REPUBLIC OF TURKEY
INONU UNIVERSITY
GRADUATE SCHOOL OF NATURE AND APPLIED SCIENCES

ENHANCEMENT OF OPTIMIZATION ALGORITHMS' PERFORMANCES IN ENGINEERING PROBLEMS VIA DIFFERENT DISTRIBUTION FUNCTIONS

MASTER THESIS

Mehmet AKPAMUKÇU

Computer Engineering Department

Thesis Supervisor: Asst. Prof. Dr. Abdullah ATEŞ

JUNE 2021

ACKNOWLEDGEMENTS AND FOREWORD

I thank to;

My supervisor Dr. Abdullah Ateş, who guided me in every aspect of this thesis, without sparing his help, advice, knowledge, experience and support;

My family who did not spare any support from me during my studies, as well as throughout my life.

DECLARATION

I hereby declare that I wrote this Master Thesis titled "Enhancement Of Optimization Algorithms' Performances In Engineering Problems Via Different Distribution Functions" in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Inonu University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

Mehmet AKPAMUKÇU

CONTENTS

TABLES

FIGURES

SYMBOLS AND ABBREVIATIONS

pdf : Probability Density Function

cdf : Cumulative Distribution Function

SMDO: Stochastics Multi-parameters Divergence Optimization

MBO: Monarchy Butterfly Optimization

BAR: Butterfly Adjusting Rate

DOF: Degree of Freedom

DE: Differential Evolution

ABC: Artificial Bee Colony

ACO: Ant Colony Optimization

BBO: Biogeography-based optimization

SGA: Standart Genetic Algorithm

DSO: Donkey and Smuggler Optimization

LQR: Linear-Quadratic Regulator

ÖZET

Yüksek Lisans Tezi

FARKLI DAĞILIM FONKSİYONLARIYLA OPTİMİZASYON ALGORİTMALARININ MÜHENDİSLİK PROBLEMLERİNDEKİ PERFORMANSININ İYİLEŞTİRİLMESİ

Mehmet AKPAMUKÇU

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

76 + XI sayfa

2021

Danışman: Dr. Öğr. Üyesi Abdullah ATEŞ

Stokastik süreç içeren optimizasyon algoritmalarının performansının iyileştirilebilmesi için farklı dağılım fonksiyonlarının rassal süreçlerde kullanılması yaklaşımı optimizasyon algoritmalarının performansını artırabilir. Çünkü literatür ve uygulama açısından, yeni stokastik yöntemlerin önerilmesinin yanı sıra mevcut yöntemlerin farklı dağılım fonksiyonları gibi analitik katkılarla geliştirilmesi ve gerçek zamanlı mühendislik problemlerinde kullanılması önem kazanmıştır. Bundan dolayı bu tez çalışmasında öncelikle mevcut nümerik optimizasyon algoritmalarının analitik katkılarla nasıl geliştirilebileceği ile ilgili metotların önerilmesi amaçlanmıştır.

Bilindiği gibi nümerik optimizasyon algoritmalarındaki en kritik yapılardan biri stokastik arayışın yönünü belirleyen adım belirleme safhasıdır. Genellikle bu yöntemlerde uniform dağılıma göre türetilen rassal değişkenler kullanılmaktadır. Fakat her durumda sabit olarak uniform dağılıma göre türetilen rassal değişkenlerin kullanılması mevcut algoritmanın dinamiğine uygun olmayabilir. Hatta tez süresince yapılan çalışmalarda algoritmanın her hareket durumunda uniform dağılımın kullanılmasının algoritmaların performansını etkilediği tespit edilmiştir. Bundan dolayı rassal adımların belirlenmesinde uniform dağılım yerine farklı dağılım fonksiyonlarının kullanılmasının algoritmanın performansına olumlu bir etki yaratacağı gözlemlenmiştir.

