Kendall, Maurice George

Kendall, Maurice was born in 1907 in Kettering, Northamptonshire, England. He studied Mathematics at St. John's College, Cambridge. After graduation as a Mathematics Wrangler in 1929, he joined the British Civil Service in the Ministry of Agriculture. He was elected a Fellow of the Society in 1934. In 1937, he worked with G. Udny Yule in the revision of his standard statistical textbook, *Introduction to the Theory of Statistics.* He also work on the rank correlation coefficient which bears his name, **Kendall's tau**, which eventually led to a monograph on *Rank Correlation* in 1948.

In 1938 and 1939 he began work, along with Bernard Babington Smith, on the problem of **random number generation**, developing both one of the first early mechanical devices to produce random digits, and formulated a series of tests such as *frequency test, serial test and a poker test,* for statistical randomness in a given set of digits.

During the war he managed to produce volume one of the Advanced Theory of Statistics in 1943 and a second volume in 1946.

In 1957, he published *Multivariate Analysis* and in the same year he also developed, with W.R. Buckland, a *Dictionary of Statistical Terms*.

In 1953, he published The Analytics of Economic Time Series, and in 1961 he left the University of London and took a position as the Managing Director of a consulting company, Scientific Control Systems, and in the same year began a two-year term as President of the Royal Statistical Society.

In 1972, he became Director of the World Fertility Survey, a project sponsored by the International Statistical Institute and the United Nations. He continued this work until 1980, when illness forced him to retire. He was knighted in 1974 for his services to the theory of statistics, and received the Peace Medal of the United Nations in 1980 in recognition for his work on the World Fertility Survey. He was also elected a fellow of the British Academy and received the highest honor of the Royal Statistical Society, the Guy Medal in Gold. At the time of his death in 1983, he was Honorary President of the International Statistical Institute.

Some principal works and articles of Kendall, Maurice George:

1938 (with Babington Smith, B.) Randomness and Random Sampling Numbers. J. Roy. Stat. Soc. **101**:1, 147–166.

1979 (with Stuart, A.) Advanced theory of Statistics. Arnold.

1957 (with Buckland, W.R.) A Dictionary of Statistical Terms. International Statistical Institute, The Hague, Netherland.

1973 Time Series, Griffin, London.

FURTHER READING

- ▶ Kendall rank correlation coefficient
- ▶ Random number generation

Kendall Rank Correlation Coefficient

The Kendall rank correlation coefficient (Kendall τ) is a nonparametric measure of correlation.

HISTORY

This rank correlation coefficient was discussed as far back as the early 20th century by Fechner, G.T. (1897), Lipps, G.F. (1906), and Deuchler, G. (1914).

Kendall, M.G. (1938) not only rediscovered it independently but also studied it using a (nonparametric) approach. His 1970 monograph contains a complete detailed presentation of the theory as well as a biography.

MATHEMATICAL ASPECTS

Consider two random **variables** (X, Y) observed on a **sample** of size n with n pairs of **observations** (X_1, Y_1), (X_2, Y_2), $\ldots$, (X_n, Y_n). An indication of the correlation between X and Y can be obtained by ordering the **values** X_i in increasing order and by counting the number of corresponding values Y_i not satisfying this order.

Q will denote the number of **inversions** among the values of Y that are required to obtain the same (increasing) order as the values of X.

Since there are $\frac{n(n-1)}{2}$ distinct pairs that can be formed, $0 \leq Q \leq \frac{n(n-1)}{2}$; the value 0 is obtained when all the **values** Y_i are already in increasing order, and the value $\frac{n(n-1)}{2}$ is reached when all the values Y_i are in inverse order of X_i, each pair having to be switched to obtain the desired order.

The Kendall rank correlation coefficient, denoted by τ, is defined by:

$$\tau = 1 - \frac{4Q}{n(n-1)} \,.$$

If all the pairs are in increasing order, then:

$$\tau = 1 - \frac{4 \cdot 0}{n(n-1)} = 1 \,.$$

If all the pairs are in reverse order, then:

$$\tau = 1 - \frac{4 \cdot \frac{1}{2} \cdot n(n-1)}{n(n-1)} = -1 \,.$$

An equivalent definition of the Kendall rank coefficient can be given as follows: two **observations** are called concording if the two members of one observation are larger than the respective members of the other observation. For example, (0.9, 1.1) and (1.5, 2.4) are two concording observations because $0.9 < 1.5$ and $1.1 < 2.4$. Two observations are said to be discording if the two members of one observation are in opposite order to the respective members of the other observation. For example, (0.8, 2.6) and (1.3, 2.1) are two discording observations because $0.8 < 1.3$ and $2.6 > 2.1$.

