ESAIM: PS 20 (2016) 510–526
DOI: 10.1051/ps/2016003

ESAIM: Probability and Statistics
www.esaim-ps.org

A TEST FOR THE EQUALITY OF MONOTONE TRANSFORMATIONS OF TWO RANDOM VARIABLES

Mohamed Boutahar[1] and Denys Pommeret[1]

Abstract. It is frequent that observations arise from a random variable modified by an unknown transformation. This problem is considered in a two-sample context when two random variables are perturbed by two unknown transformations. We propose a test for the equality of those transformations. Two cases are considered: first, the two random variables have known distributions. Second, they have unknown distributions but they are observed before transformations. We propose nonparametric test statistics based on empirical cumulative distribution functions. In the first case the asymptotic distribution of the test statistic is the standard normal distribution. In the second case it is shown that the asymptotic distribution is a convolution of exponential distributions. The convergence under contiguous alternatives is studied. Monte Carlo studies are performed to analyze the level and the power of the test. An illustration is presented through a real data set.

Mathematics Subject Classification. 62G10, 62G05.

Received March 19, 2015. Revised January 4, 2016. Accepted February 2nd, 2016.

1. Introduction

Transformations of random variables appear naturally in probability and have many applications. In the case of unknown transformations, estimators may be constructed by non-parametric methods. This situation is common in statistic when a signal is observed after being disturbed by an unknown transformation. A simple case is the linear one, and old results exist to construct estimators of the parameters (see for instance [14]). More generally, there are many situations where an unknown transformation act on a signal, as for instance in the field of acoustical or optic signals (see the examples given in [7]) or in finance (see [11], or more recently [10]) when only the transformed signal is observed. This phenomena can be modeled as follows: If Y is the original unobserved signal, denoting by g the unknown transformation acting on Y, we assume that the transformed signal $X = g(Y)$ is observed. When Y is a random variable and g is a measurable function, the observation of an i.i.d. sample from X can lead to estimation and test on g, as soon as we have information on Y. For instance if the distribution of Y is known or estimated and if $g(y) = ay + b$, the problem consists in estimating a and b. This particular situation coincides with the estimation of both the mean and the variance of $g(Y)$. But in general the form of g is not specified and the problem is non-parametric. Another situation concerns the two sample case, when two transformed signals are observed and when the problem is to estimate and to compare the two transformations.

Keywords and phrases. Empirical cumulative distribution, nonlinear transformation, nonparametric estimation.

[1] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France. denys.pommeret@univ-amu.fr

Article published by EDP Sciences

A very simple frequently encountered situation is when both transformations are translations; in this basic case the comparison of functions reduces to a comparison of means. Similarly, one can consider a comparison of variances if the transformations are scaling functions. In a more general case, at our knowledge, it seems that few works have been intended to compare general transformations of random variables. The most closed works are those of [1,2] concerning comparisons of counting processes, this of [9] comparing autocovariances functions, and this of [8] comparing variances. Another similar problem of comparison of transformations is considered in [3] within the frame of panels data, with a technic based on CUSUM statistic. See also the work of [16] testing mean functions within the frame of panel count data.

In this paper we consider a non-parametric two sample problem as follows: we assume that two transformed random variables are observed, say X and $\tilde{X}$, and are transformations of two signals Y and $\tilde{Y}$, namely

$$X = g(Y), \;\; \tilde{X} = \tilde{g}(\tilde{Y}), \tag{1.1}$$

where g and $\tilde{g}$ are continuous monotone unknown functions. Our purpose is to test the equality between $g(y)$ and $\tilde{g}(y)$ for any given value y in the support of the probability distributions of Y and $\tilde{Y}$. The choice of y can be guided by prior information, as for instance $y = 0$ if Y and $\tilde{Y}$ are known to be centered. Then we consider the hypotheses

$$H_0(y) : g(y) = \tilde{g}(y) \text{ against } H_1(y) : g(y) \neq \tilde{g}(y), \tag{1.2}$$

based on two i.i.d. samples satisfying (1.1). The problem of testing (1.2) is of interest in various applications when a signal is noised. We illustrate this situation with the Framingham study on coronary heart disease described by [4]. The data consist of measurements of systolic blood pressure obtained at two different dates, and measured twice for each individual. The two measures of the first date coincide with two different random variables Y and $\tilde{Y}$, and the two measures of the second date are considered as an unknown transformations of the first ones, that is: $X = g(Y)$ and $\tilde{X} = \tilde{g}(\tilde{Y})$. Our purpose is to test the equality of these transformations, the null hypothesis meaning that there is a stability of the transformation of the systolic blood pressure since the modifications of the measurements are both the same.

To construct our test statistic we will distinguish two important cases:

Case 1: The distributions of Y and $\tilde{Y}$ are known and we observe two samples from (1.1).This situation may be encountered when two signals are controlled in entry but observed with perturbations in exit of a system. In this case the choice of y in (1.2) can take into account the knowledge of the distributions, as for instance a mode of a distribution, or the mean.

Case 2: The distributions of Y and $\tilde{Y}$ are unknown and we first observe two independent samples permitting to estimate their distributions (similarly to [6] or [5]). Then we observe contaminated samples X and $\tilde{X}$ satisfying (1.1). This situation may be encountered when two unknown signals are observed at the same time with and without transformation.

For both cases we construct a test statistics based on non parametric empirical estimators of g and $\tilde{g}$, adapting limit results on empirical processes. In Case 1, we obtain a standard normal asymptotic distribution for the test statistic. In Case 2, the asymptotic null distribution is a convolution of exponential distributions. Our test statistics can be easily implemented and we observe through simulations that they have a good power against various alternatives.

It is clear that when the hypothesis $H_0(y)$ is rejected, then the more general hypothesis $H_0 : g = \tilde{g}$ can also be rejected. However, if $H_0(y)$ is not rejected, no general conclusion can be drawn. In this case, a resampling procedure may be used to test the global equality H_0. We will develop this approach in our simulation study. We illustrate the test procedure with a study of the Framingham dataset.

The paper is organized as follows: in Section 1 we consider the construction of the test statistics. First the case where the two original signals have known distributions is considered. Then we relax this assumption by assuming that we observe the two signals after and before perturbations. The convergence of both test statistics under the null and under contiguous alternatives are obtained. In Section 2 a simulation study is presented and a complementary bootstrap procedure if proposed. A real data set is analyzed in Section 3.

2. The test statistics

2.1. Case 1: The two signal distributions of Y and $\tilde{Y}$ are known

Let Y and $\tilde{Y}$ be two independent random variables and consider n (resp. $\tilde{n}$) i.i.d. observations $X_1, \ldots, X_n$ (resp. $\tilde{X}_1, \ldots, \tilde{X}_{\tilde{n}}$) from (1.1). We will denote by F_Y and $F_{\tilde{Y}}$ the cumulative distribution functions of Y and $\tilde{Y}$, respectively. It is assumed here that these functions are known and invertible. The cumulative distribution functions of X and $\tilde{X}$ will be denoted by F_X and $F_{\tilde{X}}$. We will denote by $\mathcal{X}$ (resp. $\tilde{\mathcal{X}}$) the support of the probability distribution of X (resp. of $\tilde{X}$). It is assumed that the transformations g and $\tilde{g}$ are monotone and, without loss of generality, that they are increasing. Note that $g(y) = F_X^{-1}(F_Y(y))$ and $\tilde{g}(y) = F_{\tilde{X}}^{-1}(F_{\tilde{Y}}(y))$. Hence natural nonparametric estimators of the contaminating functions are given by

$$\hat{g}(\cdot) = X_{([nF_Y(\cdot)]+1)} \quad \text{and} \quad \hat{\tilde{g}}(\cdot) = \tilde{X}_{([\tilde{n}F_{\tilde{Y}}(\cdot)]+1)}, \tag{2.1}$$

where $X_{(i)}$ and $\tilde{X}_{(i)}$ denote the ith order statistics, and $[x]$ denotes the integer part of the real x. A fundamental Theorem (see for instance [15]) states the following convergence in distribution

$$\sqrt{n}\big(Z_{([np]+1)} - F_Z^{-1}(p)\big) \xrightarrow{D} \mathcal{N}\left(0, \frac{p(1-p)}{f_Z^2(F_Z^{-1}(p))}\right), \ \forall p \in (0,1), \tag{2.2}$$

for any i.i.d. sample $(Z_1, \ldots, Z_n)$, where $\xrightarrow{D}$ denotes the convergence in distribution, f_Z denotes the common density of $Z_i, i = 1, \ldots, n$, and $\mathcal{N}(m, \sigma^2)$ the normal distribution with mean m and variance σ^2. We will need the following standard assumptions:

- (A_1) There exists $0 < a < 1$ such that $n/(n + \tilde{n}) \to a$.
- (A_2) For all $x \in \mathcal{X}$ there exists $c > 0$ such that $f_X(x) > c$, and for all $x' \in \tilde{\mathcal{X}}$, there exists $\tilde{c} > 0$ such that $f_{\tilde{X}}(x') > \tilde{c}$; and f_X is $\mathcal{C}^k$, $f_{\tilde{X}}$ is $\mathcal{C}^{\tilde{k}}$, for some positive integers $k, \tilde{k}$.
- (A_3) There exists constants $B, \tilde{B} > 0$ and $B', \tilde{B}' > 0$ such that, for all $x \in \mathcal{X}$, for all $x' \in \tilde{\mathcal{X}}$, $f_X(x) < B$, $f_{\tilde{X}}(x') < \tilde{B}$ and $f'_X(x) < B'$, $f'_{\tilde{X}}(x') < \tilde{B}'$, where f' denotes the derivative.
- $(A(y))$: $0 < F_Y(y) < 1$ or $0 < F_{\tilde{Y}}(y) < 1$,

We deduce a first result which is a main tool for the construction of the test statistic.

Proposition 2.1. *Let assumptions* (A_1)–(A_2) *hold. Under* $H_0(y)$*, for* y *satisfying the assumption* $(A(y))$ *we have*

$$\sqrt{\frac{n\tilde{n}}{n+\tilde{n}}}\left(\hat{g}(y) - \hat{\tilde{g}}(y)\right) \xrightarrow{D} \mathcal{N}(0, \sigma^2(y)), \ \text{as } n \to \infty, \tilde{n} \to \infty, \tag{2.3}$$

where

$$\sigma^2(y) = (1-a)\frac{F_Y(y)(1-F_Y(y))}{f_X^2(g(y))} + a\frac{F_{\tilde{Y}}(y)(1-F_{\tilde{Y}}(y))}{f_{\tilde{X}}^2(\tilde{g}(y))}\cdot$$

Proof. By choosing $(Z_1, \ldots, Z_n) = (X_1, \ldots, X_n)$ and $p = F_Y(y)$, respectively $(Z_1, \ldots, Z_{\tilde{n}}) = (\tilde{X}_1, \ldots, \tilde{X}_{\tilde{n}})$ and $p = F_{\tilde{Y}}(y)$, in (2.2) we get

$$\sqrt{n}\big(\hat{g}(y) - g(y)\big) \xrightarrow{D} \mathcal{N}\left(0, \sigma_1^2\right), \tag{2.4}$$

$$\sqrt{\tilde{n}}\big(\hat{\tilde{g}}(y) - \tilde{g}(y)\big) \xrightarrow{D} \mathcal{N}\left(0, \sigma_2^2\right), \tag{2.5}$$

where

$$\sigma_1^2 = \frac{F_Y(y)(1-F_Y(y))}{f_X^2(g(y))}, \qquad \sigma_2^2 = \frac{F_{\tilde{Y}}(y)(1-F_{\tilde{Y}}(y))}{f_{\tilde{X}}^2(\tilde{g}(y))}\cdot$$

Combining assumption (A_1), (2.4), (2.5), and the independence of the two samples, it follows that under the null $H_0(y)$:

$$\sqrt{\frac{n\tilde{n}}{n+\tilde{n}}}\left(\hat{g}(y)-\hat{\tilde{g}}(y)\right) \xrightarrow{D} \mathcal{N}(0,\sigma^2(y)), \text{ as } n\to\infty, \tilde{n}\to\infty. \qquad \square$$

Remark 2.2. Assumptions $(A_1),(A_2)$ and $(A(y))$ ensure that $\sigma^2(y)>0$ and hence the limit in (2.3) is not degenerate.

We will estimate the variance σ^2 by using a nonparametric method. Consider a kernel $K(\cdot)$ such that

- (A_4) The support $\mathcal{K}$ of K is compact and K is of the kth degree; that is, its moments satisfy: $\int_{\mathcal{K}} y^i K(\mathrm{d}y)=0$ for $i<k$ and $\int_{\mathcal{K}} y^k K(\mathrm{d}y)\neq 0$.

For instance we can choose the quartic kernel defined by $K(x) = \frac{15}{16}(1-x^2)^2\mathbf{1}_{(-1,1)}(x)$, and an associated bandwidth h_n. In the sequel, we will set $K_{h_n}(x)=K(\frac{x}{h_n})$. Write

$$\hat{f}_X(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K_{h_n}(X_i-x) \text{ and } \hat{f}_{\tilde{X}}(x) = \frac{1}{\tilde{n}h_{\tilde{n}}}\sum_{i=1}^{\tilde{n}} K_{h_{\tilde{n}}}(\tilde{X}_i-x), \tag{2.6}$$

and to avoid small values for denominators in the estimation of the variance we use

$$\hat{f}_{X,e_n}(x) = \max\left(\hat{f}_X(x), e_n\right) \text{ and } \hat{f}_{\tilde{X},e_{\tilde{n}}}(x) = \max\left(\hat{f}_{\tilde{X}}(x), e_{\tilde{n}}\right),$$

where $e_n>0$ and $e_n\to 0$, when n tend to infinity. An estimator of σ^2 is then

$$\widehat{\sigma}^2(y) = (1-a)\frac{F_Y(y)(1-F_Y(y))}{\widehat{f}^2_{X,e_n}(\widehat{g}(y))} + a\frac{F_{\tilde{Y}}(y)(1-F_{\tilde{Y}}(y))}{\widehat{f}^2_{\tilde{X},e_n}(\widehat{\tilde{g}}(y))},$$

and we consider the statistic

$$T_1(y) = \frac{n\tilde{n}}{n+\tilde{n}}\widehat{\sigma}(y)^{-2}\left(\hat{g}(y)-\hat{\tilde{g}}(y)\right)^2. \tag{2.7}$$

Proposition 2.3. *Let assumptions* (A_1)–(A_4) *hold. If* $h_n \simeq n^{-c_1}$, $e_n \simeq n^{-c_2}$ *for some positive constants* $0<c_1<1$ *and* $0<c_2<1/4$ *such that* $\frac{c_2}{k}<c_1<1-2c_2$, *then under* $H_0(y)$ *with* y *satisfying* $(A(y))$ *we have:*

$$T_1(y) \xrightarrow{D} Z_1(y), \text{ as } n\to\infty, \tilde{n}\to\infty, \tag{2.8}$$

where $Z_1(y)$ *is chi-squared distributed with one degree of freedom.*

Proof. We need the fundamental lemma (see [13]):