Bundan dolayı bu tezde, istatistiğin temel konularından olan dağılım fonksiyonları ve istatistiksel momentler detaylı olarak analiz edilmiştir. Elde edilen çıktılar optimizasyon algoritmalarının dinamiğine uyarlanarak mühendislik problemlerinde uygulanmıştır. Bu tez çalışmasında öncelikle dağılım fonksiyonlarının etkisinin ortaya çıkarılması için rassal parametre vektör optimizasyon yöntemi (SMDO) farklı dağılım fonksiyonları ile modifiye edilmiştir. Elde edilen dağılım fonksiyonu tabanlı rassal parametre vektör optimizasyonu yöntemi benchmark fonksiyonları üzerinden literatürdeki sonuçlar ile karşılaştırmalı olarak sunulmuştur. Ve elde edilen sonuçlara göre farklı dağılım fonksiyonlarının kullanılması, ilgili metodun performansını artırmıştır. Bunların yanı sıra bu yapı için kullanıcı dostu bir araç kutusu tasarımı da yapılmıştır.

Daha sonra literatürde yeni önerilmiş bir algoritma olan monarchy butterfly optimizasyon algoritması farklı dağılım fonksiyonları ile güncellenerek modifiye edilmiş monarch butterfly optimizasyon algoritması (M^2BO) önerilmiştir. M^2BO optimizasyon algoritması öncelikle benchmark fonksiyonları üstünden test edilmiş ve literatürdeki sonuçlarla karşılaştırmalı olarak sunulmuştur. Hatta dağılım fonksiyonlarının performansına etki eden parametrelerde her benchmark fonksiyonu için ayrı ayrı ayarlanmıştır. Daha sonra önerilen farklı dağılım fonksiyonu yaklaşımının mühendislik problemlerindeki performansını göstermek için 3 DOF Hover 4 motorlu helikopter prototipi üzerinde test edilmiştir. Sistemin kontrolünü sağlayan kazanç matrisinin parametreleri M^2BO algoritmasıyla tasarlanmış ve sonuçlar simülasyon ve gerçek zamanlı sistem modeli üzerinden test edilmiştir.

Böylece farklı dağılım fonksiyonlarının optimizasyon algoritmalarındaki stokastik süreçlerde uniform dağılım yerine kullanılması yaklaşımının algoritmaların performansını artırabileceği gerçek zamanlı sistem ve benchmark fonksiyonları üzerinde gösterilmiştir.

Anahtar Kelimeler: Optimizasyon, dağılım fonksiyonları, istatistiksel moment, stokastik yöntemler

ABSTRACT

Master Thesis

ENHANCEMENT OF OPTIMIZATION ALGORITHMS' PERFORMANCES IN ENGINEERING PROBLEMS VIA DIFFERENT DISTRIBUTION FUNCTIONS

Mehmet AKPAMUKÇU

Inonu University
Graduate School of Nature and Applied Sciences
Department of Computer Engineering

76 + XI pages

2021

Supervisor: Asst. Prof. Dr. Abdullah ATEŞ

In order to improve the performance of optimization algorithms containing stochastic processes, using different distribution functions in random processes can increase the performance of optimization algorithms. Because, in terms of literature and practice, it has become important to develop existing methods with analytical contributions such as different distribution functions and to use them in real-time engineering problems , besides suggesting new stochastic methods. Therefore, in this thesis, it is primarily aimed to suggest methods for how existing numerical optimization algorithms can be developed with analytical contributions.

As it is known, one of the most critical structures in numerical optimization algorithms is the step determination phase that determines the direction of the stochastic search. Generally, random variables derived from uniform distribution are used in these methods. However, in all cases, the use of random variables derived from a uniform distribution may not be suitable for the dynamics of the current algorithm. In fact, in the studies carried out during the thesis, it has been determined that the use of uniform distribution in every motion of the algorithm affects the performance of the algorithms. Therefore, it has been observed that using different distribution functions instead of uniform distribution in determining random steps will have a positive effect on the performance of the algorithm.

Therefore, in this thesis, distribution functions and statistical moments, which are among the fundamental topics of statistics, have been analyzed in detail. The obtained outputs have been applied in engineering problems by adapting them to the dynamics of optimization algorithms. In this thesis, first of all, the random parameter vector optimization method (SMDO) has been modified with different distribution functions to reveal the effect of distribution functions. The obtained distribution function based random parameter vector optimization method is presented in comparison with the results in the literature over the benchmark functions. And using different distribution functions according to the results obtained increased the performance of the related method. In addition to these, a user-friendly toolbox has been designed for this structure.

