ARO 23010.3 -MA

(2)

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143

Department of STATISTICS
Statistical Interdisciplinary
Research Laboratory
E415EP@TAMVM1.BITNET

Emanuel Parzen
Distinguished Professor

MULTI-SAMPLE FUNCTIONAL STATISTICAL DATA ANALYSIS

Technical Report #54

May 1989

Emanuel Parzen

Texas A&M Research Foundation

Project No. 5641

Sponsored by the U. S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

89 8 7

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ARO 23010.3-MA

2. GOVT ACCESSION NO.
N/A

3. RECIPIENT'S CATALOG NUMBER
N/A

4. TITLE (and Subtitle)
Multi-Sample Functional Statistical Data Analysis

5. TYPE OF REPORT & PERIOD COVERED
Technical

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Emanuel Parzen

8. CONTRACT OR GRANT NUMBER(s)
DAAL03-87-K-0003

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Texas A&M University
Department of Statistics
College Station, TX 77843-3143

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
U. S. Army Research Office
Post Office Box 12211
Research Triangle Park, NC 27709

12. REPORT DATE
May 1989

13. NUMBER OF PAGES
14

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
NA

18. SUPPLEMENTARY NOTES
The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
quantile data analysis, multi-sample nonparametric testing, Anderson Darling tests, Cramer von Mises tests, comparison density function, components

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This paper discusses a functional approach to the problem of comparison of multi-samples (two samples or c samples, where $c \geq 2$). The data consists of c random samples whose probability distributions are to be tested for equality. A diversity of statistics to test equality of c samples are presented in a unified framework with the aim of helping the researcher choose the optimal procedures which provide greatest insight about how the samples differe in their distributions.

MULTI-SAMPLE FUNCTIONAL STATISTICAL DATA ANALYSIS

Emanuel Parzen
Department of Statistics
Texas A&M University
College Station, Texas 77843-3143

ABSTRACT. **This paper discusses a functional approach to the problem of comparison of multi-samples (two samples or c samples, where $c \geq 2$). The data consists of c random samples whose probability distributions are to be tested for equality. A diversity of statistics to test equality of c samples are presented in a unified framework with the aim of helping the researcher choose the optimal procedures which provide greatest insight about how the samples differ in their distributions. Concepts discussed are: sample distribution functions; ranks; mid-distribution function; two- sample t test and nonparametric Wilcoxon test; multi-sample analysis of variance and Kruskal Wallis test; Anderson Darling and Cramer von Mises tests; components and linear rank statistics; comparison distribution and comparison density functions, especially for discrete distributions; components with orthogonal polynomial score functions; chi-square tests and their components.** (KP)

1. INTRODUCTION. **We assume that we are observing a variable Y in c cases or samples (corresponding to c treatments or c populations). The samples can be regarded as the value of c variables $Y_1, \ldots, Y_c$ with respective true distribution functions $F_1(y), \ldots, F_c(y)$ and quantile functions $Q_1(u), \ldots, Q_c(u)$. We call $Y_1, \ldots, Y_c$ the conditioned variables (the value of Y in different populations).**

The general problem of comparison of conditioned random variables is to model how their distribution functions vary with the value of the conditioning variable $k = 1, \ldots, c$, and in particular to test the hypothesis of homogeneity of distributions:

$$H_0 : F_1 = \ldots = F_c = F$$

The distribution F to which all the others are equal is considered to be the unconditional distribution of Y (which is estimated by the sample distribution of Y in the pooled sample).

2. DATA. **The data consists of c random samples**

$$Y_k(j), j = 1, \ldots, n_k$$

for $k = 1, \ldots, c$. The pooled sample, of size $N = n_1 + \ldots + n_c$, represents observations of the pooled (or unconditional) variable Y. The c samples are assumed to be independent of each other.

3. SAMPLE DISTRIBUTION FUNCTIONS. **The sample distribution functions of the samples are defined (for $-\infty < y < \infty$) by**

$$F_k\tilde{}(y) = \text{ fraction } \leq y \text{ among } Y_k(.).$$

The unconditional or pooled sample distribution of Y is denoted

$$F\hat{}(y) = \text{ fraction } \leq y \text{ among } Y_k(.), k = 1, \ldots, c.$$

We use ˆ to denote a smoother distribution to which we are comparing a more raw distribution which is denoted by a ˜. An expectation (mean) computed from a sample is denoted $E\tilde{}$.

Research Supported by the U.S. Army Research Office

4. *RANKS, MID-RANKS, AND MID-DISTRIBUTION FUNCTION.*

Nonparametric statistics use ranks of the observations in the pooled sample; let

$$R_k(t) \text{ denote the rank in the pooled sample of } Y_k(t).$$

One can define $R_k(t) = NF^{\wedge}(Y_k(t))$.

