On Formulae for Confidence Points Based on Integrals of Weighted Likelihoods

By B. L. WELCH and H. W. PEERS

University of Leeds

[Received February 1963]

SUMMARY

Lindley's problem, relating to the conditions under which there is formal mathematical equivalence between Bayesian solutions and confidence theory solutions, is discussed. A result concerning location parameters is extended. The corresponding asymptotic theory is developed and a modification of the problem is also discussed.

1. INTRODUCTION

A SET of random variables S has a probability distribution depending on a single parameter θ, the probability element being denoted by $p(S, \theta)\, dS$. We are concerned with the problem of making inferences about θ.

If θ has some known prior probability distribution $q(\theta)\, d\theta$, then the posterior distribution of θ after a sample S has been observed is

$$\frac{p(S, \theta)\, q(\theta)\, d\theta}{\int p(S, t)\, q(t)\, dt}. \tag{1}$$

If, therefore, we define the function $g(S, \alpha)$ to satisfy the equation

$$\frac{\int^{g(S,\alpha)} p(S, t)\, q(t)\, dt}{\int p(S, t)\, q(t)\, dt} = \alpha, \tag{2}$$

we shall have

$$\Pr\{\theta < g(S, \alpha) \,|\, S\} = \alpha. \tag{3}$$

In other words, for given S, $g(S, \alpha)$ will be the α-probability point for θ.

If θ does not have a known prior distribution, so that an α-probability point cannot be calculated by this method, one often seeks instead a function $h(S, \alpha)$, say, with the property that

$$\Pr\{\theta < h(S, \alpha) \,|\, \theta\} = \alpha, \tag{4}$$

whatever θ and α. In (4), θ is fixed and S is random. The probability statement relates strictly, therefore, to the situation before a realized sample S has actually come to hand. The value of $h(S, \alpha)$ computed from an observed S is described as an α-confidence point to distinguish it from probability points calculated on the basis of prior distributions.

In those circumstances where a solution of (4) exists there is usually more than one solution and main interest centres on the existence of particular types of solution.

In this connection Lindley (1958) discussed conditions under which there exists a solution of (4) satisfying, at the same time, an equation of the form

$$\frac{\int^{h(S,\alpha)} p(S,t)\,w(t)\,dt}{\int p(S,t)\,w(t)\,dt} = \alpha, \tag{5}$$

where $w(t)$ is some positive weight function to be determined. In other words, he asked when there would be a solution $h(S,\alpha)$ of (4) belonging to the class of functions which can be constructed by substituting different prior probability functions $q(t)$ in (2). We shall make reference to some of Lindley's conclusions later. Our prime object in this paper is to develop certain results in the asymptotic theory which may be associated with his problem. We shall also discuss a modification of the problem which leads to somewhat different conclusions.

Before proceeding we should note that, although the weight function $w(t)$ appearing in (5) is to be positive, we are not wishing to give it a probability interpretation. A different notation has been introduced in (5) from that used in (2) simply in order to stress this. We do not at the moment look upon (5) as doing more than impose a *mathematical form* on the solution $h(S,\alpha)$ that we are seeking.

Let us write

$$p(S,t) = \exp\{L(S,t)\}, \quad w(t) = \exp\{\psi(t)\}, \tag{6}$$

$$I(S,\theta) = \int^{\theta} \exp\{L(S,t)+\psi(t)\}\,dt, \tag{7}$$

and let us define

$$r(S,\theta) = \frac{I(S,\theta)}{I(S,\infty)} = \frac{\int^{\theta} \exp\{L(S,t)+\psi(t)\}\,dt}{\int \exp\{L(S,t)+\psi(t)\}\,dt}. \tag{8}$$

For all S, $r(S,\theta)$ is a monotonic function of θ and therefore the inequality $\theta < h(S,\alpha)$ can be written $r(S,\theta) < r\{S, h(S,\alpha)\}$. In turn, from (5), this last inequality is equivalent to $r(S,\theta) < \alpha$. We can therefore express (4) in the form

$$\Pr\{r(S,\theta) < \alpha \mid \theta\} = \alpha. \tag{9}$$

Lindley's problem is therefore equivalent to the following: Can we find a function $\psi(t)$ which, through (8), will define a pivotal quantity $r(S,\theta)$ satisfying the probability relation (9), whatever θ and α? In other words, can we in this way define a pivotal quantity $r(S,\theta)$ which in repeated sampling will have a rectangular distribution over the range $(0, 1)$?

2. ASYMPTOTIC THEORY TO $O(n^{-\frac{1}{2}})$

In a wide class of situations $L(S,\theta)$ is "in the probability sense" $O(n)$, where n is some large quantity, and, for almost all samples, there is a single maximum-likelihood estimate T which will differ from θ by $O(n^{-\frac{1}{2}})$. All statements about orders of magnitude in the present context should be qualified by the phrase "in the probability sense" but we shall not continue to repeat it. We shall, moreover, assume that the $\psi(t)$ we are seeking is $O(n^0)$. In general, then, the contribution to the integral $I(S,\theta)$

arising from values of $|t-T|$ exceeding $O(n^{-\frac{1}{2}})$ will be negligible. Thus, expanding the integrand about T, we have, retaining terms to $O(n^{-\frac{1}{2}})$,

$$I(S,\theta)=\int^{\theta}\exp\{L+\tfrac{1}{2}(t-T)^2L''+\tfrac{1}{6}(t-T)^3L'''+\psi+(t-T)\psi'+O(n^{-1})\}\,dt. \tag{10}$$

