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 $\psi(x_1, x_2, \cdots, x_n, \theta)$ of the sample values $x_1, x_2, \cdots, x_n$ and parameter θ is devised, which is a monotonic function of θ for a given sample, so that the sampling distribution $G(\psi)d\psi$ of $\psi(x_1, x_2, \cdots, x_n, \theta_0) = \psi_0$, say, in samples from the population with $\theta = \theta_0$ is independent of θ_0 and the x's, except as they enter into ψ.

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

$$P(\psi'_\alpha < \psi_0 < \psi''_\alpha \mid \theta = \theta_0) = 1 - \alpha \tag{1}$$

which can be stated in the alternative form

$$P(\underline{\theta} < \theta_0 < \bar{\theta} \mid \theta = \theta_0) = 1 - \alpha. \tag{2}$$

(d) $\underline{\theta}$ and $\bar{\theta}$ being functions of ψ'_α, ψ''_α 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 θ, whatever it may be, that is, θ_0, between them. The statement holds for all values which θ_0 may take on.

The numbers $\underline{\theta}$ and $\bar{\theta}$ are known as *fiducial* or *confidence limits* [3] of θ_0 and $(\underline{\theta}, \bar{\theta})$ a *confidence interval* for the *confidence coefficient* $1 - \alpha$. We therefore have the following rule for making inferences about the unknown number θ_0 once ψ has been found: For a given sample solve the equations

$$\psi(x_1, x_2, \cdots, x_n, \theta_0) = \psi'_\alpha, \qquad \psi(x_1, x_2, \cdots, x_n, \theta_0) = \psi''_\alpha$$

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

for θ_0. Let $\underline{\theta}$ and $\bar{\theta}$ be the two values of θ_0 formally obtained. The statement that $\underline{\theta}$ and $\bar{\theta}$ will include the value of θ in the population actually sampled, if consistently made in each of an aggregate of cases involving populations having distributions of the same functional form $f(x, \theta)$ will be correct (theoretically) in $100(1 - \alpha)$ per cent of the cases.

If ψ is a function of statistics t_1 and t_2 of two samples from a population of known functional form, which is monotonic in each t for given values of the other, then one can argue fiducially about values of one t from values of the other one.

For a finite value of n and discrete distributions $f(x, \theta)$, it is not possible to carry through steps (b), (c), (d) as they are now stated. However, under certain conditions, it is possible to carry out a procedure for the discrete case which will allow one to say

$$P(\underline{\theta} < \theta_0 < \bar{\theta} \mid \theta = \theta_0) \geq 1 - \alpha. \tag{3}$$

ψ functions which have the property that their sampling distributions are independent of θ and the x's for a given distribution $f(x, \theta)$ are not, in general, unique. The question then arises as to how (if possible) one can choose ψ functions and limits ψ'_α and ψ''_α so as to get confidence intervals for a given α, which are shortest or "best" in some sense. Neyman [4] has investigated the problem of obtaining "best" confidence intervals for the case of small samples. The object of this paper is to consider the problem for large samples. Under fairly general conditions it will be seen that a rather simple asymptotic solution exists for the large-sample case, which is connected in an essential manner with the *method of maximum likelihood*.

2. An asymptotic distribution. Suppose a population Π has a distribution function $f(x, \theta)$, where x is a random variable and θ a parameter. Actually, $f(x, \theta)$ may involve several other parameters whose values may be regarded as fixed throughout the paper. The problem of arguing fiducially about several parameters simultaneously will not be considered in this paper. In order to include the case of a discrete as well as a continuous variate x, we shall consider the cumulative distribution function (c.d.f.) $F(x, \theta)$, which is monotonic and is such that

$$F(-\infty, \theta) = 0, \qquad F(+\infty, \theta) = 1, \qquad F(x + c, \theta) \geq F(x, \theta)$$
$$F(x + 0, \theta) = F(x, \theta), \text{ for } c > 0, \text{ and } a < \theta < b.$$

Thus, $F(x', \theta) = P(x \leq x' \mid \theta)$. In the case of a continuous variate x, where $f(x, \theta)$ is a probability density function, then $dF(x, \theta) = f(x, \theta)\, dx$; in the discrete case $dF(x, \theta) = f(x, \theta)$ where $f(x, \theta)$ is the probability that variate x takes the value indicated. We shall be interested in continuous functions $\varphi(x)$ for which the integral $\int \varphi(x)\, dF(x, \theta)$ taken in the Stieljtes sense exists. Limits on integral signs are understood to be $-\infty$ and ∞.