In defining linear rank statistics one transforms the rank to a number in the open unit interval, usually $R_k(t)/(N+1)$. We recommend $(R_k(t) - .5)/N$. These concepts assume all observations are distinct, and treat ties by using average ranks. We recommend an approach which we call the "mid-rank transform" which transforms $Y_k(t)$ to $P^{\wedge}(Y_k(t))$, defining the mid-distribution function of the pooled sample Y by

$$P^{\wedge}(y) = F^{\wedge}(y) - .5p^{\wedge}(y).$$

We call

$$p^{\wedge}(y) = \text{ fraction equal to } y \text{ among pooled sample}$$

the pooled sample probability mass function.

5. *SAMPLE MEANS AND VARIANCES.*

When the random variables are assumed to be normal the test statistics are based on the sample means (for $k = 1, \ldots, c$)

$$Y_k^{-} = E^{-}[Y_k] = (1/n_k) \sum_{t=1}^{n_k} Y_k(t).$$

We interpret Y_k^{-} as the sample conditional mean of Y given that it comes from the kth population. The unconditional sample mean of Y is

$$Y^{-} = E^{-}[Y] = p_{.1}Y_1^{-} + \ldots + p_{.c}Y_c^{-},$$

defining

$$p_{.k} = n_k/N$$

to be the fraction of the pooled sample in the kth sample; we interpret it as the empirical probability that an observation comes from the kth sample.

The unconditional and conditional variances are denoted 9

$$\mathrm{VAR}^{-}[Y] = (1/N) \sum_{k=1}^{c} \sum_{j=1}^{n_k} \{Y_k(j) - Y^{-}\}^2$$

$$\mathrm{VAR}^{-}[Y_k] = (1/n_k) \sum_{j=1}^{n_k} \{Y_k(j) - Y_k^{-}\}^2$$

Note that our divisor is the sample size N or n_k rather than $N - c$ or $n_k - 1$. The latter then arise as factors used to define F statistics.

We define the pooled variance to be the mean conditional variance:

$$\sigma^{\wedge 2} = \sum_{k=1}^{c} p_{.k} \, \mathrm{VAR}^{-}[Y_k]$$

6. TWO SAMPLE NORMAL T TEST. In the two sample case the statistic to test H_0 is usually stated in a form equivalent to

$$T = \{Y_1^{-} - Y_2^{-}\}/\sigma^{\wedge}\{(N/(N-2))((1/n_1) + (1/n_2))\}^{.5}$$

We believe that one obtains maximum insight (and analogies and extensions) by expressing T in the form which compares Y_1^{-} with Y^{-}:

$$T = \{(N-2)p_{.1}/(1-p_{.1})\}^{.5}\{Y_1^{-} - Y^{-}\}/\sigma^{\wedge}$$

The exact distribution of T is $t(N-2)$, t-distribution with $N-2$ degrees of freedom.

7. TWO-SAMPLE NONPARAMETRIC WILCOXON TEST. To define the popular Wilcoxon non-parametric statistic to test H_0 we define W_k to be the sum of the n_k ranks of the Y_k values; its mean and variance are given by

$$E[W_k] = n_k(N+1)/2, \ \mathrm{VAR}[W_k] = n_1 n_2 (N+1)/12$$

The usual definition of the Wilcoxon test statistic is

$$T_k = \{W_k - E[W_k]\}/\{\mathrm{VAR}[W_k]\}^{.5}.$$

The approach we describe in this paper yields as the definition of the nonparametric Wilcoxon test statistic (which can be verified to approximately equal the above definition of T_1, up to a factor $\{1-(1/N)^2\}^{.5}$)

$$T_1 = \{12(N-1)p_{.1}/(1-p_{.1})\}^{.5}(R_1^{-} - .5\},$$

defining

$$R_1^{-} = (1/n_1)\sum_{t=1}^{n_1}(R_1(t) - .5)/N$$
$$= (W_1/n_1 N) - (1/2N)$$

One reason we prefer this form of expressing non-parametric statistics is because of its relation to mid-ranks;

$$R_k^{-} = E^{-}[P^{\wedge}(Y_k)]$$

One should notice the analogy between our expressions for the parametric test statistic T and the nonparametric test statistic T_1; the former has an exact $t(N-2)$ distribution and the latter has asymptotic distribution Normal$\{0,1\}$.

