ISSN 2395-8618

Non-Parametric *k*-Sample Tests for Comparing Forecasting Models

Dmitriy A. Klyushin

***Abstract*—Business intelligence is impossible without practical tools for assessing the quality of forecasts and comparing forecast models. A naive approach to comparing models by comparing predicted values with observable ones ignores the probabilistic nature of errors. Many models with varying degrees of accuracy are statistically equivalent. Hence, before ranking the models for accuracy, it is necessary to test the statistical hypothesis about the homogeneity of the distributions of their errors. In the presence of several models, the problems of their pairwise and group comparison arise. This chapter provides an overview of the non-parametric tests used in business analysis for pairwise and group comparisons and describes a new non-parametric statistical test that is highly reliable, sensitive, and specific. This test is based on assessing the deviation of the observed relative frequency of an event from its a priori known probability. The prior probability is given by Hill's assumption, and the confidence intervals for the binomial success rate in the Bernoulli scheme are used to estimate its difference from the observed relative frequency. The paper presents the results of computer modeling and comparison of the proposed test with the alternative Kruskal-Wallis test and the Friedman test on artificial and real examples.**

***Index Terms*—Business process modeling, error analysis, modeling and prediction, non-parametric statistics.**

1. Introduction

BUSINESS intelligence combines technologies and methods for the collection, analysis, and prediction of business information. It may be represented as a hierarchical scheme consisting of four layers.

At the foundation layer, business data are collected and analyzed by descriptive data analytics methods. Here we answer the question: "What happened?" The examples of these data are the signals from the sensors or exchange rate. Analysis of these data allows detecting the symptoms of failure of devices and identifying change-point of time series describing exchange rate.

At the second level, we analyze diagnostic information to answer the question: "Why did it happen?" This problem is harder than the first-level problem because it requests to detect and recognize patterns. Using statistical methods, we establish intrinsic relations between data and discover reasons for events. For example, we may find the correlation between events and exchange rate changes.

At the third level, we try to answer the question: "What will happen?" Using data obtained at the first two levels and methods of predictive analytics, we forecast events that may occur in the future to determine what consequences can follow. The fundamental complexity of forecasting requires the use of very complex tools of mathematical statistics, machine learning, data mining, and simulation. For example, predictive analytics can be used for forecasting the future stock price in the stock market or determine the optimal time for repairs to prevent breakdowns of process equipment.

At the fourth level using methods of prescriptive analytics, we answer the question: "What to do?" for making the optimal decision. This level demands very complex mathematical methods of optimization. For example, we can find an optimal moment for buying or selling stocks or the beginning of a repair.

The subject of this chapter is the non-parametric approach to the estimation of predictive models. Predictive analytics opens up broad perspectives in commercial, financial, and industrial applications. For example, it allows you to optimize recommendation systems in online stores, as well as segment your customer base to improve the effectiveness of direct marketing and targeted advertising. In banking and insurance, predictive analytics has become a necessary tool in assessing an applicant's creditworthiness and detecting fraud. In the financial sector, predictive analytics can improve investment performance and optimize risk management.

In industrial applications, predictive analytics allows solving the problems of forecasting product quality, optimization of repair schedules, recognition of abnormal symptoms, and many other tasks. Sophisticated high-tech enterprises use automatic control of operational processes. To optimize this control, it is necessary to organize an automated collection and analysis of indicators in order to predict resource consumption and product output. By continuously analyzing large volumes of production data, you can prevent line failures and shutdowns, minimize costs and maximize product quality.

The variety of numerical data found in all these calculations does not allow making educated assumptions about the form of their distribution; therefore, the most suitable methods for analysis in these areas are non-parametric methods, which do not imply a certain form of data distribution. These methods include non-parametric hypothesis testing methods.

Manuscript received on October 17, 2020, accepted for publication on November 24, 2020, published on December 30, 2020.

The author is with the Taras Shevchenko National University of Kyiv, 03680, Kyiv, prospect akademika Glushkova 4D, Ukraine, (e-mail: dokmed5@gmail.com).

ISSN 2395-8618

The purpose of this chapter is to survey non-parametric methods for two- and *k*-sample used in predictive analytics and propose a new approach to test homogeneity of two- and *k*-samples for estimation forecasting model effectiveness.

The paper is organized in the following way. Section 1 describes the purpose of the chapter. Section 2 describes widely used non-parametric tests for homogeneity. Subsection 2.1 contains a short survey of two-sample tests, as subsection 2.2 is devoted to *k*-sample tests. Section 3 describes the application of non-parametric tests in business intelligence. Section 4 describes the Klyushin–Petunin non-parametric test for two-sample homogeneity without ties. The two-sample case without ties is considered in Subsection 4.1, and the version with ties is described in Subsection 4.2. The extension of the above-mentioned tests on the *k*-sample case is considered in Subsection 4.3. Section 5 describes the results of numerical experiments. Section 6 contains conclusions and directions for future work.

2. Non-Parametric Homogeneity Tests

Non-parametric homogeneity tests for two and *k*-samples were studied in many scientific papers.

