



# A Web-Based Software for the Calculation of Theoretical Probability Distributions

Fatma Hilal Yagin<sup>ID</sup>, Emek Guldogan<sup>ID</sup>, Cemil Colak<sup>ID</sup>

Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Turkey. (e-mail: {hilal.yagin, emek.guldogan, cemil.colak}@inonu.edu.tr.

## ARTICLE INFO

**Received:** Feb.,11.2021  
**Revised:** April,19.2021  
**Accepted:** May.,27.2021

### Keywords:

Discrete Probability Distributions  
Continuous Probability Distributions  
Web-Based Software  
Python.

**Corresponding author:** Fatma Hilal Yagin

✉ hilal.yagin@inonu.edu.tr  
☎ +90 422 3410660

**ISSN:** 2548-0650

**DOI:** <https://doi.org/10.52876/jcs.878742>

## ABSTRACT

**Aim:** The aim of this study is to develop a public web-based theoretical probability distributions software (KODY) that can calculate probabilities for discrete and continuous distributions.

**Materials and Methods:** The Discrete Uniform, Bernoulli, Binomial, Multinomial, Poisson, Geometric, Negative Binomial, Hypergeometric and Zeta (Zipf) distributions from the discrete distributions are explained. Among the continuous distributions, The Continuous Uniform, Beta, Normal, Log-Normal, Exponential, Gamma, Weibull, Rayleigh, Logistics, Pareto, Laplace, Cauchy and Erlang distributions are elucidated. Illustrative examples are presented on hypothetical medical data. The software was developed using the MATH and DASH libraries of the Python programming language.

**Results:** When making statistical analysis, the feature of the distribution is essential. Because the descriptive and analytical statistical methods to be applied to data with different distributions are also different. Probability distributions of variables are important in the effectiveness of these methods. For this reason, it is an essential step for researchers to determine the probability distributions of their data before starting their studies. It is thought that the software developed in this study will enable researchers to make the necessary calculations in probabilistic estimates regarding the theoretical probability distributions. The developed software can be accessed at <http://biostatapps.inonu.edu.tr/KODY/>.

**Conclusion:** The open access web-based software with Turkish/English language options may guide and contribute to researchers in probabilistic estimation processes regarding theoretical distributions. In the later stages of this study, it is foreseen to develop simulation processes based on each probability distribution.

## 1. INTRODUCTION

THE science of statistics, which developed to determine the factors affecting complex events that occur as a result of the rapid developments in the fields of science and technology, to predict uncertainties and to eliminate the risks caused by uncertainty, has led to the development of probability theory [1].

A random variable is a variable whose discrete or continuous values are random, and a probability distribution for these values can be defined. In other words, this variable shows all possible measurable values in the sample space. Random variables are frequently used in probability calculations in physics, chemistry, engineering, biology, and medical sciences. Properties of probability distributions of random variables in statistical analysis and it is fundamental in choosing the analysis method and interpreting the results.

Since these analyzes are based on a specific theoretical distribution, the variable (s) we use in the analysis must fit this theoretical distribution. Theoretical distributions are also a probability distribution [2].

The most important task in applying the probability concept is to find the probability function suitable. Today, many events are different from each other and have a complex structure. Since it would be difficult for a researcher to search for a different probability function for each different event in terms of time, information and possibility, probability functions have been developed in specific patterns related to events that provide certain properties. These probability distributions are divided into two as discrete and continuous probability distributions. Discrete probability distributions describe the possible distributions in the occurrence of a random and countable event. If the element of the set consisting of the possible outcomes of an event takes a value

in a continuous range, this random variable is called a "continuous random variable". Furthermore, the probability distribution of this random variable is called the continuous probability distribution. The discrete sampling space points in the discrete probability function are based on the continuous probability function's continuous sampling space [3].

In the field of health, probability distributions are frequently used in daily life. For example, the estimation of how many patients will apply to the hospital one day, how long the disease will be treated with a particular drug, the probability of a patient being discharged, the probability of dying if the patient is not bypassed, etc. situations are used in daily life.

