PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 31, Fasc. 1 (2011), pp. 161–174

CONFORMAL MEASURES AND DENSITY ESTIMATION

BY

MANFRED DENKER* (STATE COLLEGE) AND TINA FACCA (UNIVERSITY HEIGHTS)

Abstract. The notion of conformal measures or densities goes back to Patterson's work [8]. By a theorem of Milnor and Thurston [7] a piecewise monotone and continuous map of the interval is semiconjugate to one with constant slopes. The semiconjugacies can be defined by distribution functions of conformal measures as shown in [2]. In this note we show that for some transformations the conjugacies are estimable functions and can be used to improve estimation procedures, in particular density estimations.

2000 AMS Mathematics Subject Classification: Primary: 62G05, 37E05; Secondary: 62G07, 62G09, 37A05, 37A50.

Key words and phrases: Resampling, interval map, conjugacy, density estimation.

1. CONFORMAL MEASURES

Rokhlin's theory of measurable partitions (see [9]) in ergodic theory treats, among other topics, non-invertible measure-preserving transformations and their Jacobian J. If $R : \Omega_1 \to \Omega_2$ is a Borel map which is nonsingular with respect to two measures m_i on the Borel fields Σ_i ($i = 1, 2$), and if R is at most countable to one on the support of m_1, then the Jacobian J exists, i.e. it follows that

$$m_2\big(R(A)\big) = \int\limits_A J dm_1,$$

whenever $A \in \Sigma_1$ and R is invertible on A. The Jacobian J is m_1 a.e. uniquely defined. In particular, this applies to nonsingular self-mappings $R : \Omega_1 \to \Omega_1$.

The converse problem, when a nonsingular measure exists given a nonnegative measurable function J, has a less general answer. Of course, if R is a differentiable map on the unit interval, Lebesgue measure satisfies the above equality for J being the modulus of the derivative. Also, it follows from Ruelle's Perron–Frobenius

* The author was partially supported by the National Science Foundation under Grant Number DMS-1008538.

theorem (see [1]) that for an expanding and open map R of a compact metric space and a nonnegative continuous function J there is a measure μ satisfying

$$m\big(R(A)\big) = \int_A J dm \tag{1.1}$$

whenever $A \in \Sigma$ and R is invertible on A. Such measures are called the *Gibbs* ones. As the name suggests, the existence of such measures is closely connected to thermodynamic formalism for discrete time dynamical systems. Gibbs measures are special cases of conformal measures defined by equation (1.1) without any restriction, and $\log J$ is called its *potential*.

In a certain sense, the work of Patterson on limit sets of Fuchsian groups constitutes a breakthrough. Patterson's main result in [8] states that the limit set of a Fuchsian group of the second kind has Hausdorff dimension equal to the exponent of convergence of the Poincaré series. It connects notions from algebraic number theory with those from fractal geometry. The essential tool here is as well to construct conformal measures (this time for each conformal map given by the Fuchsian group).

As noted by Sullivan, Patterson's construction of a conformal measure is also valid for hyperbolic rational functions (see [11]). In fact, this follows as well from Ruelle's Perron–Frobenius theorem. In full generality the concept of conformal measures was investigated in [3]. It leads to the construction of conformal measures for (almost) any rational map, for maps on the unit interval, for circle maps and for non-compact spaces [10], to name a few. In particular, we have

THEOREM 1.1 (Denker and Stadlbauer [2]). *Let R be a continuous and piecewise monotone map with positive topological entropy $0 < h(R) < \infty$. Then there exists a non-atomic conformal measure for the constant function* $\exp[h(R)]$.

This theorem is used to give an explicit formula for estimating a conjugacy in Section 3. In the following Section 2 we review basic notions and theorems of conjugation between dynamical systems as needed in the sequel.

2. CONJUGACIES OF MAPS OF THE INTERVAL

In this note, a dynamical system consists of a topological space Ω and a continuous map $T : \Omega \to \Omega$. Two dynamical systems $T_i : \Omega_i \to \Omega_i$ $(i = 1, 2)$ are said to be *conjugate* if there exists a homeomorphism $h : \Omega_1 \to \Omega_2$ such that

$$T_2 \circ h = h \circ T_1,$$

and *semiconjugate* if h is merely continuous and onto. In particular, if each Ω_i is the unit interval $[0, 1]$, any semiconjugacy is a continuous monotone map of the unit interval leaving the two endpoints invariant, hence is the distribution function F (or $1 - F$) of a measure with no atoms. In particular, the semiconjugacy can be estimated if there are observables with distribution F.