Let N_c and N_d denote the total number of pairs of concording and discording **observations**, respectively.

Two pairs for which $X_i = X_j$ and $Y_i = Y_j$ are neither concording nor discording and are therefore not counted either in N_c or in N_d.
With this notation the Kendall rank coefficient is given by:

$$\tau = \frac{2(N_c - N_d)}{n(n-1)} .$$

Notice that when there are no pairs for which $X_i = X_j$ or $Y_i = Y_j$, the two formulations of τ are exactly the same. In the opposite situation, the **values** given by both formulas can be different.

Hypothesis Test

The Kendall rank correlation coefficient is often used as a statistical test to determine if there is a relation between two **random variables**. The test can be a **two-sided test** or a **one-sided test**. The **hypotheses** are:

A: Two-sided case:

- $\boldsymbol{H_0}$**:** X and Y are mutually independent.
- $\boldsymbol{H_1}$**:** There is either a positive or a negative correlation between X and Y.

There is a positive correlation when the large **values** of X tend to be associated with the large values of Y and the small values of X with the small values of Y. There is a negative correlation when the large values of X tend to be associated with the small values of Y and vice versa.

B: One-sided case:

- $\boldsymbol{H_0}$**:** X and Y are mutually independent.
- $\boldsymbol{H_1}$**:** There is a positive correlation between X and Y.

C: One-sided case:

- $\boldsymbol{H_0}$**:** X and Y are mutually independent.
- $\boldsymbol{H_1}$**:** There is a negative correlation between X and Y.

The statistical test is defined as follows:

$$T = N_c - N_d .$$

Decision Rules

The decision rules are different depending on the **hypotheses** that are made. That is why there are decision rules A, B, and C relative to the previous cases.
Decision rule A
Reject H_0 at the **sigificant level** α if

$$T > t_{n,1-\frac{\alpha}{2}} \quad \text{or} \quad T < t_{n,\frac{\alpha}{2}} ,$$

where t is the **critical value** of the test given by the Kendall table; otherwise accept H_0.
Decision rule B
Reject H_0 at the **sigificant level** α if

$$T > t_{n,1-\alpha} .$$

otherwise accept H_0.
Decision rule C Reject H_0 at the **sigificant level** α if

$$T < t_{n,\alpha} .$$

otherwise accept H_0.
It is also possible to use

$$\tau = 1 - \frac{4Q}{n(n-1)}$$

as a statistical test.
When X and Y are independently distributed in a **population**, the exact distribution of τ has an **expected value** of zero and a **variance** of:

$$\sigma_\tau^2 = \frac{2(2n+5)}{9n(n-1)}$$

and tends very quickly toward a normal distribution, the approximation being good enough for $n \geq 10$.

In this case, to test **independence** at a 5% level, for example, it is enough to verify if τ is located outside the bounds

$$\pm 1.96 \cdot \sigma_\tau$$

and to reject the **independence** hypothesis if that is the case.

DOMAINS AND LIMITATIONS

The Kendall rank correlation coefficient is used as a **hypothesis test** to study the dependence between two **random variables**. It can be considered as a **test of independence**.

As a nonparametric correlation measurement, it can also be used with nominal or ordinal **data**.

A correlation measurement between two **variables** must satisfy the following points:

1. Its **values** are between -1 and $+1$.
2. There is a positive correlation between X and Y if the value of the correlation coefficient is positive; a perfect positive correlation corresponds to a value of $+1$.
3. There is a negative correlation between X and Y if the value of the correlation coefficient is negative; a perfect negative correlation corresponds to a value of -1.
4. There is a null correlation between X and Y when the correlation coefficient is close to zero; one can also say that X and Y are not correlated.

The Kendall rank correlation coefficient has the following advantages:

- The **data** can be nonnumerical **observations** as long as they can be classified according to a determined criterion.
- It is easy to calculate.
- The associated statistical test does not make a basic **hypothesis** based on the shape of the distribution of the **population** from which the **samples** are taken.

The Kendall table gives the theoretical **values** of the **statistic** τ of the Kendall rank correlation coefficient used as a statistical test under the **independence hypothesis** of two **random variables**.

A Kendall table can be found in Kaarsemaker and van Wijngaarden (1953).

Here is a sample of the Kendall table for $n = 4, \ldots, 10$ and $\alpha = 0.01$ and 0.05:

n	$\alpha = 0.01$	$\alpha = 0.05$
4	6	4
5	8	6
6	11	9
7	15	11
8	18	14
9	22	16
10	25	19

EXAMPLES

In this example eight pairs of real twins take intelligence tests. The goal is to see if there is **independence** between the tests of the one who is born first and those of the one who is born second.

The data are given in the table below; the highest scores correspond to the best results.

