

A CONDITIONAL PROPERTY OF INVARIANT CONFIDENCE AND PREDICTION REGIONS

BY PETER M. HOOPER¹

University of Alberta

This paper considers estimation and prediction problems invariant under an amenable group that is transitive on the parameter space. It is shown that an invariant confidence or prediction region does not admit super-relevant betting procedures if and only if its level of coverage is defined conditionally given the maximal invariant statistic.

1. Introduction. Robinson (1979) presents a systematic set of definitions of conditional properties for interval and point estimators. The conditional properties are formulated in terms of an estimator's ability to withstand betting procedures. This formulation follows the approach of Buehler (1959). It is assumed that the reader is familiar with these concepts. In this paper we present a theorem which guarantees the absence of super-relevant betting procedures for a large class of invariant confidence and prediction regions. The result can also be applied to invariant conditional confidence procedures; see Kiefer (1977).

The theorem is closely related to a result of Bondar (1977). Bondar introduced a consistency principle which may be restated as follows: use no set estimator that admits a super-relevant subset. Bondar proved that, under regularity conditions, in estimation problems invariant under an amenable group that is transitive on the parameter space, set estimators that are exact improper Bayes with respect to right Haar measure do not admit super-relevant subsets. Bondar's proof is easily modified to show that all super-relevant betting procedures are excluded. Bondar's Theorem is applicable to confidence sets in the sense of Neyman. Theorem 2.1 of Hora and Buehler (1966) shows that, under regularity conditions, in estimation problems invariant under a group that is transitive on the parameter space, the following two statements are equivalent for invariant set estimators: (i) the set estimator is a level α confidence set with level determined conditionally given the group orbit in the sample space; (ii) the set estimator is level α improper Bayes with respect to right Haar measure. Stein (1965, page 223) showed that (ii) implies (i). The above results combined show that, under regularity conditions, in estimation problems invariant under an amenable group that is transitive on the parameter space, invariant level α confidence sets do not admit super-relevant betting procedures provided their level is determined conditionally given the orbit in the sample space. Our theorem extends the above statement by including prediction problems, by allowing the conditional level of coverage to vary, and by requiring fewer regularity conditions.

Received July 1982; revised October 1983.

¹ Research supported in part by the Natural Sciences and Engineering Research Council of Canada.

AMS 1980 subject classifications. Primary 62A05, 62F25.

Key words and phrases. Amenable, ancillary, betting procedure, super-relevant betting procedure.

2. The theorem. Let X and Y be random variables, not necessarily independent, taking values respectively in \mathcal{X} and \mathcal{Y} , with joint distribution lying in a given family $\{P_\theta : \theta \in \Omega\}$. We define a *set predictor* for Y given X to be a pair $\langle \phi, \alpha \rangle$ of measurable functions, $\phi : \mathcal{X} \times \mathcal{Y} \rightarrow \{0, 1\}$ and $\alpha : \mathcal{X} \rightarrow [0, 1]$. This extends the definition of interval estimator given by Robinson (1979, page 744) since the problem of estimating a parametric function $\gamma(\theta)$ can be regarded formally as a prediction problem in which $P_\theta\{Y = \gamma(\theta)\} = 1$. The function ϕ determines the set $R_\phi(x) = \{y \in \mathcal{Y} : \phi(x, y) = 1\}$ used to predict Y when $X = x$. The number $\alpha(x)$ can be regarded as an estimate of the conditional probability of coverage $E_\theta\{\phi(X, Y) | X = x\}$. If α is taken to be a constant, defined by $\alpha = \inf\{E_\theta\phi(X, Y) : \theta \in \Omega\}$, then $R_\phi(X)$ is called a level α prediction region. A *betting procedure* is a bounded measurable function $s : \mathcal{X} \rightarrow \mathbb{R}$. A betting procedure s is *super-relevant* for $\langle \phi, \alpha \rangle$ if, for some $\epsilon > 0$,

$$(2.1) \quad E_\theta[\{\phi(X, Y) - \alpha(X)\}s(X)] \geq \epsilon \quad \text{for all } \theta \in \Omega.$$

