The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 833–836

NOTE ON BAYES-FIDUCIAL INTERVALS FOR PROBLEMS OF LOCATION AND SCALE

By Donald A. Pierce and David A. Bogdanoff[1]

Oregon State University

1. Introduction. It is a well-known result of Welch and Peers [4] that for location parameter problems Bayesian α-level confidence intervals based on Lebesgue measure priors have probability α of covering the true parameter. Many scale parameter problems can be easily transformed to this case. Furthermore it follows as special cases of the general theory given by Hora and Buehler [3] that this result holds for any scale parameter problem with Lebesgue measure prior on the logarithm of the parameter, and that for problems of both location and scale the result holds for confidence intervals on certain scalar functions of the two parameters, which they call invariantly estimable functions. Bayesian methods in all of these problems lead to the same result as fiducial methods [3] and are essentially frequentist methods if one adopts the principle of conditioning on ancillary statistics (see, for example, Fraser [2]).

This note extends the work of Hora and Buehler [3] by exhibiting an important class of functions of the location and scale parameters which are not invariantly estimable but for which Bayesian intervals have the nominal probability of coverage. In addition an example is given to establish that there do exist functions for which this property does not hold (Fraser [2], Hora and Buehler [3], point out that the Behrens–Fisher problem provides such an example for the two sample problem). Finally an argument is presented which is useful in comparing the Bayes-fiducial intervals with those obtained otherwise.

2. A special non-invariantly estimable function. Let $\tilde{x}$ represent the observations in a random sample from a population with unknown and unrestricted location and scale parameters θ and $\sigma > 0$. For any specified scalar function $\psi(\theta, \sigma)$ and any realization x of $\tilde{x}$ consider the α-level upper Bayesian confidence limit $\bar{\psi}(x, \alpha)$ such that

$$P_{\tilde{\theta},\tilde{\sigma}\mid x}[\bar{\psi}(x, \alpha) \geqq \psi(\tilde{\theta}, \tilde{\sigma})] = \alpha, \tag{2.1}$$

using the prior $\pi(\theta, \sigma) \propto 1/\sigma$.

Hora and Buehler classify $\psi(\theta, \sigma)$ as invariantly estimable if $\psi(\theta_1, \sigma_1) = \psi(\theta_2, \sigma_2)$ implies that $\psi(a\theta_1 + b, a\sigma_1) = \psi(a\theta_2 + b, a\sigma_2)$ for all numbers $a > 0$ and b. Such a function is $\psi(\theta, \sigma) = h(c\theta + d\sigma)$, where c and d are arbitrary numbers and h is any 1–1 function. They then prove that if (i) ψ is invariantly

Received June 23, 1970.

[1] Presently at St. Louis University. This investigation supported by PHS Training Grant GM–1736, National Institute of General Medical Sciences.

estimable, (ii) (2.1) has, for each x, a unique solution of $\bar{\psi}$, (iii) for all numbers $a > 0$ and b, $\psi(a\theta+b, a\sigma)$ increases as $\psi(\theta, \sigma)$ increases; then

$$P_{\tilde{x}\mid\theta,\sigma}[\bar{\psi}(\tilde{x}, \alpha) \geqq \psi(\theta, \sigma)] = \alpha \tag{2.2}$$

for all (θ, σ).

The primary purpose of this note is to point out that for any number t, (2.2) is satisfied with $\psi(\theta, \sigma) = (t-\theta)/\sigma$, provided that (ii) above holds. (This function is not invariantly estimable, but does satisfy a similar criterion; if $(t-\theta_1)/\sigma_1 = (t-\theta_2)/\sigma_2$ then for any $a > 0$ and b,

$$[(at+b)-(a\theta_1+b)]/a\sigma_1 = [(at+b)-(a\theta_2+b)]/a\sigma_2.$$

This form of invariance could probably be used to generalize Hora and Buehler's theorem in their more abstract setting.) This enables, among other things, determination of confidence intervals for the cdf of $\tilde{x}$ at any fixed point, the application of which to reliability problems having led to this extension.