Pair of twins	First born X_i	Second born Y_i
1	90	88
2	75	79
3	99	98
4	60	66
5	72	64
6	83	83
7	83	88
8	90	98

The pairs are then classified in increasing order for X, and the concording and discording pairs are determined. This gives:

Pair of twins (X_i, Y_i)	Concording pairs	Discording pairs
(60,66)	6	1
(72,64)	6	0
(75,79)	5	0
(83,83)	3	0
(83,88)	2	0
(90,88)	1	0
(90,98)	0	0
(99,98)	0	0
	$N_c = 23$	$N_d = 1$

The Kendall rank correlation coefficient is given by:

$$\tau = \frac{2(N_c - N_d)}{n(n-1)} = \frac{2(23-1)}{8 \cdot 7} = 0.7857\,.$$

Notice that since there are several **observations** for which $X_i = X_j$ or $Y_i = Y_j$, the value of the coefficient given by:

$$\tau = 1 - \frac{4Q}{n(n-1)} = 1 - \frac{4 \cdot 1}{56} = 0.9286$$

is different.

In both cases, we notice a positive correlation between the intelligence tests.

We will now carry out the **hypothesis test**:

H_0: There is **independence** between the intelligence tests of a pair of twins.

H_1: There is a positive correlation between the intelligence tests.

We chose a **significant level** of $\alpha = 0.05$. Since we are in case B, H_0 is rejected if

$$T > t_{8,0.95}\,,$$

where $T = N_c - N_d$ and $t_{8,0.95}$ is the **value** of the Kendall table. Since $T = 22$ and $t_{8,0.95} = 14$, H_0 is rejected.

We can then conclude that there is a positive correlation between the results of the intelligence tests of a pair of twins.

FURTHER READING

- ► **Hypothesis testing**
- ► **Nonparametric test**
- ► **Test of independence**

REFERENCES

Deuchler, G.: Über die Methoden der Korrelationsrechnung in der Pädagogik und Psychologie. Zeitung für Pädagogische Psychologie und Experimentelle Pädagogik, **15**, 114–131, 145–159, 229–242 (1914)

Fechner, G.T.: Kollektivmasslehre. W. Engelmann, Leipzig (1897)

Kaarsemaker, L., van Wijngaarden, A.: Tables for use in rank correlation. Stat. Neerland. **7**, 53 (1953)

Kendall, M.G.: A new measure of rank correlation. Biometrika **30**, 81–93 (1938)

Kendall, M.G.: Rank Correlation Methods. Griffin, London (1948)

Lipps, G.F.: Die Psychischen Massmethoden. F. Vieweg und Sohn, Braunschweig, Germany (1906)

Kiefer, Jack Carl

Kiefer, Jack Carl was born in Cincinnati, Ohio in 1924. He entered the Massachusetts Institute of Technology in 1942, but after 1 year of studying engineering and economics he left to take on war-related work during World War II. His master's thesis, *Sequential Determination of the Maximum of a Function*, was supervised by Harold Freeman. It has been the basis for his paper "Sequential minimax search for a maximum" which appeared in 1953 in the "Proceedings of the American Mathematical Society". In 1948 he went to the Department of Mathematical Statistics at Columbia University, where Abraham Wald was preeminent in a department that included Ted Anderson, Henry Scheffé, and Jack Wolfowitz. He wrote his doctoral thesis in decision theory under Wolfowitz and went to Cornell University in 1951 with Wolfowitz. In 1973 Kiefer was elected the first Horace White Professor at Cornell University, a position he held until 1979, when he retired and joined the faculty at the University of California at Berkeley. He died at the age of 57 in 1981.

Kiefer's research area was the design of experiments. Most of his 100 publications dealt with that topic. He also wrote papers on topics in mathematical statistics including decision theory, stochastic approximation, queuing theory, nonparametric inference, estimation, sequential analysis, and conditional inference.

Kiefer was a fellow of the Institute of Mathematical Statistics and the American Statistical Association and president of the Institute of Mathematical Statistics (1969–1970). He was elected to the American Academy of Arts and Sciences in 1972 and to the National Academy of Sciences (USA) in 1975.