Our invariance assumptions are similar to those of Takada (1982). Let G be an invariance group acting on $\mathcal{X} \times \mathcal{Y}$ which leaves invariant $\{P_\theta : \theta \in \Omega\}$. We assume that G is transitive on Ω . We also assume that actions are induced on \mathcal{X} ; i.e., actions on $\mathcal{X} \times \mathcal{Y}$ take the form $(x, y) \rightarrow (gx, [g, x]y)$. Let $T : \mathcal{X} \rightarrow \mathcal{T}$ be sufficient for the family $\{P_\theta^X : \theta \in \Omega\}$ of distributions of X and suppose that X and Y are conditionally independent given T ; i.e., T is an adequate statistic (Skibinsky, 1967) or equivalently $T, (T, Y)$ is sufficient and transitive for $X, (X, Y)$ (Hall et al., 1965, Theorem 4.1). We also assume that T induces actions on \mathcal{T} and $\mathcal{T} \times \mathcal{Y}$; i.e., $T(x_1) = T(x_2)$ implies $T(gx_1) = T(gx_2)$ for all $g \in G$ and actions $[g, x]$ on \mathcal{Y} depend on x only through $T(x)$. Note that one may take $T(X) = X$ if desired. Let $W = W(T)$ be a maximal invariant on \mathcal{T} . We consider invariant set predictors $\langle \phi, \alpha \rangle$ based on T ; i.e., $\phi(X, Y) = h(T(X), Y)$, with h invariant, and $\alpha(X) = k(W(T(X)))$. Of particular interest is $k(W) = E\{\phi(X, Y) | W\}$. Here k is well defined since the invariance of ϕ and W and the transitivity assumption imply that $E_\theta\{\phi(X, Y) | W\}$ is the same for all $\theta \in \Omega$. Usually it is possible to define ϕ so that $\alpha = E\{\phi(X, Y) | W\}$ is constant, in which case $R_\phi(X)$ is an exact level α prediction region.

We assume that G satisfies the Hunt-Stein condition: there exists a sequence of asymptotically right invariant probability distributions over G . For groups encountered in all current statistical applications, the Hunt-Stein condition is equivalent to amenability. Bondar and Milnes (1981) give a survey of conditions related to amenability and provide a list of amenable groups useful in statistical applications.

The following regularity conditions will be assumed. Let \mathcal{A} and \mathcal{B} denote σ -fields of, respectively, \mathcal{T} and G . We assume that \mathcal{A} is countably generated, that $\{(t, g) : gt \in A\} \in \mathcal{A} \times \mathcal{B}$ for each $A \in \mathcal{A}$, and that $Bg \in \mathcal{B}$ for each $B \in \mathcal{B}, g \in G$. Suppose there is a σ -finite measure ν on \mathcal{B} satisfying $\nu(B) = 0$ implies $\nu(Bg) = 0$ for all $B \in \mathcal{B}, g \in G$.

THEOREM. *Let G be an invariance group acting on $\mathcal{X} \times \mathcal{Y}$, with actions*

induced on \mathcal{X} , and suppose G is transitive on Ω and satisfies the Hunt-Stein Condition. Let $T: \mathcal{X} \rightarrow \mathcal{T}$ be sufficient for $\{P_\theta^X: \theta \in \Omega\}$ with X and Y conditionally independent given T and with actions induced on \mathcal{T} and $\mathcal{T} \times \mathcal{Y}$. Let $\langle \phi, \alpha \rangle$ be a set predictor of the form $\phi(X, Y) = h(T(X), Y)$, with h invariant, and $\alpha(X) = k(W(T(X)))$, where W is a maximal invariant on \mathcal{T} . Under the above regularity assumptions there are no super-relevant betting procedures for $\langle \phi, \alpha \rangle$ if and only if $k(W)$ is a version of $E\{\phi(X, Y) | W\}$.

