Stochastic Modeling and Applications
Vol.20, No. 1(June 2016), 35-53

MUK PUBLICATIONS
Open Access Publisher

A NEW TEST IN THE LONGEST INCREASING SUBSEQUENCE FAMILY OF INDEPENDENCE TESTS

JESÚS E. GARCÍA, V.A. GONZÁLEZ-LÓPEZ, MARCEL SIMIS, THAIS TERRANOVA, AND LINAMARA RIZZO BATTISTELLA

Abstract. In this paper we introduce a new discrete one-tailed hypothesis test of independence, d_n, a member from the family of *the Longest Increasing Subsequence Hypothesis Tests*. d_n allows to incorporate an alternative hypothesis which represents a decreasing (increasing) trend, according to the necessity. Through a simulation study, focusing in situations of small and moderate sample sizes and mixture of distributions, we compare the behavior of d_n with those procedures which allow one-tailed options. We introduce a correction in the p-value of d_n, in order to control the type I error and to improve the power of d_n. d_n shows outstanding performance in all the simulations, being indicated for use in mixtures, d_n shows adequate control of the level of significance and high power when compared to other tests. We use d_n to inspect data from stroke rehabilitation, d_n detects the tendency between scores taken through conventional methods and robotic devices.

1. Introduction

We should begin this article with a question, whose positive answer is already part of science fiction and promises to become part of our everyday life, in the near future. The evolution of robotic devices is the future of the medicine practiced today? The Physiatry and Rehabilitation Medicine are investing in this possibility. The features offered by robotic devices should be equated with current medical practices adopted to deal with specific problems. The comparison between the information obtained from robotic devices with the information obtained from usual methodologies, used in various medical centers, must allow the revision of the procedures looking for more accuracy. In a sense, if robotic devices are shown efficient for medical therapies, its inclusion in everyday medical practice could represent a revolution in the health industry.

This work receives motivations of two areas, one of them coming from the stroke rehabilitation field and the second is related to the construction of discrete hypothesis tests. We will begin with the description of the real problem, which will encourage the development of a discrete one-tailed test, based on concepts of random permutations, widely discussed in [5], [6] and [7]. It is estimated that there are about 62 million stroke survivors worldwide, with a great need of rehabilitation therapy. Consequently, it is essential to have a reliable evaluation of the

2010 *Mathematics Subject Classification.* 62G10; 62G30.
Key words and phrases. Longest Increasing Subsequence; Discrete tests; Mid-p value; Robotic device in stroke.

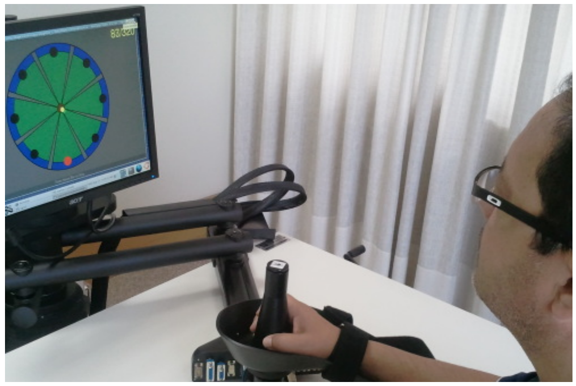


FIGURE 1. Robotic equipment running session.

treatments in order to give an objective numerical value for the abstract concept of disability and to determine the efficacy of the interventions. According to the World Health Organization International Classification of Functioning, Disability and Health (WHO-ICF) it is necessary to have an universal assessment tool for disability, although, there is no such scale, see [19]. There are several methods to assess the motor function in upper limb of stroke patients and different types of scales; however, there is no *gold standard* for the evaluation ([8]). One of the reasons for that is the limitation of the available scales, which justify the development of new tools. On the other hand, robotic devices recently have been used as systems for the evaluation since they record the kinematic and kinetic variables of upper limb movements. Nevertheless it is not clear yet the relationship between the robotic variables and the results of tools specifically designed to evaluate the motor performance of the upper limbs. In this paper we investigate the performance of a robotic equipment used for research in the *Physical and Rehabilitation Medicine Institute of the University of Sao Paulo Medical School General Hospital, Brazil.* This equipment has several available functions, i.e. devices. We see, in figure 1 an illustration of one of the sessions during its execution. These robotic devices were specifically constructed for clinical and neurological applications, see [13]. The system is formulated to maintain a low intrinsic end-point impedance, with extremely low inertia and low friction, then it is able to move smoothly and can rapidly fulfill with patient's movements. The robot sensors permit accurate and essentially continuous measurement of the variables which describe the motor behavior, namely position, velocity, and interaction forces. Through two degrees-of-freedom this device allows the movements, in the shoulder, the elbow and the hand in a gravity-eliminated plane. There is a video screen in front of the patient, it provides visual feedback of the target location and a description of the movement, see [9], [11], [10]. In the robot's evaluation, the metrics include deviation from a straight line, aim, average speed, peak speed, and the duration of the movement while attempting to reach toward different targets. These unconstrained reaching movements are decomposed into submovements to produce 6 additional kinematic metrics: number, duration, overlapping, peak, interval, and skewness of submovements. Besides, the circle-drawing activities, captures the coordination of arm's movements and the kinematic set is completed with 3 additional kinetic parameters that measure the strength of the shoulder and the ability of the patient to move against resistance or to hold against an externally applied force, leading to a total of 35 measurements for each patient. The parameter that measures shoulder strength has a ceiling effect and is measured only on the impaired side. Finally all the parameters were linearly standardized between 0 and 1, see [12] for details.

