Monographs
on Statistics and
Applied Probability 26

Density Estimation
for Statistics and Data Analysis

B.W. Silverman

MONOGRAPHS ON STATISTICS AND APPLIED PROBABILITY

General Editors

D.R. Cox, D.V. Hinkley, N. Keiding, N. Reid, D.B. Rubin and B.W. Silverman

1 Stochastic Population Models in Ecology and Epidemiology *M.S. Bartlett* (1960)
2 Queues *D.R. Cox and W.L. Smith* (1961)
3 Monte Carlo Methods *J.M. Hammersley and D.C. Handscomb* (1964)
4 The Statistical Analysis of Series of Events *D.R. Cox and P.A.W. Lewis* (1966)
5 Population Genetics *W.J. Ewens* (1969)
6 Probability, Statistics and Time *M.S. Bartlett* (1975)
7 Statistical Inference *S.D. Silvey* (1975)
8 The Analysis of Contingency Tables *B.S. Everitt* (1977)
9 Multivariate Analysis in Behavioural Research *A.E. Maxwell* (1977)
10 Stochastic Abundance Models *S. Engen* (1978)
11 Some Basic Theory for Statistical Inference *E.J.G. Pitman* (1979)
12 Point Processes *D.R. Cox and V. Isham* (1980)
13 Identification of Outliers *D.M. Hawkins* (1980)
14 Optimal Design *S.D. Silvey* (1980)
15 Finite Mixture Distributions *B.S. Everitt and D.J. Hand* (1981)
16 Classification *A.D. Gordon* (1981)
17 Distribution-free Statistical Methods, 2nd edition *J.S. Maritz* (1995)
18 Residuals and Influence in Regression *R.D. Cook and S. Weisberg* (1982)
19 Applications of Queueing Theory, 2nd edition *G.F. Newell* (1982)
20 Risk Theory, 3rd edition *R.E. Beard, T. Pentikainen and E. Pesonen* (1984)
21 Analysis of Survival Data *D.R. Cox and D. Oakes* (1984)
22 An Introduction to Latent Variable Models *B.S. Everitt* (1984)
23 Bandit Problems *D.A. Berry and B. Fristedt* (1985)
24 Stochastic Modelling and Control *M.H.A. Davis and R. Vinter* (1985)
25 The Statistical Analysis of Compositional Data *J. Aitchison* (1986)
26 Density Estimation for Statistics and Data Analysis *B.W. Silverman*
27 Regression Analysis with Applications *G.B. Wetherill* (1986)
28 Sequential Methods in Statistics, 3rd edition *G.B. Wetherill and K.D. Glazebrook* (1986)
29 Tensor Methods in Statistics *P. McCullagh* (1987)
30 Transformation and Weighting in Regression *R.J. Carroll and D. Ruppert* (1988)
31 Asymptotic Techniques for Use in Statistics *O.E. Barndorff-Nielsen and D.R. Cox* (1989)
32 Analysis of Binary Data, 2nd edition *D.R. Cox and E.J. Snell* (1989)
33 Analysis of Infectious Disease Data *N.G. Becker* (1989)
34 Design and Analysis of Cross-Over Trials *B. Jones and M.G. Kenward* (1989)
35 Empirical Bayes Methods, 2nd edition *J.S. Maritz and T. Lwin* (1989)
36 Symmetric Multivariate and Related Distributions *K.-T. Fang*

