The Annals of Statistics
1987, Vol. 15, No. 1, 432–442

MINIMAX ESTIMATION OF THE MEAN OF A GENERAL DISTRIBUTION WHEN THE PARAMETER SPACE IS RESTRICTED

BY AVRAHAM A. MELKMAN AND YA'ACOV RITOV

Ben Gurion University of the Negev and Hebrew University of Jerusalem

1. Introduction. We consider the one-dimensional additive model $Y = \vartheta + X$. If X is a (standard) normal random variable and ϑ is completely unknown then of course $\delta(y) = y$ is the minimax estimator. This same estimator is no longer minimax, however, given the added prior information $|\vartheta| \le s$. In fact, the minimax estimate is then Bayes with respect to a least favorable prior distribution that is supported on $[-s, s]$. This distribution was investigated by Casella and Strawderman (1981) for small values of s, and by Bickel (1981) and Levit (1980a–c) and Levit and Berhin (1980) for large values of s. Our interest was captured particularly by Bickel's somewhat surprising result that if the least favorable distributions are rescaled to $[-1, 1]$ then they converge weakly, as $s \to \infty$, to a distribution with density $\cos^2(\pi x/2)$ [the distribution with minimum Fisher information among all those supported on $[-1, 1]$, see Huber (1974)], and the corresponding minimax risks behave like $1 - \pi^2/s^2 + o(1/s^2)$. Moreover, Bickel produced a family of estimates that have this risk asymptotically, and proved that they have the property that $s(y - \delta(y))$ is approximately $\pi \tan(\pi y/(2s))$.

The main point of this paper is that all the above mentioned results of Bickel (1981) remain valid without the normality assumption. Namely, all that is needed is that the specified distribution of X be such that $EX = 0$, $EX^2 = 1$ and $EX^4 < \infty$. We prove actually a slightly stronger result, Theorem 2, wherein X may be any member of a family of distributions that satisfies a weakened set of requirements. We do not know, however, whether these requirements are necessary, except for the fact that some moment higher than the second has to be bounded.

Suppose $Y_i = \vartheta + X_i$, $i = 1, \dots, n$. The results of this paper can then be used if we replace the vector of observations by a one-dimensional statistic that preserves the translation structure, e.g., the sample mean. If the distribution of X_1 is specific enough we may use the best invariant estimator for ϑ, i.e., the Pitman estimator. When this estimator is also a sufficient statistic then the above reduction does not lose any information. If, however, $Y_i = \vartheta_i + X_i$, $i = 1, \dots, n$, then no easy reduction is possible and only partial results are known. Speckman (1982) restricted attention to linear estimates for the vector of means and then found the form of the minimax estimate. Melkman and Micchelli (1979)

Received February 1985; revised May 1986.
AMS 1980 *subject classifications*. 62F10, 62C20.
Key words and phrases. Minimax, Fisher information.