Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

CONFIDENCE REGIONS IN BROKEN LINE REGRESSION

BY DAVID O. SIEGMUND AND HEPING ZHANG
Stanford University and Yale University

The broken line regression model is viewed as a special case of nonlinear regression. Following the methodology of Knowles, Siegmund, and Zhang (1991), we discuss procedures for constructing confidence regions. Our method involves inversion of the likelihood ratio test. A slightly conservative bound is obtained for the level of the test given the values of statistics which are sufficient for the nuisance parameters when the parameters of interest are fixed. We use a number of published data sets and simulations to compare our method with the approximate F method, which is based on the assumption that a formal analogue of the F statistic has approximately an F distribution, and with the Bayesian method of Smith and Cook (1980).

1. Introduction. The broken line regression model

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 (x_i - \theta)^+ + \varepsilon_i, \tag{1}$$

where $a^+ = \max(a, 0)$ and ε_i $(i = 1, \cdots, m)$ are independent $N(0, \sigma^2)$, has been discussed by a number of authors. Some of these consider the model (1) as a special case of nonlinear regression (e.g., Ratkowsky (1983) p. 122 ff., and Seber and Wild (1989) p. 447 ff.), while others have addressed it directly (e.g., Hinkley (1971), Feder (1975), Smith and Cook (1980)). An interesting special case considered by Hinkley (1971) is

$$y_i = \alpha - \beta (x_i - \theta)^- + \varepsilon_i \quad (i = 1, \cdots, m), \tag{2}$$

where $a^- = -(a - a^+)$. In (2) the parameters θ and α have natural interpretations: α is the maximum mean response produced by an input x, and θ is the minimum input required to produce this expected response.

Direct asymptotic analysis of (1) leads to technical difficulties, due to the lack of smoothness of the likelihood function (Hinkley (1971), Feder (1975)); and as we show by an example below, Hinkley's asymptotic version of the

---

AMS 1991 Subject Classification: Primary 62J02; Secondary 53A07

Key words and phrases: Approximate F method, broken line regression, conditional likelihood ratio, upcrossing.