2.1 Two-Sample Tests

Let samples $x=(x_1,x_2,...,x_n)$ and $y=(y_1,y_2,...,y_m)$ be drawn from populations G_1 and G_2 which follow absolutely continuous distribution functions F_1 and F_2. The null hypothesis H_0 states that samples are homogeneous, i.e., follow the same distribution function, $F_1=F_2$, and the alternative hypothesis H_1 states the opposite ($F_1 \neq F_2$). The tests for homogeneity of two samples are subdivided into permutation tests, rank tests, randomization tests, and distance tests. Also, these tests form the group of universal tests that are valid against any pair of alternatives (e.g., the Kolmogorov–Smirnov test [1], the Kuiper test [2]), and tests that are valid against pairs of different alternatives of a particular class (Dickson [3], Wald and Wolfowitz [4], Mathisen [5], Wilcoxon [6] Mann–Whitney [6], Wilks [8], etc.). Also, the tests may be classified as pure non-parametric and conditionally non-parametric ones. The pure non-parametric tests do not depend on the assumptions of the distribution function (e.g., all the tests mentioned above). The conditionally nonparametric tests put some assumptions of distributions (Pitman [9], Lehmann [10], Rosenblatt [11], Dwass [12], Fisz [13], Barnard [14], Birnbaum [15], Jockel [16], Allen [17], Efron and Tibshirani [18], Dufour and Farhat [19]).

2.2 k-Sample Tests

Let samples $x^{(k)}=\left(x_1^{(k)},x_2^{(k)},...,x_n^{(k)}\right)\in G_k$ be drawn from populations $G_k, k=\overline{1,K}$ following distributions F_k. The null hypothesis H_0 states that samples are homogeneous, i.e. $F_1=F_2=...=F_K$, and the alternative hypothesis H_1 states the opposite, i.e., there are such i and j that $F_i \neq F_j$, $i \neq j$, $i,j=1,\ldots,K$.

k-sample tests based on the distance between empirical distributions functions include the Kolmogorov–Smirnov test [1], [2], the Cramer–Von-Mises tests [20], the Anderson–Darling test [21], etc. These tests use different distances between empirical functions, e.g. L_1 (the Manhattan distance), L_2 (the root-mean-square deviation), and L_∞ (the Chebyshev distance). The sensitivity of these tests depends on the sample size; the larger, the better. Obviously, this condition can rarely be satisfied in real applications.

k-sample tests based on the likelihood ratio are differentiated depending on a function used in the ratio: empirical distributions functions (Zhang test [22], dynamic slicing [23], energy distance [24], etc.), empirical characteristics functions (Śzekely and Rizzo [25], [26], Fernández et al. [27], Hušková and Meintanis [28], etc.), and kernel density estimations (Uña-Alvarez et al. [29], [30]).

Despite the strengths of these tests, their power and consistency depend on some parameters and assumptions that sometimes are hard to satisfy in practice. Also, to compute critical values of these tests, it is often necessary to make permutations that increase a computational burden [31].

Non-parametric rank-based tests are more powerful alternatives to the above-mentioned ones. They do not require an assumption on the distribution function (e.g., normality) to determine the p-value of the test, can be very powerful in the cases when other tests would fail and are robust to outliers [32]. That is why we shall use as benchmarks the Kruskal-Wallis test and The Friedman test [33–37] implemented in all popular statistical computing packages.

3. Non-parametric used in forecasting models estimation

Analyzing time series, the analyst has a wide range of predictive model options. The choice of the best model depends on many factors, primarily on its accuracy. Meanwhile, the accuracy of the model is a random variable that has a certain distribution. Two models having different accuracy can be compared with each other only by testing the statistical hypothesis that the error of one of them is stochastically less than the error of the other. If the distributions of errors are identical, then the models should be considered equivalent, despite the fact that they have different indicators of accuracy. Usually, the standard error (MSE), the mean absolute percentage error (MAPE) and the mean absolute deviation (MAD) are used for estimating the accuracy of forecasting models. Thus, analyzing the quality of forecasting models, we must test a hypothesis that samples of forecast errors are homogeneous as a rule model errors are supposed to be stationary and unbiased.

ISSN 2395-8618

The testing equality of forecasting model accuracy is clearly described in [38]. Let $M_j, j=1,...,m$ be forecasting models producing predictions $x_i^{(j)}$ of data sequence $x_i, i=1,...,n$ and $\varepsilon_i^{(j)}, i=1,...,n; j=1,...,m$ be errors of the model M_j following distribution F_i. Consider a loss function $g\left(\varepsilon_i^j\right)$ describing model accuracy, e.g., standard deviation. The hypothesis about similar accuracy of models M_k and M_l is equivalent to the hypothesis that the mathematical expectation of $d_i^{(k,l)} = g\left(\varepsilon_i^{(k)}\right) - g\left(\varepsilon_i^{(l)}\right)$ is zero. Thus, $E\left(g\left(\varepsilon_i^{(k)}\right)\right) = E\left(g\left(\varepsilon_i^{(l)}\right)\right)$. If a loss function is the standard deviation $\varepsilon_i^{(k)} - \varepsilon_i^{(l)}$, the problem is reduced to testing the hypothesis $E\left(\varepsilon_i^{(k)}\right) = E\left(\varepsilon_i^{(l)}\right)$. Usually, testing equality of forecasting models accuracy is reduced to the testing of this hypothesis [38]. However, the hypothesis that accuracy measures follow the same distribution is more general. Thus, to test homogeneity, it is necessary to test the null hypothesis not only in partial case $E\left(\varepsilon_i^{(k)}\right) = E\left(\varepsilon_i^{(l)}\right)$ but in the general case $F_k = F_l$.

4. Klyushin–Petunin Test for Homogeneity

The Klyushin–Petunin test has two versions: for samples without ties and for samples containing ties.

4.1 Two-sample case without ties

Let samples $x = \left(x_1, x_2, ..., x_n\right)$ and $y = \left(y_1, y_2, ..., y_m\right)$ be drawn from populations G_1 and G_2 which follows absolutely continuous distribution functions F_1 and F_2. According to Hill's assumption $A_{(n)}$ [39], if random values $x_1, x_2, ..., x_n$ are exchangeable and belong to absolutely continuous distribution, then

$$P\left(x_{n+1} \in \left(x_{(i)}, x_{(j)}\right)\right) = p_{ij} = \frac{j-i}{n+1}, \quad j > i, \tag{1}$$

where x_{n+1} is a random value following the same distribution as $x_1, x_2, ..., x_n$, and $x_{(i)}$ is the i-th value of the ordered sample. On the ground of this fact, non-parametric tests for homogeneity of samples without ties [40] and with ties [41] were developed.