A map of the unit interval is called *S-unimodal* if it is continuous and there is a partition of $[0,1]$ into S (maximal) subintervals, on which the map is monotone. A theorem of Milnor and Thurston [7] states that every unimodal map is semiconjugate to a tent map (a piecewise linear map with two monotonicity intervals). More generally, it is known that this result holds for S-unimodal maps and piecewise linear transformations. In fact, the semiconjugacy is a conjugacy if the distribution function of the associated conformal measure with respect to topological entropy is strictly monotone. As a consequence, every such S-unimodal map of the unit interval is conjugate to a piecewise monotone and continuous map of the interval with constant slopes. Hence one can produce transformations conjugate to a piecewise monotone map with constant slopes which have a given density f as its invariant density.

A different proof of the theorem of Milnor and Thurston has been given in [2] using a measure theoretic argument.

THEOREM 2.1 (Denker and Stadlbauer [2]). *Let $R : [0,1] \to [0,1]$ be a piecewise monotone and continuous transformation on the unit interval. Assume that*

$$\limsup_{n\to\infty} \frac{1}{n} \log c_n = \log M > 0,$$

where c_n denotes the number of monotonicity intervals of R^n. Then there exists a conformal measure m for the potential $\log M$ and

$$h(x) = m([0,x])$$

defines a semiconjugacy between R and a piecewise linear and continuous map T of the unit interval with slope M. The semiconjugacy h defines a conjugacy if and only if m is positive on nonempty open sets.

There has been a recent approach to formulate estimation problems for dynamical systems using the concept of computability of dynamical objects. This is understood in the present context as computable transformations on computable subsets of the real line in the sense of Turing machines. We refer to [6] for details. We also adopt the notion of an estimable parameter or function. By this we mean a sequence of statistics W_n (which is based on an i.i.d. sample of size n) which, when the sample size tends to infinity, is a consistent estimator of the parameter or the function.

An S-unimodal transformation $R : [0,1] \to [0,1]$ is called a *Bernoulli map* if each monotonicity interval is mapped onto the full unit interval. As an immediate consequence of Theorem 2.1 we obtain

COROLLARY 2.1. *Let $R : [0,1] \to [0,1]$ be a piecewise strictly monotone Bernoulli transformation with M monotonicity intervals. Then R is conjugate to*

the map $T : [0,1] \to [0,1]$ *with*

$$T(x) = \begin{cases} Mx - r + 1 & \text{if } (r-1)/M \leqslant x < r/M,\ 1 \leqslant r \leqslant M \text{ odd}, \\ r - Mx & \text{if } (r-1)/M \leqslant x < r/M,\ 1 \leqslant r \leqslant M \text{ even}. \end{cases}$$

If the inverse branches of R *and the endpoints of the monotonicity intervals of* R *are computable, so is the conjugacy. Moreover, if* R *is unknown and the inverse branches of* R *are estimable, so is the conjugacy* h*.*

P r o o f. The first assertion follows from the theorem above. Note that R is a Bernoulli map, so that the number of monotonicity intervals of R^n is $c_n = M^n$. Hence R is semiconjugate to the map T as defined above. The conjugacy is given by the conformal measure m for the map R and the potential $x \mapsto \log M$. It is easy to see that $m([a,b]) > 0$ for all $a < b$, so that, by the above theorem, h defines a conjugacy.

Let $p_0 = 0 < p_1 < \ldots < p_M = 1$ denote the endpoints of the monotonicity intervals of R. Since R is a Bernoulli map and the potential is constant, each interval $[p_i, p_{i+1}]$ maps onto $[0,1]$ and is scaled by the factor M; hence

$$h(p_i) = i/M, \quad i = 0, \ldots, M.$$

If $p_{i,n}$ denote the n-th preimages of R^n in increasing order ($i = 0, \ldots, M^n$), then the interval $[p_{i,n}, p_{i+1,n}]$ is mapped onto an interval of the form $[p_{j,n-1}, p_{j+1,n-1}]$ and scaled by the factor M. Since, by induction, $m([p_{j,n-1}, p_{j+1,n-1}]) = M^{-n+1}$, it follows that $m([p_{i,n}, p_{i+1,n}]) = M^{-n}$. Thus the preimages and their h-function values are computable. Since h is continuous, h is computable with arbitrary precision.

