SHORTEST AVERAGE CONFIDENCE INTERVALS
FROM LARGE SAMPLES

BY S. S. Wilks

1. Introduction. The method of fiducial argument [1, 2] in statistics has gained considerable prominence within the last few years as a method of inferring the values of population parameters from samples “randomly drawn” from populations having distribution laws of known functional forms. The method has also been shown to be applicable [2] to the problem of inferring the values of statistical functions in samples from samples already observed, assuming all samples to be drawn from a population with a distribution law of a given functional form.

The main ideas of a procedure which is sufficient for carrying out fiducial inference for certain cases of a single population parameter may be summed up in the following steps:

(a) A sample is assumed to be “randomly drawn” from a population with a distribution law \( f(x, \theta) \) of known functional form.

(b) A function \( \psi(x_1, x_2, \ldots, x_n, \theta) \) of the sample values \( x_1, x_2, \ldots, x_n \) and parameter \( \theta \) is devised, which is a monotonic function of \( \theta \) for a given sample, so that the sampling distribution \( G(\psi) d\psi \) of \( \psi(x_1, x_2, \ldots, x_n, \theta_0) = \psi_0 \), say, in samples from the population with \( \theta = \theta_0 \) is independent of \( \theta_0 \) and the \( x \)'s, except as they enter into \( \psi \).

(c) For a given probability \( \alpha \) a pair of numbers \( \psi_\alpha' \) and \( \psi_\alpha'' \) is chosen so that when \( \theta = \theta_0 \), the probability that \( \psi_\alpha' < \psi_0 < \psi_\alpha'' \) is \( 1 - \alpha \), or more, briefly,

\[
P(\psi_\alpha' < \psi_0 < \psi_\alpha'' | \theta = \theta_0) = 1 - \alpha
\]

which can be stated in the alternative form

\[
P(\theta < \theta_0 < \bar{\theta} | \theta = \theta_0) = 1 - \alpha.
\]

(d) \( \theta \) and \( \bar{\theta} \) being functions of \( \psi_\alpha', \psi_\alpha'' \) and the sample, are subject to sampling fluctuations and it can be stated that the probability is \( 1 - \alpha \) that they will include the true value of \( \theta \), whatever it may be, that is, \( \theta_0 \), between them. The statement holds for all values which \( \theta_0 \) may take on.

The numbers \( \theta \) and \( \bar{\theta} \) are known as fiducial or confidence limits [3] of \( \theta_0 \) and \( (\theta, \bar{\theta}) \) a confidence interval for the confidence coefficient \( 1 - \alpha \). We therefore have the following rule for making inferences about the unknown number \( \theta_0 \) once \( \psi \) has been found: For a given sample solve the equations

\[
\psi(x_1, x_2, \ldots, x_n, \theta_0) = \psi_\alpha', \quad \psi(x_1, x_2, \ldots, x_n, \theta_0) = \psi_\alpha''
\]