8. TEST OF EQUALITY OF c SAMPLES NORMAL CASE. The homogeneity of c samples is tested in the parametric normal case by the analysis of variance which starts with a fundamental identity which in our notation is written

$$\mathrm{VAR}^{-}[Y] = \sum_{k=1}^{c} p_{.k}\{Y_k^{-} - Y^{-}\}^2 + \sigma^{\wedge 2}$$

The F test of the one-way analysis of variance can be expressed as the statistic

$$T^2 = \sum_{k=1}^{c} p_{.k}|T_k|^2,$$
$$= \sum_{k=1}^{c}(1-p_{.k})|TF_k|^2,$$

defining

$$T_k = (N-c)\{Y_k^{\frown} - Y^{\frown}\}/\sigma^{\frown}$$
$$TF_k = \{(N-c)p_{.k}/(1-p_{.k})\}^{.5}\{Y_k^{\frown} - Y^{\frown}\}/\sigma^{\frown}$$

The asymptotic distribution of $T^2/(c-1)$ and TF_k^2 are $F(c-1, N-c)$ and $F(1, N-c)$ respectively.

9. TEST OF EQUALITY OF c SAMPLES NONPARAMETRIC KRUSKAL-WALLIS TEST. The Kruskal-Wallis nonparametric test of homogeneity of c samples can be shown to be

$$TKW^2 = \sum_{k=1}^{c}(1-p_{.k})|TKW_k|^2.$$
$$TKW_k = \{12(N-1)p_{.k}/(1-p_{.k})\}^{.5}\{R_k^{\frown} - .5\}$$

The asymptotic distributions of TKW^2 and TKW_k^2 are chi-squared with $c-1$ and 1 degrees of freedom respectively.

10. COMPONENTS. We have represented the analysis of variance test statistic T^2 and the Kruskal-Wallis test statistic TKW^2 as weighted sums of squares of statistics TF_k and TKW_k respectively which we call components, since their values should be explicitly calculated to indicate the source of the significance (if any) of the overall statistics. Other test statistics that can be defined can be shown to correspond to other definitions of components.

11. ANDERSON DARLING AND CRAMER VON MISES TEST STATISTICS. Important among the many test statistics which have been defined to test the equality of distributions are the Anderson-Darling and Cramer-von Mises test statistics. They will be introduced below in terms of representations as weighted sums of squares of suitable components.

12. COMPARISON DISTRIBUTION FUNCTIONS AND COMPARISON DENSITTY FUNCTIONS. We now introduce the key concepts which enable us to unify and choose between the diverse statistics available for comparing several samples. To compare two continuous distributions $F(.)$ and $H(.)$, where H is a true or smooth and F is a model or raw, we define the comparison distribution function

$$D(u) = D(u; H, F) = F(H^{-1}(u))$$

with comparison density

$$d(u) = d(u; H, F) = D'(u) = f(H^{-1}(u))/h(H^{-1}(u)).$$

Under $H_0 : H = F$, $D(u) = u$ and $d(u) = 1$. Thus testing H_0 is equivalent to testing $D(u)$ for uniformity.

Sample distribution functions are discrete. The most novel part of this paper is that we propose to form an estimator $D^{\frown}(u)$ from estimators $H^{\frown}(.)$ and $F^{\frown}(.)$ by using a general definition of $D(.)$ for two discrete distributions $H(.)$ and $F(.)$ with respective probability mass functions p_H and p_F satisfying the condition that the values at which p_H are positive include all the values at which p_F are positive.

13. COMPARISON OF DISCRETE DISTRIBUTIONS. To compare two discrete distributions we define first $d(u)$ and then $D(u)$ as follows:

$$d(u) = d(u; H, F) = p_F(H^{-1}(u))/p_H(H^{-1}(u)),$$
$$D(u) = \int_0^1 d(t)dt.$$

We apply this definition to the discrete sample distributions $F\hat{}$ and $F\tilde{}_k$ to obtain

$$d_k\tilde{}(u) = d(u; F\hat{}, F_k\tilde{})$$

and its integral $D_k\tilde{}(u)$.

We obtain the following definition of $d_k\tilde{}(u)$ for the c sample testing problem with all values distinct:

$$\begin{aligned} d_k\tilde{}(u) &= N/n_k \text{ if } (R_k(j)-1)/N < u < R_k(j)/N, j = 1, \ldots, n_k, \\ &= 0, \qquad \text{otherwise.} \end{aligned}$$

A component, with score function $J(u)$, is a linear functional

$$T_k\tilde{}(J) = \int_0^1 J(u) d_k\tilde{}(u) du$$

It equals

$$(1/n_k) \sum_{j=1}^{n_k} N \int_{(R_k(j)-1)/N}^{R_k(j)/N} J(u)du$$

which can be approximated by $E\tilde{}[J(P\hat{}(Y_k))]$.