If we write

$$u=(t-T)\left\{-\frac{\partial^2 L(S,T)}{\partial T^2}\right\}^{\frac{1}{2}}, \tag{11}$$

$$x=(\theta-T)\left\{-\frac{\partial^2 L(S,T)}{\partial T^2}\right\}^{\frac{1}{2}}, \tag{12}$$

and

$$v_j=n^{-\frac{1}{2}j}\left\{\frac{\partial^j L(S,T)}{\partial T^j}\right\}\quad (j=2,3,\ldots), \tag{13}$$

then (10) leads to the equation

$$(-nv_2)^{\frac{1}{2}}(2\pi)^{-\frac{1}{2}}I(S,\theta)\exp\{-L(S,T)-\psi(T)\}$$
$$=(2\pi)^{-\frac{1}{2}}\int^{x}\exp\{-\tfrac{1}{2}u^2+\tfrac{1}{6}u^3v_3(-v_2)^{-\frac{3}{2}}+un^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)+O(n^{-1})\}\,du \tag{14}$$

and hence, to $O(n^{-\frac{1}{2}})$,

$$r(S,\theta)=\frac{(2\pi)^{-\frac{1}{2}}\int^{x}\exp\{-\frac{1}{2}u^2+\frac{1}{6}u^3v_3(-v_2)^{-\frac{3}{2}}+un^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\}\,du}{(2\pi)^{-\frac{1}{2}}\int\exp\{-\frac{1}{2}u^2+\frac{1}{6}u^3v_3(-v_2)^{-\frac{3}{2}}+un^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\}\,du}. \tag{15}$$

The quantities u and x in (11) and (12) are $O(n^0)$, v_2 is $O(n^0)$, v_3 is $O(n^{-\frac{1}{2}})$ and generally v_j is $O(n^{-\frac{1}{2}(j-2)})$. For the remainder of this Section we shall continue to retain terms only to $O(n^{-\frac{1}{2}})$ but we shall return to the terms of $O(n^{-1})$ later. We have then, on expansion from (15),

$$r(S,\theta)=N(x)+Y(x)\{\tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(-x^2-2)-n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\}, \tag{16}$$

where

$$N(x)=(2\pi)^{-\frac{1}{2}}\int^{x}\exp\{-\tfrac{1}{2}u^2\}\,du,\quad Y(x)=(2\pi)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}x^2\} \tag{17}$$

are respectively the standard normal integral and the standard normal ordinate at x. Hence

$$r(S,\theta)=N\{x+\tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(-x^2-2)-n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\} \tag{18}$$
$$=N\{z(S,\theta)\}, \tag{19}$$

say.

The quantity $z(S,\theta)$ defined by (18) and (19) is a monotonic function of $r(S,\theta)$, being simply the abscissa of the standard normal distribution corresponding to a cumulative area $r(S,\theta)$. In discussing the distribution of $r(S,\theta)$ in repeated sampling for fixed θ it is more convenient to start with the simpler distribution of $z(S,\theta)$. For in order to make $\Pr\{r(S,\theta)<\alpha\,|\,\theta\}=\alpha$ we have simply to make $\Pr\{z(S,\theta)<\xi\,|\,\theta\}=\alpha$, where ξ is such that $N(\xi)=\alpha$. If we wish to make $\Pr\{r(S,\theta)<\alpha\,|\,\theta\}=\alpha$ for *all* α, or, equivalently, to make $r(S,\theta)$ have a $(0,1)$ rectangular distribution, then $z(S,\theta)$

will need to have a standard normal distribution. We shall proceed therefore to find the cumulant generating function of $z(S, \theta)$.

Apart from θ, z is a function of T and the derivatives of L evaluated at T. It will be convenient to express z rather in terms of the derivatives of L at θ, since often (but not invariably) the distributional properties of these are more directly known. The necessary relations are

$$\left.\begin{aligned} 0 = n^{-\frac{1}{2}}\frac{\partial L}{\partial T} &= n^{-\frac{1}{2}}\left\{\frac{\partial L}{\partial \theta}+(T-\theta)\frac{\partial^2 L}{\partial \theta^2}+\tfrac{1}{2}(T-\theta)^2\frac{\partial^3 L}{\partial \theta^3}\right\},\\ n^{-1}\frac{\partial^2 L}{\partial T^2} &= n^{-1}\left\{\frac{\partial^2 L}{\partial \theta^2}+(T-\theta)\frac{\partial^3 L}{\partial \theta^3}\right\},\\ n^{-\frac{3}{2}}\frac{\partial^3 L}{\partial T^3} &= n^{-\frac{3}{2}}\left\{\frac{\partial^3 L}{\partial \theta^3}\right\}.\end{aligned}\right\} \tag{20}$$

If we define

$$y_j = n^{-\frac{1}{2}j}\frac{\partial^j L(S,\theta)}{\partial \theta^j} \quad (j = 1, 2, \ldots), \tag{21}$$

then (20) yield, to $O(n^{-\frac{1}{2}})$,

$$x = -y_1(-y_2)^{-\frac{1}{2}}, \quad v_2 = y_2 - y_1 y_3 y_2^{-1}, \quad v_3 = y_3. \tag{22}$$