Later, the monarchy butterfly optimization algorithm, which is a newly proposed algorithm in the literature, was updated with different distribution functions and the modified monarch butterfly optimization algorithm (M^2BO) was proposed. M^2BO optimization algorithm is first tested on benchmark functions and presented in comparison with the results in the literature. In fact, the parameters that affect the performance of distribution functions are adjusted separately for each benchmark function. Then, it was tested on 3 DOF Hover 4 engine helicopter prototype to show the performance of the proposed different distribution function approach in engineering problems. The parameters of the gain matrix, which provides the control of the system, were designed with the M^2BO algorithm and the results were tested through simulation and real-time system model.

Thus, it has been shown on real-time system and benchmark functions that using different distribution functions in optimization algorithms instead of uniform distribution in stochastic processes can increase the performance of algorithms.

Keywords: Optimization, distribution functions, statistical moment, stochastic methods

1. INTRODUCTION

1.1.The Aim of the Thesis

Optimization algorithms are widely used today to solve many problems. These methods, which we can define as optimization, are algorithmic structures that allow a problem to be solved in the most efficient and accurate way in a specific solution space under certain constraints.

Engineering disciplines, in the process of solving a problem, generally use analytical methods that aim to reach a holistic solution from the solution of the parts by breaking down the problems on a mathematical basis. However, in the real world engineering problems, these solutions, which have been designed perfectly in theory, may be insufficient. Because of many disruptive factors in real world conditions, the applicability of analytical methods may be insufficient due to the dynamics of the problems. At this point, when the system and its constraints cannot be fully modeled analytically then numerical methods should be used. The best solution is tried to be reached by inspiring from the real laws of physics in the solution space of the problems with the numerical methods. These methods are developed to meet the needs of different living things and similar approaches. These methods are the process of finding or discovering the best solution within the relevant optimization process in the solution space. However, there are different uncertainties in this solution seeking process. One of them is to determine the starting point where the search for solutions will begin. The other is to determine whether this search will proceed in a deterministic or stochastic way. Deterministic steps often do not find the best solution in the search process. Besides, a deterministic process management is closer to the analytical solution understanding inherent in engineering. When the studies in the literature are examined, methods involving a stochastic process get more successful results in reaching the best solution. However, since these processes also involve a random progression, which arises from their nature, they move away from the analytics manner required by the engineering discipline and time problems may occur in reaching a solution.

In this thesis, it is aimed to increase the performance of optimization problems in engineering problems by using different distribution functions. As a result, it is aimed to make an effective contribution to the optimization problems progressing in a stochastic manner and it is aimed to increase the efficiency and effect of these optimization methods to be used in solving real engineering problems.

1.2.Subject, Scope and Summary of Literature

In this thesis, it is aimed to improve the performance of optimization problems in real engineering problems by using statistical distribution functions. From this point of view, stochastic progression processes of optimization algorithms are discussed. Analyzing these processes and improving this processes constitute the main subject of the thesis. This thesis covers optimization algorithms including stochastic processes. Further development of stochastic algorithms is mainly aimed for the above-mentioned purposes. In this thesis, as optimization algorithms that we plan to increase their effectiveness; current methods such as stochastic multi parameter divergence method (Alagoz et al., 2013)(Yeroğlu & Ateş, 2014), discrete stochastic search (Q. Wang, 2013), modified Tabu (Abdullah Ateş & Yeroglu, 2016), modified artificial physics optimization algorithm (Abdullah Ateş & Yeroğlu, 2018), cuckoo search algorithm (Rajabioun, 2011) and similar algorithms are considered. It is planned to use some distribution functions in order to increase the performance of these algorithms (*Weibull, W. (1951) A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 18, 293-297. - References - Scientific Research Publishing*, 2021). For example; by analyzing the dynamics of the progress of the specified algorithms in the solution space of functions such as gauss (Chhikara & Folks, 1974), binomial (Bahadur, 1960), geometric (*Two Characterizations of the Geometric Distribution by Record Values on JSTOR*, 2021), Poisson (*On a Characterization of the Poisson Distribution on JSTOR*, 2021), uniform (Glass & Tobler, 1971), normal (*"Normal" Distribution Functions on Spheres and the Modified Bessel Functions on JSTOR*, 2021), gamma (Stacy, 1962), exponential (Marshall & Olkin, 1967) distributions. Detailed analyzes will be made to find the most suitable matches.

