The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 157–168

ON DISTRIBUTION-FREE STATISTICAL INFERENCE WITH UPPER AND LOWER PROBABILITIES[1]

BY R. J. BERAN

University of California, Berkeley

1. Introduction. This paper sets forth a new theory of distribution-free inference for the general statistical model

$$\mathbf{x} = T_\theta^{-1}\mathbf{e}. \tag{1.1}$$

By assumption, $\{T_\theta : \theta \in \Omega\}$ is a known family of nonsingular transformations mapping R^N into R^N, Ω is a Borel subset of a real Euclidean space, $\mathbf{x} = (x_1, \cdots, x_N)$ is a sample of N observations, and the components of $\mathbf{e} = (e_1, \cdots, e_N)$ are realized values of N independent, identically distributed random variables with common continuous distribution function F on the real line.

The scientific background for model (1.1) is as follows. An experiment is performed, resulting in N measurements $\mathbf{x}$. The observed $\mathbf{x}$ are generated from underlying realized errors $\mathbf{e}$ by the transformation (1.1). While the vector $\mathbf{x}$ is an observed constant, the values of $\mathbf{e}$ and θ giving rise to $\mathbf{x}$ through (1.1) are unknown. Our main goal in this paper is to draw inferences about the unknown constant θ from the observed $\mathbf{x}$ and the model, first with no knowledge of the distribution function F, apart from continuity, and secondly under the additional assumption that F is symmetric about the origin.

The theory described in this paper both extends and applies ideas introduced by Fraser [3] and by Dempster [1]. While technically the results fall within the general framework of Dempster's upper and lower probabilities, the statistical rationale differs. Upper and lower probabilities are introduced in Section 2 to measure the reliability of certain simple decision procedures, with reliability being assessed by relative frequency of success. This approach gives a well-defined statistical interpretation to upper and lower probabilities and has the advantage of leading naturally to a decision theory.

Model (1.1) encompasses many of the common models treated in classical nonparametric theory. As examples, one and two-sample versions of location models, scale models, regression models, and auto-regressive models are analyzed in Section 3 and Section 4. In particular, an optimality property is established for the one and two-sample Wilcoxon tests, for the Hodges and Lehmann [4] location parameter estimates (based on the Wilcoxon tests), for a test of scale parameter studied by Sukhatme [6], and for a regression parameter estimate proposed by Theil [5] and Sen [7].

Received November 13, 1969; revised July 1970.

[1] This research was partially supported by a University of California Summer Faculty Fellowship and National Science Foundation Grant GP-15283.