A proof of this can be obtained through the often useful expression

$$P_{\tilde{x}\mid\theta,\sigma}[\bar{\psi}(\tilde{x}, \alpha) \geqq \psi(\theta, \sigma)] = P_{\tilde{x}\mid\theta,\sigma}\{P_{\tilde{\theta},\tilde{\sigma}\mid x}[\psi(\tilde{\theta}, \tilde{\sigma}) \leqq \psi(\theta, \sigma)] \leqq \alpha\}. \tag{2.3}$$

For any numbers c and d, $\psi(\theta, \sigma) = c\theta+d\sigma$ satisfies (i) and (iii) above. Provided that (ii) above holds, the results of Hora and Buehler yield, in conjunction with (2.3), that for all (θ, σ)

$$P_{\tilde{x}\mid\theta,\sigma}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[c\tilde{\theta}+d\tilde{\sigma} \leqq c\theta+d\sigma] \leqq \alpha\} = \alpha. \tag{2.4}$$

In particular, if $c = -1$ and $d = (\theta-t)/\sigma$,

$$P_{\tilde{x}\mid\theta,\sigma}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[(t-\tilde{\theta})/\tilde{\sigma} \leqq (t-\theta)/\sigma] \leqq \alpha\} = \alpha \tag{2.5}$$

for all (θ, σ) and thus (2.2) is satisfied with $\psi(\theta, \sigma) = (t-\theta)/\sigma$.

3. Failure of frequency property. The above results lead to consideration of whether there are functions ψ for which (2.2) does not hold. An interesting example of such failure is given by $\psi(\theta, \sigma) = \theta+\sigma^2/2$ which arises in analysis of lognormal data, ψ being the mean of a lognormal variable in terms of the mean θ and variance σ^2 of its normally distributed logarithm.

Let $\tilde{x}$ denote a random sample from a normal population with mean θ and variance σ^2. To see that (2.2) fails in this instance for $\psi(\theta, \sigma) = \theta+\sigma^2/2$, one can study the right member of (2.3) for $\theta = 0$ using the relation that for any set T in the parameter space and for any $a > 0$

$$P_{\tilde{\theta},\tilde{\sigma}\mid ax}[(\tilde{\theta}, \tilde{\sigma}) \in T] = P_{\tilde{\theta},\tilde{\sigma}\mid x}[(a\tilde{\theta}, a\tilde{\sigma}) \in T]. \tag{3.1}$$

Thus for any $\sigma > 0$

$$\begin{aligned} & P_{\tilde{x}\mid\theta=0,\sigma}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[\tilde{\theta}+\tilde{\sigma}^2/2 \leqq \sigma^2/2] \leqq \alpha\} \\ & = P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\sigma\tilde{x}}[\tilde{\theta}+\tilde{\sigma}^2/2 \leqq \sigma^2/2] \leqq \alpha\} \\ & = P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[\sigma\tilde{\theta}+\sigma^2\tilde{\sigma}^2/2 \leqq \sigma^2/2] \leqq \alpha\} \\ & = P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[\tilde{\theta}+\sigma\tilde{\sigma}^2/2 \leqq \sigma/2] \leqq \alpha\}. \end{aligned} \tag{3.2}$$

It follows from (2.4) taking $c = 1$ and $d = \sigma$, that for any $\sigma > 0$,

$$P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[\tilde{\theta}+\sigma\tilde{\sigma} \leqq \sigma] \leqq \alpha\} = \alpha. \tag{3.3}$$

Let $R_\sigma = \{(\tilde{\theta}, \tilde{\sigma}) \mid \tilde{\theta}+\sigma\tilde{\sigma}^2/2 \leqq \sigma/2\}$ and $S_\sigma = \{(\tilde{\theta}, \tilde{\sigma}) \mid \tilde{\theta}+\sigma\tilde{\sigma} \leqq \sigma\}$ and note that R_σ is a proper subset of S_σ, the boundary of S_σ being a supporting line to the strictly convex set R_σ. Thus for $\psi(\theta, \sigma) = \theta+\sigma^2/2$

$$\begin{aligned} P_{\tilde{x}\mid\theta=0,\sigma}&[\bar{\psi}(\tilde{x}, \alpha) \geqq \psi(0, \sigma)] \\ &= P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[R_\sigma] \leqq \alpha\} \\ &= P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[S_\sigma] \leqq \alpha\}+P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[S_\sigma] \\ &\qquad > \alpha, P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[R_\sigma] \leqq \alpha\} \\ &= \alpha+\varepsilon(\alpha, \sigma), \end{aligned} \tag{3.4}$$