Lemma 2.4. *Under assumptions* (A_2) *and* (A_4) *the kernel estimators given by* (2.6) *satisfy*

$$\sup_{x\in\mathcal{X}} |\hat{f}_X^2(x) - f_X^2(x)| = \mathcal{O}_{\mathbb{P}}\left(h_n^{2k} + \frac{\log n}{nh_n}\right)$$

$$\sup_{x\in\tilde{\mathcal{X}}} |\hat{f}_{\tilde{X}}^2(x) - f_{\tilde{X}}^2(x)| = \mathcal{O}_{\mathbb{P}}\left(h_{\tilde{n}}^{2k} + \frac{\log \tilde{n}}{\tilde{n}h_{\tilde{n}}}\right)\cdot$$

We can write

$$\widehat{\sigma}^2(y) = \frac{u(y)}{\widehat{f}^2_{X,e_n}(\widehat{g}(y))} + \frac{v(y)}{\widehat{f}^2_{\tilde{X},e_{\tilde{n}}}(\widehat{\tilde{g}}(y))},$$

where $u(y) = (1-a)F_Y(y)(1-F_Y(y))$ and $v(y) = aF_{\tilde{Y}}(y)(1-F_{\tilde{Y}}(y))$. Using the mean value theorem there exist $A > 0$ and $B > 0$ such that

$$\begin{aligned}\widehat{\sigma}^2(y) &= \sigma^2(y) + u(y)\Big(\widehat{f}^2_{X,e_n}(\widehat{g}(y)) - f^2_X(g(y))\Big)\Big(\frac{-1}{A^2}\Big)\\ &\quad + v(y)\Big(\widehat{f}^2_{\tilde{X},e_{\tilde{n}}}(\widehat{\tilde{g}}(y)) - f^2_{\tilde{X}}(\tilde{g}(y))\Big)\Big(\frac{-1}{B^2}\Big),\end{aligned}$$

where $A^2 \geq \min(\widehat{f}^2_{X,e_n}(\widehat{g}(y)), f^2_X(g(y)) \geq \min(e_n^2, f^2_X(g(y)))$ and $B^2 \geq \min(e_{\tilde{n}}^2, f^2_{\tilde{X}}(\tilde{g}(y)))$. We have

$$\begin{aligned}\Big|\widehat{f}^2_{X,e_n}(\widehat{g}(y)) - f^2_X(g(y))\Big| &\leq \Big|\widehat{f}^2_{X,e_n}(\widehat{g}(y)) - \widehat{f}^2_X(\widehat{g}(y))\Big|\\ &\quad + \Big|\widehat{f}^2_X(\widehat{g}(y)) - f^2_X(\widehat{g}(y))\Big| + \Big|f^2_X(\widehat{g}(y)) - f^2_X(g(y))\Big|\\ &= (I) + (II) + (III).\end{aligned}$$

We clearly have $(I) \leq 2\, e_n^2$, and

$$\frac{(I)}{A^2} \leq \frac{2e_n^2}{\max(e_n^2, f^2_X(g(y)))} = 2\mathbf{1}_{\{f_X(g(y))<e_n\}} + \frac{2e_n^2}{f^2_X(g(y))}\mathbf{1}_{\{f_X(g(y))>e_n\}} \;=\; o_{\mathbb{P}}(1), \quad \forall y \in \mathcal{Y}.$$

From Lemma 2.4, $(II) = \mathcal{O}_{\mathbb{P}}\left(h_n^{2k} + \frac{\log n}{nh_n}\right)$, and then

$$\frac{(II)}{A^2} = o_{\mathbb{P}}(1).$$

By the mean value theorem combined with (A_3) we have

$$(III) \leq 2BB'\Big|\widehat{g}(y) - g(y)\Big| \;= \mathcal{O}_{\mathbb{P}}(1/\sqrt{n}),$$

and by assumption

$$\frac{(III)}{A^2} = o_{\mathbb{P}}(1).$$

We finally have

$$\Big|\widehat{f}^2_{X,e_n}(\widehat{g}(y)) - f^2_X(g(y))\Big| = o_{\mathbb{P}}(1).$$

In the same manner we can see that $\Big|\widehat{f}^2_{\tilde{X},e_{\tilde{n}}}(\widehat{\tilde{g}}(y)) - f^2_{\tilde{X}}(\tilde{g}(y))\Big| = o_{\mathbb{P}}(1)$ and since u and v are bounded we obtain

$$\widehat{\sigma}^2(y) = \sigma^2(y) + o_{\mathbb{P}}(1). \tag{2.9}$$

To prove (2.8) we can write

$$T_1(y) = \frac{\sigma^2(y)}{\widehat{\sigma}^2(y)}\frac{n\tilde{n}}{n+\tilde{n}}\left(\frac{\hat{g}(y) - \hat{\tilde{g}}(y)}{\sigma(y)}\right)^2,$$

hence the convergence (2.8) follows from (2.9) and Proposition 2.1.

$\square$

2.2. Case 2: The two signal distributions of Y and $\tilde{Y}$ are unknown but observed before transformations

Consider n_x (resp. $\tilde{n_x}$) i.i.d. observations $X_1, \ldots, X_{n_x}$ (resp. $\tilde{X}_1, \ldots, \tilde{X}_{\tilde{n_x}}$) and n_y (resp. $\tilde{n_y}$) i.i.d. observations $Y_1, \ldots, Y_{n_y}$ (resp. $\tilde{Y}_1, \ldots, \tilde{Y}_{\tilde{n_y}}$), such that the X's and $\tilde{X}$'s satisfy (1.1). It is assumed that Y_i and $\tilde{Y}_j$ are independent for all i, j, and for simplicity of notation we set $n_x = n_y = n$ and $\tilde{n_x} = \tilde{n_y} = \tilde{n}$.

The two samples $Y_1, \ldots, Y_n$ and $\tilde{Y}_1, \ldots, \tilde{Y}_{\tilde{n}}$ can be viewed as two independent training set which permit to estimate the initial densities of the signals before perturbations. Again we want to test $H_0(y) : g(y) = \tilde{g}(y)$, but we now estimate g and $\tilde{g}$ by

$$\hat{g}(\cdot) = \widehat{F}_X^{-1}\left(\widehat{F}_Y(\cdot)\right) = X_{([n\widehat{F}_Y(\cdot)])}$$

and

$$\hat{\tilde{g}}(\cdot) = \widehat{F}_{\tilde{X}}^{-1}\left(\widehat{F}_{\tilde{Y}}(\cdot)\right) = \tilde{X}_{([\tilde{n}\widehat{F}_{\tilde{Y}}(\cdot)])}, \tag{2.10}$$

where

$$\widehat{F}_Y(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{Y_i \le y\}} \text{ and } \widehat{F}_{\tilde{Y}}(y) = \frac{1}{\tilde{n}}\sum_{i=1}^{\tilde{n}} \mathbf{1}_{\{\tilde{Y}_i \le y\}}, \tag{2.11}$$

are the empirical distribution functions of Y and $\tilde{Y}$ respectively.
Let

$$\hat{\lambda}_1(y) = \hat{f}_{\tilde{X}}(\hat{\tilde{g}}(y)) \text{ and } \hat{\lambda}_2(y) = \hat{f}_X(\hat{g}(y)),$$

where $\hat{f}_X$ and $\hat{f}_{\tilde{X}}$ are given by (2.6). When the distributions of Y and $\tilde{Y}$ are known, Proposition 2.1 states that the limiting distribution of $G(n,\tilde{n}) = \sqrt{\dfrac{n\tilde{n}}{n+\tilde{n}}}\left(\hat{g}(y) - \hat{\tilde{g}}(y)\right)$ is Gaussian. If the distributions of Y and $\tilde{Y}$ are unknown, the limiting distribution of $G(n,\tilde{n})$ is not Gaussian but a convolution of exponential distributions as follows:

Proposition 2.5. *Let assumptions* (A_1)–(A_4) *hold. Under* $H_0(y)$ *we have*

$$\frac{n\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)\hat{\lambda}_2(y)\left(\hat{\tilde{g}}(y) - \hat{g}(y)\right) \xrightarrow{D} X(y), \text{ as } n \to \infty, \tilde{n} \to \infty, \tag{2.12}$$

with

$$X(y) = (1-a)\lambda_1(y)E_1 - a\lambda_2(y)E_2, \tag{2.13}$$

where E_1, E_2 *are two independent exponential random variables with mean 1,* $\lambda_1(y) = f_{\tilde{X}}(\tilde{g}(y))$, *and* $\lambda_2(y) = f_X(g(y))$.