Finally, let μ be an R-invariant absolutely continuous measure with density f. Then its inverse branches are estimable, and so is h as shown above. ∎

PROPOSITION 2.1. *Let* $R_0 : [0,1] \to [0,1]$ *be a piecewise strictly monotone and differentiable, continuous Bernoulli map such that the inverse branches are computable. Let* X *be a random variable with values in* $[0,1]$ *and density* $f > 0$*. Let* $Y = R_0(X)$ *and* $G(x) = P(Y \leqslant x)$*. Then* $R = F^{-1} \circ G \circ R_0$ *is an invariant transformation for the absolutely continuous measure with density* f *and its inverse branches are estimable functions.*

P r o o f. Let F denote the distribution function of X. Then F is strictly increasing by assumption, whence

$$P\big(R(X) \leqslant x\big) = P\Big(R_0(X) \leqslant G^{-1}\big(F(x)\big)\Big) = F(x).$$

Since the inverse branches of R_0 are computable and since F and G and their inverses are estimable, the inverse branches of R are estimable. ∎

3. ESTIMATION OF CONJUGACIES

Let μ be an invariant probability measure for an S-unimodal transformation R on the unit interval. Denote by F the distribution function of μ. We assume that μ has no atoms and that F is strictly increasing on $[0,1]$. We are interested in estimating the conjugacy from Theorem 2.1 using a sample $X_1, \ldots, X_n$ of independent F-distributed random variables, when R is unknown.

In general, this problem cannot be solved. However, if R is of a specific form, we are able to find an estimator for the conjugacy. Assume $R = F^{-1} \circ G \circ R_0$, where $G(t) = \mu\big(\{x : R_0(x) \leqslant t\}\big)$, and where R_0 is a known S-unimodal transformation. Of course, R can be estimated by estimation of F^{-1} from the sample $X_1, \ldots, X_n$, and G can be estimated from the sample $Y_1 = R_0(X_1), \ldots, Y_n = R_0(X_n)$. The estimated transformation R will permit to estimate h, since h is well defined in terms of the constant slope $\pm M$ and the endpoints of the monotonicity intervals (see [2]).

To obtain more efficient estimators, we assume in addition that R_0 is a Bernoulli map with M monotonicity branches. In this case one easily calculates that

$$\limsup_{n\to\infty} \frac{1}{n} \log c_n = \log M > 0,$$

since there are $c_n = M^n$ monotonicity pieces for R_0^n. The conjugacy is given by the conformal measure m which scales by the factor M.

For a Bernoulli transformation we can determine the conjugacy because the preimages of 0 and 1 up to order N (denoted by $p_0 = 0 < p_1 < \ldots < p_{M^N} = 1$ in increasing order) determine the intervals which are mapped onto $[0,1]$ under R_0^N. Since the map is one-to-one on these intervals, the conformality equation (1.1) gives

$$m([p_i, p_{i+1}]) = \frac{1}{M^N},$$

which determines the distribution function h of the conformal measure m:

$$h(p_i) = \frac{i}{M^N}, \qquad i = 0, \ldots, M^N.$$

Note that the conjugacy is between the map R and the piecewise linear, continuous Bernoulli map with monotonicity intervals $[i/M, (i+1)/M]$, $i = 0, \ldots, M-1$, where the branches are increasing or decreasing corresponding to the map R.

It suffices to estimate preimages, say by $\widehat{p}_i$, $i = 0, \ldots, M^N$. We then define the estimator $\widehat{h}$ for h by linear interpolation:

$$\widehat{h}(x) = \frac{i}{M^N} + \frac{x - \widehat{p}_i}{\widehat{p}_{i+1} - \widehat{p}_i} \frac{1}{M^N}, \qquad \widehat{p}_i \leqslant x \leqslant \widehat{p}_{i+1};$$

other interpolation methods can be used as well. Let $\widehat{F}$ and $\widehat{G}$ denote the empirical distribution functions of F and G based on the samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$. Let $X_{(i)}$ and $Y_{(i)}$ $(1 \leqslant i \leqslant n)$ denote the order statistics of the X- and Y-samples.