In this paper we want to see if there is any evidence to suggest a monotonous dependence between two paired scores X and Y. X is the result of a scale manually collected by therapists of the area, to measure the average time spent on the execution of a list of tasks specified by *The Wolf Motor Function Test* [21]. Y values correspond to a given score by the robotic device, after the execution of a selection of movements. In this article we analyze three specific situations, involving just one scale collected manually $X = \mathrm{WT}$ (time spent in the execution of The Wolf Motor Function Test) in counterpoint with 3 scales coming from the robotic device: Independence (IND), Shoulder Abduction (SA) and Shoulder Flexion (SF). Thus, Y is given by the IND score in the first situation, SA, in the second situation and SF in the third situation. In figure 2 are shown, the graphs of these three situations. The treatment of stroke is still quite restricted to few patients. Given that, the application of manual scales takes a long time also it requires specialists who know the official protocols established in treatment centers. This practice also takes time from patients and their families, thereby in critical situations can be impractical. The complexity to obtain the manual scales justifies the size of the sample analyzed in this paper, $n = 41$. The utility of to discover if there is any relationship between the scales, would establish evidence to change the practice in stroke treatment centers. Seen that official measures commonly used in such centers could be efficiently obtained by robotic devices and also in the future, the practice in those centers would be modified, up itself robotic devices in the execution of these treatments. So our objective is to investigate whether there is evidence to establish a connection between manual scales and those obtained by robotic devices. To reach a conclusion regarding dependence and type of trend between X and Y we run tests of hypothesis that can respond to this problem. In this paper we propose a new statistical test, which is a test from the family of the longest increasing subsequence (LIS) tests, see [5]. This new procedure is sensitive to detect dependence between X and Y. We also investigate the behavior of the test in different simulated situations. As was reported in the literature about discrete tests (see [3]) there are weakness detected in the performance of these procedures, because as a rule the size of a discrete test and the significance level will be different. In those cases, the Neyman-Pearson theory introduces a random experiment to control the power function of the test, fitting the size of the test as a consequence. The Neyman-Pearson theory with randomisation provides a size equal to the significance level although the test will not always give the same result and as was reported in [14] inconsistent situations may appear. For example, when a single additional observation which was a favour of the hypothesis being rejected

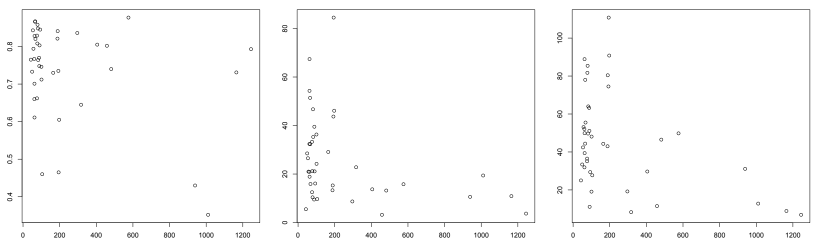


FIGURE 2. Left: x axis-time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Independence (IND). Middle: x axis-time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Shoulder Abduction (SA). Right: x axis-time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Shoulder Flexion (SF).

but which had the opposite effect, as a consequence of the auxiliary randomisation, see [14]. [14] proposed the construction of the mid-p value which is used in this paper to improve the power of the discrete test built here. The LIS family is a class of nonparametric tests for the assumption of independence between two continuous random variables X and Y. It is formulated as follows, given a sample of size n, let π be the permutation which maps the ranks of the X observations on the ranks of the Y observations. The independence assumption is identified with the uniform distribution on the space of all the permutations of samples of size n, see theorem 3.1 in [5]. To make the decision to reject (or not) the hypothesis of independence, is used the permutation defined by the sample, through the ranks, to compare it with a regular permutation under the assumption of independence. Under the conception of [5], to do that is used the longest increasing subsequence selected from the permutation. Unlike the tests already introduced in the LIS family, in this paper we see other aspects on the permutation which are more adequate to identify a kind of trend in the relation between X and Y.

We proceed as follows. In the next section we define the new one-tailed hypothesis test. In Section 3 we describe the variables that will be analyzed and we show the results obtained by this new procedure. In Section 4 we conduct a simulation study to understand the performance of this new hypothesis test. In Section 5 are shown general conclusions and the bibliography in Section 6.

2. A New Test in the Class of LIS Tests

In [5] is proposed a new class of independence tests based on the length of the longest increasing subsequence of a random permutation. The statistical tests introduced in [5] were designed to work on situations where the standard tests of independence do not work well. Furthermore, the construction of those tests could be adapted to catch other situations also very common in real data sets, as mixture of distributions. Taking into account that the data sets studied in this paper can be considered as generated by mixture of distributions, those situations will be the theoretical environment of this paper, with focus on small sample size settings. Here, we show the construction of a new one tailed test of independence in the class of the longest increasing subsequence tests, denoted by d_n. We also devote this paper to study the performance of d_n, with the correction proposed by [14], since it allows to control the Type I error and improves the power of discrete tests. To formalize the concepts behind this class, we show the main definitions.

Definition 2.1. Let $\mathcal{S}_n$ denote the group of permutations of $\{1,\dots,n\}$. If $\pi \in \mathcal{S}_n$, we say that $\pi(i_1),\dots,\pi(i_k)$ is an increasing subsequence of π if $1 \le i_1 < \ldots < i_k \le n$ and $1 \le \pi(i_1) < \pi(i_2) < \ldots < \pi(i_k) \le n$.

Example 2.2. We list here all the increasing subsequences on the permutation of $\{1,2,3,4,5,6,7\}$ given by $\pi = \{6,1,4,5,7,2,3\}$.

i. 7 of size 1: $\{6\},\{1\},\{4\},\{5\},\{7\},\{2\},\{3\}$;

ii. 10 of size 2: $\{6,7\},\{1,4\},\{1,5\},\{1,7\},\{1,2\},\{1,3\},\{4,5\},\{4,7\},\{5,7\},\{2,3\}$;

iii. 5 of size 3: $\{1,4,5\},\{1,4,7\},\{1,5,7\},\{1,2,3\},\{4,5,7\}$;

iv. 1 of size 4: $\{1,4,5,7\}$.

Definition 2.3. Given a permutation $\pi \in \mathcal{S}_n$, we call $l_n(\pi)$ the length of the longest increasing subsequence of π.

In the example 2.2 $l_7(\pi) = 4$. Define $li_n^2(\pi)$ as the length of the longest subsequence of π consisting of two (disjoint) increasing subsequences of π. In the example 2.2 $li_7^2(\pi) = 6$ which is obtained through the subsequence $\{1,4,5,7,2,3\}$, resulting from the two increasing subsequences $\{1,4,5,7\}$ and $\{2,3\}$. Check [7] on how to compute $li_n^2(\pi)$. In a similar way we can define $ld_n^2(\pi)$ as the length of the longest subsequence of π consisting of two (disjoint) decreasing subsequences of π. In the case of the example 2.2, $ld_7^2(\pi) = 5$ which is obtained through the subsequence $\{6,5,7,2,3\}$, it contains two decreasing disjoint subsequences: $\{6,5,3\}$ and $\{7,2\}$. The statistic of the test explored in this paper is based on the functions li_n^2 and ld_n^2 defined in the following way.

