Bayesian Testing of Nonparametric Hypotheses and Its Application to Global Optimization

B. Betro

Communicated by L. C. W. Dixon

Abstract. Random distribution functions are the basic tool for solving nonparametric decision-theoretic problems. In 1974, Doksum introduced the family of distributions neutral to the right, that is, distributions such that 
\[ F(t_1), \frac{F(t_2) - F(t_1)}{1 - F(t_1)}, \ldots, \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \] are independent whenever \( t_1 < \cdots < t_k \). In practice, application of distributions neutral to the right has been prevented by the lack of a manageable analytical expression for probabilities of the type \( P(F(t) < q) \) for fixed \( t \) and \( q \). A subclass of such distributions can be provided which allows for a close expression of the characteristic function of \( \log[1 - F(t)] \), given the sample. Then, the a posteriori distribution of \( F(t) \) is obtained by numerical evaluation of a Fourier integral. As an application, the global optimization problem is formulated as a problem of inference about the quantiles of the distribution \( F(y) \) of the random variable \( Y = f(X) \), where \( f \) is the objective function and \( X \) is a random point in the search domain.

Key Words. Bayesian testing, nonparametric inference, random distribution functions, global optimization.

1. Introduction

In 1974, Doksum (Ref. 1) introduced a class of random probabilities \( P \), called neutral to the right, which can be useful for solving nonparametric decision-theoretic problems. Such class, defined through the property of independence of the normalized increments

\[ F(t_1), \frac{F(t_2) - F(t_1)}{1 - F(t_1)}, \ldots, \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}, \]

1 The author thanks J. Koronacki and R. Zielinski of the Polish Academy of Sciences for their valuable criticism during the final draft of the paper.

2 Researcher, Istituto per le Applicazioni della Matematica e dell'Informatica, CNR, Milano, Italy.
for all $t_1 < \cdots < t_k$, where

$$ F(t) = P\{(-\infty, t]\}, $$

has the following nice features:

(a) random probabilities neutral to the right can be chosen so that their support is wide enough;

(b) the posterior distribution of a random probabilities neutral to the right is neutral to the right;

(c) random probabilities neutral to the right can be chosen so that the posterior probabilities of $P(A)$, given a sample $X_1, \ldots, X_N$ from $P$, depends not only on the number $N_A$ of $X$'s that fall in $A$ but also on where they fall within or outside $A$.

Within the class of probabilities neutral to the right, apart from rather trivial cases, only the Dirichlet process, introduced by Ferguson (Ref. 2), does not enjoy property (c). For this reason, the posterior distribution of this process is very easy to handle and is more widely used in the applications.

For the subclass of the so-called gamma processes, the characteristic function of the \textit{a posteriori} distribution of $\log(1 - F(t))$, given the sample, can be provided in an analytical form; hence, the \textit{a posteriori} distribution can be obtained via numerical evaluation of a Fourier integral. Once the distribution of $F(t)$, given the sample, is known, testing of nonparametric hypotheses (like hypotheses involving quantiles) can be performed in the framework of Bayesian statistics.

The interest of the author in the subject of Bayesian testing of nonparametric hypotheses was motivated by his previous studies about the global optimization problem. It has been recognized widely that the problem can be approached successfully in the framework of random sampling, but the evaluation of the goodness of the results obtained is still committed to rather empirical rules.

The formulation of the global optimization problem as a problem of Bayesian testing about quantiles of the distribution of the sampled values can provide a major step toward a more satisfactory control of global optimization algorithms.

\section{Random Distributions Neutral to the Right}

A precise definition of neutrality to the right for random distributions $F$ can be given as follows.