14. LINEAR RANK STATISTICS. The concept of a linear rank statistic to compare the equality of c samples does not have a universally accepted definition. One possible definition is

$$T_k\tilde{}(J) = (1/n_k) \sum_{j=1}^{n_k} J((R_k(j) - .5)/N)$$

However we choose the definition of a linear rank statistic as a linear functional of $d_k\tilde{}(u)$, which we call a component; it is approximately equal to the above formula.

We define

$$T_k\tilde{}(J) = ((N-1) \text{ VAR}[J(U)] p_{.k}/(1-p_{.k}))^{.5} \int_0^1 J(u)\{d_k\tilde{}(u) - 1\}du \qquad (!)$$

where U is Uniform$\{0, 1\}$, $E[J(U)] = \int_0^1 J(u)du$,

$$\text{VAR}[J(U)] = \int_0^1 \{J(u) - E[J(U)]\}^2 du.$$

Note that the integral in the definition of $T_k\tilde{}(J)$ equals

$$\int_0^1 J(u) d\{D_k\tilde{}(u) - u\}.$$

The components of the Kruskal-Wallis nonparametric test statistic TKW^2 for testing the equality of c means have score function $J(u) = u - .5$ satisfying

$$E[J(U)] = .5, \ \mathrm{VAR}[J(U)] = 1/12.$$

The components of F test statistic T^2 have score function

$$J(u) = \{Q^{\wedge}(u) - Y^{-}\}/\sigma^{\wedge}$$

where $Q^{\wedge}(u)$ is sample quantile function of the pooled sample Y.

15. GENERAL DISTANCE MEASURES. General measures of the distance of $D^{\sim}(u)$ from u and of $d^{\sim}(u)$ from 1 are provided by the integrals from 0 to 1 of

$$\{d^{\sim}(u) - 1\}^2, \quad \{D^{\sim}(u) - u\}^2, \quad \{D^{\sim}(u) - u\}^2/u(1-u), \quad \{d^{\wedge}(u) - 1\}^2$$

where $d^{\wedge}(u)$ is a smooth version of $d^{\sim}(u)$. We will see that these measures can be decomposed into components which may provide more insight; recall basic components are linear functionals defined by (!)

$$T^{\sim}(J) = \int_0^1 J(u) d^{\sim}(u) du.$$

If $\phi_i(u)$, $i = 0, 1, 2, \ldots$, are complete orthonormal functions with $\phi_0 = 1$, then H_0 can be tested by diagnosing the rate of increase (as a function of $m = 1, 2, \ldots$) of

$$\int_0^1 \{d_m(u) - 1\}^2 du = \sum_{i=1}^{m} |T^{\sim}(\phi_i|^2$$

which measure the distance from 1 of the approximating smooth densities

$$d_m(u) = \sum_{i=1}^{m} T^{\sim}(\phi_i)\phi_i(u).$$

16. ORTHOGONAL POLYNOMIAL COMPONENTS. Let $p_i(x)$ be Legendre polynomials on (-1,1):

$$\begin{aligned} p_1(x) &= x \\ p_2(x) &= (3x^2 - 1)/2, \\ p_3(x) &= (5x^3 - 3x)/2, \\ p_4(x) &= 35x^4 - 30x^2 + 3. \end{aligned}$$

Define Legendre polynomial score functions

$$\phi L_i(u) = (2i+1)^{.5} p_i(2u - 1).$$

One can show that an Anderson-Darling type statistic, denoted $AD(D^{\sim})$, can be represented

$$\begin{aligned} AD(D^{\sim}) &= \int_0^1 \{\{D^{\sim}(u) - u\}^2/u(1-u)\} du \\ &= \sum_{i=1}^{\infty} |T^{\sim}(\phi L_i)|^2/(i(i+1)) \end{aligned}$$

Define cosine score functions by

$$\phi C_i(u) = 2^{.5} cos(i\pi u).$$

One can show that a Cramer–von Mises type statistic, denoted $CM(D^{\sim})$, can be represented

$$CM(D^{\sim}) = \int_0^1 \{D^{\sim}(u) - u\}^2 du$$
$$= \sum_{i=1}^{\infty} |T^{\sim}(\phi C_i)|^2/(i\pi)^2$$

In addition to Legendre polynomial and cosine components we consider Hermite polynomial components corresponding to Hermite polynomial score functions

$$\phi H_i(u) = (i!)^{-.5} H_i(\Phi^{-1}(u))$$

where $H_i(x)$ are the Hermite polynomials:

$$\begin{aligned} H_1(x) &= x, \\ H_2(x) &= x^2 - 1, \\ H_3(x) &= x^3 - 3x, \\ H_4(x) &= x^4 - 6x^2 + 3. \end{aligned}$$