Definition 2.4. Let $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ be replications of (X,Y) with continuous marginal distributions, we denote by d_n the random variable, $d_n := li_n^2(\pi_{\mathcal{D}}) - ld_n^2(\pi_{\mathcal{D}})$ where $\mathcal{D} = \{(X_i,Y_i)\}_{i=1}^n$ and $\pi_{\mathcal{D}}$ is the permutation which assigns $\pi(rank(X_i)) = rank(Y_i)$, $i = 1,\dots,n$.

Theorem 2.5. *Let $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ be replications of (X,Y) with continuous marginal distributions, under the assumption of independence between X and Y, the random variable d_n given by definition (2.4) is symmetric around zero.*

Proof. When X and Y are independent, all possible permutations $\pi(rank(X_i)) = rank(Y_i)$, $i = 1,\dots,n$ have the same probability, i.e. $\pi_{\mathcal{D}}$ has the uniform distribution on the set of all the permutations on $\{1,2,\dots,n\}$. Define now $\mathcal{D}_- = \{(-X_i,Y_i)\}_{i=1}^n$. For the same reason, under independence, $\pi_{\mathcal{D}_-}$ has uniform distribution on the set of all the permutations on $\{1,2,\dots,n\}$. In this way we have that under the null hypothesis of independence $\pi_{\mathcal{D}}$ and $\pi_{\mathcal{D}_-}$ share the same distribution. Also, every increasing subsequence of $\pi_{\mathcal{D}}$, generated by $\{(X_{i_k}, Y_{i_k})\}_{k=1}^{m}$ corresponds to the decreasing subsequence of $\pi_{\mathcal{D}_-}$ given by $\{(-X_{i_k}, Y_{i_k})\}_{k=1}^{m}$ and then

$$li_n^2(\pi_{\mathcal{D}}) = ld_n^2(\pi_{\mathcal{D}_-}). \tag{2.1}$$

The equation (2.1) implies that $li_n^2(\pi_{\mathcal{D}})$ and $ld_n^2(\pi_{\mathcal{D}})$ have the same distribution, since $\pi_{\mathcal{D}}$ and $\pi_{\mathcal{D}_-}$ share the uniform distribution, i.e. $li_n^2(\pi_{\mathcal{D}}) = ld_n^2(\pi_{\mathcal{D}_-}) \sim ld_n^2(\pi_{\mathcal{D}})$. Using equivalent arguments we have

$$ld_n^2(\pi_{\mathcal{D}}) = li_n^2(\pi_{\mathcal{D}_-}), \tag{2.2}$$

from equations (2.1)-(2.2), follows

$$li_n^2(\pi_{\mathcal{D}}) - ld_n^2(\pi_{\mathcal{D}}) = ld_n^2(\pi_{\mathcal{D}_-}) - li_n^2(\pi_{\mathcal{D}_-}). \tag{2.3}$$

The equation (2.3) implies that $d_n = li_n^2(\pi_{\mathcal{D}}) - ld_n^2(\pi_{\mathcal{D}})$ and $-d_n = ld_n^2(\pi_{\mathcal{D}}) - li_n^2(\pi_{\mathcal{D}})$ are identically distributed and the distribution of d_n is symmetric around zero, under the independence. □

2.1. The One Tailed d_n Test. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a paired sample of size n of (X, Y) with continuous marginal distributions. The hypothesis test formulation studied here is

$$H_0 : X \text{ and } Y \text{ are independent}$$

versus

$$H_1 : X \text{ and } Y \text{ have an increasing (decreasing) relation.}$$

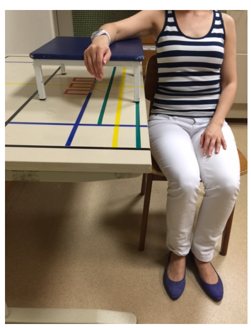
Tests of hypothesis based on measures of association/correlation quantized the trend of the relation by a coefficient ν and test $H_0 : \nu = 0$ versus $H_1 : \nu > 0$ ($\nu < 0$). This is the case of Kendall's test, Spearman's test and Pearson's test, where the coefficients used are Kendall's tau, Spearman's rho and Pearson's rho respectively. For the d_n test, it is verified if the observed value of d_n, say d_0 confirms the increasing (decreasing) trend among the ranks of the observations. The p-value of the test with null hypothesis of independence against an alternative hypothesis of increasing (decreasing) tendency between X and Y is defined as $\text{Prob}(d_n > d_0)$ ($\text{Prob}(d_n < d_0)$).

In this paper we compare the behavior of the one-tailed version of d_n with three, well known, independence tests with that possibility. Namely, Kendall's test, Spearman's test and Pearson's test. To serve the practical purpose we have set, the study will be focused on the decreasing dependence between X and Y. Being d_n a discrete test, it will suffer of the weaknesses reported for discrete tests ([14]). So, we adopt a correction in the calculation of its p-value given by the mid-p-value which is defined as $\text{Prob}(d_n < d_0) + \frac{1}{2}\text{Prob}(d_n = d_0)$. The distribution of d_n under the null hypothesis was estimated, for diverse sample sizes, through simulations.

3. Clinical Assessments versus Kinematic Scores Assessed with Robotic Devices

We begin this section showing the construction of each variable studied here. Then we show the results of the test of independence between them.