The aim of this study is to develop a public web-based theoretical probability distributions software (KODY) that can calculate probabilities for comprehensive discrete and continuous distributions. There are calculations for The Discrete Uniform, Bernoulli, Binomial, Polynomial, Poisson, Geometric, Negative Binomial, Hypergeometric and Zeta (Zipf) distributions in the section of discrete probability distributions in software. In the section of continuous probability distributions, there are The Continuous Uniform, Beta, Normal, Log-Normal, Exponential, Gamma, Weibull, Rayleigh, Logistics, Pareto, Laplace, Cauchy and Erlang distributions. The developed software can be publicly accessed at <http://biostatapps.inonu.edu.tr/KODY/>.

## 2. MATERIALS AND METHODS

### 2.1.1.1. Hypothetical examples

The calculation of illustrative examples for the usability and applications of theoretical probability distributions in Medicine and other health fields was carried out with the software proposed in this study.

### 2.1.1.2. Theoretical Distributions

Any theoretical distribution is a mathematical function defined as  $y=f(x)$ .  $y$  indicates the frequency at which  $x$  values occur.  $f(x)$  is also called the density function [3].

$f(x)$  has the following properties in case the  $x$  variable is continuous.

$$0 \leq f(x) \leq 1 \quad -\infty \leq x \leq \infty \quad \int_{-\infty}^{\infty} f(x)dx = 1$$

$f(x)$ , If the variable  $x$  is discrete, it has the following properties.

$$0 \leq f(x) \leq 1 \quad a \leq x \leq b \quad \sum_a^b f(x) = 1$$

Parameters and related explanations of the software's theoretical probability distributions are given in Tables 1 and 2.

### 2.1.1.3. Discrete Probability Distributions

The discrete probability function contains a countable infinite number or finite discrete results obtained from the sample space. The discrete probability distributions used in the study are explained below.

### 2.1.1.4. The Discrete Uniform Distribution

It is one of the simplest of discrete uniform distributions and assumes that each of the random variables has equal probability. If  $X$  is a random variable  $x_1, x_2, x_3, \dots, x_k$ . If it is

assumed that it has equal probabilities with its values, then a discrete uniform distribution is mentioned [4].

### 2.1.1.5. Bernoulli Distribution

When interested in two results, which are called success and failure in an experiment, this experiment is called the (two-result) Bernoulli experiment or the Bernoulli test. For example; Bernoulli distribution is mentioned in the event that has two-state outcomes such as successful-unsuccessful, intact-broken, positive-negative, dead-alive, patient-healthy. This distribution has repeatability under the same conditions and success does not change from experiment to experiment [5].

### 2.1.1.6. Binomial Distribution

An experiment that occurs when a Bernoulli experiment with a probability of success is repeated  $n$  times (independently) under the same conditions is called the binomial test. The binomial test must fulfill the following conditions:

- The number of subjects in the sample or the number of trials must be constant during the experiment.
- Trials are independent from each other.
- Each attempt has two possible outcomes (desired and undesired event).
- The event probability  $p$  of interest in each trial is constant. Therefore, the probability of adverse events is invariant at  $q = 1-p$ .

The binomial distribution is a discrete probability distribution. When a random variable has a binomial distribution, it is denoted by  $X \sim b(n, p)$ . This distribution is frequently used in sampling and quality control areas in statistics [6].

### 2.1.1.7. Multinomial Distribution

The Multinomial Distribution is a generalized form of the binomial distribution. If there are more than two results in each trial, if the probabilities of similar results in each experiment are equal and all trials are independent from each other, then a multinomial distribution is mentioned [7].

### 2.1.1.8. Poisson Distribution

This distribution is used in situations where the probability of an event occurring in a given time interval is minimal. The time interval given can be in terms of minutes, days, weeks or years. For example, Poisson distribution can be used for natural disasters or rare diseases that occur in a certain year. If the number of subjects,  $n$ , is large and the probability of success  $p$  is too small, the binomial distribution approaches the Poisson distribution. In general, when  $np \leq 5$ , the Poisson distribution can be used instead of the binomial distribution. In addition, there is a condition that  $n$  should be greater than 20 [8].

### 2.1.1.9. Geometric Distribution

$X$ , the number of experiments performed to obtain the first desired result (success or failure) in a Bernoulli experiment repeated  $n$  consecutive times, is called a geometric random variable. The distribution of this variable is called the geometric distribution. If the  $X$  random variable has a geometric distribution, it is shown in the format  $X \sim \text{Geo}(p)$ .

