The Annals of Statistics
1979, Vol. 7, No. 4, 742–755

CONDITIONAL PROPERTIES OF STATISTICAL PROCEDURES

BY G. K. ROBINSON

University of Sheffield

During the last twenty years several results on the conditional properties of statistical procedures have appeared in the literature. However, apart from results linking conditional properties and Bayesianity, few general results have been found. Here a systematic set of definitions of conditional properties is given and a systematic investigation of their consequences is started. An analogy between interval and point estimation is used both to extend definitions of conditional properties to point estimation and to extend the usual definition of admissibility for point estimation to provide a new definition of admissibility for interval estimation.

1. Introduction. Historically, interest in the conditional properties of statistical procedures arose from questioning of the Neyman-Pearson and fiducial theories of inference. Fisher (1956a page 55) claimed that the justification of the fiducial argument lay in the conditional properties of fiducial procedures. Neyman and Pearson, on the other hand, never claimed that their methods would ensure good conditional properties, but interest in conditional properties has been shown by their critics.

In terms of Neyman's theory of confidence intervals, conditional properties seem to be of interest because if, say,

$$P[\theta \in I(X)] = \alpha \qquad \text{for all} \quad \theta$$

while

$$P[\theta \in I(X) | X \in C] \leqslant \alpha - \varepsilon \qquad \text{for all} \quad \theta$$

for some set C and some positive number ε, then the use of α as a confidence level for the proposition "$\theta \in I(x)$" seems rather dubious, particularly whenever $x \in C$. Several writers have discussed this difficulty. Fisher (1956b) criticized Welch's test for the two means problem in a way which Buehler (1959) has shown to be equivalent to exhibiting a negatively biased relevant subset for a confidence interval based on Welch's test. Bartlett (1956), Neyman (1956) and Welch (1956) all replied to Fisher's criticism but did not discuss the major issue: whether or not they regarded conditional properties as important. Jones (1958) and Cox (1958) have made it clear they do consider conditional properties to be important. Buehler (1959) gave some definitions of conditional properties which can be used to discuss Neyman confidence intervals and that paper should be regarded as a landmark. Savage (1962), Birnbaum (1962), Dempster (1964) and Hacking (1965) have all made attempts to conceptualize the underlying weakness in Neyman's theory and I

Received September 1975; revised February 1978.
AMS 1970 *subject classifications.* Primary 62-02, 62A99, 62C15; secondary 62F10, 62F15, 62F25.
Key words and phrases. Conditional properties, betting, relevant subset, admissibility, Bayesianity.

believe that their ideas can be understood most easily using the terminology of conditional properties.

Buehler (1959) suggested, in a manner calculated not to arouse too much antagonism, that statistical inference could be viewed as a game between two players. The first player suggests a statistical procedure and the second bets against him. It is easy to see that anomalies stated in terms of conditional probabilities could always be stated in terms of betting strategies but that some betting strategies cannot be expressed in terms of conditional probabilities. Buehler's definitions of relevant and semirelevant subsets can be interpreted using either betting strategies or conditional probabilities but here, following Pierce (1973), we shall formulate more general definitions which do not necessarily have conditional probability interpretations. However, we shall still use the expression "conditional properties" to describe them.

Another restriction of Buehler's definitions is that they are confined to interval estimation and hypothesis testing. Brewster and Zidek (1974) seem to be aware that an extension to point estimation is possible, but they have not formulated definitions. Our definitions appear in Section 6.

The essential reason for investigating conditional properties of statistical procedures in a fairly general context is the search for a rationale for non-Bayesian inference. It seems natural to hope that statistical conclusions could be used for some restricted types of betting and the question which must be asked is "What sorts of betting should statistical procedures be able to withstand?" The example of Buehler and Fedderson (1963) and Brown (1967) indicates that to demand the nonexistence of relevant subsets is too stringent a condition. Robinson (1976) suggests that the nonexistence of negatively biased relevant selections is a condition which is about strong enough. The absence of super-relevant betting procedures, which is discussed below, is, perhaps, too weak a condition. The mathematical investigation of conditional properties in this paper and Robinson (1979) is directed towards answering the question.

2. Notation and terminology. We will use $\mathfrak{X}$, x and X to denote a sample space, a point in the sample space and a random variable. The symbols Θ and θ will denote a parameter space and a point in the parameter space. Integrals over $\mathfrak{X}$ and Θ will be written in the forms $\int_{\mathfrak{X}} g(x, \theta)\, dx$ and $\int_{\Theta} g(x, \theta)\, d\theta$, the measures implicit in the definitions of such integrals generally being finite dimensional Lebesgue measure.

All functions defined below (including $I(.)$) will be assumed to satisfy appropriate measurability assumptions which will not be discussed rigorously. The distribution of X given θ will usually be assumed to have a density, denoted $f(x|\theta)$. The notation used for characteristic functions of sets will be that for a set A

$$\chi_A(t) = 1 \quad \text{if} \quad t \in A$$
$$= 0 \quad \text{if} \quad t \notin A.$$

Some common statistical terms can cause confusion because they are used by different people in different ways. There are three terms that I wish to clarify.

(i) The word "confidence" will be used here in an intuitive sense, not as a technical term implying a justification in terms of Neyman's theory of confidence intervals.