LEMMA 3.1. *The preimages x_l of $t \in [0,1]$ are determined by*

$$F^{-1}\Big(G\big(R_0(x_l)\big)\Big) = t;$$

hence they can be estimated consistently by

$$x_l = R_{0,l}^{-1}(Y_{(k)}), \quad l = 1, \dots, M,$$

where $R_{0,l}^{-1}$ denotes the inverse branch of R_0 with image the l-th interval of monotonicity, and where k is the number of random variables X_l which is less than or equal to t.

Proof. $\widehat{F}(X_{(i)}) = i/n = \widehat{G}(Y_{(i)})$, and hence $G^{-1}\big(F(X_{(i)})\big)$ is estimated by $Y_{(i)}$. Since $X_{(k)}$ converges to t, where $X_{(k)} \leqslant t < X_{(k+1)}$, and $\widehat{F}$ and $\widehat{G^{-1}}$ are consistent estimators for F and G^{-1}, the claim follows. ∎

Once having estimated the inverses of the sample, we can estimate the inverses of the partition points $z_1, \dots, z_{M-1}$. We have (in case $R(0) = 0$ and $R(1) = 1$, the other cases are similar)

$$R^{-1}(1) = \{z_1, \dots, z_{M-1}, 1\},$$
$$R^{-1}(0) = \{0, z_1, \dots, z_{M-1}\}$$

and the preimages $R_l^{-1}(z_i)$ of z_i are estimated by $R_{0,l}^{-1}(Y_{(k)})$, where $X_{(k)} \leqslant z_i < X_{(k+1)}$. Linear (or other) interpolation is advisable here. In this way we can generate the tree of inverse images of the endpoints of the unit interval. Note that we need to assume that the inverse branches of R_0 are computable.

Of course, we can define the estimator of h^{-1} in an analogous manner.

REMARK 3.1. Note that the iterates of R, R^m for $m \geqslant 1$, can be estimated by the conjugacy as well:

$$\widehat{(F^{-1} \circ G \circ R_0)^m}(x) = \widehat{h^{-1}}\Big(T^m\big(\widehat{h}(x)\big)\Big), \quad 0 \leqslant x \leqslant 1.$$

This follows from the fact that $h \circ R^m = T^m \circ h$.

Resampling of order m of the sample $X_1, \dots, X_n$ is defined by

$$X_i^{(m)} = \widehat{h}^{-1}\Big(T^m\big(\widehat{h}(X_i)\big)\Big), \quad i = 1, \dots, n.$$

4. DENSITY ESTIMATION

The estimation of the conjugacy in Section 3 can be used for resampling procedures in order to improve estimators. Here we shall restrict to a discussion of density estimation using the resampling procedure explained in the previous section.

There are several procedures in the literature where resampling works well: jackknifing algorithm, bootstrapping, rank statistics for U-statistics, or log-averaging (almost sure central and local limit theorems) to name a few. While the new experiments in these cases are simple algebraic, probabilistic or analytic modifications of the data, the method here is more conceptual, since the analytic data manipulation relies on ergodic theory for interval maps.

Nonparametric density estimations are usually based on kernel estimations, orthogonal series estimations, splines or wavelet estimations. Here we restrict to kernel estimations in applications for simplicity. This is sufficient to show how the proposed method works. We will demonstrate by two examples in Sections 5 and 6 that stationarity allows one to draw more information from the data than independence alone permits. We propose the following algorithm for density estimation based on resampling; its detailed features are well known in nonparametric estimation (see, e.g., [12], [4], [5]).

As before, let $X_1, \ldots, X_n$ be independent and identically distributed random variables with values in the unit interval having density function f. In general, when the random variables take values in $\mathbb{R}$, we may transform to the unit interval using a differentiable map. Hence the above situation is general for an absolutely continuous distribution.

In particular, the choice of a bandwidth follows the general advice in the literature. Reproving the known theorems on optimal bandwidth in the present context is certainly possible (though not trivial).

We now explain the steps leading to an improved estimator based on a piecewise monotone and continuous map R of the unit interval. Some obvious variations are deduced from this algorithm.

Let $x_1, \ldots, x_n$ be a random sample taken from this sequence of random variables.

Step 0. Transform the data to the unit interval using an invertible affine transformation, also denoted by $x_1, \ldots, x_n$.