What is desired in the geometric distribution is to determine the number of experiments required to achieve the initial success. It is a distribution used for quality control purposes [9].

#### 2.1.1.10. Negative Binomial Distribution

The negative binomial distribution is one of the discrete probability distributions used to calculate the probability of occurrence of the  $r$ th success in an event in  $x$  trials. The assumptions required for the Bernoulli distribution are also valid for this distribution. This distribution forms the negative binomial regression analysis basis and is used in cases where the dependent variable shows a negative binomial distribution. The negative binomial distribution, which is the generalized form of the geometric distribution, is used to determine the number of experiments required to achieve  $k$  successes [10]. Variables with quantitative and discrete data types such as the number of COVID-19 tests performed, the number of births, the number of cases, and the number of deaths can be given as examples that show a negative binomial distribution.

#### 2.1.1.11. Hypergeometric Distribution

The hypergeometric distribution is the distribution of success numbers for an operation in which  $n$  successive objects are pulled in series without substituting them through from a finite population. The hypergeometric distribution should satisfy the following conditions:

- $N$  trials can be repeated under similar conditions.
- Each trial has two possible results.
- Non-refundable sampling is made from the finite population.
- As the sampling is non-refundable, the probability of success ( $p$ ) varies from experiment to experiment.

Hypergeometric distribution, unlike the binomial distribution, does not require independence and can be sampled without substitution. The unit tested or used in such sampling will become unusable and cannot be replaced. The hypergeometric distribution is used primarily in electronic testing and quality assurance [11].

#### 2.1.1.12. Zeta (Zipf) Distribution

The Zipf distribution is commonly used to provide a close model for the size or rank of an object randomly chosen from certain types of populations. The Zipf random variable has wide applications when a very small number of outcomes occur quite frequently but a very large number of outcomes occur quite rarely [12].

#### 2.1.1.13. Continuous Probability Distributions

If a continuous random variable can take every value in the range  $(a, b)$ , it is said that this variable has a continuous probability distribution. Continuous Probability Distributions used in the study are explained below.

#### 2.1.1.14. The Continuous Uniform Distribution

Let the range of variability of  $X$  Random variable be  $(a, b)$ . In other words,  $a$  is the minimum value that  $X$  can have and  $b$  let  $X$  be the maximum value that it can take.

If the interval  $(a, b)$  is proportional to the probability of  $X$ , this variable has a continuous uniform distribution [13].

#### 2.1.1.15. Beta Distribution

In probability theory and statistics, the beta distribution is a continuous probability distribution normalized with two positive shape parameters ( $\alpha$  and  $\beta$ ) in the range  $[0,1]$ . Beta distribution has often been applied to model random variables with finite length ranges in various domains or the random behavior of percentages and ratios [14].

#### 2.1.1.16. Normal Distribution

Most of the methods used in practical statistics are based on Normal distribution. This distribution was first discovered in 1733 by Abraham de Moivre (1667-1745) as a distribution in which the sum of variables showing the Binomial distribution converges. The normal distribution has a bell-shaped symmetrical plot. The mean (expected value =  $E(x)$ ) value for this distribution is denoted by  $\mu$ . The normal scatter plot is always symmetrical with respect to the  $\mu$  value. Calculations are made over this value.  $\mu$  is the largest value on the chart. The continuous random variable  $X$  can take all values on the real axis under a normal distribution. So, the range  $-\infty < x < +\infty$  is the range of variability. The area under the curve  $f(x)$  is always 1 [15].

#### 2.1.1.17. Log-Normal Distribution

If the logarithm  $X$  is a normally distributed  $\log X \sim N((\mu, \sigma^2))$  random variable, then the  $X$  random variable has a lognormal distribution. The probability density function of the random variable  $X$  can be obtained by applying a logarithmic transformation to the probability density function of the normal distribution. The probabilities of the lognormal distribution can be calculated using the standard normal distribution. In a lognormal distribution, the random variable can only take positive values [16].

#### 2.1.1.18. Exponential Distribution

An exponential distribution occurs naturally when modeling the time interval between independent events. Exponential distribution occurs as the time until a particular event occurs. The exponential distribution is the only continuous distribution that has memoryless property. Exponential distribution can be used in modeling lifetimes. It can be used when  $\alpha = 1$  in the gamma distribution and the geometric distribution used in discrete situations [17].