(ii) A "Bayesian" point estimator will mean the posterior expectation of some parameter. We will not be using loss functions in this context or referring to estimators which minimize posterior expected loss.

(iii) An *interval estimator* of θ will mean a function taking x into the ordered pair $\langle I(x), \alpha(x)\rangle$ where $I(x)$ specifies a subset of Θ and $\alpha(x)$ specifies a number in the unit interval for every x in $\mathcal{X}$. The interpretation of $\langle I(x), \alpha(x)\rangle$ will be the usual one that $\alpha(x)$ states a degree of belief or level of confidence in some sense in the proposition that $\theta \in I(x)$ after $X = x$ has been observed, but we will not restrict $I(x)$ to being an interval nor restrict $\alpha(x)$ to being independent of x. We will call $I(x)$ a set function and $\alpha(x)$ a confidence function.

It should also be noted that the correspondence between interval estimation and hypothesis testing appropriate for thinking about conditional properties is that of the Bayesian school of statistics, not that of the sampling-theory school. For the hypothesis $\theta = \theta_0$ to be accepted with confidence α, the point set $\{\theta_0\}$ should be an interval estimate for θ when it is associated with its degree of confidence, α. With this correspondence the conditional properties of interval estimators can be immediately applied to tests of hypotheses.

3. Definitions of conditional properties for interval estimation. Like Buehler (1959), we will state conditional properties using a betting game between two players, Peter and Paul. The interesting thing is to see to what extent statistical procedures can withstand criticism stated in terms of bets.

Peter must quote an interval estimator, $\langle I(x), \alpha(x)\rangle$, for θ. Paul's task is to bet against that estimator and his strategy can always be expressed by a real-valued function $s(x)$ such that

(i) when $s(x) > 0$, he places a bet of size $s(x)$ that $\theta \in I(x)$ at odds corresponding to Peter's quoted confidence, $\alpha(x)$, that $\theta \in I(x)$;

(ii) when $s(x) < 0$, he places a bet of size $-s(x)$ that $\theta \notin I(x)$ at odds corresponding to Peter's quoted confidence, $1 - \alpha(x)$, that $\theta \notin I(x)$; and

(iii) when $s(x) = 0$, he makes no bet.

For mathematical convenience, we define the size of a bet to be the sum of Peter's stake and Paul's stake. Now a bet of size s on an event for which Peter has quoted confidence β is such that either Paul wins βs or Peter wins $(1 - \beta)s$. Paul's gain from the above situation can therefore be written concisely as

$$\{\chi_{I(x)}(\theta) - \alpha(x)\}s(x).$$

We define six classes of betting procedures.

(i) We say that $s(x)$ is a *wide-sense betting procedure* if $s(x)$ may be unbounded but $E[|s(X)|]$ is bounded. Betting where the expected total stake may be unbounded seems to be of no statistical interest since it is possible to win against proper Bayesian procedures using such betting. (See Example 4.1). The class of wide-sense betting procedures seems to be the most general class of betting procedures which is statistically interesting, but we shall usually restrict our attention to the next class.

(ii) We refer to $s(x)$ as a *betting procedure*, with no qualification, if it is bounded as a function of x. By a change of scale we could take the bound on $|s(x)|$ as unity and interpret $|s(x)|$ as the probability that a bet of unit size is made when $X = x$ is observed. The sign of $s(x)$ would indicate the direction of the bet.

(iii) If $0 \leqslant s(x) \leqslant 1$ we call $s(x)$ a *positively biased* selection.

(iv) If $-1 \leqslant s(x) \leqslant 0$ then we call $s(x)$ a *negatively biased* selection. Sometimes $k(x) = -s(x)$ is also called a negatively biased selection.

(v) If $s(x) = \chi_c(x)$ for some subset, C, of $\mathcal{X}$, then we use the term *positively biased subset* to describe both C and $s(x)$.

(vi) If $s(x) = -\chi_c(x)$ for some subset, C, of $\mathcal{X}$, then either C or $s(x)$ may be called a *negatively biased subset*.

A conditional property of a statistical procedure is to say whether or not Paul can find a certain type of winning betting procedure when Peter adopts the given statistical procedure. For the interval estimator $\langle I(x), \alpha(x)\rangle$ the betting procedure $s(x)$ is said to be

(i) *semirelevant* if

$$E[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X)] \geqslant 0 \qquad \text{for all} \quad \theta$$

and is strictly positive for some θ;

(ii) *relevant* if for some $\varepsilon > 0$

$$E[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|] \geqslant 0 \qquad \text{for all} \quad \theta$$

and is strictly positive for some θ; and

(iii) *super-relevant* if for some $\varepsilon > 0$

$$E[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X)] \geqslant \varepsilon \qquad \text{for all} \quad \theta.$$

All super-relevant wide-sense betting procedures are relevant and all relevant wide-sense betting procedures are semirelevant. Broadly speaking, the existence of a semirelevant betting procedure is a mild criticism of an interval estimator, the existence of a super-relevant betting procedure is a severe criticism, and the existence of a relevant betting procedure is a criticism on a level which seems to be just serious enough to bother about.