3.1. Variable Description. The Wolf Motor Function Test (WMFT) is an assessment scale that quantifies upper extremity ability, through 15 timed function-based tasks and 2 strength tasks. See for illustration two different tasks in figure 3. WMFT evaluates the performance time and it allows 120 seconds per task, as well as quantifies the quality of movement through a range of functional abilities. The administration of WMFT requires minimal training for the test execution. It incorporates a camera which is placed in a predefined position/distance. The score of each task is given by the analysis of the videos, see [21] and [18]. [4] published Standard Error of Measurement (SEM) (0.2 seconds) and Minimal Detectable Changes (0.7 seconds) in chronic stroke. Another studies ([4]) show that WMFT has excellent reliability and internal consistency as well as adequate criterion validity ([21], [16], [20]). However, for the results to be reliable, it is recommend to follow the manual, as well as investing in training the therapists, see [21]. The WMFT contributes to show the patient's level of function, and potential motor recovery, as well as, allows to plan treatment for functional activities. In this work, we will use the data related to the time of execution of WMFT tasks, WT variable, which is the average time to perform all the tasks.



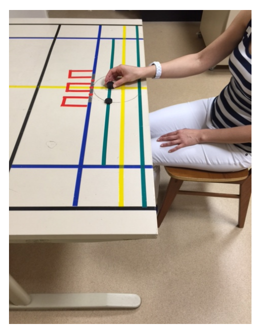


FIGURE 3. Two tasks from the Wolf Motor Function Test.

The goal-directed tasks in the robot device can be designed to measure motor impairments including poor coordination, impaired motor speed or accuracy, decreased dexterity and diminished strength. This evaluation consists of different visually guided tasks in which the patient must perform range movements, circular movements and attempts to movements against resistance. In circle-drawing tasks, the goal is to capture the coordination of the movements of the upper extremities, by the IND (Independence) variable. IND measures the patient's ability to freely coordinate their arm purposefully in all directions. A shoulder task in the robotic device, measures the patient's ability to generate a maximum force. The large range of motion, in abduction and flexion of the shoulder and in contact with the glenoid cavity, are only possible because of the movements sliding, rotating and rolling. The robot calculates the difference between the minimum and maximum of vertical forces. Applying a force with the upper extremity is essential to safe function and independence in daily living activities. See [13] for details about the formal definition of those variables. The evaluation using the robotic device can be completed in 30 minutes and at the end, it shows a graph report, with a representation of the activity and also it shows the numeric values: IND, SA and SF obtained by the patient.

3.2. Tests Results. Based on an inspection of the figure 2, the Pearson's test is excluded from the application. We show the results of the d_n for the three situations exposed in figure 2. All the tests used to compare the behavior of d_n are marginal free i.e. based on the ranks of the observations, see figure 4.

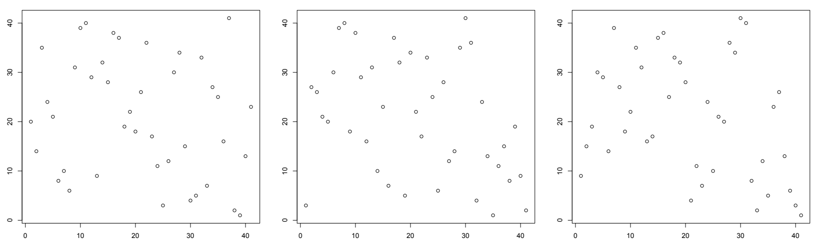


FIGURE 4. Left: x axis-Ranks of the time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Ranks of Independence (IND). Middle: x axis-Ranks of the time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Ranks of Shoulder Abduction (SA). Right: x axis-Ranks of the time spent in the execution of The Wolf Motor Function Test (WT) versus y axis-Ranks of Shoulder Flexion (SF).

TABLE 1. P-values between WT versus Y (in boldface the lower values).

Y	Spearman	Kendall	d_n
IND	0.1024	0.0773	**0.0296**
SA	0.0143	0.0157	**0.0061**
SF	**0.0255**	0.0389	0.0296

In figure 4 (right) the pairs expose a Spearman's rho=-0.3075 and a Kendall's tau=-0.1927. The three tests reject the independence assumption at level 0.03 (or 0.05), in favour of a decreasing trend. The figure 4 (left) shows the plot between the ranks of the observations of WT and the ranks of the observations of IND with Spearman's rho=-0.2019 and Kendall's tau=-0.1561. In that case, we observe that from traditional tests it is not possible to detect the negative dependence using regular significance levels. According to the d_n's test we confirm the decreasing tendency at level 0.03 or 0.05. In the case of figure 4 (middle), we have a more expressive evidence coming from the magnitude of the coefficients: Spearman's rho=-0.3434 and a Kendall's tau=-0.2341. For this case, in comparison with the other tests, the d_n's test shows a better performance for a significance level more extreme, say 0.01.

4. Simulation Study

We show in this section the performance of d_n, compared with statistical tests that allow unilateral options, focusing on scenarios produced by mixture of distributions. We consider samples which are $p\%$ independent and $(1-p)\%$ coming from a normal distribution centered in 0 with variance 1, having different degrees of correlation. These situations are inspected considering different variability on the independent proportion of the sample. In the Pearson statistical tests, the variability will cause confusion, obstructing its ability to detect the correlation of the dependent proportion of the sample. Further, we consider a displacement in x-axis, in the graphic of the independent proportion of the sample, with this we can move the proportions of samples, the dependent proportion from the independent one. We introduce these scenarios through the next equation,

$$(X,Y) \sim pN_2\left(\underline{\mu},\Sigma_1\right) + (1-p)N_2\left(\underline{0},\Sigma_2\right), \tag{4.1}$$

where $\underline{\mu} = (\mu, 0)$, $\Sigma_1 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$.

N_2 denotes the bivariate normal distribution and if necessary we will use the notation $\theta = (p,\mu,\sigma,\rho)$ for designating the parameters that can be specified in these mixtures. Figure 5 shows three situations generated by the equation (4.1),

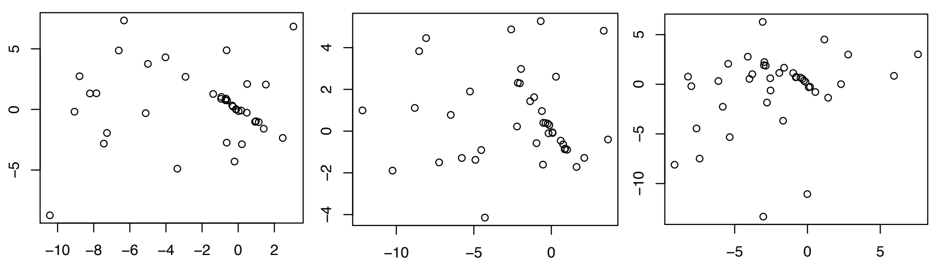


FIGURE 5. Scatterplot of simulated data following equation (4.1), sample size=40, $\mu = -4, \sigma = 4, \rho = 0.99$ Left: $p = 0.5$. Middle: $p = 0.6$. Right: $p = 0.7$.

each with a proportion $(1-p)$ of the sample, with negative correlation. We explored settings of negative correlation and small sample sizes, since the real datasets respond to these general characteristics.