17. QUARTILE COMPONENTS AND CHI-SQUARE. Quartile diagnostics of the null hypothesis H_0 are provided by components with quartile "square wave" score functions

$$\begin{aligned} SQ_1(u) &= -2^{.5}, && 0 < u < .25, \\ &= 0, && .25 < u < .75, \\ &= 2^{.5}, && .75 < u < 1; \\ SQ_2(u) &= 1, && 0 < u < .25, \\ &= -1, && .25 < u < .75, \\ &= 1, && .75 < u < 1; \\ SQ_3(u) &= 0 && \text{if } 0 < u < .25 \text{ or } .75 < u < 1, \\ &= -2^{.5}, && .25 < u < .5, \\ &= 2^{.5}, && .5 < u < .75. \end{aligned}$$

A chi-squared portmanteau statistic, which is chi-squared(3), is

$$CQ_k = (N-1)p_{.k}/(1-p_{.k}) \sum_{i=1}^{3} |T^{\sim}(SQ_i)|^2$$
$$= (N-1)p_{.k}/(1-p_{.k}) \int_0^1 \{dQ_k(u) - 1\}^2 du$$

defining the quartile density (for $i = 1,2,3,4$)

$$dQ_k(u) = 4\{D_k^{\sim}(i(.25)) - D_k^{\sim}((i-1).25), (i-1).25 < u < i(.25)$$

A pooled portmanteau chi-squared statistic is

$$CQ = \sum_{k=1}^{c} (1 - p_{.k}) CQ_k$$

18. DIVERSE STATISTICS AVAILABLE TO TEST EQUALITY OF c SAMPLES. The problem of statistical infereence is not that we don't have answers to a given question; usually we have too many answers and we don't know which one to choose. A unified framework may help determine optimum choices. To compare c samples we can compute the following functions and statistics:

1) comparison densities: $d_k^{\sim}(u)$,
2) comparison distributions $D_k^{\sim}(u)$,
3) quartile comparison density $dQ_k(u)$, quartile density chi-square

$$CQ_k = (N-1)p_{.k}/(1-p_{.k}) \int_0^1 \{dQ_k(u) - 1\}^2 du.$$

4) non-parametric regression smoothing of $d_k^{\sim}(u)$ using a boundary Epanechnikov kernel, denoted $d_k^{\wedge}(u)$,
5) Legendre components and chi-squares up to order 4 are defined using definition (!) of $T_k^{\sim}$:

$$TL_k(i) = T_k^{\sim}(\phi L_i)$$

$$CL_k(m) = \sum_{i=1}^{m} |TL_k(i)|^2$$

$$CL(m) = \sum_{k=1}^{c} (1 - p_{.k}) CL_k(m)$$

$$AD_k = \sum_{i=1}^{\infty} |TL_k(i)|^2 / i(i+1)$$

$$AD = \sum_{k=1}^{c} (1 - p_{.k}) AD_k$$

6) Cosine components and chi-squares up to order 4 are defined:

$$TC_k(i) = T_k^{\sim}(\phi C_i)$$

$$CC_k(m) = \sum_{i=1}^{m} |TC_k(i)|^2$$

$$CC(m) = \sum_{k=1}^{c} (1 - p_{.k}) CC_k(m)$$

$$CM_k = \sum_{i=1}^{\infty} |TC_k(i)|^2 / (i\pi)^2$$

$$CM = \sum_{k=1}^{c} (1 - p_{.k}) CM_k$$

7) Hermite components and chi-squares up to order 4 are defined:

$$TH_k(i) = T_k\hat{}(\phi H_i)$$

$$CH_k(m) = \sum_{i=1}^{m} |TH_k(i)|^2$$

$$CH(m) = \sum_{k=1}^{c} (1 - p_{.k}) CH_k(m)$$

8) density estimators $d_k\hat{}(u)$ computed from components up to order 4,
9) entropy measures with penalty terms which can be used to determine how many components to use in the above test statistics

19. EXAMPLES OF DATA ANALYSIS. The interpretation of the diversity of statistics available is best illustrated by examples.

In order to compare our methods with others available we consider data analysed by Boos (1986) on ratio of assessed value to sale price of residential property in Fitchburg, Mass., 1979. The samples (denoted I, II, III, IV) represent dwellings in the categories single-family, two-family, three-family, four or more families. The sample sizes (54, 43, 31, 28) are proportions .346, .276, .199, .179 of the size 156 of the pooled sample. We compute Legendre, cosine, Hermite components up to order 4 of the 4 samples; they are asymptotically standard normal. We consider components greater than 2 (3) in absolute value to be significant (very significant).