Let $A_{ij}^{(k)}$ be an event when the elements of y are greater than $x_{(i)}$ and less than $x_{(j)}$, and h_{ij} is its relative frequency. Knowing the a priori probability (1) and the observed relative frequency h_{ij}, we can estimate how much h_{ij} deviates from p_{ij} using Wilson confidence intervals for binomial proportions:

$$\begin{aligned} p_{ij}^{(1)} &= \frac{h_{ij}m + z^2/2 - z\sqrt{h_{ij}(1-h_{ij})m + z^2/4}}{m+z^2}, \\ p_{ij}^{(2)} &= \frac{h_{ij}m + z^2/2 + z\sqrt{h_{ij}(1-h_{ij})m + z^2/4}}{m+z^2}. \end{aligned} \tag{2}$$

Then, we compute the lower and upper bounds of the confidence interval $I_{ij}^{(n,m)} = \left(p_{ij}^{(1)}, p_{ij}^{(2)}\right)$ with the parameter z depending on the desired significance level. If z is equal to 3, then the significance level of $I_{ij}^{(n,m)}$ is less than 0.05 [40]. In this case, (2) is a so-called *3s-rule interval*. This rule is based on the Petunin-Vysochanskii inequality [42, 43], stating that if X be a random variable with unimodal distribution, mean μ and finite, non-zero variance σ^2, than $P\left(\left|X-\mu\right| \geq \lambda\sigma\right) \leq \frac{4}{9\lambda^2}$ for any $\lambda > \sqrt{\frac{8}{3}}$. Replacing the mean μ and the variance σ by the sample mean $\bar{x}$ and the standard deviation s and setting λ equal to 3, we obtain the *3s-rule*: $P\left(\left|X-\bar{x}\right| \geq 3s\right) \leq \frac{4}{81} \approx 0.494 < 0.05$.

Denote $N = (n-1)n/2$ and $L = \#\left\{p_{ij} = \frac{j-i}{n+1} \in I_{ij}^{(n,m)}\right\}$. Then, $h = L/N$ is a homogeneity measure of samples x and y, which we shall call *p-statistics*, and a Wilson confidence interval (2) where we set h instead of h_{ij} and N instead of m is the confidence interval for the probability $p\left(\frac{j-i}{n+1} \in I_{ij}^{(n,m)}\right)$.

The scheme of events $A_{ij}^{(k)}$ when the null hypothesis is true is called a generalized Bernoulli scheme [44], [45]. If the null hypothesis is false, this scheme is called a modified Bernoulli scheme. In the general case, when the null hypothesis can be either true or false, this scheme is called Matveichuk–Petunin scheme [46]. Thus, the test for the null hypothesis $F_1 = F_2$ with a significance level, which is less than 0.05, maybe formulated in the following way: construct the Wilson confidence interval $I_n = \left(p_1, p_2\right)$ for p-statistics; if I_n contains 0.95, the null hypothesis is accepted, else the null hypothesis is rejected.

4.2 Two-Sample Case with Ties

In practice, due to imprecise measurement, samples often contain ties, i.e., repeated elements. A sample x containing absolutely precise elements we shall call hypothetical. The sample $\tilde{x} = \left(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\right)$ containing approximations of the hypothetical elements of the hypothetical x we shall call empirical. Population $\tilde{G}$ we shall call an empirical population corresponding to the hypothetical population G. Let $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$ and $\tilde{x}_{(1)} \leq \tilde{x}_{(2)} \leq ... \leq \tilde{x}_{(n)}$ be variational series constructed by hypothetical and empirical samples.

For a sample value x^* that is drawn from G independently from x the Hill assumption holds [39]:

$$p\left(x^* \in \left[x_{(k)}, x_{(k+1)}\right)\right) = \frac{1}{n+1}, \tag{3}$$

where $k = 0,1,...,n$, $x_{(0)} = -\infty$, and $x_{(n+1)} = \infty$. Hence,

ISSN 2395-8618

$$p_{ij} = p\left(A_{ij}\right) = p\left(\tilde{x}^* \in \left[\tilde{x}_{(i)}, \tilde{x}_{(j)}\right)\right) \approx \approx \gamma_i + \gamma_{i+1} + ... + \gamma_{j-1} + \frac{j-i}{n+1}, \quad (4)$$

where $\gamma_k = \frac{t\left(\tilde{x}_{(k)}\right)-1}{n+1}$, $t\left(\tilde{x}_{(l)}\right)$ is the number of repetitions of $\tilde{x}_{(l)}$, $A_{ij} = \left\{\tilde{x}^* \in \left[\tilde{x}_{(i)}, \tilde{x}_{(j)}\right]\right\}$. If $\tilde{x} = \left(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\right)$ does not contain ties, then (4) transforms to (1).