#### 2.1.1.19. Gamma Distribution

The gamma distribution is frequently used in engineering, science, and business to model continuous variables with skewed distributions. The importance of the Gamma distribution stems mainly from its relationship with exponential and normal distributions. In this distribution, variables and results are always positive. Gamma distribution occurs naturally when a particular random process is considered over time [18]. For example, patients coming to the hospital need to line up in a network of clinical staff, x-ray machines, operating theaters, and beds, and the time spent on each can be defined by the gamma distribution.

#### 2.1.1.20. Weibull Distribution

The Weibull distribution deals with the time it takes for different systems to fail. The parameters of the distribution provide great flexibility in modeling the system. Here, the error (deterioration) numbers in the system increase, decrease, or remain the same depending on time. The Weibull

distribution is a distribution with positive random variables. The cumulative probability function for the Weibull distribution is a stretched function. Weibull distribution is used to analyze life and survival data because it is very elastic and can be changed easily. Weibull is the most used statistical distribution to determine wind energy potential [19].

2.1.1.21. Rayleigh Distribution

Rayleigh distribution, a special case of the Weibull distribution, reflects a situation in which the dimensions of a two-dimensional vector (0 origin) are normally distributed. These dimensions must also be independent and have the same variance. In this case, the size of the vector will have a Rayleigh distribution [20].

TABLE 1  
PARAMETERS AND DESCRIPTIONS FOR DISCRETE PROBABILITY DISTRIBUTIONS IN THE KODY

| Distribution         | Parameter    | Description  |
|----------------------|--------------|--|
| The Discrete Uniform | k Value      | Number of random variables   |
| Bernoulli            | p Value      | Probability of Success   |
|                      | x Value      | Random Variable  |
| Binomial             | p Value      | Probability of Success   |
|                      | N Value      | Number of Attempts   |
|                      | K Value      | Number of Success  |
| Multinomial          | p1 Value     | First Probability Value  |
|                      | p2 Value     | Second Probability Value   |
|                      | p3 Value     | Third Probability Value  |
|                      | N Value      | Number of Attempts   |
|                      | X1 Value     | First Random Variable  |
|                      | X2 Value     | Second Random Variable   |
| Geometric            | X3 Value     | Third Random Variable  |
|                      | p Value      | Probability Value  |
| Hypergeometric       | Step Value   | It is the number of attempts made until the first success is achieved. |
|                      | n Value      | Sample Observation Value   |
|                      | m Value      | Number of Members in the Population Value                              |
| Poisson              | N Value      | Number of Population Members Value                                     |
|                      | k Value      | Sample Success Value   |
| Negative Binomial    | x Value      | Random Variable  |
|                      | Lambda Value | Average Number of Events in a Range                                    |
| Zeta (Zipf)          | p Value      | Probability of Success   |
|                      | N Value      | Number of Trials   |
|                      | K Value      | Number of Success  |
|                      | x Value      | Random Variable  |
|                      | s Value      | Zeta Distribution Parameter  |

1.1.1.1. Logistic Distribution

The cumulative distribution function of the logistic distribution is a logistic function, and this function also plays a role in logistic regression and neural networks. Although the logistic distribution curve resembles the normal distribution curve in shape, it is flatter than the normal distribution curve since it has wider tails. The relevant distribution a wide range of applications from growth modeling to logistic regression analysis, from physics to hydrology, from chess score ratings to sports modeling, from logit models to artificial neural networks [21].

1.1.1.2. Pareto Distribution

There are many practical applications of the Pareto distribution in probability theory and statistics. This distribution is a continuous probability distribution or a power theory used where stability is obtained in the distribution of a "small" object to a "large" object. It is first used by Vilfredo Pareto, an Italian economist, to show the wealth distribution of individuals in economies [22].