Now consider a sample 0_n of n individuals independently drawn from Π_0, the population for which the c.d.f. is $F(x, \theta_0)$. Let the values of x in the sample be $x_1, x_2, \cdots, x_n$. The probability element associated with the sample is

$$dP_n = \prod_{i=1}^{n} dF(x_i, \theta_0). \tag{4}$$

Let $L = \log dP_n$. Then assuming that $\frac{\partial}{\partial\theta} \log d\, F(x, \theta) = g(x, \theta)$, say, exists for $\theta = \theta_0$, and for each x, (except for a set of probability 0), we have

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{n} g(x_i, \theta). \tag{5}$$

In all ordinary problems in statistics $g(x, \theta)$ reduces to $\frac{\partial f(x, \theta)}{\partial \theta} \Big/ f(x, \theta)$ where $f(x, \theta)$ is probability in the case where x is a discrete random variable and probability density in the case of a continuous random variable. Let g_0 denote $g(x, \theta_0)$ and $\left(\frac{\partial L}{\partial \theta}\right)_0$ denote $\frac{\partial L}{\partial \theta}$ with $\theta = \theta_0$. Let

$$A_0^2 = E_0[(g_0)^2] = \int g_0^2\, dF(x, \theta_0). \tag{6}$$

$E_0(\varphi)$ will be used to denote the mathematical expectation of φ in samples from Π_0, i.e. when the population distribution is $dF(x, \theta_0)$. We shall consider the sampling theory of

$$\psi_0 = \frac{\left(\frac{\partial L}{\partial \theta}\right)_0}{\sqrt{n}\, A_0} \tag{7}$$

in large samples, from Π_0.

Let $\varphi_0^{(n)}(t)$ be the characteristic function of ψ_0 for samples from Π_0; it is defined by $E_0(e^{it\psi_0})$. Then

$$\begin{aligned} \varphi_0^{(n)}(t) &= \left\{E_0\left[\exp\left(\frac{itg_0}{\sqrt{n}\, A_0}\right)\right]\right\}^n \\ &= \left\{E_0\left[1 + \frac{itg_0}{\sqrt{n}\, A_0} - \frac{t^2 g_0^2}{2nA_0^2} + \frac{1}{n^{\frac{3}{2}}} g_0^3(\phi_1 + i\phi_2)\right]\right\}^n \end{aligned} \tag{8}$$

where ϕ_1 and ϕ_2 are real functions of t, x, n and θ_0, such that if $|E_0(g_0)^3| \leq K < \infty$, (i.e. the third moment of g_0 is finite when $\theta = \theta_0$), for $a < \theta_0 < b$, then $E_0[g_0^3\phi_1]$ and $E_0[g_0^3\phi_2]$ are uniformly bounded for some t-interval δ(which includes $t = 0$ as an interior point) for n larger than some n_0 and for θ_0 on any fixed subinterval of the interval (a, b). Suppose that $F(x, \theta)$ is such that

$$\int \frac{\partial}{\partial \theta} dF(x, \theta) = \frac{\partial}{\partial \theta} \int dF(x, \theta) = 0, \qquad a < \theta < b. \tag{9}$$

This condition implies that the range of x be independent of θ.