Let the null hypothesis states that hypothetical continuous distribution functions F_1 and F_2 of hypothetical populations G_1 and G_2 are identical. Consider empirical samples $\tilde{x} = \left(\tilde{x}_1, ..., \tilde{x}_n\right) \in \tilde{G}_1$ and $\tilde{y} = \left(\tilde{y}_1, ..., \tilde{y}_m\right) \in \tilde{G}_2$, where $\tilde{G}_1$ and $\tilde{G}_2$ are the empirical populations corresponding to hypothetical populations G_1 and G_2. Suppose that $F_1 = F_2$ and denote $A_{ij}^{(k)} = \left\{\tilde{x}_k \in \left(\tilde{x}_{(i)}, \tilde{x}_{(j)}\right)\right\}$, where $\tilde{x}_{(i)}$ is the *i*-th value of the ordered sample. If $F_1 = F_2$, then the probability of $A_{ij}^{(k)}$ is equal to (4). Construct the Wilson confidence interval (2) $I_{ij}^{(n,m)} = \left(p_{ij}^{(1)}, p_{ij}^{(2)}\right)$ for the unknown probability of $A_{ij}^{(k)}$ using its observed relative frequency. The number of all confidence intervals $I_{ij}^{(n,m)} = \left(p_{ij}^{(1)}, p_{ij}^{(2)}\right)$ is equal to $N = (n-1)n/2$. Put $h = \frac{1}{N}\#\left\{\gamma_i + \gamma_{i+1} + ... + \gamma_{j-1} + \frac{j-i}{n+1} \in I_{ij}^{(n)}\right\}$. Compute a confidence interval $I^{(n,m)} = \left(p^{(1)}, p^{(2)}\right)$ for the probability $p\left(\gamma_i + \gamma_{i+1} + ... + \gamma_{j-1} + \frac{j-i}{n+1} \in I_{ij}^{(n,m)}\right)$ using (2) as described in Section 4.2. The statistics h is called the *empirical p-statistics*. It estimates the homogeneity of empirical samples $\tilde{x}$ and $\tilde{y}$, Wilson confidence interval (2), where we set *h* instead of h_{ij} and *N* instead of *m* is the confidence interval for the probability $p\left(\gamma_i + \gamma_{i+1} + ... + \gamma_{j-1} + \frac{j-i}{n+1} \in I_{ij}^{(n,m)}\right)$.

4.3 k-sample case

The two-sample Klyushin–Petunin test may be expanded on the *k*-sample case using the scheme one-vs-rest. Suppose that all *k* samples follow the same distribution. Then, if we select a sample and join other samples into one sample, we shall reduce the *k*-sample case to the two-sample case. Joining the samples following the same distribution, we obtain a sample following this distribution. Thus, if the p-statistics between selected and joined samples is greater than 0.95, the null hypothesis about identical distributions is accepted; otherwise, the *k* samples are heterogeneous.

5. Numerical Experiments

The strengths of the proposed approach are justified by numerical experiments comparing the p-statistics with the Kruskal–Wallis test and the Friedman test. We test the location shift hypothesis (samples have different means and the same variances) and the scale shift hypothesis (sample have the same means and different variances) using samples, which do not contain ties and samples with a single tie. For experiments, we selected samples from the Gaussian distributions $N(\mu, \sigma)$ with different overlapping, where μ is the mean and σ is the standard deviation.

5.1 Location Shift Hypothesis without Ties

To test a location shift hypothesis we 100 times generated by C++ pseudo random number generator 5 samples containing 10 real numbers that follows the distributions N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), and N(0.4, 1), and computed p-statistics by the scheme one-vs.-others. If the p-statistics between these samples was greater than 0.95, we concluded that they are homogeneous, otherwise the *k* samples were considered as heterogeneous. Average p-statistics was equal to 0.756. The Wilson confidence interval for the p-statistics, constructed as indicated in Section 4.1, was (0.532, 0.894). Hereinafter, Var1–Var5 denotes N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), and N(0.4, 1) in Tables 1, 3, 5, 11, 13, 15, and N(0,1), N(0, 2), N(0, 3), N(0, 4), and N(0, 5) in Tables 6, 8, 10, 16, 18 and 20.

TABLE 1
UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF THE PAIRWISE P-STATISTICS FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITHOUT TIES

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.767	0.767	0.676	0.767
Var2	–	1.000	0.907	0.801	0.907
Var3	–	–	1.000	0.864	0.921
Var4	–	–	–	1.000	0.695
Var5	–	–	–	–	1.000

TABLE 2
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITHOUT TIES

Kruskal–Wallis statistics (Observed value)	5.781
Kruskal–Wallis statistics (Critical value)	9.488
Degree of freedom	4
p-value (one-tailed)	0.216
Significance level	0.05

TABLE 3
PAIRWISE P-VALUES OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITHOUT TIES

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.974	0.505	1.000	1.000
Var2	0.974	1.000	0.984	0.653	0.974
Var3	0.505	0.984	1.000	0.110	0.505
Var4	1.000	0.653	0.110	1.000	0.700
Var5	0.789	1.000	0.921	0.700	1.000

ISSN 2395-8618

TABLE 4
SUMMARY OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITHOUT TIES

Friedman statistics (Observed value)	4.080
Friedman statistics (Critical value)	9.488
Number of degrees of freedom	4
p-value (one-tailed)	0.395
Significance level	0.05

TABLE 5
PAIRWISE P-VALUES OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITHOUT TIES

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	1.000	0.708	0.955	1.000
Var2	1.000	1.000	0.790	0.915	1.000
Var3	0.708	0.790	1.000	0.279	0.708
Var4	0.955	0.915	0.279	1.000	0.955
Var5	1.000	1.000	0.708	0.955	1.000

TABLE 6
UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF THE PAIRWISE P-STATISTICS FOR DISTRIBUTIONS N(0,1), N(0,2), N(0, 3), N(0, 4), AND N(0, 5) WITHOUT TIES

Upper bound	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.732	0.732	0.767	0.749
Var2	–	1.000	0.945	0.957	0.957
Var3	–	–	1.000	0.695	0.741
Var4	–	–	–	1.000	0.921
Var5	–	–	–	–	1.000

The results cited in Table 1 suggest that any sample was not recognized as homogeneous with others. The results cited in Table 2 and Table 3 indicate that all the samples were recognized as homogeneous with others. Thus, we may ascertain that the Kruskal–Wallis fails in this case. The results cited in Table 4 and Table 5 show that the Friedman test fails in this case.