Step 1. Find a first nonparametric density estimation $\widetilde{f}$ based on the sample $x_1, \ldots, x_n$, using one of the standard nonparametric density estimation procedures.

Step 2. Choose an integer $M \geqslant 3$ odd. Adapt a transformation R_0 to the (approximate) range of the data $[\min x_i, \max x_i]$ by setting R_0 to be the identity outside of the range, having M monotonicity intervals, and satisfying $R_0(0) = 0$. This transformation has to be a Bernoulli transformation, e.g., define the transformation on the monotonicity pieces itself piecewise linear.

Step 3. Set $y_i = R_0(x_i) \pmod 1$ for $i = 1, \ldots, n$. If the original data comes from a distribution function F, and if under the transformation the new variables y_i represent the distribution function G, then the map $R = F^{-1} \circ G \circ R_0$ of the unit interval leaves F invariant (by Proposition 2.1).

Step 4. Estimate the conjugacy h between R and the map T which is defined by the formula

$$T(x) = \begin{cases} Mx & \text{if } 0 \leqslant x \leqslant 1/M, \\ 2 - Mx & \text{if } 1/M \leqslant x \leqslant 2/M, \\ Mx - 2 & \text{if } 2/M \leqslant x \leqslant 3/M, \\ 4 - Mx & \text{if } 3/M \leqslant x \leqslant 4/M, \\ \ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\ Mx - (M-1) & \text{if } (M-1)/M \leqslant x \leqslant 1. \end{cases}$$

Step 5. Fix m. Find the resampling of order m from $x_1, \ldots, x_n$ by the estimated transformation R to obtain $z_1, \ldots, z_n$.

Step 6. Find a nonparametric density estimation $\widehat{f}$ of the transformed data from Step 5. Compare this with $\widetilde{f}$.

Step 7. Find a nonparametric density estimator $\overline{f}$ of f based on the combined sample

$$x_1, \ldots, x_n, z_1, \ldots, z_n.$$

It is clear from the foregoing discussions, if n tends to infinity, then $\widehat{f}$ and $\overline{f}$ are consistent estimators as long as the nonparametric density estimators and the transformation estimators are consistent. The bandwidth of the smoother and the second nonparametric density estimator are not found in the literature. However, the order is the same as for the independent case. Hence one may use the same as in Step 1, where a usual cross-validation provides a reasonable bandwidth.

The method here is restricted to estimation of a density on some compact interval (one-dimensional estimation). It seems to be an interesting question to extend the method to two dimensions (like the Milnor–Thurston result is extended to two-dimensional setting in [2]). This has possible application in image analysis.

5. NUMERICAL EXAMPLE: GAUSSIAN DENSITIES

We illustrate the algorithm for data of the normal density with parameters $\mu = 0.5$ and $\sigma = 0.1$. Choosing these parameters avoids transforming the data to the unit interval $[0, 1]$ since it is unlikely that an observation falls outside of this interval.

We created a sample of 40 points using the Mathematica random generator for normal distributions (see Table 1).

In this numerical example we have chosen the map R_0 piecewise linear: Let a be the minimum and b be the maximum of the data set, $\epsilon = 10^{-3}$. Define points $q_0 = 0$, $q_1 = a$, $q_2 = a + (b-a)/3 - \epsilon$, $q_3 = a + (b-a)/3$, $q_4 = a + (b-a)/3 + \epsilon$, $q_5 = a + [2(b-a)]/3 - \epsilon$, $q_6 = a + [2(b-a)]/3$, $q_7 = a + [2(b-a)]/3 + \epsilon$,

TABLE 1. A sample of 40 random numbers from $\mathcal{N}(0.5, 0.1)$

0.2793477465775748	0.3189803299987347	0.3280039351987041
0.32937649327833474	0.3341873744577941	0.3509460493227462
0.37040988475312386	0.40839417131290184	0.4278725092176636
0.4343782997566407	0.4366380504552651	0.43722272310765375
0.4383535393175012	0.44698176229713144	0.4526806329695241
0.46147553894532484	0.46581162135041665	0.4811459155174444
0.4824342046026437	0.48690570695079405	0.48758337672241336
0.4895385089720199	0.5081612541014502	0.5150922835298403
0.5152187708070246	0.5196125504711763	0.5257071417718275
0.5333088530164423	0.547680924349553	0.5556449245557475
0.5741634496492634	0.5769541655097403	0.5885220901351191
0.5910341159576358	0.5928556560241721	0.608084747824479
0.6220367003216917	0.6344369511691449	0.6444635575704425
0.6524555370984191		