If n is allowed to increase indefinitely, then we have at once that $\varphi_0^{(n)}(t)$ tends to $e^{-\frac{1}{2}t^2}$ uniformly in the interval δ. We now make use of a theorem [5] which states that if an unlimited sequence of random variables $x^{(1)}, x^{(2)}, \cdots, x^{(n)} \cdots$ with c.d.f.'s $F^{(1)}(x), F^{(2)}(x), \cdots, F^{(n)}(x) \cdots$ have corresponding characteristic functions $\varphi^{(1)}(t), \varphi^{(2)}(t) \cdots \varphi^{(n)}(t) \cdots$ then a necessary and sufficient condition for $F^{(n)}(x)$ to converge uniformly to a c.d.f. $F(x)$ at each point of continuity of $F(x)$ on the interval $(-\infty, \infty)$ is that the sequence of characteristic functions converge uniformly to a function $\varphi_1(t)$ on an interval $|t| < \epsilon$ for some $\epsilon > 0$. The characteristic function $\varphi(t)$ associated with $F(x)$ will then be identical with $\varphi_1(t)$ and the sequence $\varphi^{(1)}(t), \cdots \varphi^{(n)}(t) \cdots$ converges to $\varphi(t)$ uniformly in every finite t-interval.

From this theorem it follows at once that, since $e^{-\frac{1}{2}t^2}$ is the characteristic function of a variate distributed normally with mean 0 and variance 1, the asymptotic c.d.f. of ψ_0 for large samples is given by

$$F(\psi_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\psi_0} e^{-\frac{1}{2}x^2}\, dx. \tag{10}$$

We may conveniently summarize the foregoing results in the following

THEOREM 1. *Let $x_1, x_2, \cdots, x_n$ be the values of x in a sample of independently drawn individuals from a population Π_0 which has a c.d.f. $F(x, \theta_0)$, such that* for $a < \theta_0 < b$,

(a) $\frac{\partial}{\partial\theta} dF(x, \theta)$ *exists for all x's except possibly for a set of probability* 0;

(b) $E_0[(g_0^3)]$ *is finite; for $n > n_0$,*

(c) *condition* (9) *is satisfied.*

Then the asymptotic c.d.f. of ψ_0 for large samples defined in (7) *is given by* (10).

The statistical significance of this Theorem is that if we know the functional form $f(x, \theta)$ (for which the first derivative $f'(x, \theta)$ with respect to θ exists) of the distribution function of a population Π and if the sample $x_1, x_2, \cdots, x_n$ is "randomly drawn" from Π_0, then the quantity

$$\psi_0 = \frac{\sum_1^n \frac{f'(x_i, \theta_0)}{f(x_i, \theta_0)}}{\sqrt{n}\sqrt{E_0\left[\left(\frac{f'(x, \theta_0)}{f(x, \theta_0)}\right)^2\right]}} \tag{11}$$

is a random variable which is approximately normally distributed with mean 0 and variance 1 in repeated large samples. It will be noticed that the quantity in the numerator of (11), is simply the derivative with respect to θ, at $\theta = \theta_0$, of the logarithm of the likelihood of θ for the given sample. ψ_0 is a function of the sample O_n and the true value θ_0 of the parameter θ, and the thing that makes ψ_0 a random variable is the random nature of the sample; θ_0 is a fixed but unknown number. Thus, for example, when $1 - \alpha = .95$ in (1) and knowing that we have "randomly drawn" a large sample O_n from a population Π_0 with distribution $f(x, \theta_0)$ of known functional form, we can say that the probability is .95 that the sample will produce a value of ψ_0 in the interval -1.96 to $+1.96$ that is,

$$(12) \qquad P(-1.96 < \psi_0 < 1.96 \mid \theta = \theta_0) = .95.$$

This statement holds, whatever may be the value of the unknown θ_0. Now, the inequality $-1.96 < \psi_0 < 1.96$ is equivalent to the inequality, $\underline{\theta} < \theta_0 < \bar{\theta}$ because of the monotonic nature of ψ_0 as a function of θ_0. Hence (12) is equivalent to

$$(13) \qquad P(\underline{\theta} < \theta_0 < \bar{\theta} \mid \theta = \theta_0) = .95$$

where $\underline{\theta}$ and $\bar{\theta}$ are obtained by solving $\psi_0 = \pm 1.96$ for θ_0. The *fiducial limits* $\underline{\theta}$ and $\bar{\theta}$ will thus be functions of the sample and will be subject to sampling variations. In general, of course, one could choose any probability level $1 - \alpha$, and find ψ_α so that

$$(14) \qquad P(-\psi_\alpha < \psi_0 < \psi_\alpha \mid \theta = \theta_0) = 1 - \alpha,$$

from which fiducial limits for θ_0 can be found as before.