S. Kotz and K.W. Ng (1990)
37 Generalized Linear Models, 2nd edition *P. McCullagh and J.A. Nelder* (1989)
38 Cyclic and Computer Generated Designs, 2nd edition *J.A. John and E.R. Williams* (1995)
39 Analog Estimation Methods in Econometrics *C.F. Manski* (1988)
40 Subset Selection in Regression *A.J. Miller* (1990)
41 Analysis of Repeated Measures *M.J. Crowder and D.J. Hand* (1990)
42 Statistical Reasoning with Imprecise Probabilities *P. Walley* (1991)
43 Generalized Additive Models *T.J. Hastie and R.J. Tibshirani* (1990)
44 Inspection Errors for Attributes in Quality Control *N.L. Johnson, S. Kotz and X. Wu* (1991)
45 The Analysis of Contingency Tables, 2nd edition *B.S. Everitt* (1992)
46 The Analysis of Quantal Response Data *B.J.T. Morgan* (1993)
47 Longitudinal Data with Serial Correlation: A State-space Approach *R.H. Jones* (1993)
48 Differential Geometry and Statistics *M.K. Murray and J.W. Rice* (1993)
49 Markov Models and Optimization *M.H.A. Davis* (1993)
50 Networks and Chaos _ Statistical and Probabilistic Aspects *O.E. Barndorff-Nielsen, J.L. Jensen and W.S. Kendall* (1993)
51 Number-theoretic Methods in Statistics *K.-T. Fang and Y. Wang* (1994)
52 Inference and Asymptotics *O.E. Barndorff-Nielsen and D.R. Cox* (1994)
53 Practical Risk Theory for Actuaries *C.D. Daykin, T. Pentikäinen and M. Pesonen* (1994)
54 Biplots *J.C. Gower and D.J. Hand* (1996)
55 Predictive Inference: An Introduction *S. Geisser* (1993)
56 Model-Free Curve Estimation *M.E. Tarter and M.D. Lock* (1993)
57 An Introduction to the Bootstrap *B. Efron and R.J. Tibshirani* (1993)
58 Nonparametric Regression and Generalized Linear Models *P.J. Green and B.W. Silverman* (1994)
59 Multidimensional Scaling *T.F. Cox and M.A.A. Cox* (1994)
60 Kernel Smoothing *M.P. Wand and M.C. Jones* (1995)
61 Statistics for Long Memory Processes *J. Beran* (1995)
62 Nonlinear Models for Repeated Measurement Data *M. Davidian and D.M. Giltinan* (1995)
63 Measurement Error in Nonlinear Models *R.J. Carroll, D. Ruppert and L.A. Stefanski* (1995)
64 Analyzing and Modeling Rank Data *J.I. Marden* (1995)
65 Time Series Models - In econometrics, finance and other fields *D.R Cox, D.V. Hinkley and O.E. Barndorff Nielsen* (1996)
66 Local Polynomial Modeling and its Applications *J. Fan and I. Gijbels* (1996)
67 Multivariate Dependencies - Models, analysis and interpretation *D.R. Cox and N. Wermuth* (1996)
68 Statistical Inference - Based on the likelihood *A. Azzalini* (1996)

(Full details concerning this series are available from the Publishers).

Density Estimation
for Statistics and Data Analysis

B.W. Silverman
School of Mathematics
University of Bath, UK

CHAPMAN & HALL/CRC

Boca Raton London New York Washington, D.C.

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Originally published by Chapman & Hall
Chapman & Hall/CRC is an imprint of Taylor & Francis Group, an Informa business

Printed in the United States of America on acid-free paper
20 19 18 17 16 15 14 13 12 11 10 9
International Standard Book Number-13: 978-0-412-24620-3 (Hardcover)
Library of Congress catalog number: 99-19608

Library of Congress Cataloging-in-Publication Data

Catalog record is available from the Library of Congress

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

The recent surge of interest in the more technical aspects of density estimation has brought the subject into public view but has sadly created the impression, in some quarters, that density estimates are only of theoretical importance. I have aimed in this book to make the subject accessible to a wider audience of statisticians and others, hoping to encourage broader practical application and more relevant theoretical research. With these objects in mind, I have tried to concentrate on topics of methodological interest. Specialists in the field will, I am sure, notice omissions of the kind that are inevitable in a personal account; I hope in spite of these that they will find the treatment of the subject interesting and stimulating.

I would like to thank David Kendall for first kindling my interest in density estimation. For their useful suggestions and other help I am grateful to several people, including Adrian Baddeley, Christopher Chatfield, David Cox, Rowena Fowler and Anthony Robinson. Colleagues who have helped in various specific ways are acknowledged in the course of the text. I am also grateful to the University of Washington (Seattle) for the opportunity to present a course of lectures that formed the basis of the book.

Bernard Silverman
Bath, May 1985

CHAPTER 1

Introduction

1.1 What is density estimation?

The *probability density function* is a fundamental concept in statistics. Consider any random quantity X that has probability density function f. Specifying the function f gives a natural description of the distribution of X, and allows probabilities associated with X to be found from the relation

$$P(a < X < b) = \int_a^b f(x)\,\mathrm{d}x \qquad \text{for all } a < b.$$

Suppose, now, that we have a set of observed data points assumed to be a sample from an unknown probability density function. *Density estimation*, as discussed in this book, is the construction of an estimate of the density function from the observed data. The two main aims of the book are to explain how to estimate a density from a given data set and to explore how density estimates can be used, both in their own right and as an ingredient of other statistical procedures.

One approach to density estimation is *parametric*. Assume that the data are drawn from one of a known parametric family of distributions, for example the normal distribution with mean μ and variance σ^2. The density f underlying the data could then be estimated by finding estimates of μ and σ^2 from the data and substituting these estimates into the formula for the normal density. In this book we shall not be considering parametric estimates of this kind; the approach will be more *nonparametric* in that less rigid assumptions will be made about the distribution of the observed data. Although it will be assumed that the distribution has a probability density f, the data will be allowed to speak for themselves in determining the estimate of f more than would be the case if f were constrained to fall in a given parametric family.