We now have a two-dimensional array of concepts which can be used to discuss the conditional properties of interval estimates. The dimension which distinguishes between semirelevant, relevant and super-relevant is the more important one to understand. The distinction between relevant and super-relevant betting procedures is clear enough: a super-relevant betting procedure achieves an expected return which is bounded away from zero. The distinction between relevant and semirelevant betting procedures is more subtle. In terms of betting, a relevant betting procedure manages to achieve positive expected return although paying a fraction ε of the total stake for the privilege of betting. In terms of statistical importance, a relevant betting procedure is a more constructive criticism than a semirelevant betting procedure as we now illustrate.

Suppose that $s(x)$ is a relevant betting procedure for $\langle I(x), \alpha(x)\rangle$ with positive constant ε. Define

$$\begin{aligned} \beta(x) &= \alpha(x) - \varepsilon && \text{if } s(x) \leqslant -1 \\ &= \alpha(x) + \varepsilon s(x) && \text{if } -1 \leqslant s(x) \leqslant 1 \\ &= \alpha(x) + \varepsilon && \text{if } s(x) \geqslant 1. \end{aligned}$$

Then

$$\{\chi_{I(x)}(\theta) - \beta(x)\}s(x) \geqslant \{\chi_{I(x)}(\theta) - \alpha(x)\}s(x) - \varepsilon|s(x)|$$

so $s(x)$ is a semirelevant betting procedure for $\langle I(x), \beta(x)\rangle$. Although $\beta(x)$ may not be an optimal confidence function in any sense, it does seem to be better than $\alpha(x)$ in that the change from $\alpha(x)$ to $\beta(x)$ is in the direction suggested by $s(x)$. Thus the relevant betting procedure $s(x)$ has implicitly suggested an alternative confidence function. This is what is meant by saying it is a constructive criticism of the interval estimator.

The definitions given above are intended to be an essentially complete set. They include all those given by Buehler (1959), Wallace (1959) and Pierce (1973) except for the fact that our definitions do not demand that $\alpha(x)$ be a constant and some differences in positivity conditions.

Stone's (1976) definition of strong inconsistency corresponds, in our terminology, to the existence of a super-relevant betting procedure such that $s(x) = 1$ for all x. Bondar's (1977) consistency principle III is, in our terminology, the nonexistence of super-relevant subsets.

4. Interval estimation examples. The following examples are given to motivate the definitions of the previous section. Further examples are given by Buehler (1959).

4.1. *An example to show why unbounded betting is not considered to be statistically interesting.* Suppose that X is normally distributed with mean θ and unit variance. For the interval estimator $\langle(x - 2, x + 2), \alpha\rangle$, where α is the Neyman confidence level and also the improper Bayesian posterior confidence that $\theta \in (x - 2, x + 2)$ for a uniform prior, we consider unbounded betting procedure $s(x) = -x^2$. Paul's expected yield from betting is

$$\int_{-\infty}^{\infty}\{\alpha - \chi_{(-2,2)}(z)\}(\theta + z)^2(2\pi)^{-\frac{1}{2}}\exp\left(-\tfrac{1}{2}z^2\right) dz$$

$$= \int_0^{\infty}\{\alpha - \chi_{(0,2)}(z)\}\{(\theta + z)^2 + (\theta - z)^2\}(2\pi)^{-\frac{1}{2}}\exp\left(-\tfrac{1}{2}z^2\right) dz.$$

This can be seen to be independent of θ since $(\theta + z)^2 + (\theta - z)^2 = 2\theta^2 + 2z^2$ and the θ^2 term vanishes on integration. Paul's yield is strictly positive. As we show in Robinson (1979), no semirelevant wide-sense betting procedure exists for this interval estimator. Statistically it seems unreasonable to call $-x^2$ a super-relevant betting procedure for this eminently reasonable interval estimator.

4.2. *A situation where a semirelevant betting procedure exists but a relevant betting procedure does not exist.* This example is a special case of Buehler's Example 4.1. Again take X to be a normally distributed random variable with mean θ and unit variance. The set $[0, \infty)$ is a positively biased semirelevant subset for the interval estimator $\langle(-\infty, x), \frac{1}{2}\rangle$ since

$$P[\theta < X | X > 0] > \tfrac{1}{2} \qquad \text{for all} \quad \theta.$$

However, no relevant betting procedure exists (See Section 2 of Robinson (1979)). It seems that the confidence function $\alpha(x) = \frac{1}{2}$ is not completely satisfactory for use in betting, but that its use as a degree of confidence that $\theta \in (-\infty, x)$ is defensible.

4.3. *A situation where a relevant betting procedure exists but a super-relevant betting procedure does not.* Once more, take X to be normally distributed with mean θ and unit variance. Consider the interval estimator $\langle(x - 1.96, x + 1.96), \alpha(x)\rangle$, where

$$\begin{aligned} \alpha(x) &= 0.95 \qquad \text{if} \quad x > 10^{10} \\ &= 0.99 \qquad \text{if} \quad x \leqslant 10^{10}. \end{aligned}$$

The selection

$$\begin{aligned} s(x) &= -1 && \text{if} \quad x < 0 \\ &= -1 + 10^{-10}x && \text{if} \quad 0 \leqslant x \leqslant 10^{10} \\ &= 0 && \text{if} \quad x > 10^{10} \end{aligned}$$

can be seen to be relevant with a little effort. However no super-relevant betting procedure exists. (To see this consider a sequence of smooth proper priors which give no prior probability to $\theta < 10^{10}$ and apply Proposition 7.4.) This illustrates the point that the nonexistence of super-relevant betting procedures is a very mild restriction on interval estimators; and conversely, the existence of a super-relevant betting procedure is a severe criticism.