Selected works and publications of Jack Carl Kiefer:

1987 Introduction to Statistical Inference. Springer, Berlin Heidelberg New York

FURTHER READING

▸ **Design of experiments**

REFERENCES

Kiefer, J.: Optimum experimental designs. J. Roy. Stat. Soc. Ser. B **21**, 272–319 (1959)

L.D. Brown, I. Olkin, J. Sacks and H.P. Wynn (eds.): Jack Karl Kiefer, Collected Papers. I: Statistical inference and probability (1951–1963). New York (1985)

L.D. Brown, I. Olkin, J. Sacks and H.P. Wynn (eds.): Jack Karl Kiefer, Collected Papers. II: Statistical inference and probability (1964–1984). New York (1985)

L.D. Brown, I. Olkin, J. Sacks and H.P. Wynn (eds.): Jack Karl Kiefer, Collected Papers. III: Design of experiments. New York (1985)

Kolmogorov, Andrei Nikolaevich

Born in Tambov, Russia in 1903, Kolmogorov, Andrei Nikolaevich is one of the founders of modern probability. In 1920, he entered Moscow State University and studied mathematics, history, and metallurgy. In 1925, he published his first article in probability on the inequalities of the partial sums of random variables, which became the principal reference in the field of stochastic processes. He received his doctorate in 1929 and published 18 articles on the law of large numbers as well as on intuitive logic. He was named professor at Moscow State University in 1931. In 1933, he published his monograph on probability theory.

In 1939 he was elected member of the Academy of Sciences of the USSR. He received the Lenin Prize in 1965 and the Order of Lenin on six different occasions, as well as the Lobachevsky Prize in 1987. He was elected member of many other foreign academies including the Romanian Academy of Sciences (1956), the Royal Statistical Society of London (1956), the Leopoldina Academy of Germany (1959), the American Academy of Arts and Sciences (1959), the London Mathematical Society (1959), the American Philosophical Society (1961), the Indian Institute of Statistics(1962), the Holland Academy of Sciences (1963), the Royal Society of London (1964), the National Academy of the United States (1967), and the Académie Française des Sciences (1968).

Selected principal works of Kolmogorov, Andrei Nikolaevich:

1933 Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin Heidelberg New York.

1933 Sulla determinazione empirica di una lege di distribuzione. Giornale dell'Instituto Italiano degli Attuari, 4, 83–91 (6.1).

1941 Local structure of turbulence in incompressible fluids with very high Reynolds number. Dan SSSR, 30, 229.

1941 Dissipation of energy in locally isotropic turbulence. Dokl. Akad. Nauk. SSSR, 32, 16–18.

1958 (with Uspenskii, V.A.) K opredeleniyu algoritma. (Toward the definition of an algorithm). Uspekhi Matematicheskikh Nauk 13(4):3–28, American Mathematical Society Translations Series 2(29):217–245, 1963.

1961 (with Fomin, S.V.) Measure, Lebesgue integrals and Hilbert space. Natascha Artin Brunswick and Alan Jeffrey. Academic, New York.

1963 On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. Doklady Akademii Nauk SSR, 114, 953–956, 1957. English translation. Mathematical Society Transactions, 28, 55–59.

1965 Three approaches to the quantitative definition of information. Problems of Information Transmission, 1, 1–17. Translation of Problemy peredachi informatsii 1(1), 3–11 (1965).

1987 (with Uspenskii, V.A.) Algorithms and randomness. Teoria veroyatnostey i ee primeneniya (Probability theory and its applications), 3(32):389–412.

FURTHER READING

► **Kolmogorov–Smirnov test**

Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov test is a nonparametric **goodness-of-fit test** and is used to determine wether two distributions differ, or whether an underlying probability distribution differes from a hypothesized distribution. It is used when we have two samples coming from two populations that can be different. Unlike the **Mann–Whitney test** and the **Wilcoxon test** where the goal is to detect the difference between two means or medians, the Kolmogorov–Smirnov test has the advantage of considering the distribution functions collectively. The Kolmogorov–Smirnov test can also be used as a **goodness-of-fit test**. In this case, we have only one random sample obtained from a **population** where the distribution function is specific and known.

HISTORY

The **goodness-of-fit test** for a **sample** was invented by Andrey Nikolaevich Kolmogorov (1933).

The Kolmogorov–Smirnov test for two samples was invented by Vladimir Ivanovich Smirnov (1939).

In Massey (1952) we find a Smirnov table for the Kolmogorov–Smirnov test for two samples, and in Miller (1956) we find a Kolmogorov table for the goodness-of-fit test.

MATHEMATICAL ASPECTS

Consider two independent random samples: $(X_1, X_2, \ldots, X_n)$, a sample of size n coming from a **population** 1, and $(Y_1, Y_2, \ldots, Y_m)$, a sample of dimension m coming from a population 2. We denote by, respectively, $F(x)$ and $G(x)$ their unknown **distribution functions**.

