha$. Then $\zeta_n(E_n)$ satisfies the following two conditions:*

(a) *$P[\zeta_n(E_n)c\theta \mid \theta] = \alpha$ for all values of θ.*

(b) *For any sequence of setfunctions $\{\zeta_n'(E_n)\}$ which satisfies the condition (a), the least upper bound of*

$$P[\zeta_n(E_n)c\theta' \mid \theta''] - P[\zeta_n'(E_n)c\theta' \mid \theta'']$$

in the set $\theta'' \leq \theta'$ converges to 0 with $n \to \infty$.

THEOREM III: *Denote by $\rho_n(E_n)$ the set of all values of θ for which $| y_n(\theta, E_n) | \leq C_n(\theta)$ and $C_n(\theta)$ is chosen such that $P[| y_n(\theta, E_n) | > C_n(\theta) \mid \theta] = 1 - \alpha$. Then $\rho_n(E_n)$ satisfies the following three conditions:*

(a) *$P[\rho_n(E_n)c\theta \mid \theta] = \alpha$ for all values of θ.*

(b) *The least upper bound of $P[\rho_n(E_n)c\theta' \mid \theta'']$ converges to α with $n \to \infty$.*

(c) *For any sequence of setfunctions $\{\rho_n'(E_n)\}$ which satisfies the conditions (a) and (b), the least upper bound of*

$$P[\rho_n(E_n)c\theta' \mid \theta''] - P[\rho_n'(E_n)c\theta' \mid \theta'']$$

converges to zero with $n \to \infty$.

Now we shall investigate the question whether the sets $\xi_n(E_n)$, $\zeta_n(E_n)$ and $\rho_n(E_n)$ are intervals. For this purpose we will prove some propositions.

PROPOSITION III: *Let ϵ and D be two positive numbers such that $\epsilon < D$. Denote by $Q_n(\theta, \epsilon, D)$ the region which consists of all points E_n for which*

$$y_n(\theta + \epsilon', E_n) \leq -n^{1/4}, \qquad \text{and} \qquad y_n(\theta - \epsilon', E_n) \geq n^{1/4}$$

for all values ϵ' in the interval $[\epsilon, D]$. Then we have

$$\lim_{n=\infty} P[Q_n(\theta, \epsilon, D) \mid \theta] = 1 \tag{1}$$

uniformly in θ.

PROOF: Let $\epsilon_1, \epsilon_2, \cdots, \epsilon_r$ be a sequence of points in the interval $[\epsilon, D]$ such that $\epsilon_1 - \epsilon = \epsilon_2 - \epsilon_1 = \cdots = \epsilon_r - \epsilon_{r-1} = D - \epsilon_r = k_0$ (say), where r is chosen sufficiently large such that Assumption II holds for $| \delta | \leq k_0$. Denote by $R_n(\theta, \epsilon_i)$ the region in which

$$y_n(\theta + \epsilon_i, E_n) \leq -n^{1/4}. \tag{2}$$

We will show that

$$(3) \qquad \lim_{n=\infty} P[R_n(\theta, \epsilon_i) \mid \theta] = 1$$

uniformly in θ.

From Assumption I it follows that the greatest lower bound of

$$\left| E_\theta \frac{\partial}{\partial\theta} \log f(x, \theta + \epsilon') \right|$$

with regard to ϵ' in the interval $[\epsilon, D]$ is positive. Let this greatest lower bound be $A > 0$. Since on account of Assumption I $E_\theta \frac{\partial}{\partial\theta} \log f(x, \theta + \epsilon')$ is a continuous function of ϵ', it does not change sign in the interval $\epsilon \leq \epsilon' \leq D$. Since this is true for arbitrarily small ϵ and since $E_\theta \left[\frac{\partial}{\partial\theta} \log f(x,\theta) \right]^2 = -E_\theta \frac{\partial^2}{\partial\theta^2} \log f(x, \theta)$ has a positive lower bound (Assumption I), it follows easily on account of Assumption II that