PROOF. If $k(W)$ is not a version of $E\{\phi(X, Y) | W\}$, define $s(X) = s_1(W)$ to be 1, 0, or -1 respectively as $E\{\phi(X, Y) | W\} - k(W)$ is $>$, $=$, or $<$ 0. We have

$$E[\{\phi(X, Y) - \alpha(X)\}s(X)] = E[\{E(\phi(X, Y) | W) - k(W)\}s_1(W)] > 0,$$

which implies that s is super-relevant for $\langle \phi, \alpha \rangle$ since the expectation is free of θ .

Conversely suppose $k(W) = E\{\phi(X, Y) | W\}$ a.s. Suppose s is a betting procedure satisfying (2.1) for $\epsilon > 0$. Put $a(T, Y) = h(T, Y) - k(W)$ and $b(T) = E\{s(X) | T\}$, the latter being free of θ since T is sufficient. Observe that

$$\begin{aligned} E_\theta[\{\phi(X, Y) - \alpha(X)\}s(X)] &= E_\theta[a(T, Y)E_\theta\{s(X) | T, Y\}] \\ (2.2) \qquad \qquad \qquad &= E_\theta[a(T, Y)b(T)]. \end{aligned}$$

The second equality holds because X and Y are conditionally independent given T ; see Hall et al. (1965, Theorem 2.1). Fix $\theta^* \in \Omega$. Using (2.1), (2.2), the transitivity of G on Ω , and the invariance of a , we obtain

$$\begin{aligned} 0 < \epsilon &\leq \inf_{\theta \in \Omega} E_\theta[a(T, Y)b(T)] \\ (2.3) \qquad \qquad &= \inf_{g \in G} E_{\theta^*}[a(T, Y)b(T)] \\ &= \inf_{g \in G} E_{\theta^*}[a(T, Y)b(gT)]. \end{aligned}$$

Let $\{\nu_n\}$ denote a sequence of asymptotically right invariant probability measures on \mathcal{B} and define

$$(2.4) \qquad \qquad \psi_n(t) = \int_G b(gt)\nu_n(dg).$$

Since b is a.s. bounded, the weak compactness theorem (Lehmann, 1959, page 354) shows that there exists a subsequence $\{n_i\}$ and a bounded measurable function $\psi: \mathcal{T} \rightarrow \mathbb{R}$ such that $\lim_{i \rightarrow \infty} E_{\theta^*}[p(T)\psi_{n_i}(T)] = E_{\theta^*}[p(T)\psi(T)]$ for all functions p with $E_{\theta^*}|p(T)| < \infty$. In particular, taking $p(T) = E_{\theta^*}\{a(T, Y) | T\}$ we obtain

$$(2.5) \qquad \qquad \lim_{i \rightarrow \infty} E_{\theta^*}\{a(T, Y)\psi_{n_i}(T)\} = E_{\theta^*}\{a(T, Y)\psi(T)\}.$$

An argument essentially the same as that at Equation 15 (Lehmann, 1959, page 336) establishes that ψ is almost invariant. Then Theorem 4 (Lehmann, 1959, page 225) shows that ψ is equivalent to an invariant function; i.e., $\psi(T) = \psi_0(W)$ a.s. Using Fubini's theorem, (2.4), (2.5), and the fact that $E\{a(T, Y) | W\} = 0$

a.s., we obtain

$$\begin{aligned} \lim_{i \rightarrow \infty} \int_G E_{\theta^*} \{a(T, Y)b(gT)\} \nu_{n_i}(dg) \\ = \lim_{i \rightarrow \infty} E_{\theta^*} \{a(T, Y)\psi_{n_i}(T)\} = E\{a(T, Y)\psi_0(W)\} \\ = E[E\{a(T, Y) \mid W\}\psi_0(W)] = 0. \end{aligned}$$

However by (2.3) each term in the above sequence is bounded below by ε . This establishes the desired contradiction. \square

REMARK 1. The first part of the proof is easily modified to show that if $\alpha(X)$ is not a version of $E\{\phi(X, Y) \mid W\}$ then $\langle \phi, \alpha \rangle$ admits a super-relevant subset.