Legendre, cosine, and Hermite components are very significant only for sample I, order 1 (-4.06, -4.22, -3.56 respectively). Legendre components are significant for sample IV, orders 1 and 2 (2.19, 2.31). Cosine components are significant for sample IV, orders I and II (2.36, 2.23) and sample III, order 1 (2.05). Hermite components are significant for sample IV, orders 2 and 3 (2.7 and -2.07).

Conclusions are that the four samples are not homogeneous (have the same distributions). Samples I and IV are significantly different from the pooled sample. Estimators of the comparison density show that sample I is more likely to have lower values than the pooled sample, and sample IV is more likely to have higher values. While all the statistical measures described above have been computed, the insights are provided by the linear rank statistics of orthogonal polynomials rather than by portmanteau statistics of Cramer-von Mises or Anderson-Darling type.

20. CONCLUSIONS. The goal of our recent research (see Parzen (1979), (1983)) on unifying statistical methods (especially using quantile function concepts) has been to help the development of both the theory and practice of statistical data analysis. Our ultimate aim is to make it easier to apply statistical methods by unifying them in ways that increase understanding, and thus enable researchers to more easily choose methods that provide greatest insight for their problem. We believe that if one can think of several ways of looking at a data analysis one should do so. However to relate and compare the answers, and thus arrive at a confident conclusion, a general framework seems to us to be required.

One of the motivations for this paper was to understand two-sample tests of the Anderson-Darling type; they are discussed by Pettitt (1976) and Scholz and Stephens (1987). This paper provides new formulas for these test statistics based on our new definition of sample comparison density functions. Asymptotic distribution theory for rank processes defined by Parzen (1983) is given by Aly, Csorgo, and Horvath (1987); an excellent review of theory for rank processes is given by Shorack and Wellner (1986).

H .ever one can look at k sample Anderson-Darling statistics as a single number formed from combining many test statistics called components. The importance of components is also advocated by Boos (1986), Eubank, La Riccia, and Rosenstein (1987) and Alexander (1989). Insight is greatly increased if instead of basing one's conclusions on the values of single test statistics, one looks at the components and also at graphs of the densities of which the components are linear functionals corresponding to various score functions. The question of which score functions to use can be answered by considering the tail behavior of the distributions that seem to fit the data.

REFERENCES

Alexander, William (1989) "Boundary kernel estimation of the two-sample comparison density function" Texas A&M Department of Statistics Ph. D. thesis.

Aly, E.A.A., M. Csorgo, and L. Horvath (1987) "P-P plots, rank processes, and Chernoff-Savage theorems" in *New Perspectives in Theoretical and Applied Statistics* (ed. M.L. Puri, J.P. Vilaplann, W. Wertz) New York: Wiley 135-156.

Boos, Dennis D. (1986) "Comparing k populations with linear rank statistics" *Journal of the American Statistical Association*, 81, 1018-1025.

Eubank, R.L., V.N. La Riccia, R.B. Rosenstein (1987) "Test statistics derived as components of Pearson's Phi-squared distance measure" *Journal of the American Statistical Association*, 82, 816-825.

Parzen, E. (1979) "Nonparametric statistical data modeling" *Journal of the American Statistical Association*, 74, 105-131.

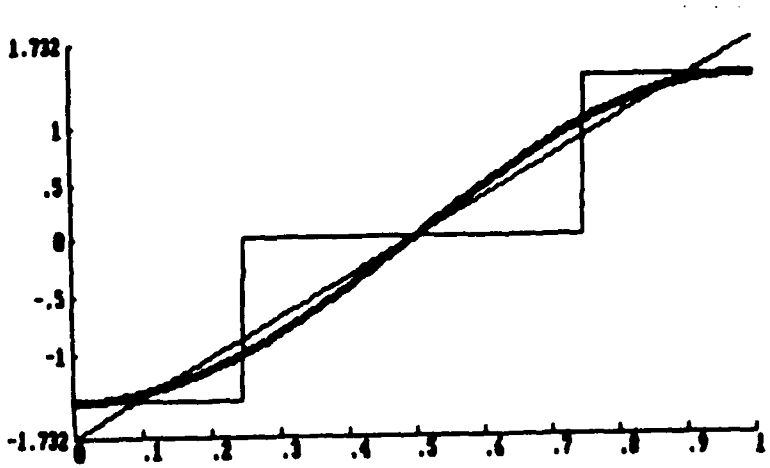
Parzen, E. (1983) "FunStat quantile approach to two-sample statistical data analysis" Texas A&M Institute of Statistics Technical Report A-21 April 1983.