Proof. We first show that

$$E_{n,1} = n\hat{\lambda}_2(y)\left(g(y) - \hat{g}(y)\right) \xrightarrow{D} E_1 \text{ as } n \to \infty. \tag{2.14}$$

Let

$$K_i = \begin{cases} X_i & \text{if } X_i \le g(y) \\ -\infty & \text{if } X_i > g(y) \end{cases}, i = 1, \ldots, n.$$

Under $H_0(y)$ we have

$$\widehat{F}_Y(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \le g(y)},$$

that we combine with (2.1) to get

$$\hat{g}(y) = \max_{1\le i\le n} K_i, \quad \text{and} \quad \hat{g}(y) \le g(y).$$

Note first that for all $z \ge g(y)$, $P(K_1 \le z) = 1$. Assume that $0 \le z \le g(y)$. Then

$$P(\hat{g}(y) \le z) = (P(K_1 \le z))^n, \text{ since the variables } K_i \text{ are i.i.d.}$$

We now use

$$\{K_1 \le z\} = \{K_1 \le z\} \cap \{X_1 \le g(y)\} \cup \{K_1 \le z\} \cap \{X_1 \ge g(y)\}.$$

If $X_1 \le g(y)$ then $\{K_1 \le z\} = \{X_1 \le z\}$ and if $X_1 > g(y)$ then $\{K_1 \le z\} = \Omega$.
Therefore we have

$$\begin{aligned} P(K_1 \le z) &= P\left(\{X_1 \le z\} \cap \{X_1 \le g(y)\}\right) + P\left(X_1 \ge g(y)\right) \\ &= F_X(z) + 1 - F_X(g(y)). \end{aligned}$$

Hence

$$\begin{aligned} F_{\hat{g}}(z) &= P(\hat{g}(y) \le z) \\ &= \{1 + F_X(z) - F_X(g(y)))\}^n, \end{aligned}$$

and

$$\begin{aligned} P(n(g(y) - \hat{g}(y)) \le z) &= 1 - F_{\hat{g}}\left(g(y) - \frac{z}{n}\right) \\ &= 1 - \left\{1 + F_X\left(g(y) - \frac{z}{n}\right) - F_X(g(y))\right\}^n. \end{aligned}$$

By Taylor's expansion we obtain

$$F_X\left(g(y) - \frac{z}{n}\right) \approx F_X(g(y)) - \frac{z}{n} f_X(g(y)).$$

It follows that

$$\left\{1 - F_X(g(y)) + F_X\left(g(y) - \frac{z}{n}\right)\right\}^n \approx \left\{1 - \frac{z}{n} f_X((g(y)))\right\}^n.$$

By Taylor's expansion again we get

$$\left\{1 - \frac{z}{n} f_X((g(y)))\right\}^n \approx \mathrm{e}^{-z f_X(g(y))}.$$

Therefore

$$P(n(g(y) - \hat{g}(y)) \le z) \approx 1 - \mathrm{e}^{-z f_X((g(y)))},$$

which means that

$$n\left(g(y) - \hat{g}(y)\right) \xrightarrow{D} E_1/\lambda_2(y) \sim \mathrm{Exp}(\lambda_2(y)).$$

To get the desired conclusion (2.14), we note that

$$\begin{aligned} P(|\widehat{\lambda}_2(y) - \lambda_2(y)| > \epsilon) &= P(|\widehat{f}_X \circ \widehat{g}(y) - f_X \circ g(y)| > \epsilon) \\ &\le P(|\widehat{f}_X \circ \widehat{g}(y) - f_X \circ \widehat{g}(y)| + |f_X \circ \widehat{g}(y) - f_X \circ g(y)| > \epsilon), \end{aligned}$$

and under (A_2), by Taylor's formula, there exists ξ such that $f_X \circ \widehat{g}(y) - f \circ g(y) = (\widehat{g}(y) - g(y))f'(\xi)$. Both estimators $\widehat{f}_X$ and $\widehat{g}$ are consistent and thus it is also satisfied for $\widehat{\lambda}_2$ which implies (2.14).

Similarly we obtain

$$E_{\tilde{n},2} = \tilde{n}\hat{\lambda}_1(y)\left(\tilde{g}(y) - \hat{\tilde{g}}(y)\right) \xrightarrow{D} E_2 \text{ as } \tilde{n} \to \infty. \tag{2.15}$$

For the proof of (2.12) we write

$$\hat{\lambda}_1(y)\hat{\lambda}_2(y)\frac{n\tilde{n}}{n+\tilde{n}}\left(\hat{\tilde{g}}(y) - \hat{g}(y)\right) = \frac{\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)E_{n,1} - \frac{n}{n+\tilde{n}}\hat{\lambda}_2(y)E_{\tilde{n},2}. \qquad \square$$

To test the null hypothesis (1.2), we will use the following statistic

$$T_2(y) = \left\{\frac{n\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)\hat{\lambda}_2(y)\left(\hat{\tilde{g}}(y) - \hat{g}(y)\right)\right\}^2. \tag{2.16}$$

The limiting distribution of $T_2(y)$ follows immediately from Proposition 2.5.

Proposition 2.6. *Let assumptions* $(\mathrm{A}_1)-(\mathrm{A}_4)$ *hold. Under* $H_0(y)$ *we have*

$$T_2(y) \xrightarrow{D} Z_2(y) = X^2(y), \ \text{ as } n \to \infty, \tilde{n} \to \infty, \tag{2.17}$$

where $X(y)$ *is given by* (2.13).

Remark 2.7. Writing $U = (1-a)\lambda_1(y)$ and $V = a\lambda_2(y)$, the cumulative distribution function of $Z_2(y)$ is given by

$$F_{Z_2}(z) = 1 - \frac{1}{U+V}\left\{U\mathrm{e}^{-\frac{1}{U}\sqrt{z}} + V\mathrm{e}^{-\frac{1}{V}\sqrt{z}}\right\}. \tag{2.18}$$

Note that $F_{Z_2}(z)$ depends on the probability densities of the output data, X_t and $\tilde{X}_t$, and also on the common transformation g. However a consistent estimator of the critical values can be constructed from data. From (2.18) it is easily seen that a consistent estimator for the p-value corresponding to the statistic $T_2(y)$ is given by

$$\widehat{pv} = 1 - \frac{1}{\widehat{U}+\widehat{V}}\left\{\widehat{U}\mathrm{e}^{-\frac{1}{U}\sqrt{z}} + \widehat{V}\mathrm{e}^{-\frac{1}{V}\sqrt{z}}\right\},$$

where $\widehat{U} = (1-a)\widehat{\lambda}_1(y)$ and $\widehat{V} = a\widehat{\lambda}_2(y)$.

2.3. Behaviour of the tests under alternatives

We will study convergence properties of the tests T_1 and T_2 under both general and local alternatives. Let assumptions $(A_1)-(A_4)$ hold.

General alternatives We consider the limit distribution under $H_1(y)$.