$$E_\theta \frac{\partial}{\partial\theta} \log f(x, \theta + \epsilon') < 0.$$

Hence

$$(4) \qquad E_\theta \frac{\partial}{\partial\theta} \log f(x, \theta + \epsilon') \leq -A < 0 \quad \text{for} \quad \epsilon \leq \epsilon' \leq D,$$

and therefore

$$(5) \qquad E_\theta y_n(\theta + \epsilon', E_n) \leq -A\sqrt{n} \qquad \text{for } \epsilon \leq \epsilon' \leq D.$$

From Assumption II it follows that the variance of $y_n(\theta + \epsilon', E_n)$ is a bounded function of θ and ϵ'. Hence

$$(6) \qquad \lim_{n=\infty} P[y_n(\theta + \epsilon_i, E_n) \leq -\tfrac{1}{2}A\sqrt{n} \mid \theta] = 1$$

uniformly in θ. The equation (3) is a consequence of (6).

Denote by $S_n(\theta, \epsilon_i)$ the region in which

$$\left| \frac{1}{n} \sum_\alpha \varphi_i(x_\alpha, \theta + \epsilon_i, k_0) \right| < C \qquad (i = 1, 2)$$

where C is greater than the least upper bound of $| E_\theta \varphi_i(x, \theta', k_0) |$ with respect to θ and θ'. Then we have on account of Assumption II:

$$(7) \qquad \lim_{n=\infty} P[S_n(\theta, \epsilon_i) \mid \theta] = 1 \qquad (i = 1, 2, \cdots, r)$$

uniformly in θ. In the region $S_n(\theta, \epsilon_i)$ we obviously have

$$(8) \qquad y_n(\theta + \epsilon_i', E_n) \leq y_n(\theta + \epsilon_i, E_n) + 2k_0\sqrt{n}C$$

for all values ϵ_i' in the interval $[\epsilon_i - k_0, \epsilon_i + k_0]$. By choosing r sufficiently large we can always achieve that

$$2k_0 C \leq \frac{A}{4}.$$

Denote by $T_n(\theta, \epsilon_i)$ the region in which

$$y_n(\theta + \epsilon_i', E_n) \leq -\frac{A}{4}\sqrt{n} \quad \text{for} \quad \epsilon_i - k_0 \leq \epsilon_i' \leq \epsilon_i + k_0. \tag{9}$$

From (6), (7) and (8) we get

$$\lim_{n=\infty} P[T_n(\theta, \epsilon_i) \mid \theta] = 1 \tag{10}$$

uniformly in θ. Let $Q_n'(\theta, \epsilon, D)$ be the common part of the r regions $T_n(\theta, \epsilon_1), \cdots, T_n(\theta, \epsilon_r)$, i.e. $Q_n'(\theta, \epsilon, D)$ is the set of all points E_n for which

$$y_n(\theta + \epsilon', E_n) \leq -\frac{A}{4}\sqrt{n}$$

for all ϵ' in the interval $[\epsilon, D]$. Since r is a fixed positive integer not depending on n, we get from (10)

$$\lim_{n=\infty} P[Q_n'(\theta, \epsilon, D) \mid \theta] = 1 \tag{11}$$

uniformly in θ.

In the same way we can prove that

$$\lim_{n=\infty} P[Q_n''(\theta, \epsilon, D) \mid \theta] = 1 \tag{12}$$

uniformly in θ, where $Q_n''(\theta, \epsilon, D)$ denotes the region in which

$$y_n(\theta - \epsilon', E_n) \geq \frac{A}{4}\sqrt{n} \quad \text{for all} \quad \epsilon' \text{ in } [\epsilon, D].$$

Proposition III follows from (11) and (12).