TABLE 2  
PARAMETERS AND DESCRIPTIONS FOR CONTINUOUS PROBABILITY DISTRIBUTIONS IN THE KODY

| Distribution           | Parameter       | Description                        |
|------------------------|-----------------|------------------------------------|
| The Continuous Uniform | x Value         | Random Variable                    |
|                        | a Value         | Minimum value of a random variable |
|                        | b Value         | Maximum value of a random variable |
| Beta                   | x Value         | Random Variable                    |
|                        | $\alpha$ Value  | Shape Parameter                    |
|                        | $\beta$ Value   | Shape Parameter                    |
| Log Normal             | $\mu$ Value     | Arithmetic Mean                    |
|                        | Sigma Value     | Standard Deviation                 |
|                        | x Value         | Random Variable                    |
| Normal                 | $\mu$ Value     | Arithmetic Mean                    |
|                        | Sigma Value     | Standard Deviation                 |
|                        | x Value         | Random Variable                    |
| Exponential            | $\mu$ Value     | Arithmetic Mean                    |
|                        | x Value         | Random Variable                    |
| Gamma Distribution     | N Value         | N Value of Gamma                   |
|                        | x Value         | Range Value                        |
|                        | Lamda Value     | Ratio Parameter ( $\lambda$ )      |
| Weibull                | Alfa Value      | Scale Parameter                    |
|                        | K Value         | Shape Parameter (or slope)         |
| Rayleigh               | x Value         | Random Variable                    |
|                        | Sigma Value     | Standard Deviation                 |
| Logistic               | x Value         | Random Variable                    |
|                        | $\mu$ Value     | Location Parameter                 |
|                        | Sigma Value     | Scale Parameter ( $\sigma$ )       |
| Pareto                 | x Value         | Random Variable                    |
|                        | Alfa Value      | Location Parameter ( $\alpha$ )    |
|                        | Beta Value      | Shape Parameter ( $\beta$ )        |
| Laplace                | x Value         | Random Variable                    |
|                        | Lamda Value     | Scale Parameter ( $\lambda$ )      |
|                        | X Value         | Random Variable                    |
| Cauchy                 | Alfa Value      | Location Parameter ( $\alpha$ )    |
|                        | Beta Value      | Scale Parameter ( $\beta$ )        |
|                        | x Value         | Random Variable                    |
| Erlang                 | x Value         | Random Variable                    |
|                        | $\lambda$ Value | Rate Parameter                     |
|                        | k Value         | Shape Parameter                    |

1.1.1.1. Laplace Distribution

In probability theory and statistics, the Laplace distribution is a continuous probability distribution named in memory of Pierre-Simon Laplace. It is also referred to as the double exponential distribution since it consists of two exponential distributions, which are glued back to each other and combined by including a position parameter. A random variable with two independent and precisely the same exponential distribution functions as a Laplace distribution. The Laplace distribution is the distribution of the differences between two independent exponentially distributed random variables [23].

1.1.1.2. Cauchy Distribution

The Cauchy distribution plays a unique role in statistical theories. The Cauchy distribution is symmetrical and shows a bell-shaped distribution in the range of  $(-\infty, \infty)$ . Although the Cauchy distribution does not look very different from the normal distribution, it contains large differences compared to the normal distribution. One of these is that the mean and moment of the Cauchy distribution are absent. The standard Cauchy distribution is the same as the Student *t*-distribution with 1 degree of freedom [24].

1.1.1.3. Erlang Distribution

The Erlang distribution is a generalization of the exponential distribution. While the exponential random variable describes the time between adjacent events, the Erlang random variable describes the time interval between any event and the *k*th following event [25].

1.1.1.4. Web-Based Software

The theoretical probability distribution software is a web-based platform and can be used free of charge from any device(s) with an internet connection (desktop computer, laptop, mobile phone, etc.). The software was developed using the math and DASH libraries in the Python programming language [26]. The software includes English and Turkish language options. The current software can be accessed at <http://biostatapps.inonu.edu.tr/KODY/>.

3. RESULTS

The following examples were illustrated for binomial and Poisson distributions in order to demonstrate the working style and principle of the software.

Examples of Discrete Probability Distributions in KODY

An example for Binomial Distribution:

The probability of a successful result of a particular surgery is 80%. What is the probability that 6 out of 10 operated patients will recover? (K: Number of patients recovering after surgery)

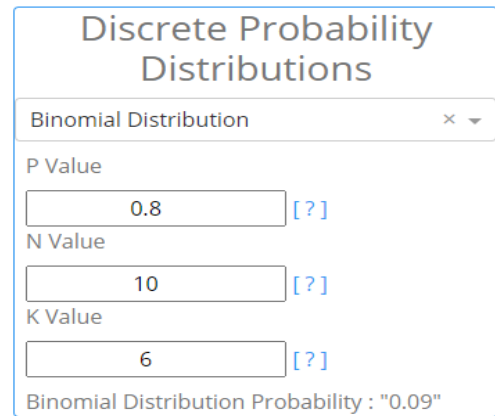


Fig.1. An example for Binominal Distribution

The average number of people who died in a week from a rare disease in a city is 4. What is the probability that three people will die from this disease in a given week?