$q_8 = b$, $q_9 = 1$. Then R_0 is defined to be linear on these intervals leaving a and b fixed. The transformation T is defined by

$$T(x) = \begin{cases} 3x & \text{if } 0 \leqslant x \leqslant \frac{1}{3}, \\ 2 - 3x & \text{if } \frac{1}{3} \leqslant x \leqslant \frac{2}{3}, \\ 3x - 2 & \text{if } \frac{2}{3} \leqslant x \leqslant 1. \end{cases}$$

Finally, we used resampling of order 1, 2, 3, and 4, and kernel density estimators with the Gaussian kernel. While the improvements for order less than or equal to 3 are about 5%, the resampled data set of order 4 gives almost the exact density. The result is shown in Figures 1 and 2 together with the true normal density. The new estimator is surprisingly close to the original density.

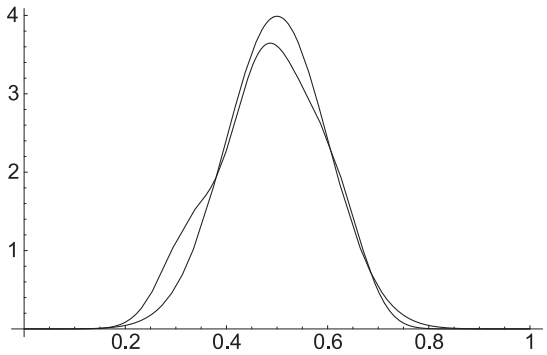

FIGURE 1. Density estimation from data of Table 1: Original data set

We also calculated the sample means of both data sets. While it is 0.486353 for the original sample, it becomes 0.492501 for the resampled data of order 4 (it is ~ 0.485 for order 2).

It should be noted that the new method relies on precise data up to the fourth digit (note that T^4 scales by the factor 81). Rounding the fifth decimal in the data

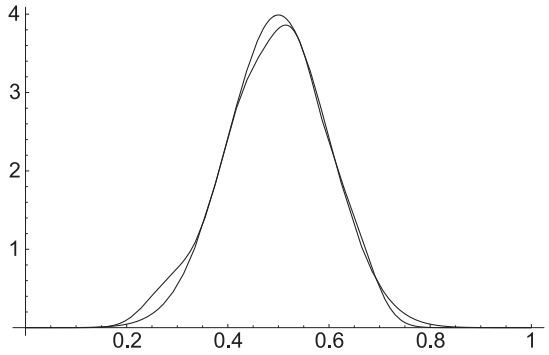

FIGURE 2. Density estimation from data of Table 1: Resampled data set of order 4

from Table 1 shows the same density as in Figure 2. Although the method needs accurate measurements, it is not demanding a practically unrealistic amount of precision (depending on the order of resampling).

REMARK 5.1. 1. We would like to point out that this example does not replace a systematic study of numerical applications of the method. We are just showing here that the new method has some potential, but it is not yet completely understood.

2. The resampled data set may be used in many ways, in particular also for density estimation: One may consider the combined data set together with a new bandwidth, or one may use the second sample to correct the density estimation from the original data.

3. Other parameters of the distribution may also be reexamined by the resampled data (to obtain a possibly biased but better estimate).

4. We did not attempt to provide a theoretical justification for the applicability of the method, except consistency (which is trivial).

6. LIFE BALANCE SCORES

In this section we discuss a data set collected in 2008 and show how the transformation method can be used to analyse a discrete data set. Strictly speaking, the method is not applicable when the distribution is not continuous. However, for distributions with sufficiently small jumps we expect the method to work. This will be demonstrated in this section using the 2008 data set of life balance scores.

In this example, seven variables, rated on 1–10 scales, were summed for an overall score, which was then scaled by 100. The data is from a "Life Balance Survey" administered *via* the Internet to 75 students at a Midwestern university in the United States in the fall of 2008. Respondents were asked to rate from seven to nine variables related to each type of balance in their lives, then an overall balance ques-

tion for the section was rated. Only the overall balance scores for each section were used in the analysis. These ratings summarized the respondents: self-perception of overall physical balance, spiritual balance, balance of mind, emotions, relationships, work, and environment.