Denote by $\left\{(X_i^j, Y_i^j)\right\}_{i=1}^n$ the j simulated sample of size $n = 10, 20, 30, 40, 50$, where $j = 1, \ldots, 5000$. Given a level α, we calculate the empirical significance level as being,

$$\frac{\#\left\{j : \text{p-value}\left(\left\{(X_i^j, Y_i^j)\right\}_{i=1}^n\right) \le \alpha\right\}}{5000}. \tag{4.2}$$

where p-value$\left(\left\{ (X_i^j, Y_i^j) \right\}_{i=1}^{n} \right)$ denotes the p-value associated with the sample j, $\left\{ (X_i^j, Y_i^j) \right\}_{i=1}^{n}$.

Consider now the situation of independence, table 2 and figure 6 expose the

TABLE 2. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50. Model generated from equation (4.1) with $\theta = (0.5, 0, 1, 0)$, *independent case*; with $\alpha = 0.05$, equation (4.2).

n	d_n	Spearman	Pearson	Kendall
10	0.068	0.058	0.055	0.040
20	0.035	0.039	0.045	0.041
30	0.054	0.042	0.043	0.039
40	0.038	0.041	0.036	0.040
50	0.043	0.043	0.046	0.044

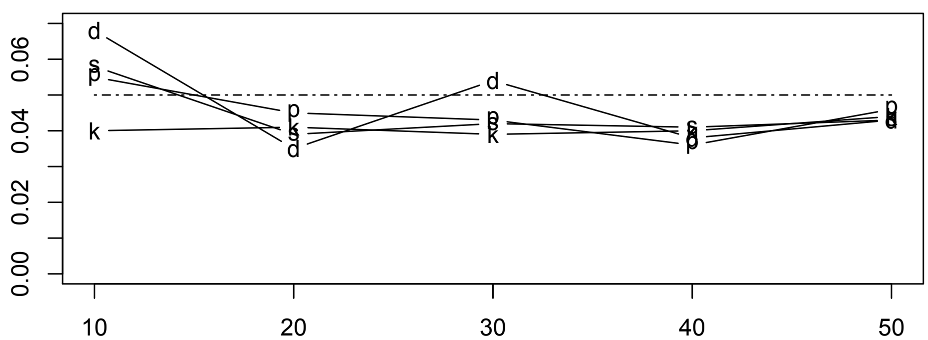


FIGURE 6. Empirical powers of one tailed statistical tests: d_n (d), Spearman (s), Pearson (p) and Kendall (k), for sample sizes: 10, 20, 30, 40 and 50. Model generated from equation (4.1) with $\theta = (0.5, 0, 1, 0)$, *independent case*; with $\alpha = 0.05$, equation (4.2). See table 2.

performance of the empirical powers of the one tailed version of the tests: d_n, Kendall, Pearson and Spearman, at significance level 0.05. Since the p-value is corrected, according to [14], as expected, the significance level of d_n is maintained.

Now we consider dependent situations, given by the equation (4.1), as those illustrated by the figure 5 showing the scatterplot of samples (with sample size=40) for three values of p. The figure 7 shows the empirical powers associated with each test, considering sample sizes from 10 to 50.

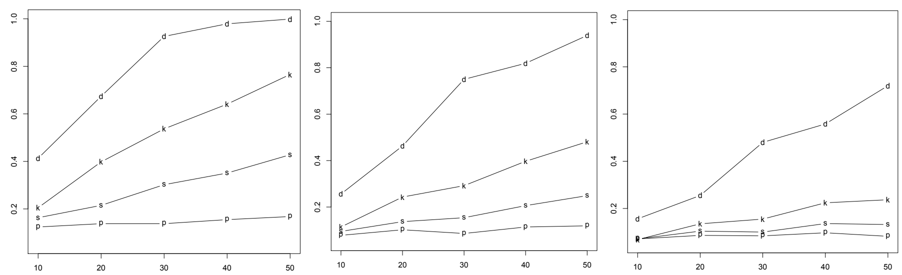


FIGURE 7. Sample sizes versus empirical power of one tailed statistical tests: d_n (d), Spearman (s), Pearson (p), Kendall (k), $\mu = -4, \sigma = 4, \rho = 0.99$ in equation (4.1). Left: $p = 0.5$, Middle: $p = 0.6$, Right: $p = 0.7$.

In the tables 3, 4 and 5 we explore situations generated from the equation (4.1) with $\mu = -4, \sigma = -4$, more precisely, we use $\theta = (0.5, -4, 4, \rho), \theta = (0.6, -4, 4, \rho)$ and $\theta = (0.7, -4, 4, \rho)$ respectively for several values of ρ. It is observed the good behavior of the test d_n, for all the values of p considered: 0.5, 0.6 and 0.7. This is, the d_n's power increases with the size of the sample and with the magnitude of the correlation coefficient ρ associated with the dependent proportion of the sample. The increase of the d_n's power, is always more pronounced than the observed in the other procedures. And, the power of d_n is weakened by the increase in the proportion of independent samples: 0.5, 0.6 and 0.7. This effect can be visualized in figure 7 from left to right. Moreover, the good performance of d_n's power is also observed for other settings. See tables 6 and 7, in which we investigate the situations $\theta = (p, -4, 2, \rho)$ and $\theta = (p, -2, 4, \rho)$ respectively, with $p = 0.5, 0.6, 0.7$ and $\rho = 0.95, 0.975, 0.99$.

TABLE 3. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50, considering different degrees of correlation ρ. Model generated from equation (4.1) with $\theta = (0.5, -4, 4, \rho)$; $\alpha = 0.05$-equation (4.2). In boldface we highlight the higher powers.