Proposition 2.8. *Let assumptions* $(\mathrm{A}_1)-(\mathrm{A}_4)$ *hold. For all* y *satisfying* $(A(y))$ *such that* $g(y) \neq \tilde{g}(y)$ *we have*

$$T_1(y) \xrightarrow{P} +\infty \text{ and } T_2(y) \xrightarrow{P} +\infty,$$

where $\xrightarrow{P}$ *denotes the convergence in probability,* *Proof.* We can write

$$T_1(y) = \left(T_{1,1}(y) + T_{1,2}(y)\right)^2, \tag{2.19}$$

where

$$T_{1,1}(y) = \frac{1}{\widehat{\sigma}(y)}\sqrt{\frac{n\tilde{n}}{n+\tilde{n}}}(\hat{g}(y) - g(y)) - \frac{1}{\widehat{\sigma}(y)}\sqrt{\frac{n\tilde{n}}{n+\tilde{n}}}(\hat{\tilde{g}}(y) - \tilde{g}(y)), \tag{2.20}$$

$$T_{1,2}(y) = \frac{1}{\widehat{\sigma}(y)}\sqrt{\frac{n\tilde{n}}{n+\tilde{n}}}(g(y) - \tilde{g}(y)). \tag{2.21}$$

Results (2.4), (2.5) and (2.9) do not depend on the null $H_0(y)$ nor the alternative $H_1(y)$, therefore we have

$$T_{1,1}(y) \xrightarrow{D} N(0,1). \tag{2.22}$$

Under the alternative $H_1(y)$ we have

$$T_{1,2}(y) \xrightarrow{P} (+\infty)\mathrm{sgn}(g(y) - \tilde{g}(y)),$$

where $sgn(x) = 1$ if $x > 0$ and $sgn(x) = -1$ if $x < 0$. It follows that

$$T_1(y) \xrightarrow{P} +\infty.$$

Similarly we can write

$$T_2(y) = \left(T_{2,1}(y) + T_{2,2}(y)\right)^2, \tag{2.23}$$

where

$$T_{2,1}(y) = \frac{n\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)\hat{\lambda}_2(y)(g(y) - \hat{g}(y)) - \frac{n\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)\hat{\lambda}_2(y)(\tilde{g}(y) - \hat{\tilde{g}}(y)), \tag{2.24}$$

$$T_{2,2}(y) = \frac{n\tilde{n}}{n+\tilde{n}}\hat{\lambda}_1(y)\hat{\lambda}_2(y)(\tilde{g}(y) - g(y)). \tag{2.25}$$

The term $T_{2,1}(y)$ converges in distribution to the finite random variable $X(y)$ given by (2.13), whereas under the alternative $H_1(y)$ the term $T_{2,2}(y)$ converges in probability to $(+\infty)sgn(\tilde{g}(y) - g(y))$, which implies that $T_2(y) \xrightarrow{P} +\infty$. □

Local alternatives Write $m = \frac{n\tilde{n}}{n+\tilde{n}}$. We consider the local alternatives

$$H_1'(y) : \tilde{g}(y) = g(y) + \frac{k(y)}{m^\beta}\cdot$$

The behaviour of the test statistics T_1 and T_2 can be described as follows:

Proposition 2.9. *Let assumptions* (A_1)–(A_4) *hold. Under* $H_1'(y)$ *for* y *satisfying* $(A(y))$ *and when* $n \to \infty$, $\tilde{n} \to \infty$, *we have:*

(a) *If* $\beta > 1/2$ *then*

$$T_1(y) \xrightarrow{D} Z_1(y),$$

(b) *If* $\beta = 1/2$ *then*

$$T_1(y) \xrightarrow{D} Z_{1,k}(y),$$

(c) *If* $0 < \beta < 1/2$ *then*

$$T_1(y) \xrightarrow{P} +\infty.$$

The behaviour of the statistic T_2 *is quite different from* T_1 *and can be described as follows:*

(d) *If* $\beta > 1$ *then*

$$T_2(y) \xrightarrow{D} Z_2(y),$$

(e) *If* $\beta = 1$ *then*

$$T_2(y) \xrightarrow{D} Z_{2,k}(y),$$

(f) *If* $\beta < 1$ *then*

$$T_2(y) \xrightarrow{P} +\infty,$$

where $Z_1(y)$ *and* $Z_2(y)$ *are given by* (2.8) *and* (2.17) *respectively.* $Z_{1,k}(y)$ *is a non central Chi-squared distributed with one degree of freedom and parameter* $k^2(y)/\sigma^2(y)$. *The cumulative distribution function of* $Z_{2,k}(y)$ *is given by*

$$F_{2,k}(z) = \begin{cases} \frac{V \mathrm{e}^{-\frac{1}{a}\lambda_1(y)k(y)}}{U+U} \left\{ \mathrm{e}^{\frac{1}{V}\sqrt{z}} - \mathrm{e}^{-\frac{1}{V}\sqrt{z}} \right\} & \text{if } \sqrt{z} \leq \lambda_1(y)\lambda_2(y)k(y) \\ \\ F_2(z) & \text{if } \sqrt{z} > \lambda_1(y)\lambda_2(y)k(y) \end{cases}$$

where

$$F_2(z) = 1 - \frac{1}{U+V} \left\{ U \mathrm{e}^{-\frac{1}{U}(\sqrt{z} - \lambda_1(y)\lambda_2(y)k(y))} + V \mathrm{e}^{-\frac{1}{V}(\sqrt{z} - \lambda_1(y)\lambda_2(y)k(y))} \right\},$$

with $U = (1-a)\lambda_1(y)$ *and* $V = a\lambda_2(y)$.

Proof. Behaviour of the statistic $T_1(y)$.

Note first that $T_1(y)$ can be written as (2.19)$-$(2.21). The first term $T_{1,2}(y)$ converges to a standard normal distribution $\mathcal{N}(0,1)$. Under the local alternative $H_1'(y)$, the second term $T_{1,2}(y)$ becomes

$$T_{1,2}(y) = -\frac{1}{\hat{\sigma}(y)} k(y) m^{1/2-\beta}.$$

Therefore

(a) If $\beta > 1/2$ then $T_{1,2}(y) \xrightarrow{P} 0$, and hence $T_1(y) \xrightarrow{D} Z_1(y)$,

(b) If $\beta = 1/2$ then $T_{1,2}(y) \xrightarrow{P} -k(y)/\sigma(y)$, and hence $T_1(y) \xrightarrow{D} Z_{1,k}(y)$,

(c) If $0 < \beta < 1/2$ then $T_{1,2}(y) \xrightarrow{P} (-\infty)sgn(k(y))$, and hence $T_1(y) \xrightarrow{P} +\infty$.

Behaviour of the statistic $T_2(y)$.

The term $T_2(y)$ can be written as (2.23)–(2.25). The term $T_{2,1}(y)$ converges in distribution to the finite random variable $X(y)$ given by (2.13), whereas under the alternative $H_1'(y)$ the second term $T_{2,2}(y)$ becomes

$$T_{2,2}(y) = \hat{\lambda}_1(y)\hat{\lambda}_2(y)k(y)m^{1-\beta}.$$

Therefore

(d) If $\beta > 1$ then $T_{2,2}(y) \xrightarrow{P} 0$, and hence $T_2(y) \xrightarrow{D} Z_2(y)$,

(e) If $\beta = 1$ then $T_{2,2}(y) \xrightarrow{P} \lambda_1(y)\lambda_2(y)k(y)$, and hence $T_2(y) \xrightarrow{D} Z_{2,k}(y)$,

(f) If $\beta < 1$ then $T_{2,2}(y) \xrightarrow{P} (+\infty)\mathrm{sgn}(k(y))$, and hence $T_2(y) \xrightarrow{P} +\infty$. $\square$

3. Simulations study

3.1. Application of the test statistics for testing $H_0(y)$

For all empirical powers or empirical levels we carry out experiments of 10000 samples and we use three different sample sizes: $n = 50$, $n = 100$, and $n = 500$. When the distributions of Y and $\tilde{Y}$ are unknown we will use the same sample size to estimate them. For each replication we compute the statistics $T_1(y)$ and $T_2(y)$ given by (2.7) and (2.16). Both signals Y_t and $\tilde{Y}_t$ are centered and we will choose $y = 0$ for testing $H_0(0)$. The bandwidth is chosen as $h_n = n^{-1/5}$ and the trimming as $e_n = n^{-1/6}$.