Hypotheses

The hypotheses to test are as follows:

A: Two-sided case:

$\boldsymbol{H_0}$: $F(x) = G(x)$ for each x

$\boldsymbol{H_1}$: $F(x) \neq G(x)$ or at least one **value** of x

B: One-sided case:

$\boldsymbol{H_0}$: $F(x) \leq G(x)$ for each x

$\boldsymbol{H_1}$: $F(x) > G(x)$ for at least one value of x

C: One-sided case:

$\boldsymbol{H_0}$: $F(x) \geq G(x)$ for each x

$\boldsymbol{H_1}$: $F(x) < G(x)$ for at least one value of x

In case A, we make the hypothesis that there is no difference between the distribution functions of these two populations. Both populations can then be seen as one population.

In case B, we make the hypothesis that the distribution function of population 1 is smaller than those of population 2. We sometimes say that, generally, X tends to be smaller than Y.

In case C, we make the hypothesis that X is greater than Y.

We denote by $H_1(x)$ the empirical distribution function of the **sample** $(X_1, X_2, \ldots, X_n)$ and by $H_2(x)$ the empirical distribution function of the sample $(Y_1, Y_2, \ldots, Y_m)$. The statistical test are defined as follows:

A: Two-tail case

The statistical test T_1 is defined as the greatest vertical **distance** between two empirical distribution functions:

$$T_1 = \sup_x |H_1(x) - H_2(x)| \, .$$

B: One-tail case

The statistical test T_2 is defined as the greatest vertical distance when $H_1(x)$ is greater than $H_2(x)$:

$$T_2 = \sup_x [H_1(x) - H_2(x)] \, .$$

C: One-tail case

The statistical test T_3 is defined as the greatest vertical distance when $H_2(x)$ is greater than $H_1(x)$:

$$T_3 = \sup_x [H_2(x) - H_1(x)] \, .$$

Decision Rule

We reject H_0 at the **significance level** α if the appropriate statistical test (T_1, T_2, or T_3) is greater than the **value** of the Smirnov table having for parameters n, m, and $1-\alpha$, which we denote by $t_{n,m,1-\alpha}$, that is, if

$$T_1 (\text{or } T_2 \text{or } T_3) > t_{n,m,1-\alpha} .$$

If we want to test the goodness of fit of an unknown distribution function $F(x)$ of a random sample from a population with a specific and known distribution function $F_o(x)$, then the hypotheses will be the same as those for testing two samples, except that $F(x)$ and $G(x)$ are replaced by $F(x)$ and $F_o(x)$.

If $H(x)$ is the empirical distribution function of a random sample, then the statistical tests T_1, T_2, and T_3 are defined as follows:

$$T_1 = \sup_x |F_o(x) - H(x)| ,$$

$$T_2 = \sup_x [F_o(x) - H(x)] ,$$

$$T_3 = \sup_x [H(x) - F_o(x)] .$$

The decision rule is as follows: reject H_0 at the significance level α if T_1 (or T_2 or T_3) is greater than the value of the Kolmogorov table having for parameters n and $1-\alpha$, which we denote by $t_{n,1-\alpha}$, that is, if

$$T_1 (\text{or } T_2 \text{or } T_3) > t_{n,1-\alpha} .$$

DOMAINS AND LIMITATIONS

To perform the Kolmogorov–Smirnov test, the following must be observed:

1. Both samples must be taken randomly from their respective **populations**.
2. There must be mutual **independence** between two samples.
3. The measure scale must be at least ordinal.
4. To perform an exact test, the random variables must be continuous; otherwise the test is less precise.

EXAMPLES

The first example treats the Kolmogorov–Smirnov test for two samples and the second one for the **goodness-of-fit test**.

In a class, we count 25 pupils: 15 boys and 10 girls. We perform a test of mental calculations to see if the boys tend to be better than the girls in this domain.

The data are presented in the following table; the highest scores correspond to the results of the test.

Boys (X_i)		Girls (Y_i)	
19.8	17.5	17.7	14.1
12.3	17.9	7.1	23.6
10.6	21.1	21.0	11.1
11.3	16.4	10.7	20.3
13.3	7.7	8.6	15.7
14.0	15.2		
9.2	16.0		
15.6			

We test the **hypothesis** according to which the distributions of the results of the girls and those of the boys are identical. This means that the **population** from which the sample of X is taken has the same **distribution function** as the population from which the sample of Y is taken. Hence the **null hypothesis**:

$$H_0 : F(x) = G(x) \quad \text{for each } x .$$

If the two-tail case is applied here, we calculate:

$$T_1 = \sup_x |H_1(x) - H_2(x)| ,$$

where $H_1(x)$ and $H_2(x)$ are the empirical distribution functions of the samples

$(X_1, X_2, \ldots, X_{15})$ and $(Y_1, Y_2, \ldots, Y_{10})$, respectively. In the following table, we have classed the observations of two samples in increasing order to simplify the calculations of $H_1(x) - H_2(x)$.