Note that y_1 and y_2 are $O(n^0)$ and the subsequent y_j are $O(n^{-\frac{1}{2}(j-2)})$. Hence we have

$$\begin{aligned} z(S,\theta) &= x+\tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(-x^2-2)-n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\\ &= -y_1(-y_2)^{-\frac{1}{2}}+\tfrac{1}{6}y_3(-y_2)^{-\frac{3}{2}}(y_1^2 y_2^{-1}-2)-n^{-\frac{1}{2}}(-y_2)^{-\frac{1}{2}}\psi'(\theta). \end{aligned} \tag{23}$$

Now the commonest situation with which the present asymptotic theory is concerned is that arising when S is a sample of n independent values from the same population, so that $L = \Sigma l$ where l relates to the log likelihood of a single individual. We denote the joint cumulants of $\partial^j l/\partial\theta^j$ $(j = 1, 2, \ldots)$ by $\kappa_{rst\ldots}$. In this notation $E(\partial^2 l/\partial\theta^2) = \kappa_{01}$, and $\text{var}(\partial l/\partial\theta) = \kappa_{20}$. However, we shorten κ_{20} to κ_2 since this is unlikely to cause any confusion. In general the κ's are functions of θ and occasionally we write, for instance, $\kappa_2(\theta)$ where it appears necessary to stress the fact. We shall also wish to refer to a function $\kappa_2(t)$ evaluated at points other than θ, as, for example, in (33) below. The joint cumulant generating function of the y's is then

$$\log E\{\exp(\Sigma t_j y_j)\} = \frac{\kappa_2 t_1^2}{2!}+\kappa_{01}t_2+n^{-\frac{1}{2}}\left(\frac{\kappa_3 t_1^3}{3!}+\kappa_{11}t_1 t_2+\kappa_{001}t_3\right)+\ldots. \tag{24}$$

The present theory applies also to a somewhat wider class of situations where the joint cumulants of the y's are given by (24) although the κ's may depend on n. It is essential, however, that the κ's are $O(n^0)$.

It will be recalled that there are in general relations between the κ's such as the well-known relation

$$\kappa_2+\kappa_{01} = 0, \tag{25}$$

and the further relations, due to Bartlett (1953),

$$\left.\begin{aligned} \kappa_3+3\kappa_{11}+\kappa_{001} &= 0,\\ \frac{\partial \kappa_2}{\partial \theta} &= \kappa_3+2\kappa_{11}. \end{aligned}\right\} \tag{26}$$

From (24) the joint moment generating function of the y's about their means is

$$M(t_1, t_2, \ldots) = \exp\left\{\frac{\kappa_2 t_1^2}{2!} + n^{-\frac{1}{2}}\left(\frac{\kappa_3 t_1^3}{3!} + \kappa_{11} t_1 t_2\right) + \ldots\right\}. \tag{27}$$

The moment generating function of $z(S, \theta)$ is then given symbolically (compare Welch, 1947) by

$$E\{\exp(tz)\} = M(D_1, D_2, \ldots) \exp [t\{-w_1(-w_2)^{-\frac{1}{2}} + \tfrac{1}{6} w_3(-w_2)^{-\frac{3}{2}}(w_1^2 w_2^{-1} - 2) - n^{-\frac{1}{2}}(-w_2)^{-\frac{1}{2}} \psi'(\theta)\}], \tag{28}$$

where the D's denote derivatives with respect to the corresponding w's and, after differentiation, we must set $w_1 = 0, w_2 = \kappa_{01}, w_3 = n^{-\frac{1}{2}} \kappa_{001}$. The final result of carrying out this operation and of making use of the relations (25) and (26) is, to $O(n^{-\frac{1}{2}})$,

$$E\{\exp(tz)\} = \exp\left[tn^{-\frac{1}{2}}\left\{\tfrac{1}{2}\kappa_2^{-\frac{3}{2}} \frac{\partial \kappa_2}{\partial \theta} - \kappa_2^{-\frac{1}{2}} \psi'(\theta)\right\} + \tfrac{1}{2} t^2\right]. \tag{29}$$

To this order, then, z will have the standard normal distribution if $\psi(\theta)$ is chosen to satisfy

$$\psi'(\theta) = \tfrac{1}{2}\kappa_2^{-1} \frac{\partial \kappa_2}{\partial \theta} = \frac{1}{2} \frac{\partial \log \kappa_2}{\partial \theta}. \tag{30}$$

Hence

$$\psi(\theta) = \tfrac{1}{2} \log \kappa_2 + \text{constant} \tag{31}$$

and

$$w(\theta) = \text{constant} \times \kappa_2^{\frac{1}{2}}. \tag{32}$$

To this order the quantity

$$r(S, \theta) = \frac{\int^{\theta} p(S, t)\{\kappa_2(t)\}^{\frac{1}{2}}\, dt}{\int p(S, t)\{\kappa_2(t)\}^{\frac{1}{2}}\, dt} \tag{33}$$

will therefore have a (0, 1) rectangular distribution in the *repeated sampling* sense. The corresponding asymptotic formula for the quantity z having a standard normal distribution is, from (23),

$$z(S, \theta) = x - \tfrac{1}{6} v_3(-v_2)^{-\frac{3}{2}}(x^2 + 2) - \tfrac{1}{2} n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}} \frac{\partial \log \kappa_2(T)}{\partial T}. \tag{34}$$