In the literature, it has been determined that statistical distribution functions and statistical moments are used in various fields. First of all, in the field of pharmacokinetics, the use of statistical moments in the analysis of the process of the presence of drugs in the human body has been examined (Yamaoka et al., 1978). Here, the analyzes made by associating the statistical distribution with the process of entry, distribution and excretion of drugs, which is a very dynamic process, are extremely interesting. Later, the use of statistical moments together with distribution functions in bearing defect detection was investigated (Martin & Honarvar, 1995). In this study, it is seen that statistical distribution functions can contribute to the solving process of a real engineering problem under the guidance of statistical moments. In (ATEŞ & ALAGÖZ, 2019), which we examined in the literature review, the effects of different probability distribution functions on stochastic based optimization algorithms were examined through controller design. In this study, analysis of the effect of distribution functions is presented without specifying analytical justifications.

In this thesis, it is aimed to develop the methodology of determining the appropriate distribution functions analytically in a certain discipline for the reasons of using the distribution functions used. In (A. Ates et al., 2019), the random parameter vector optimization method was modified with different distribution functions and a fractional order controller was designed for a fractional order system. In this study, clear effects of different distribution functions on the operation of different algorithms in random number generation were observed, and it is thought that a detailed study in this area will have positive effects on increasing the performance of optimization algorithms. In addition, the study in which statistical moments are used in character recognition is presented in (Chim et al., 1999). Analyzes on statistical results of stochastic optimization problems in (Shapiro, 2000) show the relationship between stochastic optimization algorithms and statistics. The study in which the constrained optimization problem, which is an analytical optimization problem, is solved with the particle swarm optimization algorithm using the Gaussian distribution function is presented in (Krohling & Dos Santos Coelho, 2006). It is shown in the study (Yang et al., 2014) that different probability distribution functions are used in the chaos optimization algorithm. In (Lee & Yao, 2004), the use of the levy distribution function in evolutionary calculation processes is presented. As can be understood from these studies, the distribution functions used are only used instead of the relevant stochastic search. No analytical reason for use is given. Therefore, in this thesis, it is aimed to find these relationships analytically and to use them in engineering problems.

2. THE CHARACTERISTICS OF DISTRIBUTION FUNCTIONS AND STATISTICAL MOMENTS

2.1. Probability Distribution

A probability distribution; defines the values and probabilities for a random event to occur. These values should include all possible outcomes for the event, and the sum of the probabilities must be exactly one or 100 percent. For example, let's take a single random event that a coin is thrown into the air and fell to the ground; values are 'tails' or 'heads', or 0 (tails) or 1 (heads) if they are expressed in a nominal variable scale; probabilities will be for both values. Thus, two values and two associated probabilities for a single throw event of a coin becomes the probability distribution of this random event (*Olasılık Dağılımı*, 2020)

The probability distribution and the random variables it defines are important sub-sections of the branches of mathematics science, probability theory and statistics. Probability distributions are models used for probability analysis and defining the probability of events. According to the probability theory, a function is connected to each random variable in the state space, which is defined as a probability distribution. This probability function has determined a probability for every subset (every measurable subset) in each state space, in accordance with the probability axioms (*Olasılık Dağılımı*, 2020)

2.1.1. Distribution of discrete random variables

2.1.1.1. Distribution of discrete random variables

Let X be a discrete random variable that can take the finite number $x_1, x_2, x_3, \ldots, x_N$ with the probabilities $f(x_i) = P(X = x_i)$, $i = 1,2,\ldots, N$. In this case, the function f(x) satisfying the following conditions is called the probability function of X (Akdeniz, 2018)

a) $f(x) >= 0$, for all x
b) $\sum_{i=1}^{N} f(x_i) = 1$

X=x	x_1 x_2 x_3 …………… x_N
f(x) = P(X=x)	$f(x_1)$ $f(x_2)$ $f(x_3)$ ………….. $f(x_N)$

Table 2.1 Probability Distribution

2.1.1.2.Distribution function

The distribution function of a random variable X is denoted by F (x) and it is the probability that X is equal to or less than x (Akdeniz, 2018).