X_i	Y_i	$H_1(x) - H_2(x)$
	7.1	$0 - 1/10 = -0.1$
7.7		$1/15 - 1/10 = -0.0333$
	8.6	$1/15 - 2/10 = -0.1333$
9.2		$2/15 - 2/10 = -0.0667$
10.6		$3/15 - 2/10 = 0$
	10.7	$3/15 - 3/10 = -0.1$
	11.1	$3/15 - 4/10 = -0.2$
11.3		$4/15 - 4/10 = -0.1333$
12.3		$5/15 - 4/10 = -0.0667$
13.3		$6/15 - 4/10 = 0$
14.0		$7/15 - 4/10 = 0.0667$
	14.1	$7/15 - 5/10 = -0.0333$
15.2		$8/15 - 5/10 = 0.0333$
15.6		$9/15 - 5/10 = 0.1$
	15.7	$9/15 - 6/10 = 0$
16.0		$10/15 - 6/10 = 0.0667$
16.4		$11/15 - 6/10 = 0.1333$
17.5		$12/15 - 6/10 = 0.2$
	17.7	$12/15 - 7/10 = 0.1$
17.9		$13/15 - 7/10 = 0.1667$
19.8		$14/15 - 7/10 = 0.2333$
	20.3	$14/15 - 8/10 = 0.1333$
	21.0	$14/15 - 9/10 = 0.0333$
21.1		$1 - 9/10 = 0.1$
	23.6	$1 - 1 = 0$

We have then:

$$T_1 = \sup_x |H_1(x) - H_2(x)| = 0.2333 .$$

The **value** of the Smirnov table for $n = 15$, $m = 10$, and $1 - \alpha = 0.95$ equals $t_{15,10,0.95} = 0.5$.

Thus $T_1 = 0.2333 < t_{15,10,0.95} = 0.5$, and H_0 cannot be rejected. This means that there is no significant difference in the level of mental calculations of girls and boys.

Consider the following random sample of dimension 10: $X_1 = 0.695$, $X_2 = 0.937$, $X_3 = 0.134$, $X_4 = 0.222$, $X_5 = 0.239$, $X_6 = 0.763$, $X_7 = 0.980$, $X_8 = 0.322$, $X_9 = 0.523$, $X_{10} = 0.578$.

We want to verify by the Kolmogorov–Smirnov test if this sample comes from a **uniform distribution**. The distribution function of the uniform distribution is given by:

$$F_o(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{otherwise} . \end{cases} .$$

The **null hypothesis** H_0 is then as follows, where $F(x)$ is the unknown distribution function of the population associated to the sample:

$$H_0 : F(x) = F_o(x) \quad \text{for each } x .$$

If the two-tail case is applied, we calculate:

$$T_1 = \sup_x |F_o(x) - H(x)| ,$$

where $H(x)$ is the empirical distribution function of the sample $(X_1, X_2, \ldots, X_{10})$.

In the following table, we class the 10 observations in increasing order to simplify the calculation of $F_0(x) - H(x)$.

X_i	$F_o(x)$	$H(x)$	$F_o(x) - H(x)$
0.134	0.134	0.1	$0.134 - 0.1 = 0.034$
0.222	0.222	0.2	$0.222 - 0.2 = 0.022$
0.239	0.239	0.3	$0.239 - 0.3 = -0.061$
0.322	0.322	0.4	$0.322 - 0.4 = -0.078$
0.523	0.523	0.5	$0.523 - 0.5 = 0.023$
0.578	0.578	0.6	$0.578 - 0.6 = -0.022$
0.695	0.695	0.7	$0.695 - 0.7 = -0.005$
0.763	0.763	0.8	$0.763 - 0.8 = -0.037$
0.937	0.937	0.9	$0.937 - 0.9 = 0.037$
0.980	0.980	1.0	$0.980 - 1.0 = -0.020$

We obtain then:

$$T_1 = \sup_x |F_o(x) - H(x)| = 0.078\,.$$

The value of the Kolmogorov table for $n = 10$ and $1 - \alpha = 0.95$ is $t_{10,0.95} = 0.409$.
If T_1 is smaller than $t_{10,0.95}$ $(0.078 < 0.409)$, then H_0 cannot be rejected. That means that the random sample could come from a uniformly distributed population.