Density estimates of the kind discussed in this book were first proposed by Fix and Hodges (1951) as a way of freeing discriminant analysis from rigid distributional assumptions. Since then, density estimation and related ideas have been used in a variety of contexts, some of which, including discriminant analysis, will be discussed in the final chapter of this book. The earlier chapters are mostly concerned with the question of how density estimates are constructed. In order to give a rapid feel for the idea and scope of density estimation, one of the most important applications, to the exploration and presentation of data, will be introduced in the next section and elaborated further by additional examples throughout the book. It must be stressed, however, that these valuable exploratory purposes are by no means the only setting in which density estimates can be used.

1.2 Density estimates in the exploration and presentation of data

A very natural use of density estimates is in the informal investigation of the properties of a given set of data. Density estimates can give valuable indication of such features as skewness and multimodality in the data. In some cases they will yield conclusions that may then be regarded as self-evidently true, while in others all they will do is to point the way to further analysis and/or data collection.

An example is given in Fig. 1.1. The curves shown in this figure were constructed by Emery and Carpenter (1974) in the course of a study of sudden infant death syndrome (also called 'cot death' or 'crib death'). The curve A is constructed from a particular observation, the degranulated mast cell count, made on each of 95 infants who died suddenly and apparently unaccountably, while the cases used to construct curve B were a control sample of 76 infants who died of known causes that would not affect the degranulated mast cell count. The investigators concluded tentatively from the density estimates that the density underlying the sudden infant death cases might be a mixture of the control density with a smaller proportion of a contaminating density of higher mean. Thus it appeared that in a minority (perhaps a quarter to a third) of the sudden deaths, the degranulated mast cell count was exceptionally high. In this example the conclusions could only be regarded as a cue for further clinical investigation.

Another example is given in Fig. 1.2. The data from which this figure was constructed were collected in an engineering experiment

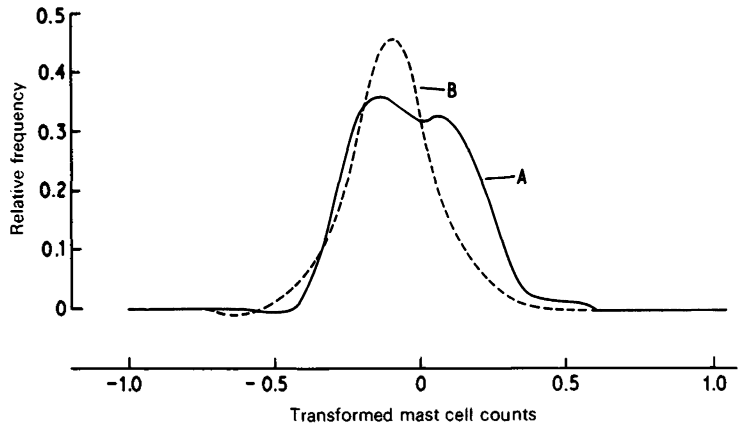


Fig. 1.1 *Density estimates constructed from transformed and corrected degranulated mast cell counts observed in a cot death study. (A, Unexpected deaths; B, Hospital deaths.) After Emery and Carpenter (1974) with the permission of the Canadian Foundation for the Study of Infant Deaths. This version reproduced from Silverman (1981a) with the permission of John Wiley & Sons Ltd.*

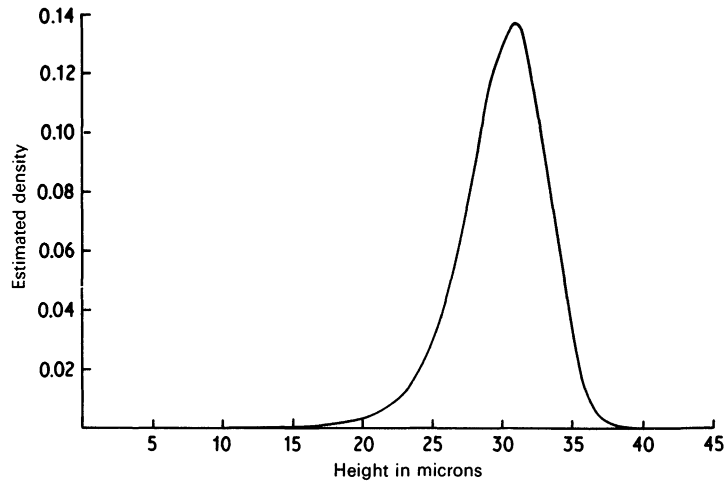


Fig. 1.2 *Density estimate constructed from observations of the height of a steel surface. After Silverman (1980) with the permission of Academic Press, Inc. This version reproduced from Silverman (1981a) with the permission of John Wiley & Sons Ltd.*

described by Bowyer (1980). The height of a steel surface above an arbitrary level was observed at about 15 000 points. The figure gives a density estimate constructed from the observed heights. It is clear from the figure that the distribution of height is skew and has a long lower tail. The tails of the distribution are particularly important to the engineer, because the upper tail represents the part of the surface which might come into contact with other surfaces, while the lower tail represents hollows where fatigue cracks can start and also where lubricant might gather. The non-normality of the density in Fig. 1.2 casts doubt on the Gaussian models typically used to model these surfaces, since these models would lead to a normal distribution of height. Models which allow a skew distribution of height would be more appropriate, and one such class of models was suggested for this data set by Adler and Firman (1981).