The extension of Theorem 1 to the case in which the distribution function of the population Π involves several parameters $\theta_1, \theta_2, \cdots \theta_h$ having values in some region R of the space of θ's, is immediate. Π_0 in this case would be specified by the values $\theta_{10}, \theta_{20}, \cdots \theta_{h0}$. In fact, we can state the situation as

THEOREM 1′: *Let $F(x, \theta_1, \theta_2, \cdots, \theta_h)$ denote the c.d.f. of x and (allowing i, j, k to take on values $1, 2, \cdots, h$) let*

$$\psi_{i0} = \frac{1}{\sqrt{n}}\left(\frac{\partial L}{\partial \theta_i}\right)_0 \quad \textit{where} \quad L = \sum_{j=1}^{n} \log dF(x_t, \theta_{10}, \theta_{20}, \cdots, \theta_{h0}),$$

$$g_i = \frac{\partial}{\partial \theta_i}[\log dF(x, \theta_1, \cdots, \theta_h)],$$

$$A_{ij} = E_0[g_{i0}g_{j0}] \textit{ where } g_{i0} = g_i \textit{ with } \theta_i = \theta_{i0}.$$

If, in R,

(a) $\frac{\partial}{\partial \theta_i} dF(x, \theta_1, \cdots, \theta_h)$ *exists for all x's except possibly for a set of probability* 0;

(b) $E_0(g_{i0}g_{j0}g_{k0})$ *are all finite*;

(c) $\frac{\partial}{\partial \theta_i}\int dF(x, \theta_1, \theta_2, \cdots, \theta_h) = \int \frac{\partial}{\partial \theta_i} dF(x, \theta_1, \theta_2, \cdots, \theta_h) = 0$

(d) $\|A_{ij}\|$ *is non-singular*;

then the asymptotic distribution of the ψ_{i0} in large samples from Π_0 is a normal multivariate distribution with matrix $\| A_{ij} \|$ of variances and covariances, and zero means.

A similar theorem holds for the case in which Π is a multivariate population in addition to having several parameters.

The question now arises: In what sense is the *confidence interval* between $\underline{\theta}$ and

$\bar{\theta}$ as determined from ψ_0 "best"? It will be shown that the average rate of change of ψ_0 with respect to θ at $\theta = \theta_0$ is greater than that for a rather broad class of functions of the ψ type, that is functions of the observations and θ which are asymptotically normally distributed. Since we are dealing with large samples, we are only interested in values of θ in the neighborhood of θ_0, for which ψ as a function of θ is approximately linear, and demonstrating the property just stated regarding the average rate of change of ψ with respect to θ at $\theta = \theta_0$ is equivalent to showing that the two "values" of θ_0 for which $\psi_0 = \pm\psi_\alpha$, will, on the average be closer together than those computed from any other ψ function than ψ_0 of the class of functions to be considered. This class of functions will be designated as belonging to *class* C, and will now be more accurately defined.