5. Interval estimation viewed as point estimation. In this section we note that interval estimation can be viewed as a slightly peculiar form of point estimation. For an interval estimator $\langle I(x), \alpha(x)\rangle$, we can regard the set function, $I(x)$, as fixed and regard the confidence function, $\alpha(x)$, as a point estimator of $\chi_{I(x)}(\theta)$. The function to be estimated depends on x, so we do not have point estimation of the usual sort, but the analogy is useful nevertheless. The primary purpose of the analogy is to enable conditional properties to be defined for point estimators using the definitions of conditional properties for interval estimators. In addition, it leads to a new definition of admissibility for interval estimation.

NOTE. Most of the properties of interval estimators discussed in this paper may be viewed as properties of the confidence function for a fixed set function. These properties are referred to as properties of the interval estimator (rather than as properties of the confidence function which would be more natural from a purely mathematical point of view) because the interval estimator is the object of statistical interest. Properties of interval estimators which might be viewed as properties of a set function for a given confidence function will also be referred to as properties of interval estimators.

6. Definitions of conditional properties for point estimation. Using the analogy between interval and point estimation of Section 5, reconsider Buehler's betting game between Peter and Paul. Peter quotes a point estimator, $\alpha(x)$, of the function $\chi_{I(X)}(\theta)$ and Paul tries to guess when $\alpha(x)$ is too high and when it is too low.

Peter and Paul can play a similar betting game for point estimation. Peter quotes a point estimator, $T(x)$, of some function $\phi(\theta)$ and then Paul tries to guess when it is too high or too low. We say that the betting procedure $s(x)$ is *semirelevant* if

$$E[\{\phi(\theta) - T(X)\}s(X)] \geqslant 0 \qquad \text{for all} \quad \theta$$

and is strictly positive for some θ.

The definitions of relevant and super-relevant for interval estimation contain a quantity ε which is intended to be small but not negligible. The quantities $\chi_{I(x)}(\theta)$ and $\alpha(x)$ are always between 0 and 1 so a small positive constant is small but not negligible. For point estimation the situation is not so simple. If $\phi(\theta)$ is a scale parameter then a small positive constant would be negligible as $\phi(\theta) \to \infty$ but all-important as $\phi(\theta) \to 0$. We therefore use $\varepsilon\phi(\theta)$ as the small quantity in our definitions in this case.

For a point estimator $T(x)$ of a location parameter $\phi(\theta)$, we say that the betting procedure $s(x)$ is *relevant* if for some $\varepsilon > 0$

$$E[\{\phi(\theta) - T(X)\}s(X) - \varepsilon|s(X)|] \geqslant 0 \qquad \text{for all} \quad \theta$$

and is strictly positive for some θ; and we say that $s(x)$ is *super-relevant* if for some $\varepsilon > 0$

$$E[\{\phi(\theta) - T(X)\}s(X)] \geqslant \varepsilon \qquad \text{for all} \quad \theta.$$

For a point estimator $T(x)$ of a scale parameter $\phi(\theta)$, we say that $s(x)$ is *relevant* if for some $\varepsilon > 0$

$$E[\{\phi(\theta) - T(X)\}s(X) - \varepsilon\phi(\theta)|s(X)|] \geqslant 0 \qquad \text{for all} \quad \theta$$

and is strictly positive for some θ; and we say that $s(x)$ is *super-relevant* if for some $\varepsilon > 0$

$$E[\{\phi(\theta) - T(X)\}s(X)] \geqslant \varepsilon\phi(\theta) \qquad \text{for all} \quad \theta.$$

The classification of betting procedures into subsets, selections, etc., given in

Section 3, can still be applied when the betting concerns point estimation, so a wide range of concepts is again available.

For interval estimation the classification of betting procedures which was given in Section 3 seems to be fairly complete. However, for point estimation, completeness is more difficult to achieve. The essential thing is that a betting procedure is semirelevant if it wins (in the sense of achieving positive expected return for all θ and strictly positive expected return for some θ); relevant if it wins although paying a nontrivial amount related to the frequency of betting; and super-relevant if it wins a nontrivial amount unconditionally.

All these ideas can be extended from real point estimation to the multidimensional case. However, instead of saying that an estimate is too high or too low, a betting procedure must specify a direction in the Euclidean space containing $\boldsymbol{\phi}(\theta)$ and $\mathbf{T}(x)$. Therefore $\mathbf{s}(x)$ must be allowed to be vector-valued and the definitions above can be used provided that the product of $\boldsymbol{\phi}(\theta) - \mathbf{T}(x)$ and $\mathbf{s}(x)$ is understood to be an inner product.

7. Basic mathematical results. We now commence a survey of the conditional properties of statistical procedures. We will concentrate on interval estimation and leave many of the results on point estimation as conjectures. In this section we concentrate on the definitions and their links with Bayesianity.