3.1.1. Study of the empirical levels.

We will denote by $\mathcal{N}(0,1)$ the standard normal distribution with mean zero and variance 1. We assume that Y_t and $\tilde{Y}_t$ are independent and $\mathcal{N}(0,1)$ distributed.

To study the empirical levels of T_1 and T_2 we choose

$$g(y) = \tilde{g}(y) = 3y + 5,$$

and we fix a theoretical level $\alpha = 5\%$. Table 1 shows empirical levels of the two tests under $H_0(0)$. It can be observed that for small sample size T_1 seems to over estimate to the theoretical asymptotic value and T_2 seems to provide smaller levels.

Remark 3.1. To evaluate the effect of the bandwidth we used two different values for h_n. We chose $h_n = an^{-1/5}$, with $a = 0.5$ and 2. Empirical levels are given in Table 4. Results with $a = 2$ are similar to those obtained with $a = 1$. With $a = 0.5$ we obtained results not as good as with $a = 2$ or $a = 1$.

Table 1. Empirical levels of T_1 and T_2 (in %) for a theoretical level $\alpha = 5\%$.

	$n=50$	$n=100$	$n=500$
T_1	7.96	6.88	5.72
T_2	2.76	3.53	4.45

TABLE 2. Empirical levels of T_1 and T_2 (in %) for a theoretical level $\alpha = 5\%$ with $h_n = an^{-1/5}$.

		$n = 50$	$n = 100$	$n = 500$
T_1	$a = 0.5$	11.02	9.22	6.35
	$a = 2$	6.10	5.62	4.82
T_2	$a = 0.5$	3.93	3.46	3.96
	$a = 2$	2.74	3.57	4.68

TABLE 3. Empirical powers of T_1 and T_2 (in %) for a theoretical level $\alpha = 5\%$.

	T_1	T_2	T_1	T_2	T_1	T_2
	$\tilde{g}_1$	$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{g}_3$
$n = 50$	90.35	89.54	99.96	99.25	81.15	82.44
$n = 100$	93.06	92.53	99.98	99.62	85.79	89.01
$n = 500$	96.18	98.83	100	99.94	93.29	98.43

TABLE 4. Empirical powers of T_1 (in %) for a theoretical level $\alpha = 5\%$ under local alternative $\tilde{g}_4$.

	$\beta = 1/4$	$\beta = 1/2$	$\beta = 4$
$n = 50$	94.67	71.19	8.08
$n = 100$	95.70	71.83	6.54
$n = 500$	97.67	73.71	6.06

TABLE 5. Empirical powers of T_2 (in %) for a theoretical level $\alpha = 5\%$ under local alternative $\tilde{g}_4$.

	$\beta = 1/2$	$\beta = 1$	$\beta = 4$
$n = 50$	87.72	13.98	2.46
$n = 100$	91.53	15.66	3.34
$n = 500$	96.40	16.51	4.37

3.1.2. Study of the empirical powers

We consider the model where Y_t and $\tilde{Y}_t$ are independent and $\mathcal{N}(0,1)$ distributed. To study the empirical powers of T_1 and T_2 we consider $g(y) = 3y + 5$ and the three following transformations:

$$\tilde{g}_1(y) = \exp((y+3)/(y+5)), \quad \tilde{g}_2(y) = -(y+11)/(y+5), \quad \tilde{g}_3(y) = y/2 + 3,$$

and we also study local alternatives by considering:

$$\tilde{g}_4(y) = g(y) + \frac{2(y+5)}{n^\beta}\cdot$$

Tables 3–5 present empirical powers for T_1 and T_2 under fixed and local alternatives, for a theoretical level α equal to 5%.From Table 3, it seems that the empirical powers of T_1 and T_2 are very close. The test statistic T_2 provides slightly better power than T_1 for the alternative $\tilde{g}_3(y)$ which is not far from the null hypothesis. Table 4 indicates that T_1 provides good power for $\beta \leq 1/2$. For $\beta > 1/2$ the power converges to the theoretical level α. Table 5 indicates that T_2 provides good power for $\beta \leq 1$. For $\beta > 1$ the power converges to the theoretical level α. These comments are in accordance with the theoretical results stated in Proposition 2.9.

3.2. Resampling procedure

When the hypothesis $H_0(y)$ is not rejected a bootstrap procedure can be used for the global comparison $H_0 : g = \tilde{g}$. We propose to use a grid $y_1, \ldots, y_M$, where we will compare g and $\tilde{g}$. The associated test statistics are defined by:

Case 1 (the distributions of Y and $\tilde{Y}$ are known)

$$\mathcal{T}_1 = \frac{1}{M} \sum_{i=1}^{M} T_1(y_i). \tag{3.1}$$

Case 2 (the distributions of Y and $\tilde{Y}$ are estimated)

$$\mathcal{T}_2 = \frac{1}{M} \sum_{i=1}^{M} T_2(y_i), \tag{3.2}$$

where T_1 and T_2 are given by (2.7) and (2.16), respectively. From Propositions 2.1 and 2.3, under $H_0 : g = \tilde{g}$, $\mathcal{T}_1$ is asymptotically distributed as the sum of dependent chi-squared random variables and $\mathcal{T}_2$ as the sum of dependent mixtures of exponential random variables. In both cases the asymptotic distribution is not explicit. To overcome this problem we construct a naive bootstrap statistic. The test statistic $\mathcal{T}_j, j = 1, 2$, is compared to the empirical bootstrap distribution of $(\mathcal{T}_j{}^{*b})_{b=1,\ldots,B}$, where each $\mathcal{T}_j{}^{*b}, b = 1, \ldots, B$, is constructed as follows: From the sample $\mathbf{E} = \Big((X_1, Y_1), \ldots, (X_n, Y_n), (\tilde{X}_1, \tilde{Y}_1), \ldots, (\tilde{X}_{\tilde{n}}, \tilde{Y}_{\tilde{n}})\Big)$,

- Draw randomly with replacement two bootstrap samples of size n and $\tilde{n}$, respectively.
- Consider the first bootstrap sample $\mathbf{E}_1^*$, and for each $i = 1, \ldots, M$ compute $\widehat{g_1}^*(y_i) = p\widehat{g_1}(y_i) + (1-p)\widehat{\tilde{g}_1}(y_i)$, where p (resp. (1-p)) is the observed proportion of variables X's (resp of $\tilde{X}$'s) in $\mathbf{E}_1^*$, $\widehat{g_1}$ and $\widehat{\tilde{g}_1}$ are estimators obtained by (2.1) and (2.10) based on the (X, Y)'s and the $(\tilde{X}, \tilde{Y})$'s belonging to $\mathbf{E}_1^*$.
- Similarly construct an estimator $\widehat{\tilde{g}_2}^*(y_i) = q\widehat{g_2}(y_i) + (1-q)\widehat{\tilde{g}_2}(y_i)$ from the second bootstrap sample $\mathbf{E}_2^*$.
- Using $\widehat{g_1}^*(y_i)$ and $\widehat{\tilde{g}_2}^*(y_i)$ construct the bootstrap statistic $T_1^{*b}(y_i)$ and $T_2^{*b}(y_i)$ as follows