ρ	n	d_n	Spearman	Pearson	Kendall
0.85	10	**0.207**	0.138	0.135	0.147
	20	**0.295**	0.191	0.105	0.276
	30	**0.517**	0.251	0.120	0.373
	40	**0.592**	0.289	0.125	0.433
	50	**0.741**	0.325	0.139	0.534
0.90	10	**0.248**	0.141	0.133	0.157
	20	**0.380**	0.195	0.140	0.307
	30	**0.632**	0.266	0.148	0.427
	40	**0.701**	0.283	0.145	0.512
	50	**0.819**	0.372	0.149	0.586
0.95	10	**0.297**	0.157	0.117	0.188
	20	**0.471**	0.214	0.142	0.350
	30	**0.744**	0.268	0.149	0.462
	40	**0.848**	0.319	0.141	0.576
	50	**0.934**	0.379	0.140	0.667
0.975	10	**0.357**	0.143	0.118	0.195
	20	**0.561**	0.215	0.134	0.371
	30	**0.862**	0.299	0.126	0.505
	40	**0.940**	0.349	0.141	0.647
	50	**0.979**	0.399	0.158	0.721
0.99	10	**0.413**	0.163	0.124	0.205
	20	**0.674**	0.215	0.138	0.398
	30	**0.926**	0.302	0.138	0.537
	40	**0.979**	0.351	0.155	0.641
	50	**0.999**	0.428	0.168	0.766

TABLE 4. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50, considering different degrees of correlation ρ. Model generated from equation (4.1) with $\theta = (0.6, -4, 4, \rho)$; $\alpha = 0.05$-equation (4.2). In boldface we highlight the higher powers.

ρ	n	d_n	Spearman	Pearson	Kendall
0.85	10	**0.162**	0.079	0.088	0.079
	20	**0.195**	0.122	0.101	0.176
	30	**0.367**	0.153	0.106	0.219
	40	**0.382**	0.189	0.106	0.286
	50	**0.510**	0.174	0.099	0.285
0.9	10	**0.173**	0.100	0.093	0.106
	20	**0.247**	0.118	0.089	0.188
	30	**0.427**	0.149	0.094	0.234
	40	**0.487**	0.171	0.109	0.297
	50	**0.599**	0.202	0.106	0.365
0.95	10	**0.209**	0.097	0.089	0.101
	20	**0.289**	0.144	0.106	0.220
	30	**0.502**	0.166	0.089	0.262
	40	**0.620**	0.177	0.097	0.322
	50	**0.766**	0.227	0.108	0.418
0.975	10	**0.267**	0.113	0.096	0.126
	20	**0.370**	0.140	0.108	0.225
	30	**0.647**	0.170	0.105	0.298
	40	**0.744**	0.194	0.092	0.358
	50	**0.876**	0.214	0.096	0.435
0.99	10	**0.256**	0.095	0.078	0.114
	20	**0.463**	0.137	0.102	0.243
	30	**0.750**	0.154	0.086	0.292
	40	**0.819**	0.206	0.114	0.397
	50	**0.939**	0.249	0.119	0.481

TABLE 5. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50, considering different degrees of correlation ρ. Model generated from equation (4.1) with $\theta = (0.7, -4, 4, \rho)$; $\alpha = 0.05$-equation (4.2). In boldface we highlight the higher powers.

ρ	n	d_n	Spearman	Pearson	Kendall
0.85	10	**0.111**	0.066	0.061	0.057
	20	**0.138**	0.070	0.059	0.090
	30	**0.216**	0.083	0.074	0.122
	40	**0.234**	0.109	0.077	0.164
	50	**0.317**	0.121	0.085	0.177
0.9	10	**0.114**	0.072	0.079	0.068
	20	**0.165**	0.092	0.067	0.113
	30	**0.266**	0.108	0.087	0.137
	40	**0.257**	0.087	0.070	0.145
	50	**0.374**	0.124	0.080	0.188
0.95	10	**0.140**	0.071	0.072	0.058
	20	**0.194**	0.103	0.079	0.143
	30	**0.344**	0.108	0.074	0.161
	40	**0.350**	0.107	0.082	0.164
	50	**0.475**	0.122	0.080	0.209
0.975	10	**0.136**	0.086	0.080	0.079
	20	**0.227**	0.075	0.059	0.119
	30	**0.401**	0.110	0.074	0.165
	40	**0.444**	0.137	0.091	0.201
	50	**0.587**	0.127	0.077	0.227
0.99	10	**0.156**	0.071	0.071	0.068
	20	**0.255**	0.104	0.086	0.135
	30	**0.480**	0.100	0.084	0.155
	40	**0.558**	0.136	0.097	0.224
	50	**0.720**	0.132	0.082	0.237

TABLE 6. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50, considering different degrees of correlation ρ. Model generated from equation (4.1) with $\theta = (p, -4, 2, \rho)$; $\alpha = 0.05$-equation (4.2). In boldface we highlight the higher powers.