DEFINITION. The betting procedure $s(x)$ is nontrivial if $E[|s(X)|] > 0$ for some θ. Using this concept the following proposition provides an alternative definition of relevant betting procedures for interval estimation. Similar statements hold for relevant point estimators. All are easy to prove.

PROPOSITION 7.1. *The betting procedure $s(x)$ is relevant for the interval estimator $\langle I(x), \alpha(x)\rangle$ if and only if it is nontrivial and for some $\varepsilon > 0$*

$$E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|\right] \geqslant 0 \qquad \text{for all} \quad \theta.$$

The next three propositions show the extent to which proper Bayesianity is sufficient to ensure that statistical procedures have good conditional properties for interval estimators. Note that by proper Bayesian procedures we mean ones based on prior measures which are probability distributions.

PROPOSITION 7.2. *There is no wide-sense relevant betting procedure for a proper Bayesian interval estimator which is well-defined throughout the sample space in the sense that*

$$\int_\Theta f(x|\theta)\, dG(\theta) > 0 \qquad \textit{for all} \quad x \tag{7.1}$$

where $G(\theta)$ is the prior probability distribution function.

PROOF. Suppose that the interval estimator $\langle I(x), \alpha(x)\rangle$ for θ is Bayesian with respect to the prior distribution function $G(\theta)$. That is,

$$\alpha(x) = \int_{I(x)} f(x|\theta)\, dG(\theta) / \int_\Theta f(x|\theta)\, dG(\theta). \tag{7.2}$$

Suppose $s(x)$ is a nontrivial wide-sense betting procedure such that for some $\varepsilon > 0$

$$E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|\right] \geqslant 0 \qquad \text{for all} \quad \theta.$$

From (7.2)

$$\int_\Theta \{\chi_{I(x)}(\theta) - \alpha(x)\} f(x|\theta)\, dG(\theta) = 0;$$

so, using the properness of the prior and the fact that $s(x)$ is a wide-sense betting procedure to justify the change of order of integration,

$$\begin{aligned} 0 &\leqslant \int_\Theta E[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|]\, dG(\theta) \\ &= \int_{\mathfrak{X}} \int_\Theta [\{\chi_{I(x)}(\theta) - \alpha(x)\}s(x) - \varepsilon|s(x)|] f(x|\theta)\, dG(\theta)\, dx \\ &= -\varepsilon \int_{\mathfrak{X}} \int_\Theta f(x|\theta)\, dG(\theta) |s(x)|\, dx. \end{aligned}$$

The nontriviality of $s(x)$ implies that $\int_{\mathfrak{X}} |s(x)|\, dx > 0$; so, using (7.1),

$$\int_{\mathfrak{X}} \int_\Theta f(x|\theta)\, dG(\theta) |s(x)|\, dx > 0;$$

in contradiction with the preceding inequality.

PROPOSITION 7.3. *If an interval estimator $\langle I(x), \alpha(x)\rangle$ is a proper Bayesian confidence region for θ then the following four conditions are sufficient to ensure that no semirelevant wide-sense betting procedure exists.*

(i) *The prior distribution has a strictly positive density, say $g(\theta)$.*
(ii) *The density, $f(x|\theta)$, is continuous as a function of θ.*
(iii) *The measure of $J(\theta)\Delta J(\theta')$ tends to zero as $\theta \to \theta'$ where $J(\theta) = \{x \in \mathfrak{X} : \theta \in I(x)\}$ and Δ denotes the symmetric difference.*
(iv) *For any θ' there is a neighbourhood, $N(\theta')$, of θ' and a number M such that $|f(x|\theta) - f(x|\theta')| < M$ for all $x \in \mathfrak{X}$ and all $\theta \in N(\theta')$.*

REMARKS. Although Theorems 2 and 3 of Wallace (1959) are incorrect, his Theorem 1 did establish the essential links between proper Bayesianity and conditional properties. Our Propositions 7.2 and 7.3 are similar to the two parts of his Theorem 1, but our definitions of conditional properties are different from his so we require different regularity conditions.

Condition (iii), above, is rather like a continuity condition on $\chi_{I(x)}(\theta)$. To see that it is necessary consider an otherwise well-behaved Bayesian interval estimator with a single point θ_0 adjoined to the confidence set for all x. The simple betting function $s(x) = 1$ has zero expected return for all θ except θ_0 where the expected return is strictly positive. Thus $s(x) = 1$ is a semirelevant betting procedure.

PROOF OF PROPOSITION 7.3. For an arbitrary wide-sense betting procedure, $s(x)$, consider the function of θ

$$A(\theta) = E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X)\right].$$

Since

$$\alpha(x) = \int_\Theta \chi_{I(x)}(\theta) g(\theta) f(x|\theta)\, d\theta \Big/ \int_\Theta g(\theta) f(x|\theta)\, d\theta,$$

$$\text{(7.3)} \quad \int_\Theta A(\theta) g(\theta)\, d\theta = \int_{\mathfrak{X}} \int_\Theta \{\chi_{I(x)}(\theta) - \alpha(x)\} f(x|\theta) g(\theta)\, d\theta\, s(x)\, dx = 0.$$