In that case;

$$F(x) = P(X<=x) = \sum_{xi \leq x}^{N} \boldsymbol{f(x_i)}$$

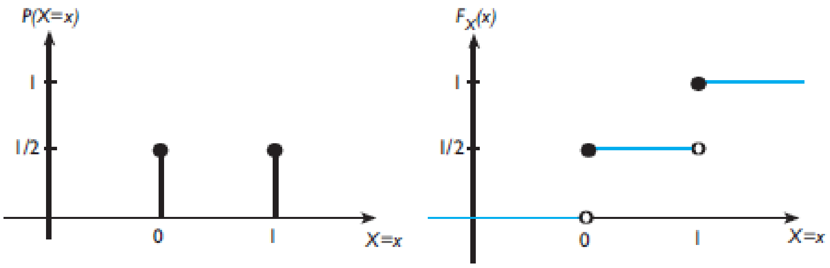

Figure 2.1 Distribution Function of Discrete Random Variable (Yalçın, 2020)

2.1.2. Distribution of a continuous random variable

2.1.2.1. Probability density function

Let X be a continuous random variable defined in the interval (-∞, + ∞). The function f (x) satisfying the following conditions is called the probability density function of the X random variable (Akdeniz, 2018).

a) f(x) >= 0, -∞ < x < +∞

b) $\int_{-\infty}^{+\infty} f(x)dx = 1$, f(x) The area under the curve and bounded by the x-axis is equal to 1.

2.1.2.2. Distribution function

Let X be a continuous random variable with a probability density function f (x). Distribution function of x is defined as (Akdeniz, 2018).

$$F(x) = P(X<=x) = \int_{-\infty}^{x} f(s)ds$$

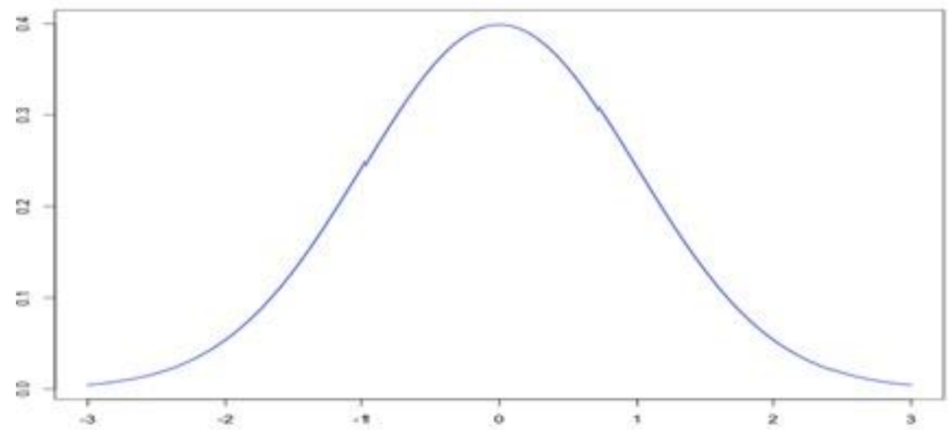

Figure 2.2 Distribution Function of Continuous Variable (Lentini,2020)

2.2.Statistical Moments

The concept of moment in mathematics has been developed from the concept of moment introduced for physics (*Moment*, 2020). In mathematics, the numerical measure that describes what a set of points looks like is called a moment. Moment is the arithmetic mean of the forces of the difference of the observation values in a series from zero or the arithmetic mean. These measurements are used to determine the shape of the frequency distribution of the series. Hence, a probability distribution can be summed up by a set of moments belonging to that distribution (Yalta, 2020).