3. **Functions of class C and their asymptotic distributions.** Following an argument similar to that used in proving Theorem 1, we can readily prove

THEOREM 2: *Let $h(x, \theta)$ be a function in which x has the c.d.f. $F(x, \theta)$, and which satisfies the following conditions for $a < \theta_0 < b$:*

$$\text{(a)}\quad E_0[h(x, \theta_0)] = 0; \tag{15}$$

$$\text{(b)}\quad E_0[\{h(x, \theta_0)\}^3] \text{ is finite, for } n > n_0.$$

Let

$$A_0^* = E_0\,[\{h(x, \theta_0)\}^2] \tag{16}$$

and for a sample of values $x_1, x_2, \cdots, x_n$ let

$$\psi_0^* = \frac{\sum_1^n h(x_i, \theta_0)}{\sqrt{n}\,A_0^*}. \tag{17}$$

Then the asymptotic c.d.f. of ψ_0^* for large samples from Π_0 is given by

$$F(\psi_0^*) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\psi_0^*} e^{-\frac{1}{2}x^2}\,dx.$$

We shall designate as belonging to class C any function ψ_0^* made up according to the rule expressed by (17), of functions $h(x, \theta)$ satisfying (a) and (b) in (15) and such that ψ_0^* is asymptotically normally distributed with zero mean and unit number. Clearly, ψ_0 as defined by (7) belongs to class C.

4. **Comparison of average confidence intervals computed from ψ_0 and ψ_0^*.** We shall now show that for each fixed value θ_0 of θ the average rate of change of ψ with respect to θ is greater than that of ψ^* for any $h(x, \theta)$ which is not a constant multiple of $g(x, \theta)$. Consider $\frac{\partial\psi}{\partial\theta}$ and $\frac{\partial\psi^*}{\partial\theta}$ for a given n. We have,

$$\frac{\partial\psi}{\partial\theta} = \frac{1}{\sqrt{n}A}\left\{\sum_1^n \frac{\partial g(x_i, \theta)}{\partial\theta} - \frac{1}{A}\sum_1^n g(x_i, \theta)\frac{\partial A}{\partial\theta}\right\} \tag{18}$$

$$\frac{\partial\psi^*}{\partial\theta} = \frac{1}{\sqrt{n}A^*}\left\{\sum_1^n \frac{\partial h(x_i, \theta)}{\partial\theta} - \frac{1}{A^*}\sum_1^n h(x_i, \theta)\frac{\partial A^*}{\partial\theta}\right\} \tag{19}$$

Now

$$\frac{\partial g(x_i, \theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left\{\frac{\partial[dF(x_i, \theta)]}{\partial \theta} \bigg/ dF(x_i, \theta)\right\} = \frac{\frac{\partial^2}{\partial \theta^2}[dF(x_i, \theta)]}{dF(x_i, \theta)} - [g(x_i, \theta)]^2.$$

Assuming that

$$\int \frac{\partial^2}{\partial \theta^2} dF(x, \theta) = \frac{\partial^2}{\partial \theta^2}\int dF(x, \theta) = 0 \tag{20}$$

and remembering that $E_0[g(x_i, \theta_0)] = 0$, we have

$$E_0\left[\left(\frac{\partial \psi}{\partial \theta}\right)_0\right] = -\sqrt{n}\, A_0 = \Delta_1 \tag{21}$$

and

$$E_0\left[\left(\frac{\partial \psi^*}{\partial \theta}\right)_0\right] = \frac{\sqrt{n}}{A_0^*} E_0\left[\left(\frac{\partial h(x, \theta)}{\partial \theta}\right)_0\right] = \Delta_2. \tag{22}$$

Now, since

$$\int h(x, \theta)\, dF(x, \theta) = 0 \tag{23}$$

and assuming that (23) can be differentiated under the integral sign, we have

$$E_0\left[\left(\frac{\partial h(x, \theta)}{\partial \theta}\right)_0\right] = \left[-\int h(x, \theta) \frac{\partial}{\partial \theta} dF(x, \theta)\right]_{\theta=\theta_0}. \tag{24}$$

For the difference $\Delta_1^2 - \Delta_2^2$ in samples from populations with $\theta = \theta_0$, we have

$$\frac{n}{(A_0^*)^2}\left\{\int \left(\frac{\frac{\partial}{\partial \theta} dF(x, \theta)}{dF(x, \theta)}\right)^2 dF(x, \theta) \cdot \int (h(x, \theta))^2\, dF(x, \theta) - \left[\int (h(x, \theta)\sqrt{dF(x, \theta)})\left(\frac{\frac{\partial}{\partial \theta} dF(x, \theta)}{\sqrt{dF(x\ \theta)}}\right)\right]^2\right\}_{\theta=\theta_0} \tag{25}$$