5.2 Scale Shift Hypothesis without Ties

To test a scale hypothesis, we 100 times generated five samples containing ten real numbers that follow the distributions N(0,1), N(0, 2), N(0, 3), N(0, 4), and N(0, 5), and computed p-statistics by the scheme one-vs-others. If the p-statistics between these samples is greater than 0.95, we concluded that they are homogeneous; otherwise, the *k* samples were considered heterogeneous. The average p-statistics was equal to 0.822. The confidence interval was (0.603, 0.934). This interval does not contain 0.95; thus, the samples in whole may be considered heterogeneous. The corresponding results are provided in Tables 6–10. Note that the Klyushin–Petunin test recognized the samples as heterogeneous in 8 cases of 10. Meantime, the Kruskal–Wallis and the Friedman tests failed.

TABLE 7
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITHOUT TIES

Kruskal–Wallis statistics (Observed value)	5.512
Kruskal–Wallis statistics (Critical value)	9.488
Number of degrees of freedom	4
Kruskal–Wallis statistics (Observed value)	0.239
Kruskal–Wallis statistics (Critical value)	0.05

TABLE 8
PAIRWISE P-VALUES OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITHOUT TIES

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.155	1.000	0.700	0.961
Var2	0.155	1.000	0.213	0.984	0.996
Var3	1.000	0.213	1.000	0.700	0.974
Var4	0.700	0.984	0.700	1.000	1.000
Var5	0.961	0.996	0.974	1.000	1.000

TABLE 9
SUMMARY OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITHOUT TIES

Friedman statistics (Observed value)	7.600
Friedman statistics (Critical value)	9.488
Number of degrees of freedom	4
p-value (one-tailed)	0.107
Significance level	0.05

TABLE 10
PAIRWISE P-VALUES OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITHOUT TIES

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.118	0.993	0.279	0.527
Var2	0.118	1.000	0.279	0.993	0.915
Var3	0.993	0.279	1.000	0.527	0.790
Var4	0.279	0.993	0.527	1.000	0.993
Var5	0.527	0.915	0.790	0.993	1.000

5.3 Location shift Hypothesis with Ties

To test a shift hypothesis for samples with ties, we 100 times generated five samples containing ten real numbers with a single tie that follow the distributions N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), and N(0.4, 1), were two sample values were the same, and compute p-statistics by the scheme one-vs-others. The average p-statistics was equal to 0.533. The Wilson confidence interval constructed using the 3s-rule was (0.324, 0.731). This interval does not contain 0.95; thus, the sample in whole may be considered heterogeneous.

The results cited in Table 11 suggest that only 3 cases were recognized as homogeneous with others. The results cited in Tables 12–15 indicate that all the samples were recognized as homogeneous. Thus, we may ascertain that the Kruskal–Wallis test and the Friedman test fail in this case.

ISSN 2395-8618

TABLE 11
UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF THE PAIRWISE P-STATISTICS FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.988	0.986	0.968	0.992
Var2	–	1.000	0.998	0.921	0.997
Var3	–	–	1.000	0.849	0.992
Var4	–	–	–	1.000	0.934
Var5	–	–	–	–	1.000

TABLE 12
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITH SINGLE TIE

Kruskal–Wallis statistics (Observed value)	5,808
Kruskal–Wallis statistics (Critical value)	9,488
Degree of freedom	4
p-value (one-tailed)	0,214
Significance level	0,05

TABLE 13
PAIRWISE P-VALUES OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.984	0.505	1.000	0.745
Var2	0.984	1.000	0.984	0.652	0.998
Var3	0.505	0.984	1.000	0.110	0.921
Var4	1.000	0.652	0.110	1.000	0.700
Var5	0.745	0.998	0.921	0.700	1.000

TABLE 14
SUMMARY OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITH SINGLE TIE

Friedman statistics (Observed value)	4,080
Friedman statistics (Critical value)	9,488
Number of degrees of freedom	4
p-value (one-tailed)	0,395
Significance level	0,05

TABLE 15
PAIRWISE P-VALUES OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0.1, 1), N(0.2, 1), N(0.3, 1), AND N(0.4, 1) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	1.000	0.708	0.955	1.000
Var2	1.000	1.000	0.708	0.955	1.000
Var3	0.708	0.708	1.000	0.279	0.790
Var4	0.955	0.955	0.279	1.000	0.915
Var5	1.000	1.000	0.790	0.915	1.000

TABLE 16
UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF THE PAIRWISE P-STATICTICS FOR DISTRIBUTIONS N(0,1), N(0,2), N(0, 3), N(0, 4), AND N(0, 5) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.921	0.985	0.946	0.921
Var2	–	1.000	0.894	0.993	0.879
Var3	–	–	1.000	0.948	0.934
Var4	–	–	–	1.000	0.946
Var5	–	–	–	–	1.000

TABLE 17
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITH SINGLE TIE

Kruskal–Wallis statistics (Observed value)	6,145
Kruskal–Wallis statistics (Critical value)	9,488
Degree of freedom	4
p-value (one-tailed)	0,189
Significance level	0,05

TABLE 18
PAIRWISE P-VALUES OF THE KRUSKAL–WALLIS TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.155	0.999	0.700	0.921
Var2	0.155	1.000	0.182	0.974	0.999
Var3	0.999	0.182	1.000	0.652	0.863
Var4	0.700	0.974	0.652	1	1.000
Var5	0.921	0.999	0.863	1.000	1.000

