A NOTE ON CONFIDENCE INTERVALS AND INVERSE PROBABILITY

By Albert Wertheimer

The object of this note is to discuss a certain property of confidence intervals from the point of view of inverse probability. We shall not go into detailed applications, but merely into fundamental ideas, so we shall work with distribution functions that are continuous and satisfy conditions which are sufficient to insure the validity of the mathematical steps used.

A clear and concise statement of the subject is given in a paper by Neyman,[1] and we shall use it as the basis for our discussion. His presentation can be summarized as follows: Let x be a sample statistic having a distribution function

$$p(x, \theta) \qquad \begin{matrix} x_1 \leq x \leq x_2 \\ \theta_1 \leq \theta \leq \theta_2 \end{matrix}$$

where θ is a parameter of the population. Now define two monotonic functions

$$x = f(\theta); \qquad x = g(\theta) \qquad \begin{matrix} x_1 \leq x \leq x_2 \\ \theta_1 \leq \theta \leq \theta_2 \end{matrix}$$

such that $f(\theta) < g(\theta)$, and

$$\int_{f(\theta)}^{g(\theta)} p(x, \theta)\, dx = 1 - \epsilon, \qquad \text{for all } \theta. \tag{1}$$

Let the prior distribution function of θ be

$$\psi(\theta) \qquad \theta_1 \leq \theta \leq \theta_2 .$$

It then follows directly that the probability for any pair of values (x, θ) lying within the region enclosed by the curves is given by

$$\int_{\theta_1}^{\theta_2} \psi(\theta)\, d\theta \int_{f(\theta)}^{g(\theta)} p(x, \theta)\, dx = 1 - \epsilon. \tag{2}$$

regardless of the prior function $\psi(\theta)$. His conclusion then is this: Stating that

$$g^{-1}(x) \leq \theta \leq f^{-1}(x) \tag{3}$$

every time the observation gives us a value of x equal to that given in (3) we may in any one instance be wrong; this will happen only if the pair (x, θ) for this observation lies outside the region enclosed by the curves; but from (2) the probability for this to happen is ϵ. This statement is equivalent to saying that

[1] *Journal of the Royal Statistical Society*, Vol. 97, part IV, 1934; pp. 589–93.

if for every observed x we write the inequality (3), then for a large number of samples, the fraction $1 - \epsilon$ of the inequalities will be found correct.

We note here that this is true only if in the inequality (3) x is presumed to range over its entire interval of definition. But if for an observation $x = x'$, we mean to consider the corresponding inequality

$$g^{-1}(x') \leq \theta \leq f^{-1}(x') \tag{4}$$

as one member of the class of inequalities that could be written just for those samples that had $x = x'$, then we can not assert that the inequality (4) has a probability of $1 - \epsilon$ of being correct. In fact, any probability statement dealing with this class must involve the prior distribution function $\psi(\theta)$; and if it is not given, then we do not know in what percent of cases the restricted inequality (4) will be found correct.

Let us nevertheless approach the problem from the viewpoint of inverse probability. Having observed $x = x'$, the posterior probability of inequality (4) being correct is

$$\eta(x') = \frac{\int_{g^{-1}(x')}^{f^{-1}(x')} \psi(\theta)p(x', \theta)\, d\theta}{\int_{\theta_1}^{\theta_2} \psi(\theta)p(x', \theta)\, d\theta} \tag{5}$$

the numerator being the probability for the simultaneous occurrence of

$$x = x'; \qquad g^{-1}(x') \leq \theta \leq f^{-1}(x'),$$

and the denominator the probability[2] that $x = x'$, θ lying anywhere between θ_1 and θ_2.

As long as $\psi(\theta)$ is unknown $\eta(x')$ cannot be evaluated; however its average value $\bar{\eta}(x)$ with respect to x can be evaluated. By definition of an average,

$$\bar{\eta}(x) = \int_{x_1}^{x_2} \eta(x)\, dx \int_{\theta_1}^{\theta_2} \psi(\theta)p(x, \theta)\, d\theta \tag{6}$$

From (5) we have

$$\int_{g^{-1}(x)}^{f^{-1}(x)} \psi(\theta)p(x, \theta)\, d\theta = \eta(x) \int_{\theta_1}^{\theta_2} \psi(\theta)p(x, \theta)\, d\theta \tag{7}$$

Integrating both sides of (7) over the entire range of x we get

$$\int_{x_1}^{x_2} dx \int_{g^{-1}(x)}^{f^{-1}(x)} \psi(\theta)p(x, \theta)\, d\theta = \int_{x_1}^{x_2} \eta(x)\, dx \int_{\theta_1}^{\theta_2} \psi(\theta)p(x, \theta)\, d\theta$$

$$= \bar{\eta}(x)$$

[2] When we say probability that $x = x'$, we mean the probability that x will lie in the internal $x \pm \frac{1}{2}dx$ to within terms of order dx.

Interchanging the order of integration, as is permissible under the assumptions, we get

$$\bar{\eta}(x) = \int_{\theta_1}^{\theta_2} \psi(\theta)\, d\theta \int_{f(\theta)}^{g(\theta)} p(x, \theta)\, dx$$

But since

$$\int_{f(\theta)}^{g(\theta)} p(x, \theta)\, dx = 1 - \epsilon, \qquad \text{for all } \theta$$

we finally get

$$\bar{\eta}(x) = 1 - \epsilon$$

Thus when approached from the standpoint of inverse probability we see that the average value of the posterior probability of the inequality (4) is precisely the quantity $1 - \epsilon$ regardless of the prior distribution function $\psi(\theta)$.

In conclusion it is a pleasure to thank Dr. Deming for the criticisms and suggestions which he has made in connection with this note.

Navy Department,
Washington.