Furthermore

$$|A(\theta) - A(\theta')|$$
$$= |\int_{\mathfrak{X}}\{\chi_{I(x)}(\theta) - \chi_{I(x)}(\theta')\}s(x)f(x|\theta)\,dx$$
$$+ \int_{\mathfrak{X}}\{\chi_{I(x)}(\theta') - \alpha(x)\}s(x)\{f(x|\theta) - f(x|\theta')\}\,dx|$$
$$\leqslant \int_{\mathfrak{X}}|\chi_{I(x)}(\theta) - \chi_{I(x)}(\theta')|s(x)f(x|\theta)\,dx + \int_{\chi}|f(x|\theta) - f(|\theta')|s(x)\,dx$$
$$\to 0 \quad \text{as} \quad \theta \to \theta'$$

by the dominated convergence theorem, the fact that $s(x)$ is a wide-sense betting procedure and conditions (ii), (iii) and (iv). Thus $A(\theta)$ is continuous, so (7.3) and the positivity of $g(\theta)$ imply that if $A(\theta) > 0$ for some value of θ then $A(\theta) < 0$ for some other value of θ. Hence $s(x)$ cannot be semirelevant.

PROPOSITION 7.4. *If an interval estimator* $\langle I(x), \alpha(x)\rangle$ *is such that for every* $\varepsilon > 0$ *there is a proper prior distribution* $G_\varepsilon(\theta)$ *such that*

$$\int_\Theta \int_{w(\varepsilon)} f(x|\theta)\,dx\,dG_\varepsilon(\theta) < \varepsilon \tag{7.4}$$

where

$$w(\varepsilon) = \left\{x \in \mathfrak{X}: \left|\alpha(x) - \frac{\int_{I(x)} f(x|\theta)\,dG_\varepsilon(\theta)}{\int_\Theta f(x|\theta)\,dG_\varepsilon(\theta)}\right| \geqslant \varepsilon\right\}$$

then there is no super-relevant betting procedure for $\langle I(x), \alpha(x)\rangle$.

PROOF. Suppose that $s(x)$ is a super-relevant betting procedure for an interval estimator satisfying the condition of the proposition. For some $\varepsilon > 0$ (7.4) holds and

$$\int_{\mathfrak{X}}\{\chi_{I(x)}(\theta) - \alpha(x)\}s(x)f(x|\theta)\,dx \geqslant \varepsilon \qquad \text{for all} \quad \theta. \tag{7.5}$$

Using (7.4),

$$\int_\Theta\int_{\mathfrak{X}}\{\chi_{I(x)}(\theta) - \alpha(x)\}s(x)f(x|\theta)\,dx\,dG_\varepsilon(\theta)$$
$$= \int_{w(\varepsilon)}\int_\Theta\{\chi_{I(x)}(\theta) - \alpha(x)\}f(x|\theta)\,dG_\varepsilon(\theta)s(x)\,dx$$
$$+ \int_{\mathfrak{X}\setminus w(\varepsilon)}\int_\Theta\{\chi_{I(x)}(\theta) - \alpha(x)\}f(x|\theta)\,dG_\varepsilon(\theta)s(x)\,dx$$
$$< \varepsilon + \int_{\mathfrak{X}\setminus w(\varepsilon)}\int_\Theta\left\{\chi_{I(x)}(\theta) - \frac{\int_{I(x)} f(x|\theta)\,dG_\varepsilon(\theta)}{\int_\Theta f(x|\theta)\,dG_\varepsilon(\theta)} + \varepsilon\right\} f(x|\theta)dG_\varepsilon(\theta)s(x)\,dx$$
$$\leqslant \varepsilon.$$

However, from (7.5)

$$\int_\Theta\int_{\mathfrak{X}}\{\chi_{I(x)}(\theta) - \alpha(x)\}s(x)f(x|\theta)\,dx\,dG_\varepsilon(\theta) \geqslant \varepsilon.$$

This is the required contradiction.

For point estimation it seems that the conditional properties of Bayesian point estimators are similar to those of Bayesian interval estimators. However, different conditions are required. A condition which is usually required in order to prove any conditional property for Bayesian point estimators is that the prior expectation of the parameter to be estimated must be finite. We give one result on the conditional properties of Bayesian point estimators.

PROPOSITION 7.5. *Suppose that $T(x)$ is a Bayesian point estimator of $\phi(\theta)$ with respect to a strictly positive prior density, $g(\theta)$. Suppose $\phi(\theta)$ is continuous and $f(x|\theta)$ is continuous as a function of θ. If $\int_\Theta|\phi(\theta)|\, g(\theta) d\theta$ is finite then there is no semirelevant betting procedure for $T(x)$.*

PROOF. For an arbitrary betting procedure, $s(x)$, consider the function of θ

$$A(\theta) = E[\{\phi(\theta) - T(X)\}s(X)].$$

Since

$$T(x) = \int_\Theta \phi(\theta)g(\theta)f(x|\theta)\, d\theta / \int_\Theta g(\theta)f(x|\theta)\, d\theta$$

and the existence of $\int_\Theta|\phi(\theta)|/g(\theta)\, d\theta$ justifies the change of order of integration, we find that

$$\int_\Theta A(\theta)g(\theta)\, d\theta = 0.$$

That $A(\theta)$ is continuous and that, therefore, there is no semirelevant betting procedure for $T(x)$, can be shown as in the proof of Proposition 7.3.