TABLE 19
SUMMARY OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITH SINGLE TIE

Friedman statistics (Observed value)	8,400
Friedman statistics (Critical value)	9,488
Number of degrees of freedom	4
p-value (one-tailed)	0,078
Significance level	0,05

TABLE 20
PAIRWISE P-VALUES OF THE FRIEDMAN TEST FOR DISTRIBUTIONS N(0,1), N(0, 2), N(0, 3), N(0, 4), AND N(0, 5) WITH SINGLE TIE

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.118	0.993	0.279	0.527
Var2	0.118	1.000	0.279	0.993	0.915
Var3	0.993	0.279	1.000	0.527	0.790
Var4	0.279	0.993	0.527	1.000	0.993
Var5	0.527	0.915	0.790	0.993	1.000

ISSN 2395-8618

TABLE 21
RESULTS OF ANN AND ARIMA MODELS FOR DELL STOCK INDEX [47]

Forecast error			Forecast error		
Time	ARIMA	ANN	Time	ARIMA	ANN
01.03.2010	0.0302	0.0162	18.03.2010	0.0206	−0.0110
02.03.2010	0.0095	0.0161	19.03.2010	0.0090	−0.0278
03.03.2010	0.0007	0.0110	22.03.2010	−0.0034	−0.0198
04.03.2010	0.0088	0.0000	23.03.2010	0.0019	0.01380
05.03.2010	0.0252	0.0072	24.03.2010	0.0220	−0.0080
08.03.2010	0.0086	0.0093	25.03.2010	−0.006	−0.0229
09.03.2010	0.0183	0.0134	26.03.2010	0.0160	−0.02135
10.03.2010	0.0325	0.0154	29.03.2010	0.0047	−0.0294
11.03.2010	0.0021	0.0007	26.03.2010	0.0160	−0.02135
12.03.2010	−0.0035	−0.0028	29.03.2010	0.0047	−0.0294
15.03.2010	0.0077	−0.0098	30.03.2010	−0.003	−0.0354
16.03.2010	−0.0133	−0.0147	31.03.2010	0.0033	−0.0386
17.03.2010	0.0062	−0.0014			

5.4 Scale Shift Hypothesis with Ties

To test a scale hypothesis, we 100 times generated five samples containing ten real numbers that follow the distributions N(0, 1), N(0, 2), N(0, 3), N(0, 4), and N(0, 5) and contain a single tie, and then compute p-statistics by the scheme one-vs-others. If the p-statistics between these samples was greater than 0.95, we concluded that they are homogeneous; otherwise, the *k* samples were considered as not homogeneous. The average p-statistics was equal to 0.777. The confidence interval constructed using 3s-rule is (0.555, 0.907). The interval does not contain 0.95; thus, the samples in whole may be considered as heterogeneous. The results provided in Table 16–20 show that the Kruskal–Wallis and Friedman tests failed.

5.3 Two-Sample Tests for Dell Stock Index

Let us apply the proposed test to data from [47] (Table 21). The authors of this paper investigated the accuracy of the ARIMA predictive model and artificial neural networks ANN using the Dell stock index collected over 23 days from the New York Stock Exchange during the period from August 17 (1988) to February 25 (2011) and containing 5680 observations. Here ARIMA(p,d,q) denotes the model of the autoregressive integrated moving average, where p is the number of time lags, d is the number of times the data have had past values subtracted, and q is the order of the moving-average model. Hereinafter, Var1–Var5 denotes ARIMA(1,0,0)–ARIMA(5,0,0) models.

Comparing the accuracy indicators, the authors of the paper concluded that ANN neural networks are more accurate than the autoregressive integrated moving average model in terms of relative forecast errors. Using the Klyushin-Petunin test, the Kruskal-Wallis test, and the Friedman test, we can draw different conclusions. According to the Klyushin-Petunin test, the errors of the two considered models are not statistically different since the upper bound of the confidence interval for the p-statistic (0.96) is greater than 0.95. Therefore, these models can be considered statistically equivalent, opposite to Kruskal-Wallis and Friedman tests (Tables 22–23).

TABLE 22
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR DELL STOCK INDEX

Kruskal–Wallis statistics (Observed value)	6.607
Kruskal–Wallis statistics (Critical value)	3.841
Degree of freedom	1
p-value (one-tailed)	0.010
Significance level	0.05

TABLE 23
SUMMARY OF THE FRIEDMAN TEST FOR DELL STOCK INDEX

Friedman statistics (Observed value)	7.348
Friedman statistics (Critical value)	3.841
Number of degrees of freedom	1
p-value (one-tailed)	0.007
Significance level	0.05

TABLE 24
UPPER BOUNDS OF THE CONFIDENCE INTERVALS OF THE PAIRWISE P-STATICTICS FOR MODELS ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), AND ARIMA(5,0,0)

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.953	0.929	0.927	0.921
Var2	–	1.000	0.899	0.892	0.902
Var3	–	–	1.000	0.905	0.899
Var4	–	–		1.000	0.939
Var5	–	–	–	–	1

TABLE 25
SUMMARY OF THE KRUSKAL–WALLIS TEST FOR MODELS ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), AND ARIMA(5,0,0)

Kruskal–Wallis statistics (Observed value)	1.000
Kruskal–Wallis statistics (Critical value)	9.488
Degree of freedom	4
p-value (one-tailed)	0.910
Significance level	0.05

5.4 k-Sample Tests for Dell Stock Index

To extend this experiment on the *k*-sample case, let us consider 5 ARIMA models predicting the Dell Stock Index: ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), and ARIMA(5,0,0) using the training sample published in [47] The results are provided in Tables 24–28. According to the Klyushin-Petunin test, the errors of the five considered ARIMA models are not statistically different in total (homogeneous) since the upper bound of the confidence interval for the

ISSN 2395-8618

p-statistic (0.99) is greater than 0.95. Thus, these models can be considered statistically equivalent. Meantime, the Kruskal-Wallis (Table 25, 26) and Friedman tests (Table 27, 28) lead to opposite conclusions. Pairwise comparisons show that the Klyushin–Petunin test recognizes different samples in almost all the cases.