The overall scores are reproduced in Table 2, which is rounded off from the original data. We have chosen a random sample from this population of size 17. This sample is reproduced in Table 3.

TABLE 2. Balance Score Data

0.114	0.3	0.414	0.5	0.614	0.614	0.657	0.657
0.671	0.686	0.686	0.686	0.686	0.686	0.686	0.7
0.7	0.714	0.714	0.729	0.743	0.743	0.743	0.743
0.743	0.757	0.757	0.764	0.764	0.764	0.764	0.764
0.764	0.764	0.764	0.764	0.764	0.764	0.764	0.764
0.764	0.764	0.764	0.764	0.771	0.771	0.786	0.786
0.8	0.8	0.8	0.8	0.814	0.814	0.814	0.814
0.829	0.857	0.857	0.857	0.871	0.871	0.886	0.886
0.886	0.886	0.914	0.914	0.929	0.943	0.943	0.971
0.971	0.999	0.999					

TABLE 3. Balance Score Data

0.414	0.614	0.657	0.686	0.686	0.686	0.686	0.743
0.764	0.764	0.764	0.8	0.8	0.814	0.929	0.999
0.999							

Our goal is to show how well we can recover the information in the grand data set from the sample of 17, which is roughly 20% of the observations. To this end we first choose representations for the data: One is obviously the mean, the other one a continuous representation of the data given by a smoothed density obtained from the full data set using a Gaussian kernel smoother and the traditional bandwidth $O(n^{1/5})$. This gives a representation of the data set in form of some type of density function. Figure 3 shows this representation derived from the full data set.

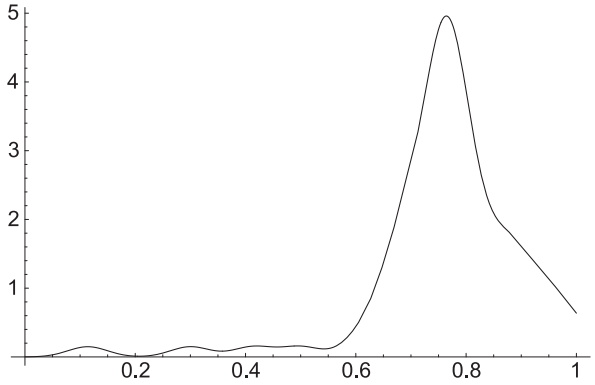

FIGURE 3. Gauss kernel representation for the life balance data

Next we produced the Gaussian kernel representation for the random sample in Table 3 using the same smoother as before. Figure 4 shows this representation. The agreement with Figure 3 is not too good. We repeated the last procedure with

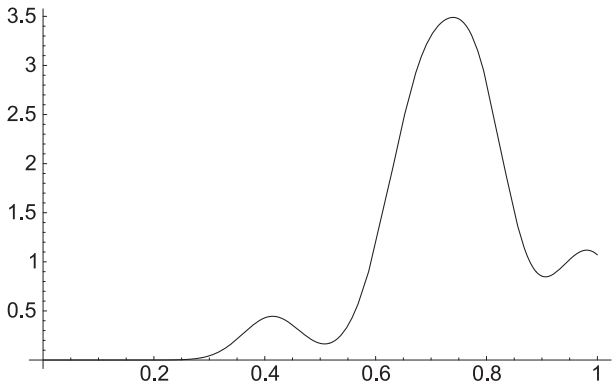

FIGURE 4. Gauss kernel representation for the random sample

the transformed data, which was combined with the original sample of 17 to make a sample of size 34. We used the order two for the resampling algorithm. The results are shown in Figure 5. Again, the agreement is not overwhelming.