The confidence point $h(S, \alpha)$ is that value of θ which makes $z = \xi$. But in the right-hand side of (34) θ is contained only in $x = n^{\frac{1}{2}}(\theta - T)(-v_2)^{\frac{1}{2}}$. This suggests that we first solve (34) for x giving

$$x = n^{-\frac{1}{2}}(\theta - T)(-v_2)^{\frac{1}{2}} = z + \tfrac{1}{6} v_3(-v_2)^{-\frac{3}{2}}(z^2 + 2) + \tfrac{1}{2} n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}} \frac{\partial \log \kappa_2(T)}{\partial T} + O(n^{-1}). \tag{35}$$

Hence

$$n^{\frac{1}{2}}\{h(S, \alpha) - T\}(-v_2)^{\frac{1}{2}} = \xi + \tfrac{1}{6} v_3(-v_2)^{-\frac{3}{2}}(\xi^2 + 2) + \tfrac{1}{2} n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}} \frac{\partial \log \kappa_2(T)}{\partial T} + O(n^{-1}) \tag{36}$$

and therefore, from (13),

$$h(S,\alpha) = T+\xi\left(-\frac{\partial^2 L}{\partial T^2}\right)^{-\frac{1}{2}}+\tfrac{1}{6}(\xi^2+2)\left(\frac{\partial^3 L}{\partial T^3}\right)\left(-\frac{\partial^2 L}{\partial T^2}\right)^{-2}$$
$$+\frac{1}{2}\frac{\partial \log \kappa_2(T)}{\partial T}\left(-\frac{\partial^2 L}{\partial T^2}\right)^{-1}+O(n^{-\frac{3}{2}}). \quad (37)$$

Alternatively, theoretically at least, we need not proceed from (34) to (37) by this manipulation of series. From (33) directly, the equation $r(S,\theta)=\alpha$ can, subject only to continuity, be solved numerically to give a unique $\theta = h(S,\alpha)$. Indeed, in the discussion in Section 4 below, it is strictly this value of $h(S,\alpha)$ that we shall be referring to rather than the truncated series in (37).

3. ASYMPTOTIC THEORY TO $O(n^{-1})$

If we carry the analysis of the preceding Section to a further order in $n^{-\frac{1}{2}}$ we find, instead of (16),

$$\begin{aligned} r(S,\theta) = N(x)+Y(x)\,[&-\tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(x^2+2)-n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\\ &-\tfrac{1}{24}v_4(-v_2)^{-2}(x^2+3x)-\tfrac{1}{2}xn^{-1}(-v_2)^{-1}\psi''(T)\\ &-\tfrac{1}{72}v_3^2(-v_2)^{-3}(x^5+5x^3+15x)-\tfrac{1}{2}xn^{-1}(-v_2)^{-1}\{\psi'(T)\}^2\\ &-\tfrac{1}{6}n^{-\frac{1}{2}}v_3(-v_2)^{-2}(x^3+3x)\psi'(T)]. \end{aligned} \quad (38)$$

Instead of (18) we have, for the standard normal abscissa corresponding to $r(S,\theta)$,

$$\begin{aligned} z(S,\theta) = x&-\tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(x^2+2)-n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\psi'(T)\\ &-\tfrac{1}{24}v_4(-v_2)^{-2}(x^3+3x)-\tfrac{1}{72}v_3^2(-v_2)^{-3}(x^3+11x)\\ &-\tfrac{1}{6}n^{-\frac{1}{2}}xv_3(-v_2)^{-2}\psi'(T)-\tfrac{1}{2}n^{-1}x(-v_2)^{-1}\psi''(T). \end{aligned} \quad (39)$$

Instead of (22) we have the formulae

$$\begin{aligned} x &= -y_1(-y_2)^{-\frac{1}{2}}+\tfrac{1}{8}y_1^3y_3^2(-y_2)^{-\frac{9}{2}}+\tfrac{1}{12}y_1^3y_4(-y_2)^{-\frac{7}{2}},\\ v_2 &= y_2-y_1y_3y_2^{-1}-\tfrac{1}{2}y_1^2y_3^2y_2^{-3}+\tfrac{1}{2}y_1^2y_4y_2^{-2},\\ v_3 &= y_3-y_1y_4y_2^{-1}, \quad v_4 = y_4. \end{aligned} \quad (40)$$

Substituting these in (39) we obtain the alternative expression

$$\begin{aligned} z(S,\theta) = &-y_1(-y_2)^{-\frac{1}{2}}-\tfrac{1}{3}y_3(-y_2)^{-\frac{3}{2}}-\tfrac{1}{6}y_1^2y_3(-y_2)^{-\frac{5}{2}}\\ &-n^{-\frac{1}{2}}(-y_2)^{-\frac{1}{2}}\psi'(\theta)-\tfrac{1}{3}n^{-\frac{1}{2}}y_1y_3(-y_2)^{-\frac{5}{2}}\psi'(\theta)\\ &-\tfrac{1}{2}n^{-1}y_1(-y_2)^{-\frac{3}{2}}\psi''(\theta)-\tfrac{5}{24}y_1^3y_4(-y_2)^{-\frac{7}{2}}-\tfrac{5}{24}y_1y_4(-y_2)^{-\frac{5}{2}}\\ &-\tfrac{25}{72}y_1y_3^2(-y_2)^{-\frac{7}{2}}-\tfrac{17}{72}y_1^3y_3^2(-y_2)^{-\frac{9}{2}}. \end{aligned} \quad (41)$$