TABLE 26
PAIRWISE P-VALUES OF THE KRUSKAL–WALLIS TEST FOR MODELS ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), AND ARIMA(5,0,0)

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.155	1.000	0.700	0.961
Var2	0.155	1.000	0.213	0.984	0.996
Var3	1.000	0.213	1.000	0.700	0.974
Var4	0.700	0.984	0.700	1.000	1.000
Var5	0.961	0.996	0.974	1.000	1.000

TABLE 27
SUMMARY OF THE FRIEDMAN TEST FOR MODELS ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), AND ARIMA(5,0,0)

Friedman statistics (Observed value)	18.400
Friedman statistics (Critical value)	9.488
Number of degrees of freedom	4
p-value (one-tailed)	0.001
Significance level	0.05

TABLE 28
PAIRWISE P-VALUES OF THE FRIEDMAN TEST FOR MODELS ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(4,0,0), AND ARIMA(5,0,0)

	Var1	Var2	Var3	Var4	Var5
Var1	1.000	0.744	0.744	0.629	0.044
Var2	0.744	1.000	0.112	1.000	0.508
Var3	0.744	0.112	1.000	0.072	0.001
Var4	0.629	1.000	0.072	1.000	0.629
Var5	0.044	0.508	0.001	0.629	1.000

6. CONCLUSION

Direct comparison of forecasts using accuracy indicators without taking into account their stochastic nature is incorrect. Before comparing the accuracy, it is necessary to test the hypothesis about the identity of the distribution functions of different prediction models. To solve this problem, non-parametric methods are widely used, in particular, the Kruskal–Wallis and Friedman tests. We have proposed a new test that is effective for comparing predictive models. The level of asymptotic significance of this test does not exceed 0.05. The Klyushin-Petunin test is more universal than the Kruskal–Wallis and Friedman tests; it allows ordering pairs of samples and is easy to calculate. The practical usefulness of the proposed test is illustrated by examples. Further work will focus on the study of the theoretical properties of the proposed criterion.

REFERENCES

[1] N.V. Smirnov, "Estimate of difference between empirical distribution curves in two independent samples," *Byull. Mosk. Gos. Univ*., vol. 2, no. 2, pp. 3–14, 1939.

[2] N.H. Kuiper, "Tests concerning random points on a circle," *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A*. vol. 63, pp. 38–47, 1960.

[3] W.G. Dixon, "A criterion for testing the hypothesis that two samples are from the same population," *Ann. Math. Stat*., vol. 11, pp. 199–204, 1940.

[4] A. Wald, J. Wolfowitz, "On a test whether two samples ate from the same population," *Ann. Math. Stat.*, vol. 11, pp. 147–162, 1940.

[5] H.C. Mathisen, "A method of testing the hypothesis that two samples are from the same population," *Ann. Math. Stat.*, vol. 14. pp. 188–194, 1943.

[6] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrika*, vol. 1. pp. 80–83, 1945.

[7] H.B. Mann, D.R. Whitney, "On a test of whether one of the random variables is stochastically larger than other," *Ann. Math. Stat.*, vol. 18. pp. 50–60, 1947.

[8] S.S. Wilks, "A combinatorial test for the problem of two samples from continuous distributions," *Proc. Fourth Berkeley Symp. on Math. Stat. and Prob.*, vol. 1, pp 707–717, 1961.

[9] E.J.G. Pitman, "Significance tests which may be applied to samples from any populations," *J. Royal. Stat. Soc. Ser. A*, vol. 4, pp.119–130, 1937.

[10] E.L. Lehmann, "Consistency and unbiasedness of certain non-parametric tests," *Ann. Math. Stat.*, vol. 22, pp. 165–179, 1947.

[11] M. Rosenblatt, "Limit theorems associated with variants of the von Mises statistic," *Ann. Math. Stat.*, vol. 23, pp. 617–623, 1952.

[12] M. Dwass, "Modified randomization tests for non-parametric hypotheses," *Ann. Math. Stat.*, vol. 28, pp. 181–187, 1957.

[13] M. Fisz, "On a result be M. Rosenblatt concerning the Mises–Smirnov test," *Ann. Math. Stat.*, vol. 31, pp. 427–429, 1960.

[14] G.A. Barnard, "Comment on "The spectral analysis of point processes" by M.S. Bartlett," *J. Royal. Stat. Soc. Ser. B*, vol. 25, p. 294, 1963.

[15] Z.W. Birnbaum, "Computers and unconventional test-statistics," *Reliability and Biometry*, F. Prochan and R.J. Serfling (eds), Philadelphia, PA: SIAM, pp 441–458, 1974.

[16] K.-H. Jockel, "Finite sample properties and asymptotic efficiency of Monte Carlo tests," *Ann. Stat.*, vol. 14, pp. 336–347, 1986.

[17] D.L. Allen, "Hypothesis testing using L1-distance bootstrap," *Am. Stat.*, vol. 51, pp. 145–150, 1997.

[18] B. Efron, R.J. Tibshirani, "An Introduction to the Bootstrap". *Monographs on Statistics and Applied Probability*, Vol. 57, New York: Chapman-Hall, 1993.