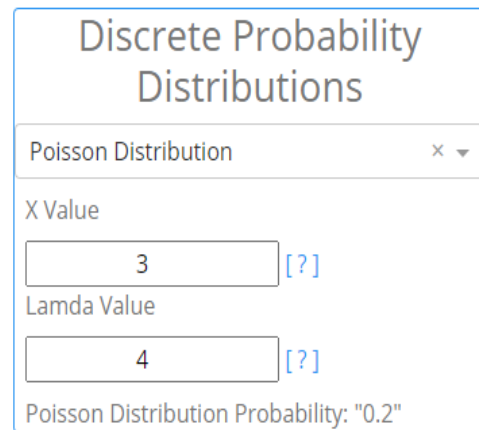


Fig.2. Examples of Poisson Distributions

An example for Normal distribution:

The average length of stay of the patients in Farabi Hospital is = 9 days,  $\sigma = 4$  days. What is the probability that a randomly selected inpatient will be discharged on the 5th day?

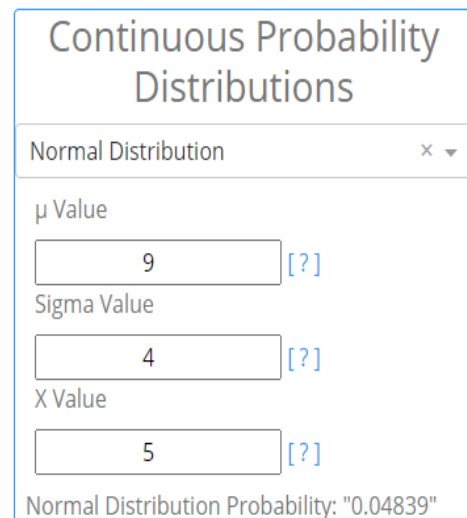


Fig.3. An example for normal distribution

It is known that the operating time (in hours) of electronic devices used in a hospital conforms to an exponential distribution and the average non-defect working

time is calculated to be 24 hours. Accordingly, what is the probability that a randomly selected device will non-defect work for 36 hours maximum?

Continuous Probability Distributions

Exponential Distribution x ▾

μ Value

24 [?]

X Value

36 [?]

Exponential Distribution Probability: "0.78"

Fig.4. An example for Exponential distribution

#### 4. DISCUSSION

Probability is one of the methods used to generalize the information obtained by examining the frequency of occurrence of an event through examples. The frequency and rate of occurrence of random events in society are called probability. When we look at the health problems examined in society, we observe that different results occur in the units under similar conditions. For this reason, health issues are random events that give different results even under stable conditions. Individuals with health problems have biological variability. Results obtained from health problems arise by chance. Thus, the concepts of probability and probability distributions are essential in medicine [27]. Therefore, in this study, a theoretical probability distribution software that includes probability calculations for comprehensive discrete and continuous probability distributions has been developed using Python programming language.

In virtually every aspect of life, randomness, and uncertainty, which always go synonymously, exist. To this effect, almost everyone, through intuition or experience, has a basic understanding of the term probability. The analysis of probability derives from the study of certain games of chance. Probability is the measure of the chance of an event occurring and, as such, finds applications in uncertain disciplines. Probability theory is used extensively, to name just a few, in a host of fields in science, engineering, medicine, and business. Probability theory is nothing but common sense reduced to calculation, as claimed by Pierre-Simon Laplace, a prominent French scholar. A probabilistic model is required in order to account for uncertainties in a random experiment. As a simplified approximation to an actual random experiment, a probability model provides sufficient details to include all significant aspects of the random phenomenon. Models of probability are generally based on the fact that averages obtained from random experiments in long sequences of independent trials almost always give rise to the same value. In many instances, this property, known as a statistical regularity, is an experimentally verifiable phenomenon. As the relative frequency of the event, the ratio representing the number of times a particular event happens over the number of times the trial has been repeated is defined. When infinity is approached by the number of times the experiment is repeated, the relative frequency of the event, which