FURTHER READING

- ▸ **Goodness of fit test**
- ▸ **Hypothesis testing**
- ▸ **Nonparametric test**

REFERENCE

Kolmogorov, A.N.: Sulla determinazione empirica di una legge di distribuzione. Giornale dell'Instituto Italiano degli Attuari **4**, 83–91 (6.1) (1933)

Massey, F.J.: Distribution table for the deviation between two sample cumulatives. Ann. Math. Stat. **23**, 435–441 (1952)

Miller, L.H.: Table of percentage points of Kolmogorov statistics. J. Am. Stat. Assoc. **31**, 111–121 (1956)

Smirnov, N.V.: Estimate of deviation between empirical distribution functions in two independent samples. (Russian). Bull. Moscow Univ. **2**(2), 3–16 (6.1, 6.2) (1939)

Smirnov, N.V.: Table for estimating the goodness of fit of empirical distributions. Ann. Math. Stat. **19**, 279–281 (6.1) (1948)

Kruskal-Wallis Table

The Kruskal–Wallis table gives the theoretical values of the **statistic** H of the **Kruskal–Wallis test** under the **hypothesis** that there is no difference among the k $(k \geq 2)$ populations that we want to compare.

HISTORY

See **Kruskal–Wallis test**.

MATHEMATICAL ASPECTS

Let k be the number of samples of probably different sizes $n_1, n_2, \ldots, n_k$. We designate by N the total number of observations:

$$N = \sum_{i=1}^{k} n_i\,.$$

We class the N observations in increasing order without taking into account which samples they belong to. We then give rank 1 to the smallest **value**, rank 2 to the next greatest value, and so on until rank N, which is given to the greatest value.
We denote by R_i the sum of the ranks given to the observations of sample i:

$$R_i = \sum_{j=1}^{n_i} R\left(X_{ij}\right)\,, \quad i = 1, 2, \ldots, k\,,$$

where X_{ij} represents observation j of sample i and $R(X_{ij})$ the corresponding rank. When many observations are identical and of the same rank, we give them a mean rank (see **Kruskal–Wallis test**). If there are no mean ranks, the statistical test is defined in the following way:

$$H = \left(\frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i}\right) - 3\,(N+1)\,.$$

On what to do if there are mean ranks, see **Kruskal–Wallis test**.
The Kruskal–Wallis table gives the values of the **statistic** H of the Kruskal–Wallis test in the case of three samples, for different values of n_1, n_2, and n_3 (with $n_1, n_2, n_3 \leq 5$).

DOMAINS AND LIMITATIONS

The Kruskal–Wallis table is used for nonparametric tests that use ranks and particularly for tests with the same name.

When the number of samples i is greater than 3, we can make an approximation of the **value** of the Kruskal–Wallis table by the **chi-square table** with $k-1$ degrees of freedom.

EXAMPLES

See Appendix D.

For an example of the use of the Kruskal–Wallis table, see **Kruskal–Wallis test**.

FURTHER READING

- ▶ **Chi-square table**
- ▶ **Kruskal-Wallis test**
- ▶ **Statistical table**

REFERENCES

Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. **47**, 583–621 and errata, ibid. **48**, 907–911 (1952)

Kruskal-Wallis Test

The Kruskal–Wallis test is a **nonparametric test** that has as its goal to determine if all k populations are identical or if at least one of the populations tends to give observations that are different from those of other populations.

The test is used when we have k samples, with $k \geq 2$, coming from k populations that can be different.

HISTORY

The Kruskal–Wallis test was developed in 1952 by Kruskal, W.H. and Wallis, W.A.

MATHEMATICAL ASPECTS

The data are represented in k samples. We designate by n_i the dimension of the **sample** i, for $i = 1, \ldots, k$, and by N the total number of observations:

$$N = \sum_{i=1}^{k} n_i \,.$$

We class the N observations in increasing order without taking into account whether or not they belong to the same samples. We then give rank 1 to the smallest **value**, rank 2 to the next greatest value, and so on until N, which is given to the greatest value.

Let X_{ij} be the jth observation of sample i, and set $i = 1, \ldots, k$ and $j = 1, \ldots, n_i$; we then denote the rank given to X_{ij} by $R\left(X_{ij}\right)$.

If many observations have the same value, we give them a mean rank. The sum of the ranks given to the observations of sample i is denoted by R_i, and we have:

$$R_i = \sum_{j=1}^{n_i} R\left(X_{ij}\right), \quad i = 1, \ldots, k \,.$$

If there are no mean ranks (or if there is a limited number of them), then the statistical test is defined as follows:

$$H = \left(\frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} \right) - 3\,(N+1) \,.$$

If, on the contrary, there are many mean ranks, it is necessary to make a correction and to calculate:

$$\widetilde{H} = \frac{H}{1 - \dfrac{\sum_{i=1}^{g}\left(t_i^3 - t_i\right)}{N^3 - N}} \,,$$

where g is the number of groups of mean ranks and t_i the dimension of ith such group.