[19] J.-M. Dufour, A. Farhat, "Exact non-parametric two-sample homogeneity tests for possibly discrete distributions". Preprint 2001-23, Center for Interuniversity research in Quantitative Economics (CIREQ), California Press, pp. 707–717, 2001.

[20] J. Kiefer, "k-Sample analogues of the Kolmogorov–Smirnov, Cramér–V. Mises test," *Ann. Math. Stat*., vol. 30, pp. 420–447, 1959.

[21] F.W. Scholz, M.A. Stephens, "k-Samples Anderson–Darling test," J. *Am. Stat. Assoc.*, vol. 82, pp. 918–924, 1987.

[22] J. Zhang, Y. Wu, "k-Sample test based on the likelihood ratio," *Comput. Stat. Data. Anal.*, vol. 51, pp. 4682–4691, 2007.

[23] B. Jiang, Y. Chao, J.S. Liu, "Nonparametric K-Sample Tests via Dynamic Slicing," *J. Amer. Stat. Assoc.*, vol. 110, no. 510, pp. 642-653, 2015.

[24] Y. Sang, X. Dang, Z. Yichuan, "Jackknife Empirical Likelihood Approach for K-sample Tests," arXiv:1908.00477v1, 2019.

[25] G.J. Śzékely, M.L. Rizzo, "Energy statistics: A class of statistics based on distances," *J. Stat. Plan. Infer.*, vol. 143, pp. 1249-1272, 2013/

[26] G.J. Śzékely, M.L. Rizzo, "The energy of data," *Ann. Rev. Stat. Appl.*, vol. 4, no. 1, pp. 447-479, 2017.

ISSN 2395-8618

[27] V.A. Fernández, M.D. Jiménez-Gamero, J. Garcìa, “A test for the two-sample problem based on empirical characteristic functions,” *Comput. Stat. Data Anal.*, vol. 52, pp. 3730–3748, 2008.

[28] M. Hušková, S.G. Meintanis, "Tests for the multivariate k-sample problem based on the empirical characteristic function,” J. Nonparametr. Stat.vol. 20, pp. 263–277, 2008.

[29] P. Camblor, J. de Uña-Alvarez, N. Corral, "k-Sample test based on the common area of kernel density estimator,” J. Stat. Plan. Inference, vol. 138, pp. 4006–4020, 2008.

[30] P. Martínez-Camblor, J. de Uña-Alvarez, "Non-parametric k-sample tests: density functions vs. distribution functions,” *Comput. Stat. Data Anal.*, vol. 53, pp. 3344–3357, 2009.

[31] S. Chen, "A new distribution-free k-sample test: Analysis of kernel density functionals,” *Canadian J. Stat.*, vol. 48, No. 2, pp. 167-186, 2020.

[32] T.P. Hettmansperger, “Nonparametric Rank Tests”. *International Encyclopedia of Statistical Science*, M. Lovric (eds) , Berlin, Heidelberg: Springer, 2011.

[33] W.H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis,” *J. Amer. Stat. Assoc*., vol. 47, no. 260, pp. 583–621, 1952.

[34] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance,” *J. Amer. Stat. Assoc.*, vol. 32, no. 200, pp. 675–701, 1937.

[35] M. Friedman, "A correction: The use of ranks to avoid the assumption of normality implicit in the analysis of variance”, *J. Amer. Stat. Assoc.*. vol. 34, no. 205: p. 109, 1939.

[36] M. Friedman, "A comparison of alternative tests of significance for the problem of m rankings,” Ann. Math. Stat., vol. 11, no. 1, pp. 86–92, 1940.

[37] K.M. Wittkowski, "Friedman-Type statistics and consistent multiple comparisons for unbalanced designs with missing data,” J. Amer. Stat. Assoc, vol. 83, no. 404, pp. 1163–1170, 1988.

[38] F.X. Diebold, R.S. Mariano, "Comparing predictive accuracy,” J. Bus. Econ. Stat., vol. 20, no. 1, pp. 134–144, 2002.

[39] B.M. Hill, "Posterior distribution of percentiles: Bayes' theorem for sampling from a population,” *J. Am. Stat. Assoc.*, vol. 63, pp. 677–691, 1968.

[40] D.A. Klyushin, Yu.I. Petunin, "A Nonparametric Test for the Equivalence of Populations Based on a Measure of Proximity of Samples,” *Ukrainian Math. J*., vol. 55, no. 2, pp. 181–198, 2003.

[41] R.I. Andrushkiw, D.A. Klyushin, Yu.I. Petunin, "Proximity measure between samples with repetition factor greater than one,” CAMS Research Report:0708-12, New Jersey Institute of Technology, 2008.

[42] D.F. Vysochanskij, YI Petunin "Justification of the 3σ rule for unimodal distributions,” *Theory of Probability and Mathematical Statistics*, vol. 21, pp. 25–36, 1980.

[43] F. Pukelsheim The Three Sigma Rule. *The American Statistician*, vol. 48(2), pp. 88-91, 1994.

[44] S.A. Matveichuk, Yu.I. Petunin, "Generalization of Bernoulli schemes that arise in order statistics, I,” *Ukrainian. Math. J.,* vol. 42, no. 4, pp. 459–466, 1990.

[45] S.A. Matveichuk, Yu.I. Petunin, "Generalization of Bernoulli schemes that arise in order statistics. II,” *Ukrainian Math. J.*, vol. 43, no. 6, pp. 728–734, 1991.

[46] N. Johnson, S. Kotz, "Some generalizations of Bernoulli and Polya-Eggenberger contagion models,” *Statist. Paper*, vol. 32, pp. 1–17, 1991.

[47] Adebiyi A, Adewumi A, Ayo C (2014) Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. J App Math 2014:1-7.