A third example is given in Fig. 1.3. The data used to construct this curve are a standard directional data set and consist of the directions in which each of 76 turtles was observed to swim when released. It is clear that most of the turtles show a preference for swimming approximately in the 60° direction, while a small proportion prefer exactly the opposite direction. Although further statistical model-

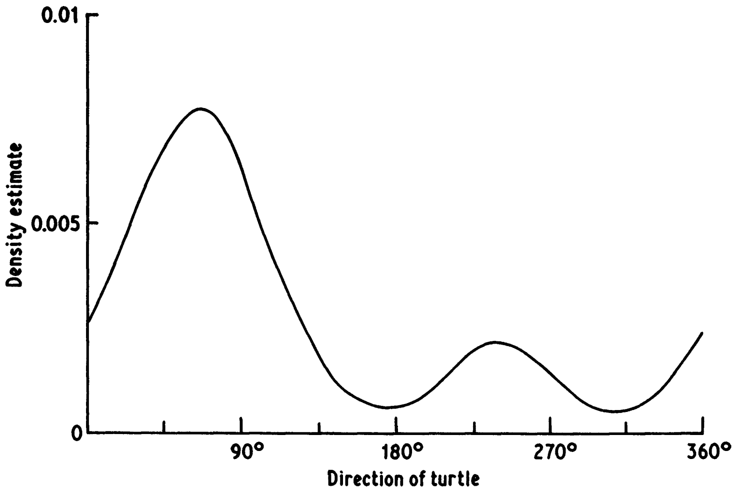
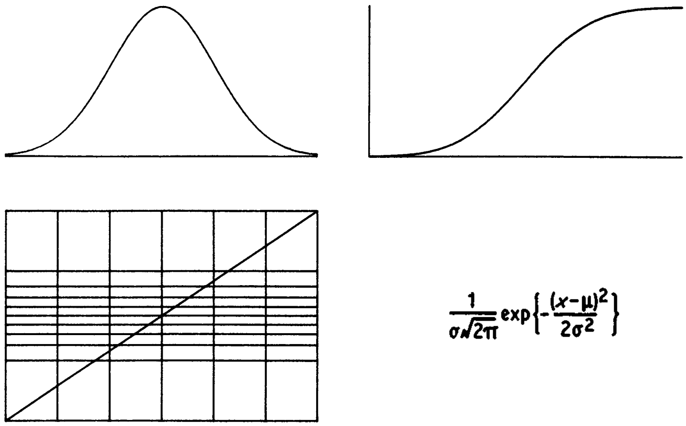


Fig. 1.3 *Density estimate constructed from turtle data. After Silverman (1978a) with the permission of the Biometrika Trustees. This version reproduced from Silverman (1981a) with the permission of John Wiley & Sons Ltd.*



$$\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

Fig. 1.4 *Four ways of explaining the normal distribution: a graph of the density function; a graph of the cumulative distribution function; a straight line on probability paper; the formula for the density function.*

ling of these data is possible (see Mardia, 1972) the density estimate really gives all the useful conclusions to be drawn from the data set.

An important aspect of statistics, often neglected nowadays, is the presentation of data back to the client in order to provide explanation and illustration of conclusions that may possibly have been obtained by other means. Density estimates are ideal for this purpose, for the simple reason that they are fairly easily comprehensible to non-mathematicians. Even those statisticians who are sceptical about estimating densities would no doubt explain a normal distribution by drawing a bell-shaped curve rather than by one of the other methods illustrated in Fig. 1.4. In all the examples given in this section, the density estimates are as valuable for explaining conclusions as for drawing these conclusions in the first place. More examples illustrating the use of density estimates for exploratory and presentational purposes, including the important case of bivariate data, will be given in later chapters.

1.3 Further reading

There is a vast literature on density estimation, much of it concerned with asymptotic results not covered in any detail in this book.

Prakasa Rao's (1983) book offers a comprehensive treatment of the theoretical aspects of the subject. Journal papers providing surveys and bibliography include Rosenblatt (1971), Fryer (1977), Wertz and Schneider (1979), and Bean and Tsokos (1980). Tapia and Thompson (1978) give an interesting perspective paying particular attention to their own version of the penalized likelihood approach described in Sections 2.8 and 5.4 below. A thorough treatment, rather technical in nature, of a particular question and its ramifications is given by Devroye and Györfi (1985). Other texts on the subject are Wertz (1978) and Delecroix (1983). Further references relevant to specific topics will be given, as they arise, later in this book.