REMARK 2. If X and Y are independent then X and Y are conditionally independent given T , for any function T of X . In this situation the actions $[g, x]$ on \mathcal{Y} usually do not involve x . If we have $P_\theta\{Y = \gamma(\theta)\} = 1$ for all $\theta \in \Omega$, which is the case in estimation problems, then the assumption that the action of G on $\mathcal{X} \times \mathcal{Y}$ leaves invariant $\{P_\theta : \theta \in \Omega\}$ is equivalent to the assumption that the action of G on \mathcal{X} leaves invariant $\{P_\theta^X : \theta \in \Omega\}$ and that $\gamma(\theta_1) = \gamma(\theta_2)$ implies $\gamma(g\theta_1) = \gamma(g\theta_2)$ for all $g \in G$; i.e., $\gamma : \Omega \rightarrow \mathcal{Y}$ induces an action of G on \mathcal{Y} defined by $\gamma(g\theta) = g\gamma(\theta)$. Hora and Buehler (1966) call a parametric function with this property invariantly estimable.

3. Applications. Before illustrating the use of the theorem in several examples, we mention a point of possible confusion. In multivariate problems one often uses procedures which are invariant under the group $GL(p)$ of $p \times p$ invertible matrices. The theorem cannot be applied using this group since $GL(p)$ is not amenable. However the subgroup $LT(p)$ of $p \times p$ lower triangular matrices with positive diagonal elements is amenable. Often the theorem can be applied to the class of $LT(p)$ -invariant procedures, which of course includes the class of $GL(p)$ -invariant procedures.

EXAMPLE 1. *Location and scale family.* Let $X_1, \dots, X_n, n \geq 2$, be independent identically distributed random variables with common density $\sigma^{-1}f((x - \gamma)/\sigma)$, where f is known but $\gamma \in \mathbb{R}$ and $\sigma > 0$ are unknown. Let G be the affine group with actions $x_i \rightarrow ax_i + b, \gamma \rightarrow a\gamma + b, \sigma \rightarrow a\sigma$, for $a > 0, b \in \mathbb{R}$. The vector of order statistics $(X_{(1)}, \dots, X_{(n)})$ is sufficient and a maximal invariant is $W = ((X_{(2)} - X_{(1)})/(X_{(n)} - X_{(1)}), \dots, (X_{(n-1)} - X_{(n-2)})/(X_{(n)} - X_{(1)}))$. The theorem shows that an invariant level α confidence region for γ and/or σ based on the order statistics does not admit super-relevant betting procedures if and only if its conditional level given W is identically α . The same holds true concerning invariant level α prediction regions for independent future observations.

EXAMPLE 2. *Multivariate regression.* Let $M(m, p)$ denote the set of $m \times p$ matrices. Suppose we have $Z \sim N_{m \times p}(C\beta, I_m \otimes \Sigma)$, with $C \in M(m, q)$ known,

$\beta \in M(q, p)$ unknown, and Σ an unknown $p \times p$ positive definite symmetric matrix. Thus the rows of Z are independent p -variate normal random vectors with common covariance matrix Σ . Partition $Z = (Z_{ij})$, with $Z_{ij} \in M(m_i, p_j)$, $m_1 + m_2 = m$, $p_1 + p_2 = p$, and $C' = (C'_1 : C'_2)$, with $C_i \in M(m_i, q)$. We assume that C_1 has full rank q . Suppose we wish to predict $Y = Z_{22}$ having observed $X = (Z_{11}, Z_{12}, Z_{21})$. The problem is invariant under the amenable group $LT(p) \times M(q, p)$, with actions defined by $Z \rightarrow ZA' + Cb$, $\beta \rightarrow \beta A' + b$, $\Sigma \rightarrow A \Sigma A'$ for $A \in LT(p)$ and $b \in M(q, p)$. Define

$$B = [B_1 : B_2] = (C'_1 C_1)^{-1} C'_1 [Z_{11} : Z_{12}]$$

and

$$S = (S_{ij}) = [Z_{11} : Z_{12}]' (I_{m_1} - C_1 (C'_1 C_1)^{-1} C'_1) [Z_{11} : Z_{12}].$$