$$T_1^{*b}(y_i) = \frac{n\tilde{n}}{n+\tilde{n}} \widehat{\sigma}^{*-2}(y_i) \left(\widehat{g_1}^*(y_i) - \widehat{\tilde{g}_2}^*(y_i) \right)^2,$$

where $\widehat{\sigma}^{*2}(y_i) = (1-a)s_1^2 + as_2^2$ with

$$s_1^2 = p^2 \frac{F_Y(y_i)(1-F_Y(y_i))}{\widehat{f}^2_{X,e_n}(\widehat{g_1}^*(y_i))} + (1-p)^2 \frac{F_{\tilde{Y}}(y_i)(1-F_{\tilde{Y}}(y_i))}{\widehat{f}^2_{\tilde{X},e_n}(\widehat{\tilde{g}_1}^*(y_i))},$$

$$s_2^2 = q^2 \frac{F_Y(y_i)(1-F_Y(y_i))}{\widehat{f}^2_{X,e_n}(\widehat{g_2}^*(y_i))} + (1-q)^2 \frac{F_{\tilde{Y}}(y_i)(1-F_{\tilde{Y}}(y_i))}{\widehat{f}^2_{\tilde{X},e_n}(\widehat{\tilde{g}_2}^*(y_i))}.$$

$$T_2^{*b}(y_i) = \left\{ \frac{n\tilde{n}}{n+\tilde{n}} \hat{\lambda}_1^*(y) \hat{\lambda}_2^*(y_i) \left(\widehat{g_1}^*(y_i) - \widehat{\tilde{g}_2}^*(y_i) \right) \right\}^2,$$

where $\hat{\lambda}_1(y_i) = \hat{f}_{\tilde{X}}(\widehat{\tilde{g}_2}^*(y_i))$ and $\hat{\lambda}_2(y_i) = \hat{f}_X(\widehat{g_1}^*(y_i))$.
- Compute $\mathcal{T}_j^{*b} = \frac{1}{M} \sum_{i=1}^{M} T_j^{*b}(y_i)$, $j = 1, 2$.

We reject H_0 as soon as the test statistic is larger than the empirical bootstrap threshold. We did not study the convergence of this bootstrap method. One possibility would be to follow the approach of [12], but it exceeds the scope of this paper. Eventually, our test procedure is based on the resampling method, as in permutation tests theory, excepted that we did not use all permutations to keep a reasonable calculation time and make the test applicable. Thus, we can see this procedure as a permutation test, but with a limited number of resampling.

Table 6 shows empirical powers with the bootstrap procedure under alternatives $\tilde{g}_1, \tilde{g}_2$ and $\tilde{g}_3$. It can be observed that all alternatives are very well detected by the bootstrap approach.

TABLE 6. Empirical powers of $\mathcal{T}_1$ and $\mathcal{T}_2$ (in %) for a theoretical level $\alpha = 5\%$.

	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_1$	$\mathcal{T}_2$
	$\tilde{g}_1$	$\tilde{g}_1$	$\tilde{g}_2$	$\tilde{g}_2$	$\tilde{g}_3$	$\tilde{g}_3$
$n = 50$	100	89.54	100	99.25	96.90	100
$n = 100$	100	92.53	100	99.62	100	100
$n = 500$	100	98.83	100	99.94	100	100

4. Real example: Framingham data

We consider the Framingham Study on coronary heart disease described in [4] (see more recently [17]). The data consist of measurements of systolic blood pressure (SBP) obtained at two different examinations in 1615 males on an 8-year follow-up. At each examination, the SBP was measured twice for each individual. The four variables of interest are:

$Y =$ the first SBP at examination 1,

$\tilde{Y} =$ the second SBP at examination 1,

$X =$ the first SBP at examination 2,

$\tilde{X} =$ the second SBP at examination 2.

Our purpose is to examine whether the distribution of the SBP changed during time, and which type of transformation it underwent. Following our notations, we will study the transformation between Y and X and also the one between $\tilde{Y}$ and $\tilde{X}$.

Table 7 indicates that all the distributions of X, Y, $\tilde{X}$ and $\tilde{Y}$ are skewed to the right and are leptokurtic, KS is the Kolomogorov–Smirnov statistic, the associated p-values are lesser than 2.210^{-6} and hence the normality assumption is strongly rejected. Figure 1 represents nonparametric estimations of the probability densities of X, Y, $\tilde{X}$ and $\tilde{Y}$.

From Figure 1, it seems that the distributions of the variables Y and X have a similar shape. However, from Table 7 we observe a noticeable decrease in the mean and an increase in the variance. Figure 2 represents $\widehat{g}$, $\widehat{\tilde{g}}$, and $\widehat{g}_0$, the common estimator under H_0 obtained by aggregation of the two previous ones. Based on these

TABLE 7. Descriptive statistics of Framingham data

Y						X					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
80.0	120.0	130.0	132.8	142.0	230.0	88.0	118.0	128.0	131.2	142.0	260.0
Var.	Skewness.	Kurtosis.	KS.			Var.	Skewness.	Kurtosis.	KS.		
419.12	1.27	7.79	0.0119			439.11	1.39	6.65	0.1125		
$\tilde{Y}$						$\tilde{X}$					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
75.0	118.0	128.0	130.2	140.0	270.0	85.0	115.0	125.0	128.8	138.0	270.0
Var.	Skewness.	Kurtosis.	KS.			Var.	Skewness.	Kurtosis.	KS.		
409.97	1.46	7.25	0.1171			410.21	1.47	7.10	0.1117		

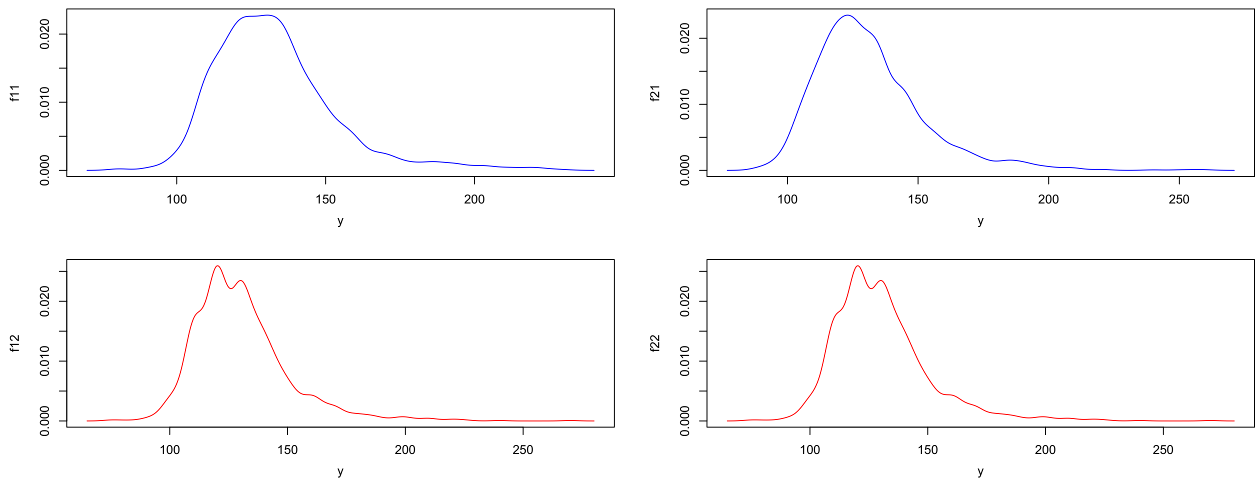

FIGURE 1. Kernel estimates of the probability densities of $X, Y, \tilde{X}, \tilde{Y}$. In the top panel: f11 (resp. f21) is the Kernel estimate of the density of Y (resp. of X). In the bottom panel: f12 (resp. f22) is the Kernel estimate of the density of $\tilde{Y}$ (resp. of $\tilde{X}$).

nonparametric estimators, we can postulate that only the location and the scale are affected by time, therefore, the transformation g should be linear; that is, $g(y) = ay + b$. Similarly the distributions of the variables $\tilde{Y}$ and $\tilde{X}$ should be linked by $\tilde{g}(y) = \tilde{a}y + \tilde{b}$.