References

Abramson, I.S. (1982). On bandwidth variation in kernel estimates a square root law. Ann. Statist., 10 , 12171223.
Adler, R.J. and Firman, D. (1981). A non-Gaussian model for random surfaces. Phil. Trans. Roy. Soc. Lond. A, 303 , 433462.
Aitchison, J. and Aitken, C.G.G. (1976). Multivariate binary discrimination by the kernel method. Biometrika, 63 , 413420.
Anderson, J.A. , Whaley, K. , Williamson, J. and Buchanan, W.W. (1972). A statistical aid to the diagnosis of Keratoconjunctivitis sicca. Quart. J. Med., 41 , 175189.
Anderson, J.A. and Senthilselvan, A. (1980). Smooth estimates for the hazard function. J. Roy. Statist. Soc. B., 42 , 322327.
Bartlett, M.S. (1963). Statistical estimation of density functions. Sankhya Ser. A, 25 , 245254.
Bean, S.J. and Tsokos, C.P. (1980). Developments in nonparametric density estimation. Int. Stat. Rev., 48 , 267287.
Bertrand-Retali, M. (1978). Convergence uniforme dun estimateur de la densit par la mthode de noyau. Rev. Roumaine Math. Pures. Appl, 23 , 361385.
Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviation of density function estimates. Ann. Statist., 1 , 10711095.
Bock, H.H. (1984). Statistical testing and evaluation methods in cluster analysis. Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, 116146.
Bock, H.H. (1985). On some significance tests in cluster analysis. J. Classification, 2 , 77108.
Boneva, L.I. , Kendall, D.G. and Stefanov, I. (1971). Spline transformations: three new diagnostic aids for the statistical data-analyst (with Discussion). J. Roy. Statist. Soc. B, 33 , 170.
Bowman, A.W. (1984). An alternative method of cross-validation for the smoothing of density estimates. Biometrika, 71 , 353360.
Bowman, A.W. (1985). A comparative study of some kernel-based nonparametric density estimators. J. Statist. Comput. Simul., 21 , 313327.
Bowman, A.W. , Hall, P. and Titterington, D.M. (1984). Cross-validation in 160nonparametric estimation of probabilities and probability densities. Biometrika, 71 , 341351.
Bowyer, A. (1980). Experiments and computer modelling in stick-slip. Ph.D. Thesis, University of London.
Breiman, L. , Meisel W. and Purcell, E. (1977). Variable kernel estimates of multivariate densities. Technometrics, 19 , 135144.
Breiman, L. , Friedman, J.H. , Olshen, R.A. and Stone, C.J. (1984). Classification and Regression Trees. Belmont, California: Wadsworth.
Cacoullos, T. (1966). Estimation of a multivariate density. Ann. Inst. Statist. Math., 18 , 179189.
encov, N.N. (1962). Evaluation of an unknown distribution density from observations. Soviet Math., 3 , 15591562.
Chatfield, C. and Collins, A.J. (1980). Introduction to Multivariate Analysis. London: Chapman and Hall.
Chow, Y.S. , Geman, S. and Wu, L.D. (1983). Consistent cross-validated density estimation. Ann. Statist., 11 , 2538.
Copas, J.B. and Fryer, M.J. (1980). Density estimation and suicide risks in psychiatric treatment. J. Roy. Statist. Soc. A, 143 , 167176.
Cormack, R.M. (1971). A review of classification (with Discussion). J. Roy. Statist. Soc. A, 134 , 321367.
Courant, R. and Hilbert, D. (1953). Methods of Mathematical Physics, Volume II. New York: Interscience.
Cover, T.M. and Hart, P.E. (1967). Nearest neighbor pattern classification. IEEE Trans. Inf. Thy., IT-13 , 2127.
Cox, D.R. (1966). Notes on the analysis of mixed frequency distributions. Brit. J. Math Statist. Psychol., 19 , 3947.
Cox, D.R. and Hinkley, D.V. (1974). Theoretical Statistics. London: Chapman and Hall.
Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data. London: Chapman and Hall.
Deheuvels, P. (1977). Estimation nonparametrique de la densit par histogrammes generaliss. Rev. Statist. Appl., 35 , 542.
Delecroix, M. (1983). Histogrammes et Estimation de la Densit. Que sais-je? no. 2055. Paris: Presses Universitaires de France.