For improper Bayesian interval estimators the change of order of integration used to prove Propositions 7.2 to 7.5 cannot be justified by Fubini theorems and conditional properties can only be established in special cases. This is a large part of the subject of Robinson (1979).

It seems to be generally true that limiting Bayesianity of some form is necessary for the nonexistence of the various types of betting procedures, but only Pierce (1973) has made much progress in this direction. His Theorem 1 shows that for finite parameter spaces only Bayesian interval estimators do not allow semirelevant betting procedures. For general parameter spaces, his Theorem 3 shows that only interval estimators which are "α-level weak Bayes in mean" do not allow super-relevant, possibly unbounded, betting procedures. The condition "α-level weak Bayes in mean" is equivalent to the condition imposed in Proposition 7.4 (Pierce's definition and theorem can be easily changed to allow α to depend on x), so Proposition 7.4 is almost the converse of Pierce's Theorem 3.

The final proposition in this section links relevant betting schemes and uniform convergence. It is straightforward to prove.

PROPOSITION 7.6. *If there is no relevant betting procedure for any interval estimator $\langle I(x), \alpha_n(x)\rangle$ in a sequence, and $\alpha_n(x) \to \alpha(x)$ uniformly in x as $n \to \infty$, then there is no relevant betting procedure for the limiting interval estimator $\langle I(x), \alpha(x)\rangle$.*

8. The relationship between conditional and admissibility properties. No review of admissibility results will be given here since these results are well known and adequately reviewed by Zacks (1971). It is generally true that admissibility is a property intermediate between the absence of semirelevant betting procedures and the absence of relevant betting procedures. This link seems to me to be of help in understanding both conditional and admissibility properties. In at least one case the link is easy to establish.

PROPOSITION 8.1. *For a point estimator of a location parameter, $\phi(\theta)$, the absence of wide-sense semirelevant betting procedures implies admissibility with respect to squared-error loss, and admissibility with respect to squared-error loss implies the absence of relevant betting procedures.*

PROOF. If $s(x)$ is a relevant betting procedure for an estimator, $T(x)$, of $\phi(\theta)$ then for some $\varepsilon > 0$

$$E[\{\phi(\theta) - T(X)\}s(X) - \varepsilon|s(X)|] \geqslant 0 \qquad \text{for all} \quad \theta,$$

and is strictly positive for some θ. Defining $R(x) = T(x) + \varepsilon s(x)$,

$$\begin{aligned} E[\{\phi(\theta) - T(X)\}^2] - E[\{\phi(\theta) - R(X)\}^2] \\ &= E[2\varepsilon\phi(\theta)s(X) - 2\varepsilon T(X)s(X) - \varepsilon^2 s(X)^2] \\ &\geqslant 2\varepsilon E[\{\phi(\theta) - T(X)\}s(X) - \varepsilon|s(X)|]. \end{aligned}$$

Hence the estimator $T(x)$ is inadmissible with respect to squared error loss.

For arbitrary estimators $T(x)$ and $R(x)$ of $\phi(\theta)$,

$$\begin{aligned} E[\{\phi(\theta) - T(X)\}^2] - E[\{\phi(\theta) - R(X)\}^2] \\ &= 2E[\{\phi(\theta) - T(X)\}\{R(X) - T(X)\} - \{T(X) - R(X)\}^2] \\ &\leqslant 2E[\{\phi(\theta) - T(X)\}\{R(X) - T(X)\}]. \end{aligned}$$

If $T(x)$ is inadmissible then there must be an estimator $R(x)$ with smaller risk and so the wide-sense betting procedure $R(x) - T(x)$ must be semirelevant.

It is natural to ask whether a similar result is true for interval estimation. To answer this question we must first formulate a new definition of admissibility for interval estimation. We say that the interval estimator $\langle I(x), \alpha(x)\rangle$ is *admissible with respect to squared-error loss* if there is no other confidence function, $\beta(x)$, such that

$$E[\{\chi_{I(x)}(\theta) - \alpha(X)\}^2] \geqslant E[\{\chi_{I(x)}(\theta) - \beta(X)\}^2]$$

for all θ with strict inequality for some θ.

As well as being the definition which enables a result like Proposition 8.1 to be proved most easily, this definition is the natural result of applying the analogy of Section 5 between interval and point estimation to the usual definition of admissibility for point estimation. If we think of $\alpha(x)$ as a point estimator of $\chi_{I(x)}(\theta)$ then its squared error loss is $\{\chi_{I(x)}(\theta) - \alpha(x)\}^2$.

Godambe (1961) gave a definition of admissibility which is substantially different from the one used here. Essentially, he says a confidence region is inadmissible if there is another confidence region which is smaller but which has a higher probability of covering the true parameter value. Godambe's concept is concerned with the size of confidence regions, not with the correctness of the confidence function for the given set function. Proper Bayesian interval estimators which do not quote highest posterior density regions for the parameter are inadmissible in his sense. With our definition only regularity conditions are required for admissibility of proper Bayesian interval estimators.

We now state the relationship between conditional and admissibility properties for interval estimation.