p	ρ	n	d_n	Spearman	Pearson	Kendall
0.5	0.95	10	**0.260**	0.141	0.112	0.145
		20	**0.413**	0.231	0.161	0.356
		30	**0.725**	0.271	0.162	0.449
		40	**0.801**	0.354	0.194	0.577
		50	**0.919**	0.431	0.204	0.695
0.5	0.975	10	**0.339**	0.147	0.124	0.164
		20	**0.519**	0.218	0.159	0.371
		30	**0.817**	0.318	0.174	0.522
		40	**0.918**	0.373	0.186	0.639
		50	**0.973**	0.454	0.251	0.760
0.5	0.99	10	**0.388**	0.139	0.115	0.169
		20	**0.642**	0.233	0.152	0.395
		30	**0.901**	0.309	0.182	0.545
		40	**0.967**	0.346	0.195	0.665
		50	**0.997**	0.450	0.236	0.796
0.6	0.95	10	**0.200**	0.098	0.103	0.094
		20	**0.247**	0.126	0.096	0.176
		30	**0.491**	0.172	0.114	0.258
		40	**0.554**	0.216	0.144	0.355
		50	**0.707**	0.234	0.157	0.385
0.6	0.975	10	**0.207**	0.099	0.091	0.090
		20	**0.333**	0.122	0.089	0.184
		30	**0.617**	0.184	0.120	0.291
		40	**0.699**	0.223	0.148	0.368
		50	**0.829**	0.226	0.142	0.401
0.6	0.99	10	**0.234**	0.095	0.093	0.097
		20	**0.434**	0.161	0.111	0.242
		30	**0.720**	0.183	0.124	0.296
		40	**0.813**	0.228	0.142	0.410
		50	**0.913**	0.241	0.143	0.457
0.7	0.95	10	**0.116**	0.049	0.053	0.038
		20	**0.152**	0.091	0.067	0.105
		30	**0.286**	0.095	0.077	0.117
		40	**0.302**	0.095	0.082	0.132
		50	**0.414**	0.130	0.097	0.182
0.7	0.975	10	**0.135**	0.072	0.069	0.057
		20	**0.177**	0.082	0.075	0.100
		30	**0.361**	0.099	0.090	0.139
		40	**0.420**	0.116	0.093	0.167
		50	**0.538**	0.141	0.111	0.208
0.7	0.99	10	**0.142**	0.073	0.068	0.061
		20	**0.235**	0.088	0.070	0.113
		30	**0.471**	0.099	0.077	0.138
		40	**0.516**	0.116	0.082	0.169
		50	**0.677**	0.130	0.111	0.220

TABLE 7. Empirical powers of one tailed statistical tests: d_n, Spearman, Pearson and Kendall, for sample sizes: 10, 20, 30, 40 and 50, considering different degrees of correlation ρ. Model generated from equation (4.1) with $\theta = (p, -2, 4, \rho)$; $\alpha = 0.05$-equation (4.2). In boldface we highlight the higher powers.

p	ρ	n	d_n	Spearman	Pearson	Kendall
		10	**0.269**	0.142	0.151	0.176
		20	**0.441**	0.237	0.162	0.387
0.5	0.95	30	**0.734**	0.303	0.150	0.493
		40	**0.826**	0.377	0.177	0.590
		50	**0.938**	0.422	0.172	0.698
		10	**0.324**	0.159	0.149	0.206
		20	**0.519**	0.255	0.161	0.397
0.5	0.975	30	**0.846**	0.323	0.173	0.529
		40	**0.907**	0.388	0.171	0.645
		50	**0.971**	0.413	0.170	0.702
		10	**0.373**	0.166	0.142	0.212
		20	**0.638**	0.238	0.144	0.422
0.5	0.99	30	**0.910**	0.322	0.151	0.585
		40	**0.968**	0.411	0.175	0.673
		50	**0.993**	0.444	0.180	0.767
		10	**0.182**	0.104	0.109	0.111
		20	**0.273**	0.155	0.121	0.234
0.6	0.95	30	**0.522**	0.187	0.104	0.284
		40	**0.609**	0.233	0.136	0.383
		50	**0.749**	0.245	0.116	0.448
		10	**0.210**	0.123	0.111	0.140
		20	**0.343**	0.172	0.114	0.252
0.6	0.975	30	**0.602**	0.186	0.116	0.294
		40	**0.673**	0.193	0.104	0.342
		50	**0.840**	0.277	0.140	0.461
		10	**0.244**	0.114	0.107	0.132
		20	**0.392**	0.151	0.116	0.237
0.6	0.99	30	**0.732**	0.188	0.124	0.334
		40	**0.817**	0.235	0.118	0.429
		50	**0.931**	0.262	0.126	0.472
		10	**0.155**	0.095	0.082	0.094
		20	**0.186**	0.119	0.102	0.158
0.7	0.95	30	**0.321**	0.121	0.081	0.182
		40	**0.357**	0.130	0.100	0.196
		50	**0.462**	0.143	0.096	0.219
		10	**0.152**	0.100	0.097	0.093
		20	**0.198**	0.110	0.093	0.143
0.7	0.975	30	**0.362**	0.120	0.085	0.174
		40	**0.435**	0.136	0.088	0.223
		50	**0.566**	0.151	0.095	0.280
		10	**0.164**	0.099	0.092	0.094
		20	**0.253**	0.102	0.088	0.153
0.7	0.99	30	**0.446**	0.132	0.094	0.190
		40	**0.516**	0.141	0.089	0.226
		50	**0.666**	0.160	0.097	0.265

5. Conclusion

We devote this paper to introduce the one-tailed hypothesis test d_n, a member of the LIS family of tests. It is noteworthy that the tests introduced in [5] which are coming from the LIS family, show evidence of being suitable to detect dependence when the sample is coming from mixture of distributions, a challenge for the majority of the independence tests. For that reason this issue is explored in depth in this article, under the formulation of this new test d_n. In this paper is explored the possibility to catch the dependence in mixture of distributions with the presence of a decreasing tendency. In relation to [5], in this paper is incorporated an innovation in the construction of d_n, since is not only used the concept of the longest increasing subsequence defined by the permutation π which maps the ranks of the X observations in the ranks of the Y observations. In fact d_n is defined using increasing and decreasing subsequences, over π, which allows to catch a decreasing trend in mixture of distributions. In the computation of the p-values of d_n we incorporate the correction proposed by [14], since it allows to control the Type I error and improves the power of the test. As can be concluded from the simulation study, the empirical power of the test d_n, excels in comparison with the tests of Kendall, Pearson and Spearman, in mixture of distributions, as the situations investigated in this paper. This procedure and its outstanding performance, emphasizes the role of the *Longest Increasing Subsequence* and related concepts in the development of new tests of hypothesis.

In the application to real data, we see that the test d_n corroborates that the time (WT) used in executing the *The Wolf Motor Function Test* has a negative dependence when compared with the 3 variables provided by the robotic device: Independence (IND), Shoulder Abduction (SA) and Shoulder Flexion (SF). For IND, d_n provides significant results, which can not be said about Kendall and Spearman procedures. For SA, the test d_n, allows to use very strict significance levels, when compared to those permitted by Kendall and Spearman. The present study contributes to the rehabilitation field since it helps to clarify the dependence between variables measured in conventional methods and new instruments (robotic devices). Besides, it reinforce the importance of further studies with a larger sample size in order to validate the robotic assessment for the general stroke population and to define its psychometric properties.