Instead of (24) we have, for the joint cumulant generating function of the y's,

$$\log E\{\exp(\Sigma t_j y_j)\} = \frac{\kappa_2 t_1^2}{2!}+\kappa_{01}t_2+n^{-\frac{1}{2}}\left(\frac{\kappa_3 t_1^3}{3!}+\kappa_{11}t_1t_2+\kappa_{001}t_3\right)$$
$$+n^{-1}\left(\frac{\kappa_4 t_1^4}{4!}+\frac{\kappa_{02}t_2^2}{2!}+\kappa_{101}t_1t_3+\frac{\kappa_{21}t_1^2t_2}{2!}+\kappa_{0001}t_4\right). \quad (42)$$

Between the κ's defined in (42) there exist the following relations additional to those due to Bartlett (1953).

$$\left.\begin{aligned} &3\kappa_{02}+\kappa_4+6\kappa_{21}+4\kappa_{101}+\kappa_{0001}=0,\\ &\kappa_3'=\frac{\partial\kappa_3}{\partial\theta}=\kappa_4+3\kappa_{21},\\ &\kappa_2''=\frac{\partial^2\kappa_2}{\partial\theta^2}=2\kappa_{02}+5\kappa_{21}+2\kappa_{101}+\kappa_4.\end{aligned}\right\} \tag{43}$$

By carrying (28) to $O(n^{-1})$ and making use of these extra relations we obtain finally, instead of (29),

$$\begin{aligned} E\{\exp(tz)\} = \exp[&tn^{-\frac{1}{2}}\{\tfrac{1}{2}\kappa_2^{-\frac{3}{2}}\kappa_2'-\kappa_2^{-\frac{1}{2}}\psi'(\theta)\}+\tfrac{1}{2}t^2\\ &+\tfrac{1}{2}t^2n^{-1}\{\tfrac{1}{3}\kappa_2^{-3}\kappa_3\kappa_2'+\tfrac{3}{4}\kappa_2^{-3}(\kappa_2')^2-\tfrac{1}{6}\kappa_2^{-2}\kappa_3'\\ &-\tfrac{1}{2}\kappa_2^{-2}\kappa_2''+\kappa_2^{-1}\psi''(\theta)-\tfrac{1}{6}\kappa_2^{-2}\psi'(\theta)(3\kappa_2'+\kappa_3)\}].\end{aligned} \tag{44}$$

To order n^{-1} the third and fourth cumulants of z both vanish but it is not possible by suitable choice of $\psi(\theta)$ to make the mean zero and the variance unity simultaneously. To this order, therefore, our problem has no solution in general. If, as before, we make the mean zero by writing

$$\psi'(\theta)=\tfrac{1}{2}\kappa_2^{-1}\kappa_2'=\frac{1}{2}\frac{\partial\log\kappa_2}{\partial\theta}, \tag{45}$$

then (44) gives

$$E\{\exp(tz)\}=\exp\left\{\tfrac{1}{2}t^2-\tfrac{1}{12}t^2n^{-1}\kappa_2^{-\frac{1}{2}}\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}})\right\}, \tag{46}$$

and then z will have the standard normal distribution to order n^{-1}, only if the skewness $\kappa_3\kappa^{-\frac{3}{2}}$ of the distribution of $\partial L/\partial\theta$ is independent of θ.

4. DISCUSSION

Our motive in discussing the present problem is probably different from that of some other writers on the topic. We certainly do not wish to imply that the only allowable confidence theory solutions are ones possessing some formal mathematical equivalence with Bayesian solutions. But some of the properties of the quantity $r(S,\theta)$ defined in (33) seem to be worth summarizing quite apart from this equivalence. For instance, we may note:

(i) $r(S,\theta)$ is a rectangularly distributed pivotal quantity at least to $O(n^{-\frac{1}{2}})$.

(ii) $r(S,\theta)$ is a monotonic function of θ and hence $r(S,\theta)=\alpha$ may be solved uniquely to give $\theta=h(S,\alpha)$. Samples for which $r(S,\theta)<\alpha$ correspond with those for which $\theta<h(S,\alpha)$. Several writers have drawn attention to the desirability of this property; see, for example, Kendall and Stuart (1961, p. 109).

(iii) Also from the form of $r(S,\theta)$ it follows that $h(S,\alpha_1)<h(S,\alpha_2)$ for $\alpha_1<\alpha_2$. This is a necessary property if we wish to talk of a confidence "distribution" of θ.

(iv) If we approach confidence theory through the initial consideration of a test of some statistical hypothesis (as is commonly done) a weighted likelihood ratio criterion for testing $\theta < \theta_0$ against $\theta > \theta_0$ might be defined by

$$u(S, \theta_0) = \frac{\int_{-\infty}^{\theta_0} p(S, t)\{\kappa_2(t)\}^{\frac{1}{2}}\, dt}{\int_{\theta_0}^{\infty} p(S, t)\{\kappa_2(t)\}^{\frac{1}{2}}\, dt}. \tag{47}$$

Since $r(S, \theta)$ is a monotonic function of $u(S, \theta)$ a link-up with the usual confidence theory is therefore possible.