where $\varepsilon(\alpha, \sigma)$ is defined as the second summand in the preceding line. Thus the probability of coverage is at least α for all $\sigma > 0$ (it is easy to verify that this also holds for all θ). That for any given $\sigma > 0$ there exists an α such that $\varepsilon(\alpha, \sigma) > 0$ is a consequence of the relation $P_{\tilde{x}\mid\theta=0,\sigma=1}\{P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[R_\sigma] > P_{\tilde{\theta},\tilde{\sigma}\mid\tilde{x}}[S_\sigma]\} > 0$, which certainly holds for the case of normality or any other setting in which the posterior density is everywhere positive. It is intuitively apparent that in this case $\varepsilon(\alpha, \sigma) > 0$ for all (α, σ) but this seems difficult to prove.

4. Comparison with other methods. Suppose that $\bar{\psi}^*(x, \alpha)$, defined for all x, is an α-level upper confidence limit, meaning only that it satisfies (2.2). To avoid more clumsy notation consider α as fixed in the following. Suppose further that $\bar{\psi}^*(x, \alpha)$ is invariant, as is $\bar{\psi}(x, \alpha)$, in the following sense. Let $ax+b$ represent a transformation of scale and location of the data and for each x let (θ_x, σ_x) represent any solution to $\psi(\theta, \sigma) = \bar{\psi}^*(x, \alpha)$. If ψ is invariantly estimable require that $\bar{\psi}^*(ax+b, \alpha) = \psi(a\theta_x+b, a\sigma_x)$ for all x; if $\psi = (t-\theta)/\sigma$ require that $\bar{\psi}^*(ax+b, \alpha) = (t-a\theta_x-b)/a\sigma_x$. This restriction on $\bar{\psi}^*$, along with the fact that $\bar{\psi}(x, \gamma)$ satisfies the same condition for all γ, implies that for any γ the event $A_\gamma = \{x \mid \bar{\psi}^*(x, \alpha) = \bar{\psi}(x, \gamma)\}$ has the property that x belongs to A_γ if and only if $ax+b$ belongs for all $a > 0$ and b.

It follows that if $\tilde{x}$ has a location and scale parameter distribution then the conditional distribution of $\tilde{x}$ given A_γ is also of this form. Also $\bar{\psi}(x, \gamma)$ remains the Bayesian confidence limit for the conditional problem, and (2.2) holds for the conditional distribution of $\tilde{x}$. Thus, for all (θ, σ),

$$P_{\tilde{x}\mid\theta,\sigma}[\bar{\psi}^*(\tilde{x}, \alpha) \geqq \psi(\theta, \alpha) \mid \bar{\psi}^*(\tilde{x}, \alpha) = \bar{\psi}(\tilde{x}, \gamma)] = \gamma. \tag{4.1}$$

It would seem, then, that when an $\tilde{x}$ in A_γ is obtained the appropriate "confidence level" of $\bar{\psi}^*(\tilde{x}, \alpha)$ is not α but γ. A cogent discussion of this notion, which has been considered by many writers, is given in Buehler [1], where the A_γ would be called relevant subsets of the sample space. It should not be supposed that (4.1) presents an objection to orthodox theory, however, because when ψ is invariantly estimable there seems to be no $\bar{\psi}^*$ different from $\bar{\psi}$ which satisfies any kind of frequentist "optimality" considerations. The intent of (4.1) is to provide a cogent argument for rejection of any *ad hoc* $\bar{\psi}^*$ such as those based on standard pivotal quantities, $(\hat{\theta}-\theta)/\hat{\sigma}$ and $\hat{\sigma}/\sigma$, when there is no sufficient reduction of the data. Presumably a more appealing frequentist approach in such problems is to condition on the ancillary statistics, which leads to the Bayes-fiducial solution.

REFERENCES

[1] BUEHLER, R. J. (1959). Some validity criteria for statistical inferences. *Ann. Math. Statist.* **30** 845–863.

[2] FRASER, D. A. S. (1961). The fiducial method and invariance. *Biometrika* **48** 261–280.

[3] HORA, R. B. and BUEHLER, R. J. (1966). Fiducial theory and invariant estimation. *Ann. Math. Statist.* **37** 643–656.

[4] WELCH, B. L. and PEERS, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *J. Roy. Statist. Soc. Ser B* **25** 318–329.