PROPOSITION 8.2. *For an interval estimator $\langle I(x), \alpha(x) \rangle$ of a parameter θ, the absence of semirelevant betting procedures implies admissibility with respect to squared-error loss, and admissibility with respect to squared-error loss implies the absence of relevant betting procedures.*

PROOF. If $s(x)$ is relevant, define $\beta(x) = \alpha(x) + \varepsilon s(x)$.

$$E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}^2\right] - E\left[\{\chi_{I(X)}(\theta) - \beta(X)\}^2\right]$$
$$\geqslant 2\varepsilon E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|\right].$$

Hence $\langle I(x), \alpha(x) \rangle$ is inadmissible.

If $\langle I(x), \alpha(x) \rangle$ is inadmissible, note that

$$E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}^2\right] - E\left[\{\chi_{I(X)}(\theta) - \beta(X)\}^2\right]$$
$$\leqslant 2E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}\{\beta(X) - \alpha(X)\}\right].$$

The betting procedure $\beta(x) - \alpha(x)$ is semirelevant.

One way of viewing the relationship between admissibility with respect to squared-error loss and conditional properties for interval estimation is to consider a concept which includes both of them. Suppose that for some γ there is a number $\varepsilon > 0$ and a betting procedure $s(x)$ such that

$$E\left[\{\chi_{I(X)}(\theta) - \alpha(X)\}s(X) - \varepsilon|s(X)|^\gamma\right] \geqslant 0$$

for all θ with strict inequality for some θ. For $\gamma = 0$ we have the condition that $s(x)$ is super-relevant. For $\gamma = 1$ we have that $s(x)$ is relevant. For $\gamma = 2$ we have that $\langle I(x), \alpha(x) \rangle$ is inadmissible with respect to squared-error loss (since $\beta(x) = \alpha(x) + s(x)$ has smaller squared-error loss). For $\gamma = \infty$, interpreting this to mean that $|s(x)|^\gamma = 0$ whenever $|s(x)| < 1$, we have that a semirelevant betting procedure exists. The generalisation seems to have little statistical interest other than in these special cases.

REFERENCES

[1] BARTLETT, M. S. (1956). Comment on Sir Ronald Fisher's paper. *J. Roy Statist. Soc. Ser. B* **18** 295–296.

[2] BIRNBAUM, A. (1962). On the foundations of statistical inference. *J. Amer. Statist. Assoc.* **57** 269–306.

[3] BONDAR, J. V. (1977). A conditional confidence principle. *Ann. Statist.* **5** 881–891.

[4] BREWSTER, J. F. and ZIDEK, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38.

[5] BROWN, L. D. (1967). The conditional level of Student's test. *Ann. Math. Statist.* **38** 1068–1071.

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

[7] BUEHLER, R. J. and FEDDERSON, A. P. (1963). Note on a conditional property of Student's *t*. *Ann. Math. Statist.* **34** 1098–1100.

[8] COX, D. R. (1958). Some problems connected with statistical inference. *Ann. Math. Statist.* **29** 357–372.

[9] DEMPSTER, A. P. (1964). On the difficulties inherent in Fisher's fiducial argument. *J. Amer. Statist. Assoc.* **59** 56–66.

[10] FISHER, R. A. (1956a). *Statistical Methods and Scientific Inference.* Oliver and Boyd, Edinburgh.

[11] FISHER, R. A. (1956b). On a test of significance in Pearson's *Biometrika* Table No. 11. *J. Roy Statist. Soc. Ser. B* **18** 56–60.

[12] GODAMBE, V. P. (1961). Bayesian shortest confidence intervals and admissibility. *Bull. Inst. Internat. Statist. 33rd Session (Paris)* **125** 1–5.

[13] HACKING, I. (1965). *Logic of Statistical Inference.* Cambridge University.

[14] JONES, H. L. (1958). Inadmissible samples and confidence limits. *J. Amer. Statist. Assoc.* **53** 482–490.

[15] NEYMAN, J. (1956). Note on an article by Sir Ronald Fisher. *J. Roy Statist. Soc. Ser. B* **18** 288–294.

[16] PIERCE, D. A. (1973). On some difficulties in a frequency theory of inference. *Ann. Statist.* **1** 241–250.

[17] ROBINSON, G. K. (1976). Conditional properties of Student's *t* and of the Behrens-Fisher solution to the two means problem. *Ann. Statist.* **4** 963–971.

[18] ROBINSON, G. K. (1979). Conditional properties of statistical procedures for location and scale parameters. *Ann. Statist.* **7** 756–771.

[19] SAVAGE, L. J. (1962). *The Foundations of Statistical Inference: A Discussion.* Methuen, London.

[20] STONE, M. (1976). Strong inconsistency from uniform priors. *J. Amer. Statist. Assoc.* **71** 114–116.

[21] WALLACE, D. L. (1959). Conditional confidence level properties. *Ann. Math. Statist.* **30** 864–876.

[22] WELCH, B. L. (1947). Generalization of Student's problem. *Biometrika* **34** 28–35.

[23] WELCH, B. L. (1956). Note on some criticisms made by Sir Ronald Fisher. *J. Roy Statist. Soc. Ser. B* **18** 297–302.

[24] ZACKS, S. (1971). *The Theory of Statistical Inference.* Wiley, New York.

AUSTRALIAN ROAD RESEARCH BOARD
P.O. BOX 156, BAG 4,
NUNANADING 3131
VICTORIA, AUSTRALIA