PROPOSITION IV: *Denote by $V_n(\theta, \epsilon)$ the region in which*

$$\frac{\partial}{\partial\theta} y_n(\theta', E_n) < -n^{\frac{1}{2}}$$

for all values θ' in the interval $[\theta - \epsilon, \theta + \epsilon]$. There exists a positive ϵ such that

$$\lim_{n=\infty} P[V_n(\theta, \epsilon) \mid \theta] = 1$$

uniformly in θ.

PROOF: Since the least upper bound of $E_\theta\varphi_2(x, \theta, 0)$ is <0, we get from Assumption II that the least upper bound of $E_\theta\varphi_2(x, \theta, \epsilon)$ is <0 for sufficiently small $\epsilon > 0$. Denote the least upper bound of $E_\theta\varphi_2(x, \theta, \epsilon)$ by $-B$ and let the region in which

$$\frac{1}{n}\sum_{\alpha} \varphi_2(x_\alpha, \vartheta, \epsilon) < -\tfrac{1}{2}B$$

be denoted by $W_n(\theta, \epsilon)$. From Assumption II it follows that

$$\lim_{n=\infty} P[W_n(\theta, \epsilon) \mid \theta] = 1$$

uniformly in θ. Since for almost all n $W_n(\theta, \epsilon)$ is a subset of $V_n(\theta, \epsilon)$, Proposition IV is proved.

PROPOSITION V: *Let $A_n(\theta)$, $B_n(\theta)$, $C_n(\theta)$ be the functions as defined in Theorems* I-III. *There exists a finite value G such that*

$$| A_n(\theta) | < G, \qquad | B_n(\theta) | < G \quad \text{and} \quad | C_n(\theta) | < G$$

for all θ and all n.

Proposition V follows easily from the fact that the variance of $y_n(\theta, E_n)$ is a bounded function of n and θ.

Let D be an arbitrary positive number and denote by $W_n(\theta, D)$ the region consisting of all points E_n for which the following conditions are fulfilled:

(a) The equation $y_n(\theta', E_n) = A_n(\theta')$ has exactly one root in θ' which lies in the interval $[\theta - D, \theta + D]$.

(b) The equation $y_n(\theta', E_n) = B_n(\theta')$ has exactly one root in θ' which lies in the interval $[\theta - D, \theta + D]$.

(c) The equation $y_n(\theta', E_n) = C_n(\theta')$ has exactly one root in θ' which lies in the interval $[\theta - D, \theta + D]$.

(d) The equation $y_n(\theta', E_n) = -C_n(\theta')$ has exactly one root in θ' which lies in the interval $[\theta - D, \theta + D]$.

(e) The common part of $[\theta - D, \theta + D]$ and $\xi_n(E_n)$ is the interval $[\theta'_n(E_n), D]$ where $\theta'_n(E_n)$ denotes the root of the equation in (a).

(f) The common part of $\zeta_n(E_n)$ and $[\theta - D, \theta + D]$ is the interval $[-D, \theta''_n(E_n)]$ where $\theta''_n(E_n)$ denotes the root of the equation in (b).

(g) The common part of $\rho_n(E_n)$ and $[\theta - D, \theta + D]$ is the interval $[\underline{\theta}_n(E_n), \bar{\theta}_n(E_n)]$ where $\underline{\theta}_n(E_n)$ denotes the root of the equation in (c) and $\bar{\theta}_n(E_n)$ denotes the root of the equation in (d).

From Propositions III–V follows easily the following

PROPOSITION VI: *For any positive value D*

$$\lim_{n=\infty} P[W_n(\theta, D) \mid \theta] = 1,$$

uniformly in θ, provided that the functions $A_n(\theta)$, $B_n(\theta)$ and $C_n(\theta)$ are continuous and of bounded variation in any finite interval.