Devroye, L. and Gyrf, L. (1985). Nonparametric Density Estimation: The L1 View. New York: Wiley.
Duin, R.P.W. (1976). On the choice of smoothing parameters for Parzen estimators of probability density functions. IEEE Trans. Comput., C-25 , 11751179.
Efron, B. (1981). Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. Biometrika, 68 , 589599.
Efron, B. (1982). The Jackknife, the Bootstrap and other Resampling Plans. Philadelphia: SIAM.
Emery, J.L. and Carpenter, R.G. (1974). Pulmonary mast cells in infants and 161their relation to unexpected death in infancy. In Robinson, R.R. (ed.), SIDS 1974 Proceedings of the Francis E. Camps International Symposium on Sudden and Unexpected Deaths in Infancy. Toronto: Canadian Foundation for the Study of Infant Deaths, pp. 719.
Epanechnikov, V.A. (1969). Nonparametric estimation of a multidimensional probability density. Theor. Probab. Appl., 14 . 153158.
Feller, W. (1966). An Introduction to Probability Theory and its Applications, Volume II. New York: Wiley.
Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. Ann. Eugen., 7 , 179188.
Fix, E. and Hodges, J.L. (1951). Discriminatory analysis, nonparametric estimation: consistency properties. Report No. 4, Project no. 21-49-004, USAF School of Aviation Medicine, Randolph Field, Texas.
Friedman, J.H. , Baskett, F. and Shustek, L.J. (1975). An algorithm for finding nearest neighbors. IEEE Trans. Comput., 24 , 10001006.
Friedman, J.H. , Bentley, J.L. and Finkel, R.A. (1977). An algorithm for finding best matches in logarithmic expected time. ACM Trans. on Math. Software, 3 , 209226.
Friedman, J.H. and Tukey, J.W. (1974). A projection pursuit algorithm for exploratory data analysis. IEEE Trans. Comput., C-23 , 881889.
Fryer, M.J. (1976). Some errors associated with the nonparametric estimation of density functions. J. Inst. Maths. Applics., 18 , 371380.
Fryer, M.J. (1977). A review of some nonparametric methods of density estimation. J. Inst. Maths. Applies., 20 , 335354.
Fukunaga, K. (1972). Introduction to Statistical Pattern Recognition. New York: Academic Press.
Fukunaga, K. and Hostetler, L.D. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. Info. Thy., IT-21 , 3240.
Geisser, S. (1975). The predictive sample reuse method with applications. J. Amer. Statist. Assoc., 70 , 320328.
Ghorai, J. and Rubin, H. (1979). Computational procedure for maximum penalized likelihood estimate. J. Statist. Comput. Simul., 10 , 6578.
Good, I.J. and Gaskins, R.A. (1971). Nonparametric roughness penalties for probability densities. Biometrika, 58 , 255277.
Good, I.J. and Gaskins, R.A. (1980). Density estimation and bump-hunting by the penalized likelihood method exemplified by scattering and meteorite data. J. Amer. Statist. Assoc., 75 , 4273.
Gordon, A.D. (1981). Classification. London: Chapman and Hall.
Habbema, J.D.F. , Hermans, J. and van der Broek, K. (1974). A stepwise discrimination program using density estimation. In Bruckman, G. (ed.), Compstat 1974. Vienna: Physica Verlag, pp. 100110.
Habbema, J.D.F. , Hermans, J. and Remme, J. (1978). Variable kernel density estimation in discriminant analysis. Compstat 1978, Proceedings in 162 Computational Statistics. Vienna: Physica Verlag.
Hall, P. (1983). Large sample optimality of least squares cross-validation in density estimation, Ann. Statist., 11 , 11561174.
Hand, D.J. (1981). Discrimination and Classification. Chichester: Wiley.
Hand, D.J. (1982). Kernel Discriminant Analysis. Chichester: Research Studies Press.
Hartigan, J.A. and Hartigan, P.M. (1985). The dip test of unimodality. Ann. Statist., 13 , 7084.
Hellman, M.E. (1970). The nearest neighbor classification rule with a reject option. IEEE Trans. Sys. Sci. Cyb., SSC-6 , 179185.
Hermans, J. , Habbema, J.D.F. , Kasanmoentalib, T.K.D. and Raatgever, J.W. (1982). Manual for the ALLOC80 discriminant analysis program. Dept. of Medical Statistics, University of Leiden, The Netherlands.