Pettitt, A.N. (1976) "A two-sample Anderson Darling statistic" *Biometrika*, 63, 161-168.

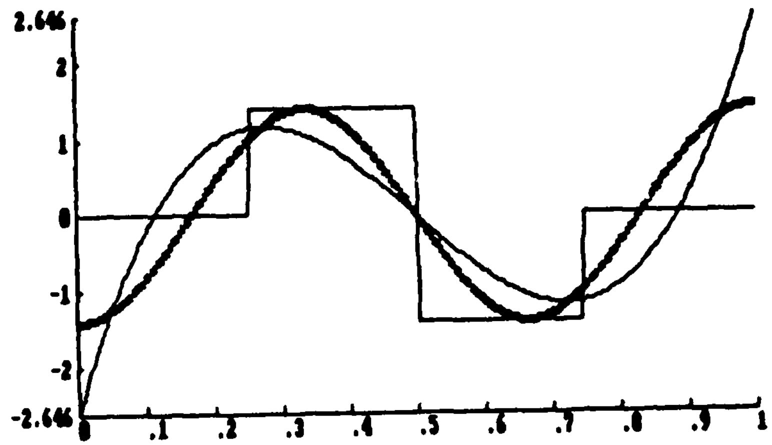
Scholz, F.W, and M.A.Stephens (1987) "k-sample Anderson- Darling tests" *Journal of the American Statistical Association*, 82, 918-924.

Shorack, Galen and Jon Wellner (1986) *Empirical Processes With Applications to Statistics* New York: Wiley.

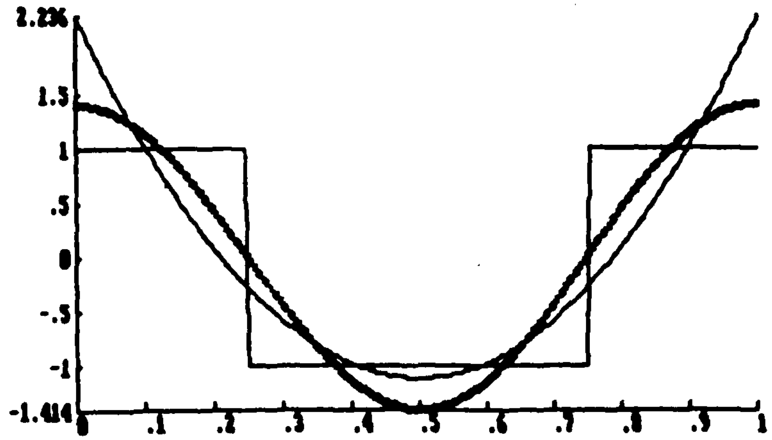
Score functions for location: SQ1 (square wave), phiL1 (light curve), phiC1 (dark curve).



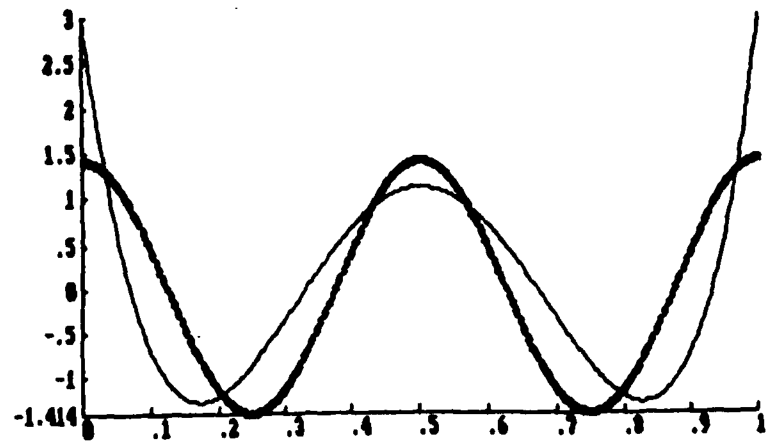
Score functions for skewness: SQ3 (square wave), phiL3 (light curve), phiC3 (dark curve).