We observe that $T = (B, S, Z_{21})$ is sufficient for the family of distributions of X . The actions on T and Y are: $B \rightarrow BA' + b$, $S \rightarrow ASA'$, $Z_{21} \rightarrow Z_{21}A'_{11} + C_2 b_1$, $Y = Z_{22} \rightarrow Z_{21}A'_{21} + Z_{22}A'_{22} + C_2 b_2$ for $A = (A_{ij}) \in LT(p)$ and $b = [b_1 : b_2] \in M(q, p)$. A maximal invariant function of T is given by $W = (Z_{21} - C_2 B_1) L'_{11}{}^{-1}$, where $S_{11} = L_{11} L'_{11}$, $L_{11} \in LT(p_1)$. The theorem shows that an invariant level α prediction region for Y based on T does not admit super-relevant betting procedures if and only if its conditional level given W is identically α .

EXAMPLE 3. *The general multivariate analysis of variance model.* Sufficiency and invariance reductions for the GMANOVA confidence estimation problem are described in Hooper (1982a). Corresponding reductions under an amenable subgroup show that, if only invariant regions are of interest, it suffices to consider a simpler analysis of covariance model. The reader is referred to the notation of Hooper (1982b, Section 5), where optimal confidence regions are derived. The theorem is applied using the amenable group $G_1 = M(m, p) \times \mathcal{A}_1$. A maximal invariant function of (X_1, X_2, S) is given by $W = U_2 = X_2 L'_{22}{}^{-1}$. Thus an invariant level α confidence region does not admit super-relevant betting procedures if and only if its conditional level given W is identically α . This is the case for all confidence regions for M based on $U_1(X, M)$ only, since $U_1(X, M)$ and U_2 are independent under $\theta = (M, \Sigma)$. In particular, this is true of confidence regions based on $T_1 = U_1 U'_1$. In the MANOVA problem we have $q = 0$ and so all invariant confidence regions do not admit super-relevant betting procedures.

REFERENCES

- BONDAR, J. V. (1977). A conditional confidence principle. *Ann. Statist.* **5** 881–891.
 BONDAR, J. V. and MILNES, P. (1981). Amenability: a survey for statistical applications of Hunt-Stein and related conditions on groups. *Z. Wahrsch. verw. Gebiete* **57** 103–128.
 BUEHLER, R. J. (1959). Some validity criteria for statistical inference. *Ann. Math. Statist.* **30** 845–863.
 HALL, W. J., WIJSMAN, R. A., and GHOSH, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.
 HOOPER, P. M. (1982a). Sufficiency and invariance in confidence set estimation. *Ann. Statist.* **10** 549–555.

- HOOPER, P. M. (1982b). Invariant confidence sets with smallest expected measure. *Ann. Statist.* **10** 1283–1294.
- HORA, R. B. and BUEHLER, R. J. (1966). Fiducial theory and invariant estimation. *Ann. Math. Statist.* **37** 643–656.
- KIEFER, J. (1977). Conditional confidence statements and confidence estimators. *J. Amer. Statist. Assoc.* **72** 789–827.
- LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.
- ROBINSON, G. K. (1979). Conditional properties of statistical procedures. *Ann. Statist.* **7** 742–755.
- SKIBINSKY, M. (1967). Adequate subfields and sufficiency. *Ann. Math. Statist.* **8** 155–161.
- STEIN, C. M. (1965). The approximation of improper prior measures by prior probability measures. *Bernoulli-Bayes-Laplace Festschrift* 217–240. Springer, Berlin.
- TAKADA, Y. (1982). A comment on best invariant predictors. *Ann. Statist.* **10** 971–978.

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA T6G 2G1
CANADA