approaches a limit due to statistical regularity, is called the definition of probability of relative frequency. Note that this limit, based on an a posteriori approach, cannot really exist, because the number of times repeated in a physical experiment can be huge, but always finite [12]. In addition to applying many statistical and data science models, the theoretical distributions are frequently used in the verification of different methods in which hypothesis tests are evaluated for various assumptions or propositions. Knowing the structure of the distribution of the properties studied and predicting the occurrence of the event of interest is widely applied in biomedical research.

When the literature is examined, a web-based tool called MATCH was developed in a study to determine probability distributions about uncertain model parameters. The tool is free to use and includes five methods for extracting univariate probability distributions. These are normal, student-t, scaled beta, gamma, log-normal distributions [28]. When the MATCH tool is compared with the theoretical probability distribution software developed in this study, MATCH tool includes restricted distributions and makes parameter estimation by examining the distributions graphically. But in KODY, probability calculations for The Discrete Uniform, Bernoulli, Binomial, Multinomial, Poisson, Geometric, Negative Binom, Hypergeometric, Zeta (Zipf), The Continuous Uniform, Beta, Normal, Log Normal, Exponential, Gamma, Weibull, Rayleigh, Logistics, Pareto, Laplace, Cauchy and Erlang distributions can be performed quite easily. Probabilities for various theoretical distributions can be calculated in package programs such as Microsoft Excel [29], IBM SPSS Statistics [30], MedCalc [31], Statistica [32]. However, although these distributions are limited in the tools, most of these software are licensed/paid and cannot be easily used by researchers regardless of the platform, since they operate depending on the operating system.

In the literature, various parameters belonging to each probability distribution are used. These parameters are shown in different places in the literature with different symbols and different explanations. For this reason, we added the description of the relevant parameter for each parameter in KODY as tooltip. In this way, researchers will obtain results without error by using these tooltips while performing probabilistic calculations regarding each probability distribution in the software.

As a result, the open-access web-based software with Turkish/English language options may guide and contribute to researchers in probabilistic estimation processes regarding theoretical distributions. In the later stages of this study, it is foreseen to develop simulation processes based on each probability distribution.

#### REFERENCES

- [1] Akdeniz, F. (1995). *Olasılık ve istatistik*. AÜ FF Döner Sermaye İşletmesi Yayınları. Beaumont, G. P. (1983). Introductory applied probability. *JOHN WILEY & SONS, INC., 605 THIRD AVE., NEW YORK, NY 10158, USA, 1983, 200.*
- [2] Beaumont, G. P. (1983). Introductory applied probability. *JOHN WILEY & SONS, INC., 605 THIRD AVE., NEW YORK, NY 10158, USA, 1983, 200.*
- [3] Ghahramani, S. (2000). *Fundamentals of probability* (Vol. 2). New Jersey: Prentice Hall.
- [4] Sandhya, E., & Prasanth, C. B. (2014). Marshall-Olkin discrete uniform distribution. *Journal of probability, 2014.*
- [5] Dai, B., Ding, S., & Wahba, G. (2013). Multivariate bernoulli distribution. *Bernoulli, 19(4), 1465-1483.*