We should, of course, have liked $r(S, \theta)$ to be *exactly* pivotal in the repeated sampling sense but in general, as we have seen, we cannot secure this within the present limitation of form. There is, however, one particular case where further progress can be made and to this we shall now turn.

5. LOCATION PARAMETERS

Lindley (1958) discussed the present problem firstly in the situation where S consisted of a single observation x following a probability distribution $p(x, \theta)\, dx$, say. He showed that equivalence of mathematical form between a confidence theory distribution and some posterior Bayesian distribution for θ could only hold if x could be monotonically transformed to a new variable x', whose distribution depended on a location parameter θ' which was a monotonic function of θ. It would be necessary to have $p(x, \theta)\, dx = f(x' - \theta')\, dx'$, say. He observed also that, even if S consists of more than one observation, then equivalence of the required nature would still hold if a sufficient statistic x' existed with distribution of the form $f(x' - \theta')\, dx'$.

We can go even further than this as the following remarks will show. Suppose that S consists of observations x_i $(i = 1, 2, \ldots, m)$ and that θ is a parameter of location, so that

$$p(S, \theta)\, dS = f(x_1 - \theta, x_2 - \theta, \ldots)\, dx_1\, dx_2 \ldots dx_m \tag{48}$$

but that there is no sufficient statistic for θ. We can transform the m x's to new variables of which the first $m-1$ are configurational, such as the successive differences $x_1 - x_2, x_2 - x_3$, etc., and the remaining variable a quantity such as $\bar{x}$, which has a distribution depending on θ as a location parameter. Denoting the configurational variables collectively by C, we can write

$$\begin{aligned} p(S, \theta)\, dS &\equiv p(C)\, dC\, p(\bar{x} \mid C, \theta)\, d\bar{x} \\ &= p(C)\, dC\, f_1(\bar{x} - \theta, C)\, d\bar{x}, \end{aligned} \tag{49}$$

say. If then we take a constant for the weight function $w(t)$ we obtain from (8)

$$r(S, \theta) = \frac{\int^{\theta} p(C) f_1(\bar{x} - t, C)\, dt}{\int p(C) f_1(\bar{x} - t, C)\, dt} = \frac{\int^{\theta} f_1(\bar{x} - t, C)\, dt}{\int f_1(\bar{x} - t, C)\, dt}. \tag{50}$$

The integrals may be transformed by writing $\bar{x} - t = u - \theta$ to give

$$r(S, \theta) = \frac{\int_{\bar{x}}^{\infty} f_1(u - \theta, C)\, du}{\int f_1(u - \theta, C)\, du}. \tag{51}$$

But since $f_1(\bar{x}-\theta, C)\,d\bar{x}$ is the probability distribution of $\bar{x}$ for given C, the denominator of (51) is unity and the numerator, considered as a function of $\bar{x}$, is simply an example of the probability integral transformation. For given C, $r(S,\theta)$ has therefore a rectangular distribution over the range $(0,1)$. In repeated sampling with C not fixed the distribution, of course, remains the same. The problem posed by Lindley is therefore also capable of solution in this situation (and also more generally if some monotonic function θ' of θ is a location parameter). It appears that in these cases the restriction of $h(S,\alpha)$ to the class of functions satisfying (5) is practically equivalent in its effect to demanding that

$$\Pr\{\theta < h(S,\alpha) \mid C, \theta\} = \alpha, \tag{52}$$

that is, to imposing a probability requirement conditional on fixed C, as was advocated by Fisher (1934).

6. Weight Functions Dependent on the Probability Level

Reverting again to the general case where a pivotal quantity of the form (8) does not exist exactly, there is a modification of Lindley's problem which may still be worth· considering in relation to the points (i)–(iv) of Section 4. Suppose that, instead of $r(S,\theta)$, we define a quantity $r(S,\theta,\alpha)$ by

$$r(S,\theta,\alpha) = \frac{\int^{\theta} p(S,t)\,w(t,\alpha)\,dt}{\int p(S,t)\,w(t,\alpha)\,dt} = \frac{\int^{\theta} \exp\{L(S,t)+\psi(t,\alpha)\}\,dt}{\int \exp\{L(S,t)+\psi(t,\alpha)\}\,dt}, \tag{53}$$

where the weight function is now dependent on α. We stress again that our weights are not to be regarded in any sense as prior probabilities. We may then consider the possibility of finding $w(t,\alpha)$ such that

$$\Pr\{r(S,\theta,\alpha) < \alpha \mid \theta\} = \alpha \tag{54}$$

for a *particular* α, or, equivalently,

$$\Pr\{z(S,\theta,\alpha) < \xi \mid \theta\} = \alpha, \tag{55}$$

where $z(S,\theta,\alpha)$ and ξ are the respective standard normal abscissae corresponding to $r(S,\theta,\alpha)$ and α.

Equation (54) defines an integral equation with arbitrary function $w(t,\alpha)$ containing one current variable t. The equation has to be satisfied for all θ and there will usually be some prospect of a solution. Where asymptotic theory is applicable we can at least find a *series* solution, although the exact status of this solution may not be easy to determine. The issue here is indeed similar to that arising in the two means problem (Welch, 1947; Trickett and Welch, 1954) where the relation between an asymptotic series solution and a direct numerical attack on an integral equation has been discussed in some detail.