6. Acknowledgments

The authors gratefully acknowledge the support for this research provided by FAPESP's project "Research, Innovation and Dissemination Center for Neuromathematics" 2013/ 07699-0, (S. Paulo Research Foundation). The authors also wish to express their gratitude to two referees and the editor for their helpful comments on an earlier draft of this paper.

References

[1] Culham, E. and Peat, M.: Functional anatomy of the shoulder complex, *Journal of Orthopaedic & Sports Physical Therapy* **18** (1993)(1) 342-350.

[2] Escamilla, R. F., Yamashiro, K., Paulos, L. and Andrews, J. R.: Shoulder muscle activity and function in common shoulder rehabilitation exercises, *Sports Medicine* **39** (2009) (8) 663-685.

[3] Fisher, R. A.: The Logical Inversion of the Notion of the Random Variable, *Sankhya: The Indian Journal of Statistics* (1945) 129-132.

[4] Fritz, S. L., Blanton, S., Uswatte, G., Taub, E. and Wolf, S. L.: Minimal detectable change scores for the Wolf Motor Function Test, *Neurorehabilitation and Neural Repair* **23** (2009) 662-667.

[5] García, Jesús E. and González-López, V.A.: Independence tests for continuous random variables based on the longest increasing subsequence, *Journal of Multivariate Analysis* **127** (2014) 126-146.

[6] García, Jesús E. and González-López, V.A.: Independence Test for Sparse Data, in: *International Conference of Numerical Analysis and Applied Mathematics 2015 (ICNAAM 2015)*, **1738** (2016) (1) 140002, AIP Publishing.

[7] Greene, C.: An extension of Schensted's theorem, *Advances in Mathematics* **14** (1974) (2) 254-265.

[8] Harrison, J. K., McArthur, K. S. and Quinn, T. J.: Assessment scales in stroke: clinimetric and clinical considerations, *Clinical interventions in aging* **8** (2013) 201-211.

[9] Hogan, N. and Krebs, H. I.: Interactive robots for neuro-rehabilitation, *Restorative neurology and neuroscience* **22** (2004) (3-5), 349-358.

[10] Hogan, N., Krebs, H. I., Sharon, A. and Charnnarong, J.: *U.S. Patent No. 5466213*, Washington, DC: U.S. Patent and Trademark Office. 1995.

[11] Krebs, H. I., Hogan, N., Aisen, M. L. and Volpe, B. T.: Robot-aided neurorehabilitation, *IEEE Transactions on rehabilitation engineering* **6** (1998) (1) 75-87.

[12] Krebs, H. I., Krams, M., Agrafiotis, D. K., DiBernardo, A., Chavez, J. C., Littman, G. S., Yang, E., Byttebier, G., Dipietro, L., Rykman, A., McArthur, K., Hajjar, K., Lees, K. R. and Volpe, B. T.: Robotic measurement of arm movements after stroke establishes biomarkers of motor recovery, *Stroke* **45** (2014) (1) 200-204.

[13] Krebs, H. I., Palazzolo, J. J., Dipietro, L., Ferraro, M., Krol, J., Rannekleiv, K., Volpe, B. T. and Hogan, N.: Rehabilitation robotics: Performance-based progressive robot-assisted therapy, *Autonomous robots* **15** (2003) (1) 7-20.

[14] Lancaster, H. O.: Significance Tests in Discrete Distributions, *Journal of the American Statistical Association* **56** (1961) (294) 223-234.

[15] McMullen, J. and Uhl, T. L.: A kinetic chain approach for shoulder rehabilitation, *Journal of athletic training* **35** (2000) (3) 329-337.

[16] Morris, D. M., Uswatte, G., Crago, J. E., Cook, E. W. and Taub, E.: The reliability of the Wolf Motor Function Test for assessing upper extremity function after stroke, *Archives of physical medicine and rehabilitation* **82** (2001) (6) 750-755.

[17] Norrving, B. and Kissela, B.: The global burden of stroke and need for a continuum of care, *Neurology* **80** (2013)(3 Supplement 2) S5-S12.

[18] Pereira, N. D., Michaelsen, S. M., Menezes, I. S., Ovando, A. C., Lima, R. and Teixeira-Salmela, L. F.: Reliability of the Brazilian version of the Wolf Motor Function Test in adults with hemiparesis, *Brazilian Journal of Physical Therapy* **15** (2011) (3) 257-265.

[19] Quinn, T. J., Dawson, J., Walters, M. R. and Lees, K. R.: Functional outcome measures in contemporary stroke trials, *International journal of stroke: official journal of the International Stroke Society* **4** (2009) (3) 200-205.

[20] Whitall, J., Savin, D. N., Harris-Love, M. and Waller, S. M.: Psychometric properties of a modified Wolf Motor Function test for people with mild and moderate upper-extremity hemiparesis, *Archives of physical medicine and rehabilitation* **87** (2006) (5) 656-660.

[21] Wolf, S. L., Catlin, P. A., Ellis, M., Archer, A. L., Morgan, B. and Piacentino, A.: Assessing Wolf motor function test as outcome measure for research in patients after stroke, *Stroke* **32** (2001) (7) 1635-1639.

Jesús E. García: Department of Statistics, University of Campinas, Campinas, SP, CEP 13083-859, Brazil
E-mail address: jg@ime.unicamp.br

V.A. González-López: Department of Statistics, University of Campinas, Campinas, SP, CEP: 13083-859, Brazil
E-mail address: veronica@ime.unicamp.br

Marcel Simis: Physical and Rehabilitation Medicine Institute of the University of Sao Paulo Medical School General Hospital, CEP: 05403-010, Brazil
E-mail address: marcel.simis@hc.fm.usp.br

Thais Terranova: Physical and Rehabilitation Medicine Institute of the University of Sao Paulo Medical School General Hospital, CEP: 05403-010, Brazil
E-mail address: thais.terranova@hc.fm.usp.br

Linamara Rizzo Battistella: Physical and Rehabilitation Medicine Institute of the University of Sao Paulo Medical School General Hospital, CEP: 05403-010, Brazil
E-mail address: linamara.battistella@hc.fm.usp.br