For probability theory and statistics, the functions that moments are related to the probability density function for a random variable. It is the mathematical expectation of the nth moment X^N of a probability density function around zero. Moments around mean μ are called central moments; they describe the form of a function (*Moment*, 2020). The nth moment around the c value of f (x), the real-valued function of a real variable, is expressed as follows.

$$\mu_n^{'} = \int_{-\infty}^{\infty} (x-c)^n f(x)dx$$

Moment relative to zero is the average difference of observation values in a series from zero to various degrees. The first moment around zero, if significant, is the mathematical expectation of X, that is the mean of the probability distribution of X written as μ. For higher orders, central moments are more interesting than moments around zero (*Moment*, 2020).

As mentioned earlier, a moment in mathematics is a specific quantitative measure of the shape of a function. If the function is a probability distribution; then

a) The zeroth moment is the total probability (i.e. equals one),
b) First moment: The expected value,
c) Second central moment: Variance,
d) Third moment: Skewness,
e) Fourth moment: It is kurtosis.

2.2.1. Expected value

The X random variable is called the mean or population mean. Let X be a discrete random variable with the probability function in Table 2.1. The expected value of X, represented by E(X), is defined as follows.

$$E(X) = x_1.f(x_1) + x_2.f(x_2) + \ \dots + x_N.f(x_N) = \sum_{i=1}^{N} x_i f(x_i)$$

Let X be a continuous one-dimensional random variable. Where f (x) is the probability density function of X, the expected value of X, E(X) is defined as follows:

$$E(X)=\int_{-\infty}^{+\infty} x.f(x)dx \quad -\infty < x < +\infty$$

The value or mean of a random variable gives us information about the center of the probability function. However, the mean value does not give information about the distribution, change or spread of the values of the random variable from one experiment to another (Akdeniz, 2018).

2.2.2. Variance

Let the probability function X be a discrete random variable given as in Table 2.1. If the mean of X is E(X) = μ, the variance of X, Var (X) or σ_x^2 is defined as follows.

$$\sigma_x^2 = Var(X) = E\,[(X- \mu)^2]$$

a) If X is a discrete random variable;

σ_x^2 = Var(X) = $\sum_{i=1}^{N}(x_i - \mu)^2 f(x_i)$

b) If X is a continuous random variable;

σ_x^2 = Var(X) = $\int_{-\infty}^{+\infty}(x - \mu)^2 . f(x). dx$

Standard deviation; Let X be a discrete or continuous random variable with μ mean. The standard deviation of x, σ_x, is the square root of the variance (Akdeniz, 2018).
Variance measures how far a set of (random) numbers is spread from their mean value.

2.2.3. Skewness

The third moment with respect to the mean, μ^3 is called the measure of non-symmetry. That is, skewness is the degree of departure from symmetry in a distribution. If the frequency curve of the distribution has a longer extension to the right of the central maximum, the distribution is called right skewed. If there is a reverse situation, it is said to be skewed to the left (Akdeniz, 2018).

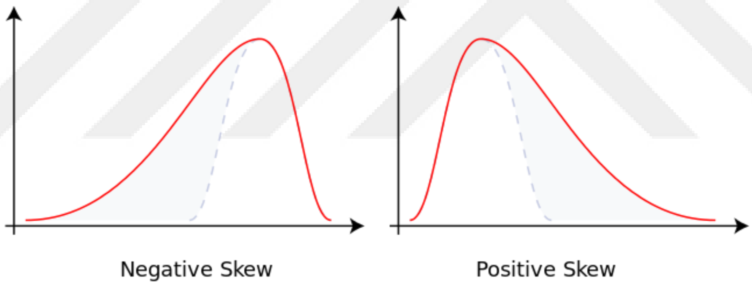

Figure 2.3 Skewness (*Skewness*, 2020)

However, the third moment does not always give clues about the shape of the distribution.