The asymptotic theory to $O(n^{-1})$ may be deduced from the expression for $E[\exp\{tz(S,\theta,\alpha)\}]$ on the right-hand side of (44) with $\psi'(\theta)$ replaced by $\psi'(\theta,\alpha)$. If we write

$$\psi'(\theta,\alpha) = \tfrac{1}{2}\kappa_2^{-1}\kappa_2' + n^{-\frac{1}{2}} j(\theta,\alpha), \tag{56}$$

where $n^{-\frac{1}{2}}j(\theta,\alpha)$ is a corrective term, we find

$$E[\exp\{tz(S,\theta,\alpha)\}] = \exp\left\{-tn^{-1}\kappa_2^{-\frac{1}{2}}j(\theta,\alpha)+\tfrac{1}{2}t^2-\tfrac{1}{12}t^2 n^{-1}\kappa_2^{-\frac{1}{2}}\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}})\right\}, \tag{57}$$

so that the mean and variance of $z(S,\theta,\alpha)$ are respectively

$$-n^{-1}\kappa_2^{-\frac{1}{2}}j(\theta,\alpha),\quad 1-\tfrac{1}{6}n^{-1}\kappa_2^{-\frac{1}{2}}\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}}), \tag{58}$$

the higher cumulants vanishing to $O(n^{-1})$.

To satisfy (55) we must therefore have

$$\xi = -n^{-1}\kappa_2^{-\frac{1}{2}}j(\theta,\alpha)+\xi\left\{1-\tfrac{1}{12}n^{-1}\kappa_2^{-\frac{1}{2}}\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}})\right\} \tag{59}$$

and so

$$j(\theta,\alpha) = -\tfrac{1}{12}\xi\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}}). \tag{60}$$

Thus

$$\psi'(\theta,\alpha) = \tfrac{1}{2}\kappa_2^{-1}\kappa_2'-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\frac{\partial}{\partial\theta}(\kappa_3\kappa_2^{-\frac{3}{2}}), \tag{61}$$

so that

$$\psi(\theta,\alpha) = \tfrac{1}{2}\log\kappa_2-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\kappa_3\kappa_2^{-\frac{3}{2}}+\text{constant}, \tag{62}$$

and

$$w(\theta,\alpha) = \text{constant}\times\kappa_2^{\frac{1}{2}}\exp(-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\kappa_3\kappa_2^{-\frac{3}{2}}). \tag{63}$$

To $O(n^{-1})$, therefore, the quantity

$$r(S,\theta,\alpha) = \frac{\int^{\theta} p(S,t)\{\kappa_2(t)\}^{\frac{1}{2}}\exp[-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\kappa_3(t)\{\kappa_2(t)\}^{-\frac{3}{2}}]\,dt}{\int p(S,t)\{\kappa_2(t)\}^{\frac{1}{2}}\exp[-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\kappa_3(t)\{\kappa_2(t)\}^{-\frac{3}{2}}]\,dt} \tag{64}$$

is such that $\Pr\{r(S,\theta,\alpha)<\alpha\,|\,\theta\} = \alpha$.

We may now note the following points.

(i) Although $r(S,\theta,\alpha)$ is not a pivotal quantity in the usual sense of having a repeated sampling distribution independent of θ, it is approximately pivotal in the more restricted sense that $\Pr\{r(S,\theta,\alpha)<\alpha\,|\,\theta\} = \alpha$ at least to $O(n^{-1})$ for *particular* α. Moreover, the iterative process leading to (64) can be carried to higher orders to define an $r(S,\theta,\alpha)$ which will be pivotal in this restricted sense to any prescribed order of $n^{-\frac{1}{2}}$.

(ii) From (64) we can solve the equation $r(S,\theta,\alpha) = \alpha$ numerically to give a unique value $\theta = h(S,\alpha)$. Samples for which $r(S,\theta,\alpha)<\alpha$ will still correspond to those for which $\theta<h(S,\alpha)$, so that one of our important monotonicity requirements is still met.

(iii) In general, however, we cannot still say that $h(S,\alpha_1)<h(S,\alpha_2)$ for all $\alpha_1<\alpha_2$, the reason being that $r(S,\theta,\alpha)$ depends on α. Where a solution of the integral (54) exists and can be obtained without recourse to asymptotic series, the corresponding $h(S,\alpha)$ may, of course, prove to be monotonic in α but it is not obvious that this is so. Whilst it may be possible to calculate valid confidence *points* for different α, it will remain for investigation in any particular instance whether these lead on naturally to the idea of a confidence *distribution*.

In (ii) and (iii) we have assumed that $h(S,\alpha)$ is obtained by solving exactly for θ the equation $r(S,\theta,\alpha)=\alpha$, where $r(S,\theta,\alpha)$ is given by the right-hand side of (64) or some similar expression. In the asymptotic theory we can, of course, go to the further stage of expressing $h(S,\alpha)$ in a series as we did in (37) above.