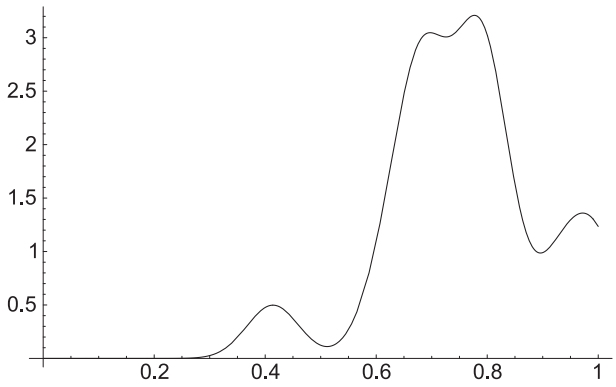

FIGURE 5. Gauss kernel representation for the transformed random sample

There are different ways to produce a new estimator of the representation for the grand data set. One estimator is that one based on the combined sample as shown in Figure 5. The other one is constructed as follows: Let f denote the representation of the grand data set, $\hat{f}$ the corresponding estimator of the random sample and $\tilde{f}$ the estimator based on the transformed data. Then $\hat{f} - \tilde{f}$ is a consistent es-

timator of $f - \hat{f}$, and hence

$$\hat{\hat{f}} = 2\hat{f} - \tilde{f}$$

is a consistent estimator of f. This estimator is shown in Figure 6 together with the representation of the grand data set.

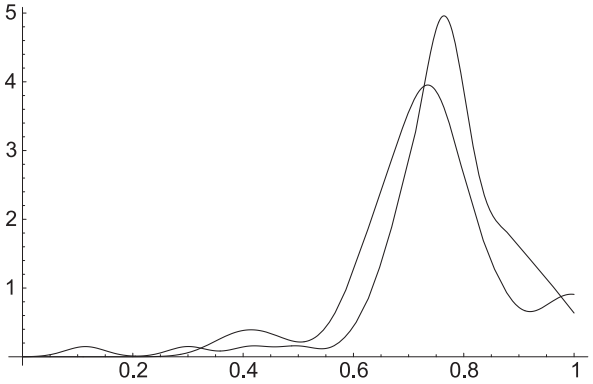

FIGURE 6. Gauss kernel representation of life balance data set *vs*. random sample data

The first representation of the data by their means results in the following estimates. The means of the three samples are as follows:

- Mean of the life balance score data: 0.763627.
- Mean of the random sample: 0.753235.
- Mean of the transformed random sample combined with the random sample: 0.760736.

As a conclusion of this discussion note that both characteristics can be estimated more accurately using the transformed data.

REFERENCES

[1] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math. Vol. 470, Springer, 1975.

[2] M. Denker and M. Stadlbauer, *Semiconjugacies for skew products of interval maps*, in: *Dynamics of Complex Systems*, RIMS Kokyuroku 1404 (2004), pp. 12–20.

[3] M. Denker and M. Urbański, *On the existence of conformal measures*, Trans. Amer. Math. Soc. 328 (1991), pp. 563–587.

[4] S. Efromovich, *Nonparametric Curve Estimation*, Springer, New York 1999.

[5] J. J. Faraway, *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Chapman & Hall, Boca Raton–London–New York 2006.

[6] S. Galatolo, M. Hoyrup and C. Rojas, *A constructive Borel–Cantelli lemma. Constructing orbits with required statistical properties*, Theoret. Comput. Sci. 410 (2009), pp. 2207–2222.

[7] J. Milnor and W. Thurston, *On iterated maps of the interval*, in: *Dynamical Systems (Maryland 1986–1987)*, Lecture Notes in Math. Vol. 1342, Springer, 1988, pp. 465–563.

[8] S. J. Patterson, *The limit set of a Fuchsian group*, Acta Math. 136 (1976), pp. 241–273.

[9] V. A. Rokhlin, *On the fundamental ideas of measure theory*, Mat. Sb. 25 (1949), pp. 107–150. Translation: Amer. Math. Soc. Transl. 71 (1952), pp. 1–54.
[10] M. Stadlbauer, *On random topological Markov chains with big images and preimages*, Stochastics & Dynamics 10 (2010), pp. 77–95.
[11] D. Sullivan, *Conformal dynamical systems*, in: *Geometric Dynamics*, Lecture Notes in Math. Vol. 1007, Springer, 1983, pp. 725–752.
[12] L. Waterman, *All of Nonparametric Statistics*, Springer, New York 2006.

Mathematics Department
Pennsylvania State University
McAllister Building
State College, PA 16802, USA
E-mail: denker@math.psu.edu

Management, Marketing and Logistics
Boler School of Business, #209
John Carroll University
University Heights, OH 44118, USA
E-mail: tfacca@jcu.edu

Received on 6.8.2010;
revised version on 23.2.2011