Hodges, J.L. and Lehmann, E.L. (1956). The efficiency of some nonparametric competitors of the t-test. Ann. Math. Statist., 27 , 324335.
Huber, P.J. (1985). Projection pursuit. Ann. Statist., 13 , 435475.
Jones, M.C. (1983). The projection pursuit algorithm for exploratory data analysis. Ph.D. Thesis, University of Bath.
Jones, M.C. and Lotwick, H.W. (1983). On the errors involved in computing the empirical characteristic function. J. Statist. Comput. Simul., 17 , 133149.
Jones, M.C. and Lotwick, H.W. (1984). A remark on Algorithm AS 176. Kernel density estimation using the fast Fourier transform. Remark AS R50. Appl. Statist., 33 , 120122.
Jones, M.C. and Sibson, R. (1987). What is projection pursuit? J. Roy. Statist. Soc. A, 150 , 136.
Kendall, M.G. and Stuart, A. (1973). The Advanced Theory of Statistics, Volume 2, 3rd Edition. London: Griffin.
Kent, J.T. (1976). Contribution to the discussion of a paper by P.R. Freeman. J. Roy. Statist. Soc. A, 139 , 3940.
Kittler, J. (1976). A locally sensitive method for cluster analysis. Pattern Recognition, 8 , 2333.
Koontz, W.L.G. , Narendra, P.M. and Fukunaga, K. (1976). A graph-theoretic approach to nonparametric cluster analysis. IEEE Trans. Comput., C-25 , 936943.
Kreider, D.L. , Kuller, R.G. , Ostberg, D.R. and Perkins, F.W. (1966) An Introduction to Linear Analysis. Reading, Mass: Addison-Wesley.
Leonard, T. (1978). Density estimation, stochastic processes and prior information (with Discussion). J. Roy. Statist. Soc. B, 40 , 113146.
Loftsgaarden, D.O. and Quesenberry, C.P. (1965). A nonparametric estimate of a multivariate density function. Ann. Math. Statist., 36 , 10491051.
Lubischew, A.A. (1962). On the use of discriminant functions in taxonomy. Biometrics, 18 , 455477.
163 Mack, Y.P. and Rosenblatt, M. (1979). Multivariate K-nearest neighbor density estimates. J. Multivariate Anal, 9 , 115.
Mardia, K.V. (1972). Statistics of Directional Data. London: Academic Press.
Mardia, K.V. , Kent, J.T. and Bibby, J.M. (1979). Multivariate Analysis. London: Academic Press.
Monro, D.M. (1976). Real discrete fast Fourier transform. Statistical Algorithm AS 97, Appl. Statist., 25 , 166172.
de Montricher, G.M. , Tapia, R.A. and Thompson, J.R. (1975). Nonparametric maximum likelihood estimation of probability densities by penalty function methods. Ann. Statist., 3 , 13291348.
Mller, H.G. (1984). Smooth optimum kernel estimators of densities, regression curves and modes. Ann. Statist., 12 , 766774.
Nadaraya, E.A. (1965). On nonparametric estimates of density functions and regression curves. Theor. Probab. Appl., 10 , 186190.
Parzen, E. (1962). On estimation of a probability density function and mode. Ann. Math. Statist., 33 , 10651076.
Parzen, E. (1979). Nonparametric statistical data modeling. J. Amer. Statist. Assoc., 74 , 105131.
Prakasa Rao, B.L.S. (1983). Nonparametric Functional Estimation. New York: Academic Press.
Remme, J. , Habbema, J.D.F. and Hermans, J. (1980). A simulative comparison of linear, quadratic and kernel discrimination. J. Statist. Comput. Simul., 11 , 87106.
Rice, J. (1975). Statistical methods of use in analysing sequences of earthquakes. Geophys. J.R. Astr. Soc., 42 , 671683.
Rice, J. and Rosenblatt, M. (1976). Estimation of the log survivor function and hazard function. Sankhya Ser. A, 38 , 6078.
Ripley, B.D. (1983). Computer generation of random variables: a tutorial. Int. Stat. Rev., 51 , 301319.
Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. Ann. Math. Statist., 27 , 832837.
Rosenblatt, M. (1971). Curve estimates. Ann. Math. Statist., 42 , 18151842.
Rosenblatt, M. (1979). Global measures of deviation for kernel and nearest neighbor density estimates. In Gasser, T. and Rosenblatt, M. (eds), Smoothing Techniques for Curve Estimation. Lecture Notes in Mathematics, 757. Berlin: Springer-Verlag, pp. 181190.

Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. Scand. J. Statist., 9 , 6578.
Schucany, W.R. and Sommers, J.P. (1977). Improvement of kernel type density estimators. J. Amer. Statist. Assoc., 72 , 420423.
Schuster, E.F. and Gregory, C.G. (1981). On the nonconsistency of maximum likelihood nonparametric density estimators. In Eddy, W.F. (ed.) Computer Science and Statistics: Proceedings of the 13th Symposium on the 164Interface. New York: Springer-Verlag, pp. 295298.
Scott, D.W. (1979). On optimal and data-based histograms. Biometrika, 66 , 605610.
Scott, D.W. (1982). Review of some results in bivariate density estimation. In Guseman, L.F. (ed.), Proceedings of the NASA Workshop on Density Estimation and Function Smoothing March 1982. Dept of Mathematics, Texas A & M University, College Station, Texas, pp. 165194.
Scott, D.W. and Factor, L.E. (1981). Monte Carlo study of three data-based nonparametric density estimators. J. Amer. Statist. Assoc., 76 , 915.
Scott, D.W. , Gotto, A.M. , Cole, J.S. and Gorry, G.A. (1978). Plasma lipids as collateral risk factors in coronary heart disease a study of 371 males with chest pain. J. Chronic Diseases, 31 , 337345.
Scott, D.W. , Tapia, R.A. and Thompson, J.R. (1977). Kernel density estimation revisited. Nonlinear Analysis, 1 , 339372.
Scott, D.W. , Tapia, R.A. and Thompson, J.R. (1980). Nonparametric probability density estimation by discrete maximum penalized-likelihood criteria. Ann. Statist., 8 , 820832.
Scott, D.W. and Thompson, J.R. (1983). Probability density estimation in higher dimensions. In Gentle, J.E. (ed.), Computer Science and Statistics: Proceedings of the Fifteenth Symposium on the Interface. Amsterdam: North-Holland, pp. 173179.
Sibson, R. and Thomson, G.D. (1981). A seamed quadratic element for contouring. Comput. J., 24 , 378382.
Silverman, B.W. (1978a). Choosing the window width when estimating a density. Biometrika, 65 , 111.
Silverman, B.W. (1978b). Weak and strong uniform consistency of the kernel estimate of a density function and its derivatives. Ann. Statist., 6 , 177184. (Addendum 1980, Ann. Statist., 8, 11751176.)
Silverman, B.W. (1978c). Density ratios, empirical likelihood and cot death. Applied Statistics, 27 , 2633.
Silverman, B.W. (1980). Density estimation: Are theoretical results useful in practice? In Chakravarti, I.M. (ed.), Asymptotic Theory of Statistical Tests and Estimation. New York: Academic Press, pp. 179203.
Silverman, B.W. , (1981a). Density estimation for univariate and bivariate data. In Barnett, V. (ed.), Interpreting Multivariate Data. Chichester: Wiley, pp. 3753.
Silverman, B.W. (1981b). Using kernel density estimates to investigate multimodality. J. Roy. Statist. Soc. B, 43 , 9799.
Silverman, B.W. (1982a). Kernel density estimation using the fast Fourier transform. Statistical Algorithm AS 176. Appl. Statist., 31 , 9397.
Silverman, B.W. (1982b). On the estimation of a probability density function by the maximum penalized likelihood method. Ann. Statist., 10 , 795810.
Silverman, B.W. (1983). Some properties of a test for multimodality based on 165kernel density estimates. In Kingman, J.F.C. and Reuter, G.E.H. (eds), Probability, Statistics and Analysis. Cambridge: Cambridge University Press, pp. 248259.
Silverman, B.W. (1984a). Spline smoothing: the equivalent variable kernel method. Ann. Statist., 12 , 898916.
Silverman, B.W. (1984b). A fast and efficient cross-validation method for smoothing parameter choice in spline regression. J. Amer. Statist. Assoc., 79 , 584589.
Silverman, B.W. (1985a). Some aspects of the spline smoothing approach to nonparametric regression curve fitting (with Discussion). J. Roy. Statist. Soc. B, 46 , 152.
Silverman, B.W. (1985b). Penalized maximum likelihood estimation. In Kotz, S. and Johnson, N.L. (eds), Wiley Encyclopedia of Statistical Sciences, Volume 6. New York: Wiley, pp. 664667.
Stone, C.J. (1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., 12 , 12851297.
Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with Discussion). J. Roy. Statist. Soc. B, 36 , 111147.

Tapia, R.A. and Thompson, J.R. (1978). Nonparametric Probability Density Estimation. Baltimore: Johns Hopkins University Press.
Thomson, G.D. (1984). Automatic contouring by piecewise quadratic approximation. Ph.D. Thesis, University of Bath.
Titterington, D.M. , Murray, G.D. , Murray, L.S. , Spiegelhalter, D.J. , Skene, A.M. , Habbema, J.D.F. and Gelpke, G.J. (1981). Comparison of discrimination techniques applied to a complex data set of head injured patients. J. Roy. Statist. Soc. A, 144 , 145174.
Tukey, P.A. and Tukey, J.W. (1981). Graphical display of data sets in 3 or more dimensions. In Barnett, V. (ed.), Interpreting Multivariate Data. Chichester: Wiley, pp. 189275.
Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis I. Biometrika, 51 , 175184.
Wegman, E.J. and Wright, I.W. (1983). Splines in statistics. J. Amer. Statist. Assoc., 78 , 351365.
Weisberg, S. (1980). Applied Linear Regression. New York: Wiley.
Wertz, W. (1978). Statistical Density Estimation: A Survey. Gttingen: Vandenhoeck and Ruprecht.
Wertz, W. and Schneider, B. (1979). Statistical density estimation: a bibliography. Int. Stat. Rev., 47 , 155175.
Whittle, P. (1958). On the smoothing of probability density functions. J. Roy. Statist. Soc. B, 20 , 334343.
Woodroofe, M. (1970). On choosing a delta-sequence. Ann. Math. Statist., 41 , 16651671.