We start with (39), replacing $\psi'(T)$ now by

$$\psi'(T,\alpha)=\frac{1}{2}\frac{\partial\log\kappa_2(T)}{\partial T}-\tfrac{1}{12}n^{-\frac{1}{2}}\,\xi\frac{\partial}{\partial T}\{\kappa_3(T)\,\kappa_2^{-\frac{3}{2}}(T)\}. \tag{65}$$

Solving (39) for x we then obtain, as in (35),

$$\begin{aligned}x = z &+ \tfrac{1}{6}v_3(-v_2)^{-\frac{3}{2}}(z^2+2)+\tfrac{1}{2}n^{-\frac{1}{2}}(-v_2)^{-\frac{1}{2}}\frac{\partial\log\kappa_2(T)}{\partial T}\\ &-\tfrac{1}{12}n^{-1}\,\xi(-v_2)^{-\frac{1}{2}}\frac{\partial}{\partial T}\{\kappa_3(T)\,\kappa_2^{-\frac{3}{2}}(T)\}\\ &+\tfrac{1}{72}v_3^2(-v_2)^{-3}(5z^3+19z)+\tfrac{1}{4}n^{-\frac{1}{2}}zv_3(-v_2)^{-2}\frac{\partial\log\kappa_2(T)}{\partial T}\\ &+\tfrac{1}{24}v_4(-v_2)^{-2}(z^3+3z)+\tfrac{1}{4}n^{-1}z(-v_2)^{-1}\frac{\partial^2\log\kappa_2(T)}{\partial T^2}\\ &+O(n^{-\frac{3}{2}}),\end{aligned} \tag{66}$$

and therefore

$$\begin{aligned}h(S,\alpha) = T &+ \xi\left(-\frac{\partial^2L}{\partial T^2}\right)^{-\frac{1}{2}}+\tfrac{1}{6}(\xi^2+2)\left(\frac{\partial^3L}{\partial T^3}\right)\left(-\frac{\partial^2L}{\partial T^2}\right)^{-2}\\ &+\frac{1}{2}\left(-\frac{\partial^2L}{\partial T^2}\right)^{-1}\frac{\partial\log\kappa_2(T)}{\partial T}-\tfrac{1}{12}n^{-\frac{1}{2}}\xi\left(-\frac{\partial^2L}{\partial T^2}\right)^{-1}\frac{\partial}{\partial T}\left(\kappa_3(T)\,\kappa_2^{-\frac{3}{2}}(T)\right)\\ &+\tfrac{1}{72}(5\xi^3+19\xi)\left(\frac{\partial^3L}{\partial T^3}\right)^2\left(-\frac{\partial^2L}{\partial T^2}\right)^{-\frac{7}{2}}+\tfrac{1}{4}\xi\left(\frac{\partial^3L}{\partial T^3}\right)\left(-\frac{\partial^2L}{\partial T^2}\right)^{-\frac{5}{2}}\frac{\partial\log\kappa_2(T)}{\partial T}\\ &+\tfrac{1}{4}\xi\left(-\frac{\partial^2L}{\partial T^2}\right)^{-\frac{3}{2}}\frac{\partial^2\log\kappa_2(T)}{\partial T^2}+\tfrac{1}{24}(\xi^3+3\xi)\left(\frac{\partial^4L}{\partial T^4}\right)\left(-\frac{\partial^2L}{\partial T^2}\right)^{-\frac{5}{2}}\\ &+O(n^{-2}).\end{aligned} \tag{67}$$

7. CONCLUSION

As has been stated in Section 4, our only purpose here has been to explore the possibility that solutions which mathematically, although not logically, are of Bayesian form may be adapted to provide confidence points in the sense of J. Neyman, E. S. Pearson and other writers on this subject. We have not discussed whether such solutions are "best" in any sense and indeed do not know of any generally accepted definition of "best" in this connection. Much used alternative methods are those based solely on the distribution of the maximum-likelihood estimate T, or solely on the distribution of $\partial L/\partial\theta$. The asymptotic theory associated with the latter method has been discussed by Bartlett (1953), and a way in which corrections of different orders of $n^{-\frac{1}{2}}$ may be supplied for the theory of maximum-likelihood estimates was outlined by Welch (1939). To order n^0 in probability, of course, all these methods are equivalent and the question really is at what order of $n^{-\frac{1}{2}}$ they diverge.

We propose elsewhere to supply some comparisons of the results of the present paper with those following from other methods of computing confidence points and also to extend some of our present work to the situation where there are several population parameters.

REFERENCES

BARTLETT, M. S. (1953), "Approximate confidence intervals", *Biometrika*, **40**, 12–19.

FISHER, R. A. (1934), "Two new properties of mathematical likelihood", *Proc. roy. Soc.* A, **144,** 285–307.

KENDALL, M. G. and STUART, A. (1961), *The Advanced Theory of Statistics*, **2**. London: Griffin.

LINDLEY, D. V. (1958), "Fiducial distributions and Bayes' theorem", *J. R. statist. Soc.* B, **20,** 102–107.

TRICKETT, W. H. and WELCH, B. L. (1954), "On the comparison of two means: further discussion of iterative methods for calculating tables", *Biometrika*, **41**, 361–374.

WELCH, B. L. (1939), "On the distribution of maximum likelihood estimates", *Biometrika*, **31,** 187–190.

—— (1947), "The generalization of 'Student's' problem when several different population variances are involved", *Biometrika*, **34**, 28–35.