Making use of Schwartz' inequality which states that

$$\int g^2(x)\, dx \cdot \int h^2(x)\, dx \geq \left[\int g(x)\, h(x)\, dx\right]^2,$$

where the equality sign holds only if $g(x) \equiv K\, h(x)$, K being a constant, it is evident that independently of n, $\Delta_1^2 \geq \Delta_2^2$, and furthermore, the only condition under which $\Delta_1^2 = \Delta_2^2$ is that

$$h(x, \theta)\sqrt{dF(x, \theta)} \equiv K \frac{\frac{\partial}{\partial \theta} dF(x, \theta)}{\sqrt{dF(x, \theta)}},$$

that is,

$$h(x, \theta) \equiv K\, g(x, \theta). \tag{26}$$

Therefore we have

THEOREM 3. *If $g(x, \theta)$ and $h(x, \theta)$ satisfy the conditions of Theorems* 1 *and* 2 *respectively and furthermore, if* (20) *is satisfied and if the expression on the left in* (23) *can be differentiated under the integral sign with respect to θ, then the average rate of change of ψ with respect to θ for each fixed value θ_0 of θ is greater than that of ψ^* (for which $h(x, \theta) \not\equiv Kg(x, \theta)$ with respect to θ, when $\theta = \theta_0$ in samples from Π_0.*

This Theorem simply means that when computed from ψ_0 the fiducial limits for the true but unknown value θ_0 of the parameter θ, whatever value θ_0 may have on the interval $a < \theta_0 < b$ of possible values, are (for large samples) closer together *on the average* than those computed from any other ψ_0^* of class C. There is no function ψ_0^* which is more *efficient*, as it were, for determining confidence intervals for θ_0 than the particular ψ_0 given by (7) which is ψ_0^* with $h(x, \theta)$ replaced by $g(x, \theta)$, that is, $\frac{\partial}{\partial\theta} \log dF(x, \theta)$. The actual manner in which the fiducial limits for θ_0 are found for a given confidence coefficient $1 - \alpha$, is to set

$$\frac{\left(\frac{\partial L}{\partial \theta}\right)_0}{\sqrt{n}\, A_0} = \pm\psi_\alpha \tag{27}$$

and solve formally, for θ_0, where ψ_α is the value for which $\frac{2}{\sqrt{2\pi}} \int_{\psi_\alpha}^{\infty} e^{-\frac{1}{2}x^2}\, dx = \alpha$, which can be found from normal probability tables. The two values of θ_0 thus found are the fiducial limits $\underline{\theta}$ and $\bar{\theta}$ for the true value θ_0 and we can state that the probability is $1 - \alpha$ that $\underline{\theta}$ and $\bar{\theta}$ will include the true value θ_0 between them. This statement is valid whatever may be the value of θ_0 between a and b. This rule consistently followed for large samples, will produce fiducial limits $\underline{\theta}$ and $\bar{\theta}$ which are closest together *on the average*, for *each fixed value* of the probability level α between 0 and 1. It should be observed that no assumptions have been made regarding the existence of sufficient statistics.

5. **Examples.** EXAMPLE 1. Suppose a large sample of n individuals to be drawn from a population known to have the Poisson distribution law

$$f(x, m) = \frac{m^x e^{-m}}{x!}.$$

We have

$$L = -\log\left(\prod_1^n x_i!\right) + (\Sigma\, x_i) \log m - nm$$

$$\left(\frac{\partial L}{\partial m}\right)_0 = \frac{\Sigma\, x_i}{m_0} - n$$

$$A_0^2 = E_0\left[\left(\frac{\partial \log f}{\partial m}\right)_0^2\right] = E_0\left[\left(\frac{x}{m_0} - 1\right)^2\right] = \frac{1}{m_0}.$$

