# 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 $\theta$. 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 $\theta'$ under the hypothesis that $\theta''$ 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 $\theta$ corresponding to the confidence coefficient $\alpha$ $(0 < \alpha < 1)$ if $P[\delta(E_n)c\theta \mid \theta] = \alpha$ for all values of $\theta$.

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

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

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

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

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

(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 $\theta'$ and $\theta''$.

For any relation $R$ we shall denote by $P(R \mid \theta)$ the probability that $R$ holds under the hypothesis that $\theta$ 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.