We will show that Proposition VI remains valid for $D = +\infty$, if we make the following

ASSUMPTION V: *Denote by* $\psi(x, \theta, D)$ *the least upper bound of* $\frac{\partial}{\partial \theta} \log f(x, \theta')$ *with respect to* θ' *where* $\theta' \geq \theta + D$. *Denote furthermore by* $\psi^*(x, \theta, D)$ *the greatest lower bound of* $\frac{\partial}{\partial \theta} \log f(x,\theta')$ *with respect to* θ' *where* $\theta' \leq \theta - D$. *There exists a positive* D *such that the least upper bound of* $E_\theta \psi(x, \theta, D)$ *with respect to* θ *is negative, the greatest lower bound of* $E_\theta \psi^*(x, \theta, D)$ *with respect to* θ *is positive, and the variances of* $\psi(x, \theta, D)$ *and* $\psi^*(x, \theta, D)$ *are bounded functions of* θ. *(The variances are calculated under the assumption that* θ *is the true value of the parameter.)*

It follows easily from Assumption V that

$$\lim_{n=\infty} P\left[\frac{1}{\sqrt{n}} \sum_\alpha \psi(x_\alpha, \theta, D) < -n^{\frac{1}{4}} \mid \theta\right] = \lim_{n=\infty} P\left[\frac{1}{\sqrt{n}} \sum_\alpha \psi^*(x_\alpha, \theta, D) > n^{\frac{1}{4}} \mid \theta\right] = 1$$

uniformly in θ.

Since

$$\frac{1}{\sqrt{n}} \sum_\alpha \psi(x_\alpha, \theta, D) \geq y_n(\theta', E_n) \quad \text{for} \quad \theta' \geq \theta + D$$

and

$$\frac{1}{\sqrt{n}} \sum_\alpha \psi^*(x_\alpha, \theta, D) \leq y_n(\theta', E_n) \quad \text{for} \quad \theta' \leq \theta - D,$$

Proposition VI remains valid if we substitute $+\infty$ for D.

Hence we obtain the following

COROLLARY: *If the assumptions* I–V *are fulfilled and if* $A_n(\theta)$, $B_n(\theta)$ *and* $C_n(\theta)$ *are continuous and of bounded variation in any finite interval, then*

(*a*) *The root* $\theta_n'(E_n)$ *of the equation* $y_n(\theta, E_n) = A_n(\theta)$ *in* θ *is an asymptotically best lower estimate of* θ.

(*b*) *The root* $\theta_n''(E_n)$ *of the equation* $y_n(\theta, E_n) = B_n(\theta)$ *in* θ *is an asymptotically best upper estimate of* θ.

(*c*) *The interval* $[\underline{\theta}_n(E_n), \bar{\theta}_n(E_n)$ *is an asymptotically shortest unbiased confidence interval of* θ, *where* $\underline{\theta}_n(E_n)$ *denotes the root of the equation* $y_n(\theta, E_n) = +C_n(\theta)$, *and* $\bar{\theta}_n(E_n)$ *denotes the root of the equation* $y_n(\theta, E_n) = -C_n(\theta)$.

5. Some Remarks. 1. I should like to make a few remarks about the relationship of these results to those obtained by S. S. Wilks.[5] The definition of a shortest confidence interval underlying Wilks' investigations is somewhat different from that of Neyman's which has been used in this paper. According to Wilks, a confidence interval $\delta(E_n)$ is called shortest in the average if the expected

[5] S. S. WILKS, "Shortest average confidence intervals from large samples," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 166–175.

value of the length of $\delta(E_n)$ is a minimum. The main result obtained by Wilks can be formulated as follows: The confidence interval $[\underline{\theta}_n(E_n), \bar{\theta}_n(E_n)]$ given in our Corollary is asymptotically shortest in the average compared with all confidence intervals computed on the basis of functions belonging to a certain class C. In the present paper no restriction to a certain class of functions has been made.

2. If the parameter space Ω is not the whole real axis, but an open subset of it, and if the assumptions I–V are fulfilled when θ can take only values in Ω, the previously proved Corollary remains valid. If Ω is a bounded set, Assumption V is a consequence of Assumptions I–IV.