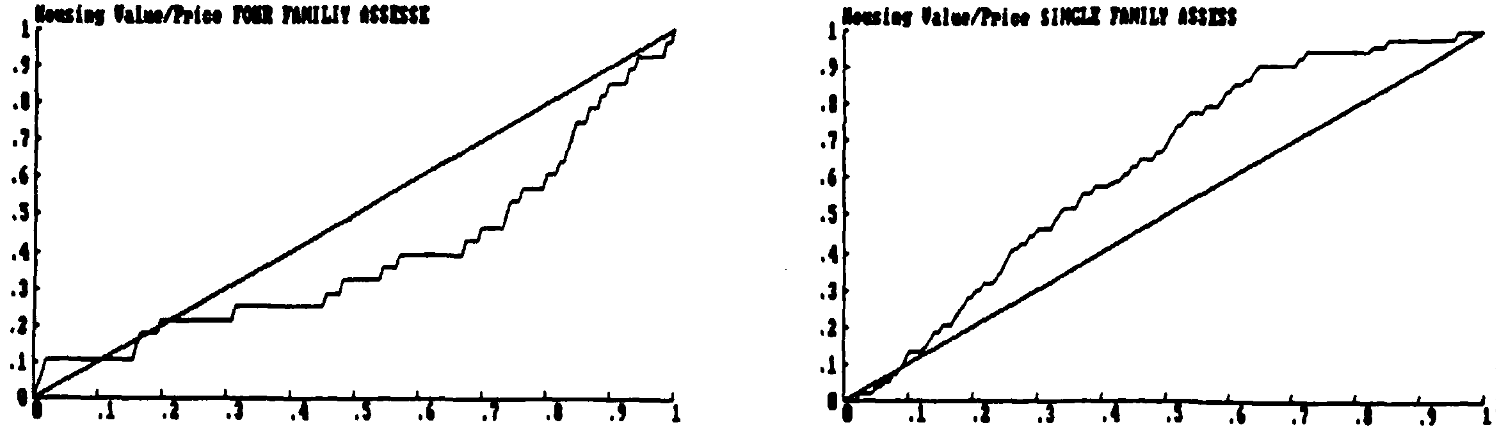
Score functions for scale: SQ2 (square wave), phiL2 (light curve), phiC2 (dark curve).



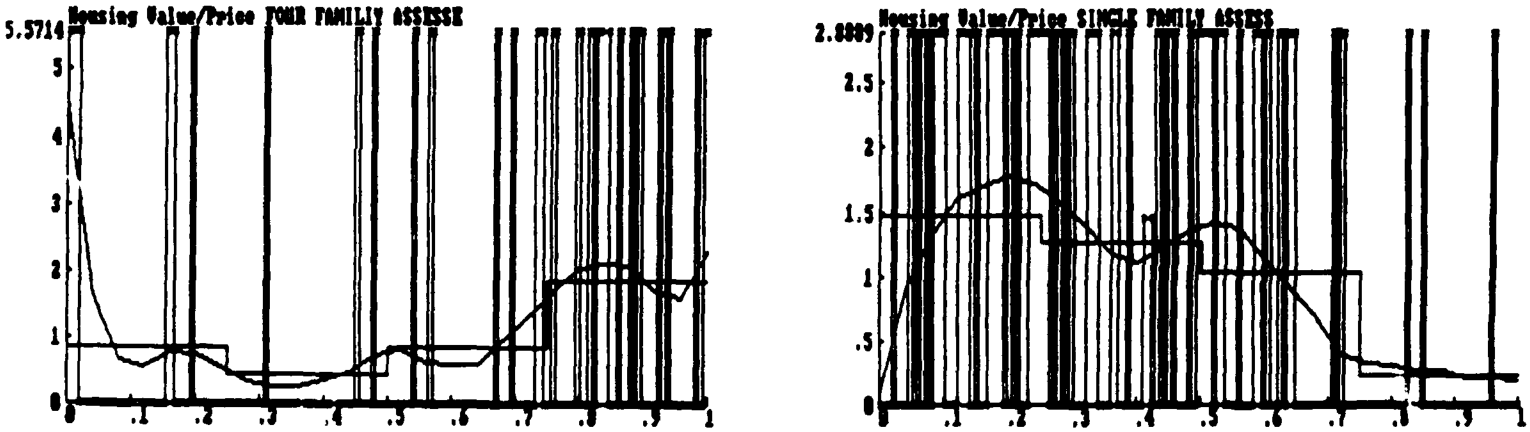
Score functions for kurtosis: phiL1 (light curve), phiC1 (dark curve).



For samples I and IV, sample comparison distribution function $D\tilde{}(u)$



For samples I and IV, sample comparison density $d\tilde{}(u)$, sample quartile density $dQ\tilde{}(u)$ (square wave), nonparametric density estimator $\hat{d}(u)$



For samples I and IV, Legendre, cosine, and Hermite orthogonal polynomial estimator of order 4 of the comparison density, denoted $d_4(u)$, compared to sample quartile density $dQ\tilde{}(u)$.