

## BOUNDS FOR THE FREQUENCY OF MISLEADING BAYES INFERENCE

BY D. KERRIDGE

*University of Sheffield*

A misleading Bayes inference is one which either assigns a low probability to a true hypothesis, or a high probability to a false hypothesis. If a Bayesian assumes that each of a finite number of simple hypotheses has equal prior probability, then an upper bound exists for the frequency with which his inferences will be misleading.

The result to be discussed may be regarded as a generalisation of Wald's [2], p. 41, bounds for the frequency of errors of each kind in a sequential probability ratio test. Here, however, in addition to allowing any finite number of hypotheses, the stopping rule is completely general: it need not even have probability one of terminating.

Suppose that there are  $k$  simple hypotheses, a true hypothesis  $T$  and false hypotheses  $F_i (i = 1, 2, \dots, k - 1)$ . Let  $X_n$  denote the complete set of observations available at the  $n$ th stage of sampling: if the sampling is terminated at the  $r$ th stage,  $X_n \equiv X_r$  for all  $n \geq r$ .  $\sum^*$  denotes summation only over those  $X_n$  for which both (i) the sampling has terminated, and (ii) the posterior probability of  $T$  is  $p$  or less. Then if  $R_n = \sum_i P(X_n | F_i) / P(X_n | T)$ , applying Bayes Theorem,  $R_n = \sum_i P(F_i | X_n) / P(T | X_n)$ , since the prior probabilities are assumed equal. Hence  $\sum^* R_n P(X_n | T) \geq [(1 - p) / p] \sum^* P(X_n | T)$ . But  $\sum^* R_n P(X_n | T) = \sum_i \{ \sum^* P(X_n | F_i) \} \leq k - 1$ . Hence  $\sum^* P(X_n | T) \leq (k - 1)p / (1 - p)$ : but  $\sum^* P(X_n | T)$  is a monotone increasing function of  $n$ . Hence  $\lim_{n \rightarrow \infty} \sum^* P(X_n | T)$  exists and  $\leq (k - 1)p / (1 - p)$ , that is: *The frequency with which, at the termination of sampling the posterior probability of the true hypothesis is  $p$  or less cannot exceed  $(k - 1)p / (1 - p)$ .*

Since a probability of  $p$  or more for a false hypothesis implies  $(1 - p)$  or less for the true hypothesis, an upper bound of  $(k - 1)(1 - p) / p$  for the frequency of misleading inferences of this kind is an obvious corollary: this may be a more useful form in practice, although the inequality is weaker.

For  $k = 2$ , the strictness of the inequality is quite remarkable, especially in view of the completely general stopping rule which is allowed. For example, the frequency with which a true hypothesis is found to have posterior probability  $\frac{1}{100}$  or less cannot be more than  $\frac{1}{99}$ . This remains true even if the stopping rule were biased against the true hypothesis by requiring, to take an extreme case, that the sampling should continue until its posterior probability becomes  $\frac{1}{100}$  or less if it ever does: in such a case the probability of terminating is not greater than  $\frac{1}{99}$ . A particular example of this type has been considered by Savage [1], p. 72.

For larger  $k$  the inequality is less sharp, but it still remains true, at any rate

---

Received December 3, 1962.