The fiducial limits $\underline{m}$ and $\bar{m}$ for $\alpha = .05$, that is, the 95 per cent fiducial limits, are found by formally solving the equations

$$\frac{\left(\frac{\partial L}{\partial m}\right)_0}{\sqrt{n}\, A_0} \equiv \frac{\left(\frac{\Sigma\, x_i}{m_0} - n\right)\sqrt{m_0}}{\sqrt{n}} = \pm 1.96$$

for m_0. The fiducial limits are found to be

$$\bar{x} + \frac{1.92}{n} \pm \sqrt{\frac{3.82}{n}\,\bar{x} + \frac{3.69}{n^2}}$$

Example 2. Consider a large sample of n individuals known to be from a binomial population having the two classes A and B. Let p denote the probability of an individual's belonging to A, and $q = l - p$ that of belonging to B. Let x denote the number of individuals belonging to A in one drawing from the population; x will take on only two possible values, 1 and 0, with probabilities p and q respectively. The population distribution is thus

$$f(x, p) = p^x(1 - p)^{1-x}.$$

We have

$$L = (\Sigma x_i) \log p + \Sigma(1 - x_i) \log (1 - p)$$

$$\left(\frac{\partial L}{\partial p}\right)_0 = \frac{m}{p_0} - \frac{n - m}{1 - p_0} = \frac{m - np_0}{p_0(1 - p_0)}$$

where m is the number of individuals belonging to A in the sample. Furthermore

$$A_0^2 = E_0\left[\left(\frac{x}{p_0} - \frac{1 - x}{1 - p_0}\right)^2\right] = [p_0(1 - p_0)]^{-1}.$$

95 per cent fiducial limits for p_0 are got by solving the following equation for p_0

$$\frac{m - np_0}{\sqrt{n}\sqrt{p_0(1 - p_0)}} = \pm 1.96.$$

It will be seen that situations, such as frequently occur in genetics, where p may be a function of some other parameter θ, say $p = u(\theta)$, can be handled by simply replacing p_0 by $u(\theta_0)$ and solving for θ_0.

Example 3. Let the form of the distribution function be $\theta e^{-\theta x}$, where $0 \leq x < \infty$. For a sample of individuals,

$$L = n \log \theta - \theta \Sigma x_i$$

$$\left(\frac{\partial L}{\partial \theta}\right)_0 = \frac{n}{\theta_0} - \Sigma x_i$$

$$A_0^2 = E_0\left[\left(\frac{1}{\theta_0} - x\right)^2\right] = \frac{1}{\theta_0^2}.$$

The 95 per cent fiducial limits $\underline{\theta}$ and $\bar{\theta}$ are given by solving the equations

$$\frac{\frac{n}{\theta_0} - \Sigma x_i}{\sqrt{n}\,(1/\theta_0)} = \pm 1.96$$

for θ_0. We get

$$\underline{\theta} = \frac{1 - 1.96/\sqrt{n}}{\bar{x}}; \qquad \bar{\theta} = \frac{1 + 1.96/\sqrt{n}}{\bar{x}}$$

where $\bar{x}$ is the mean of the sample.

PRINCETON UNIVERSITY.

REFERENCES

[1] R. A. FISHER, "The Concepts of Inverse Probability and Fiducial Probability Referring to Unknown Parameters," *Proc. Royal Society of London*, Series A, vol. 139 (1933), pp. 343–348.

[2] R. A. FISHER, "The Fiducial Argument in Statistical Inference," *Annals of Eugenics*, vol. 6 (1935), pp. 391-398.

[3] J. NEYMAN, "On the Two Different Aspects of the Representative method: the Method of Stratified Sampling and the Method of Purposive Selection," *Royal Statistical Society*, vol. 97, 1934, pp. 558–625.

[4] J. NEYMAN, "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability," *Phil. Trans. Roy. Soc. London*, Series A, vol. 236 (1937), pp. 333–380.

[5] H. CRAMÉR, *Random Variables and Probability Distributions*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 36, Cambridge University Press, 1937.