Hypotheses

The goal of the Kruskal–Wallis test is to determine if all the populations are identical or if at least one of the populations tends to give observations different from other populations. The hypotheses are as follows:

***H_0*:** There is no difference among the k populations.

***H_1*:** At least one of the populations differs from the other populations.

Decision Rule

If there are 3 samples, each having a dimension smaller or equal to 5, and if there are no mean ranks (that is, if H is calculated), then we use the **Kruskal–Wallis table** to test H_0. The decision rule is the following: We reject the **null hypothesis** H_0 at the **significance level** α if T is greater than the value of the table with parameters n_i, $k-1$, and $1-\alpha$, denoted $h_{n_1,n_2,n_3,1-\alpha}$, and if there is no available exact table or if there are mean ranks, we can make an approximation of the value of the Kruskal–Wallis table by the distribution of the chi-square with $k-1$ degrees of freedom (**chi-square distribution**), that is, if:

$$H > h_{n_1,n_2,n_3,1-\alpha} \quad (\textbf{Kruskal–Wallis table})$$

$$\text{or } H > \chi^2_{k-1,1-\alpha} \quad (\textbf{chi-square table})\,.$$

The corresponding decision rule is based on $\widetilde{H} > \chi^2_{k-1,1-\alpha}$ (chi-quare table).

DOMAINS AND LIMITATIONS

The following rules should be respected to make the Kruskal–Wallis test:

1. All the samples must be random samples taken from their respective **populations**.
2. In addition to the **independence** inside each sample, there must be mutual independence among the different samples.
3. The scale of measure must be at least ordinal.

If the Kruskal–Wallis test makes us reject the **null hypothesis** H_0, we can use the **Wilcoxon test** for all the samples taken in pairs to determine which pairs of populations tend to be different.

EXAMPLES

We cook potatoes in 4 different oils. We want to verify if the quantity of fat absorbed by potatoes depends on the type of oil used. We conduct 5 different experiments with oil 1, 6 with oil 2, 4 with oil 3, and 5 with oil 4, and we obtain the following data:

Type of oil			
1	2	3	4
64	78	75	55
72	91	93	66
68	97	78	49
77	82	71	64
56	85		70
	77		

In this example, the number of samples equals 4 ($k = 4$) with the following respective dimensions:

$$n_1 = 5\,,$$
$$n_2 = 6\,,$$
$$n_3 = 4\,,$$
$$n_4 = 5\,.$$

The number of observations equals:

$$N = 5 + 6 + 4 + 5 = 20\,.$$

We class the observations in increasing order and give them a rank from 1 to 20 taking into account the mean ranks. We obtain the following table with the rank of the observations in parentheses and at the end of each sample the sum of the ranks given to the corresponding sample:

Type of oil			
1	2	3	4
64 (4.5)	78 (14.5)	75 (11)	55 (2)
72 (10)	91 (18)	93 (19)	66 (6)
68 (7)	97 (20)	78 (14.5)	49 (1)
77 (12.5)	82 (16)	71 (9)	64 (4.5)
56 (3)	85 (17)		70 (8)
	77 (12.5)		
$R_1 = 37$	$R_2 = 98$	$R_3 = 53.5$	$R_4 = 21.5$

We calculate:

$$H = \left(\frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} \right) - 3(N+1)$$
$$= \frac{12}{20(20+1)} \cdot \left(\frac{37^2}{5} + \frac{98^2}{6} + \frac{53.5^2}{4} + \frac{21.5^2}{5} \right) - 3(20+1)$$
$$= 13.64 .$$

If we make an adjustment to take into account the mean ranks, we get:

$$\widetilde{H} = \frac{H}{1 - \frac{\sum_{i=1}^{g} \left(t_i^3 - t_i\right)}{N^3 - N}}$$
$$= \frac{13.64}{1 - \frac{8-2+8-2+8-2}{20^3 - 20}}$$
$$= 13.67 .$$

We see that the difference between H and $\widetilde{H}$ is minimal.

The hypotheses are as follows:

H_0: There is no difference between the four oils.

H_1: At least one of the oils differs from the others.

The decision rule is as follows: Reject the **null hypothesis** H_0 at the **significance level** α if

$$H > \chi^2_{k-1,1-\alpha} ,$$

where $\chi^2_{k-1,1-\alpha}$ is the **value** of the **chi-square table** at level $1-\alpha$ and $k-1=3$ degrees of freedom.

If we choose $\alpha = 0.05$, then the value of $\chi^2_{3,0.95}$ is 7.81, and if H is greater than $\chi^2_{3,0.95}$ ($13.64 > 7.81$), then we reject the hypothesis H_0.

If H_0 is rejected, we can use the procedure of multiple comparisons to see which pairs of oils are different.

FURTHER READING

- Hypothesis testing
- Nonparametric test
- Wilcoxon test

REFERENCE

Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. **47**, 583–621 and errata, ibid. **48**, 907–911 (1952)