- [6] Altham, P. M. (1978). Two generalizations of the binomial distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 27(2), 162-167.
- [7] Mosimann, J. E. (1962). On the compound multinomial distribution, the multivariate  $\beta$ -distribution, and correlations among proportions. *Biometrika*, 49(1/2), 65-82.
- [8] Consul, P. C., & Jain, G. C. (1973). A generalization of the Poisson distribution. *Technometrics*, 15(4), 791-799.
- [9] Philippou, A. N., Georghiou, C., & Philippou, G. N. (1983). A generalized geometric distribution and some of its properties. *Statistics & Probability Letters*, 1(4), 171-175.
- [10] Fisher, P. (1941). Negative Binomial Distribution. *Annals of Eugenics*, 11, 182-787.
- [11] Harkness, W. L. (1965). Properties of the extended hypergeometric distribution. *The Annals of Mathematical Statistics*, 36(3), 938-945.
- [12] Grami, A. (2019). *Probability, random variables, statistics, and random processes: Fundamentals & applications*. John Wiley & Sons.
- [13] Kuipers, L., & Niederreiter, H. (2012). *Uniform distribution of sequences*. Courier Corporation.
- [14] Gupta, A. K., & Nadarajah, S. (2004). *Handbook of beta distribution and its applications*. CRC press.
- [15] Ahsanullah, M., Kibria, B. G., & Shakil, M. (2014). Normal distribution. In *Normal and Student t Distributions and Their Applications* (pp. 7-50). Atlantis Press, Paris.
- [16] Mitchell, R. L. (1968). Permanence of the log-normal distribution. *JOSA*, 58(9), 1267-1272.
- [17] Balakrishnan, N. (1995). AP Basu: The Exponential Distribution: Theory, Method and Applications.
- [18] Thom, H. C. (1958). A note on the gamma distribution. *Monthly weather review*, 86(4), 117-122.
- [19] Rinne, H. (2008). *The Weibull distribution: a handbook*. CRC press.
- [20] Kundu, D., & Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimations. *Computational statistics & data analysis*, 49(1), 187-200.
- [21] Balakrishnan, N. (1991). *Handbook of the logistic distribution*. CRC Press.
- [22] Arnold, B. C. (2014). Pareto distribution. *Wiley StatsRef: Statistics Reference Online*, 1-10.
- [23] Kotz, S., Kozubowski, T., & Podgorski, K. (2012). *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance*. Springer Science & Business Media.
- [24] Arnold, B. C., & Beaver, R. J. (2000). The skew-Cauchy distribution. *Statistics & probability letters*, 49(3), 285-290.
- [25] Ibe, O. (2013). *Markov processes for stochastic modeling*. Newnes.
- [26] Sanner, M. F. (2004). Using the Python programming language for bioinformatics. *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*.
- [27] Grimmett, G., & Stirzaker, D. (2020). *Probability and random processes*. Oxford university press.
- [28] Morris, D. E., Oakley, J. E., & Crowe, J. A. (2014). A web-based tool for eliciting probability distributions from experts. *Environmental Modelling & Software*, 52, 1-4.
- [29] Fylstra, D., Lasdon, L., Watson, J., & Ware, A. (1998). Design and use of the Microsoft Excel Solver. *Interfaces*, 28(5), 29-55.
- [30] Nie, N. H., Bent, D. H., & Hull, C. H. (1975). *SPSS: Statistical package for the social sciences* (Vol. 227). New York: McGraw-Hill.
- [31] Schoonjans, F. R. A. N. K., Zalata, A., Depuydt, C. E., & Comhaire, F. H. (1995). MedCalc: a new computer program for medical statistics. *Computer methods and programs in biomedicine*, 48(3), 257-262.
- [32] StatSoft, I. N. C. (2001). STATISTICA (data analysis software system), version 6. *Tulsa, USA*, 150, 91-94.

University in 2005, and Ph.D. degrees in biostatistics and medical informatics from the Inonu University in 2017. He is currently working as an assistant professor of the Department of Biostatistics and Medical Informatics at Inonu University and as the information processing manager at Turgut Özal Medical Center. His research interests are cognitive systems, data mining, machine learning, and deep learning.

**Cemil Çolak** obtained his BSc. degree in Statistics from Ondokuz Mayıs University in 1999. He received MSc. degree in Biostatistics from the Inonu University in 2001, and a PhD degree in the Graduate Department of Biostatistics and Medical Informatics of Ankara University in 2005. His research interests are cognitive systems, data mining, reliability, and biomedical system, genetics, and bioengineering. In 2016, he joined the Department of Biostatistics and Medical Informatics at Inonu University as a Professor, where he is presently a professor. He is active in teaching and research in general image processing, artificial intelligence, data mining, and analysis.

## BIOGRAPHIES

**Fatma Hial Yağın** obtained her BSc. degree in Statistics from Gazi University in 2017. She received MSc. degree in biostatistics and medical informatics from Inonu University in 2020. She currently continues her Ph.D. education in biostatistics and medical informatics from Inonu University. In 2019, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning, and image processing.

**Emek GÜldoğan** obtained his BSc. degree in Computer Engineering from Middle East Technical University in 2001. He received MSc. degree in biostatistics and medical informatics from the Inonu