

Discussion of Dr. Silverman's Paper

by

R. L. Eubank

Technical Report No. SMU-DS-TR-190
Department of Statistics ONR Contract

December 1984

Research sponsored by the Office of Naval Research
Contract N00014-82-K-0207
Project NR 042-479

Reproduction in whole or in part is permitted for
any purpose of the United States Government

This document has been approved for public
release and sale; its distribution is unlimited.

Department of Statistics
Southern Methodist University
Dallas, Texas 75275

Contribution to the
Discussion of Dr. Silverman's Paper

October 1984

R. L. Eubank

Southern Methodist University

Let me begin by thanking Professor Silverman for an interesting and thought-provoking paper. His use of regression type diagnostics for smoothing splines coincides with many of my own thoughts on this subject (Eubank 1984a, b). The close connection between smoothing splines and polynomial regression leads one to believe that diagnostics appropriate for use with smoothing splines should resemble those currently in use by regression analysts. In this regard, it is well known that diagnostic procedures should include information about the design as well as the fit. Design diagnostics for smoothing splines are provided by the leverage values, $A_{ii}(\alpha)$. It can be shown that $0 \leq A_{ii}(\alpha) \leq 1$ and that a leverage value too near one indicates a sensitive point in the design where an observation will tend to dominate its own fit. A diagnostic which encompasses both information about an observation's leverage as well as its fit is (in Silverman's notation)

$$DFITS_i = |A_{ii}(\alpha)/(1-A_{ii}(\alpha))|^{1/2} |r_i|, \quad i=1, \dots, n.$$

This particular diagnostic indicator can be motivated from analogous quantities used in regression analysis and can provide valuable information over that available from measures focusing on residuals alone. Many other regression type diagnostics can also be suggested.

Concerning interval estimation, there are several alternatives to Silverman's method based on sample estimates of the influence curve for smoothing splines. One of these (c.f. Wold 1971) can be described as follows. For simplicity assume that $w_i=1, i=1, \dots, n$, and let $\hat{g}^{[i]}$ and $\hat{v}^{[i]}$ denote the smoothing spline estimate and vector of coefficient estimates, under the B-spline basis, when (t_i, Y_i) has been deleted from the data. It can be shown that

$$\hat{v}^{[i]} = [\alpha \Omega + B^T B]^{-1} B^T \left[\frac{Y - e_i}{1 - A_{ii}(\alpha)} \frac{(Y_i - \hat{g}(t_i))}{1 - A_{ii}(\alpha)} \right],$$

where e_i is the i th column of the $n \times n$ identity matrix, which gives $\hat{g}^{[i]}(t) = \sum_{j=1}^n \hat{v}_j^{[i]} \beta_j(t)$. Given a functional Ψ we then define pseudo-values

$$P_i(\Psi) = \Psi(\hat{g}) - (n-1)(\Psi(\hat{g}) - \Psi(\hat{g}^{[i]})), \quad i=1, \dots, n,$$

and obtain the jackknife variance estimate (Efron 1982)

$$S_{\Psi}^2 = [n(n-1)]^{-1} \sum_{i=1}^n (P_i(\Psi) - \bar{P}(\Psi))^2,$$

where $\bar{P}(\Psi) = n^{-1} \sum_{i=1}^n P_i(\Psi)$. An approximate 95% confidence interval for $\Psi(g)$ is provided by $\Psi(\hat{g}) \pm 2S_{\Psi}$. The computation of S_{Ψ}^2 simplifies considerably when Ψ is linear. Professor Silverman's approximations for the $A_{ij}(\alpha)$ have some obvious applications to jackknife interval estimation.

Jackknife confidence intervals do not require the assumption of normal errors and are more computationally expedient than the Bayesian approach when Ψ is nonlinear. It should also be noted (see e.g. Hinkley 1977) that

jackknife methods might be expected to be robust against nonhomogenous error variances. It would be interesting to compare jackknife methods to Silverman's approach with estimated weights in this setting.

REFERENCES

- Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans, SIAM monograph no. 38, CBMS-NSF.
- Eubank, R. L. (1984a). Approximate regression models and splines. Comm. Statist. - Theor. Meth., A13(4), 433-484.
- Eubank, R. L. (1984b). The hat matrix for smoothing splines. Statist. and Prob. Letters, 2, 9-14.
- Hinkley, D. V. (1977). Jackknifing in unbalanced situations. Technometrics, 19, 285-292.
- Wold, S. (1974). Spline functions in data analysis. Technometrics, 16, 1-11.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER SMU-DS-TR-190	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Contribution to the Discussion of Dr. Silverman's Paper		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Randall L. Eubank		8. CONTRACT OR GRANT NUMBER(s) N00014-82-K-0207
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Southern Methodist University Dallas, Texas 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-479
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE
		13. NUMBER OF PAGES 4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of The United States Government.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary, and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary, and identify by block number)		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)