By applying our test statistic $T_2(y)$ given by (2.16) to the central value $y = (\overline{Y} + \overline{\tilde{Y}})/2 = 131.5$, we obtain a p-value equal to 1, and hence we can consider that $g(y) = \tilde{g}(y)$, that is the two transformations coincide for the central value y. This is in agreement with Figure 2.

We thus consider the more global hypothesis $H_0 : g = \tilde{g}$ and then we apply the bootstrap procedure described in Section 3.2. We use the grid $\mathcal{G} = \{y_i, i = 1, \ldots, M\}$ with $y_i = c + (d - c)i/M$, belonging to the interval $[c, d]$ where

$$c = \max(\min(Y_i), \min(\tilde{Y}_j)) \text{ and } d = \min(\max(Y_i), \max(\tilde{Y}_j)).$$

For the Framingham data we obtain $c = 80$ and $d = 230$. We compute the approximate bootstrap significance level which is defined by

$$ASLboot = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}_{\{\mathcal{T}_2^{*b} \geq \mathcal{T}_{2,obs}\}},$$

where $\mathcal{T}_{2,obs}$ is the observed value of the statistic $\mathcal{T}_2$, for a grid of size $M \in \{10, 50, 100\}$ and for a number of replications $B \in \{100, 500, 1000\}$.

The $ASLboot$ is close to 10%, which is very different from the p-value obtained with the first method for testing $g(y) = \tilde{g}(y)$. When $M = 100$ the $ASLboot$ is less than 10%, what should contradict the equality of both

TABLE 8. Approximate bootstrap significance level.

	M	10	50	100
B				
100		0.13	0.09	0.08
500		0.092	0.132	0.086
1000		0.089	0.124	0.078

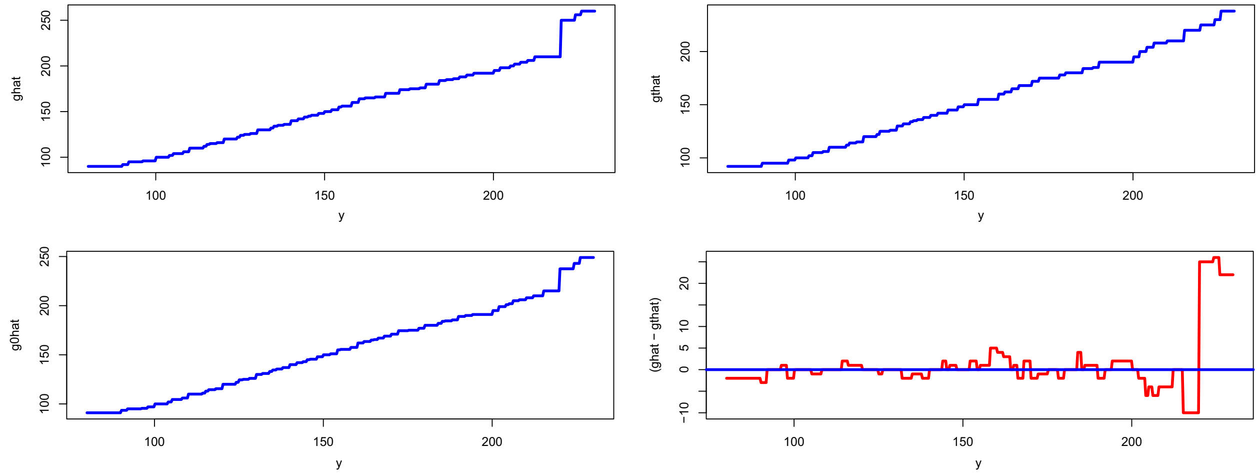

FIGURE 2. Nonparametric estimators of g and $\tilde{g}$ and the aggregated estimator on the interval $[c, d]$: *ghat* (resp. *gthat* and *g0hat*) denotes $\widehat{g}$ (resp. $\widehat{\tilde{g}}$ and $\widehat{g_0}$).

functions. In Figure 2 we observe that $\widehat{g}$, $\widehat{\tilde{g}}$ and $\widehat{g}_0$ seem to be linear on the interval $[c, d]$. However in the border (near c and d) the approximations are not good implying that $\widehat{g}$ and $\widehat{\tilde{g}}$ should be different where there are not enough observations. Despite this difference, the equality between g and $\tilde{g}$ is not rejected at 5% level of significance on the whole interval $[c, d]$.

Acknowledgements. We would like to thank the referee and an associate editor for making numerous suggestions which greatly improved the readability of the paper.

References

[1] N. Balakrishnan and B. Xingqiu Zhao, New multi-sample nonparametric tests for panel count data. *Ann. Stat.* **37** (2009) 1112–1149.

[2] N. Balakrishnan and B. Xingqiu Zhao, A nonparametric test for the equality of counting processes with panel count data. *Comput. Stat. Data Anal.* **54** (2010) 135–142.

[3] M. Boutahar, B. Ghattas and D. Pommeret, Nonparametric comparison of several transformations of distribution functions. *J. Nonparametric Stat.* **25** (2013) 619–633.

[4] R.J. Carroll, D. Ruppert, L.A. Stefanski and C. Crainiceanu, *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd edition. Chapman Hall, New York (2006).

[5] L. Cavalier and N. Hengartner, Adaptive estimation for inverse problems with noisy operators. *Inverse Probl.* **21** (2005) 1345–1361.

[6] S. Efromovich and V. Koltchinskii, On inverse problems with unknown operators. *IEEE Trans. Inf. Theory* **47** (2001) 2876–2893.

[7] R. Friden, Probability, Statistical Optics and Data Testing: A Problem Solving Approach. *Springer Series in Information Science* (2001).

[8] L. Jin, A data-driven test to compare two or multiple time series. *Comput. Stat. Data Anal.* **55** (2011) 2183–2196.

[9] H. Jin and J. Zhang, Subsampling tests for variance changes in the presence of autoregressive parameter shifts. *J. Multivariate Anal.* **101** (2010) 2255–2265

[10] J. Meyer, Stochastic Dominance and Transformations of Random Variables, in *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, edited by T.B. Fomby and T.K. Seo. Springer Science and Business Media (2012) 45–58.

[11] J. Meyer and M.B. Ormiston, Deterministic transformations of random variables and the comparative statics of risk. *J. Risk Uncertain.* **2** (1989) 179–188.

[12] M.H. Neumann and E. Paparoditis, On bootstrapping L2-type statistics in density testing. *Stat. Probab. Lett.* **50** (2000) 137–147

[13] P. Sarda and P. Vieu, Kernel regression, in Smoothing and Regression Approaches, computation and application, edited by M.G. Schimek. *Wiley Series in Probability and Statistics* (2000) 43–70.

[14] S.L. Sclove and J.V. Ryzin, Estimating the Parameters of a Linear Function of a Random Variable. *J. R. Stat. Soc. Ser. B* **32** (1970) 362–368.

[15] P.K. Sen, Limiting behavior of regular functionals of empirical distributions for stationary mixing processes. *Probab. Theory Relat. Fields* **25** (1972) 71–82.

[16] J. Sun and J.D. Kalbfleisch, Estimation of the mean function of point processes based on data panel count data. *Stat. Sin.* **5** (1995) 279–290.

[17] X.F. Wang and B. Wang, Deconvolution Estimation in Measurement Error Models: The R Package decon. *J. Stat. Softw.* **39** (2011) 8060.