2.2.4. Kurtosis

It is used as the fourth moment in relation to the mean. It is the degree of flatness near the center of the graph of the density function. μ^4/σ^4- 3 is called the kurtosis coefficient. For K<0, the curve near the center is too flat compared to the normal distribution curve. For K> 0, the curve near the center is narrower and higher than the normal distribution curve. K = 0 for normal distribution (Akdeniz, 2018). The fourth central moment is a measure of whether the distribution is thin and pointed or thick and flat, and a comparison is made with a normal distribution showing the same variance to distinguish this property.

Since the fourth central moment is the mathematical expectation of a quadruple exponential, it always takes a positive value (*Moment*, 2020) if it can be defined (excluding only the degenerate point distribution). "Flatness" (kurtosis) is used to examine flatness. To understand whether a random variable fits the normal distribution, one can look at the skewness and kurtosis values.

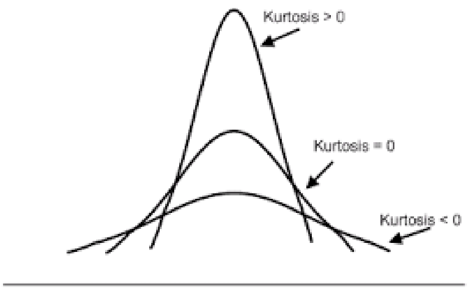

Figure 2.4 Kurtosis (*Coefficient of Kurtosis*, 2020)

2.3. Probability Distribution Functions

In this section, the properties of some distribution functions will be analyzed. These functions are: Normal Distribution, Geometric Distribution, Binomial Distribution, Beta Distribution, Weibull Distribution, Poisson Distribution, Gama Distribution, Uniform Distribution, Exponential Distribution, Rayleigh Distribution, Lognormal Distribution, Logistics Distribution, Chi-Square Distribution, Loglogistics Distribution, Student's T Distribution.

2.3.1. Normal distribution

The normal distribution, called the Gaussian distribution, is a family of two-parameter curves. The rationale for using the normal distribution for modeling is the Central Limit Theorem, which states that the finite mean and variance and the sum of samples independent of any distribution approach the normal distribution as the sample size goes to infinity (*Probability Density Function*, 2020). Normal distribution uses the following parameters:

Parametre	Definition	Explanation
mu (μ)	Mean	$-\infty<\mu<\infty$
sigma (σ)	Standart Deviation	$\sigma \geq 0$

Table 2.2 Normal Distribution Function

Properties of Normal Distribution:

a) Normal Distribution is symmetrical with respect to the x = μ line, that is, the graph of the normal distribution is the same on the left and right of the line x = μ.
b) Mean (μ) is in the middle and divides the area into two equal parts.
c) The area under the F(x) curve and above the x-axis is equal to 1.

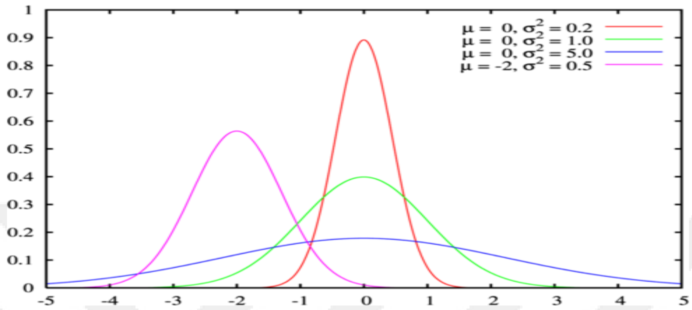

Figure 2.5 Normal Distribution

2.3.2. Geometric distribution

The geometric distribution is a single parameter curve family that models the number of failures before a success in a series of independent trials where each trial results in success or failure, and the probability of success in any individual trial is constant (Probability Density Function, 2020). The geometric distribution uses the following parameters:

Parameter	Definition	Explanation
p	Success Probability	$0 \leq p \leq 1$

Table 2.3 Geometric Distribution Parameters

Probability density function of the geometric distribution (pdf):

$y = f(x|p) = p(1-p)^x$; x = 0,1,2,…,

Cumulative distribution function of geometric distribution (cdf):

$y = F(x|p) = 1 - p(1-p)^{x+1}$; x = 0,1,2,…,

Mean of the geometric distribution: mean = (1-p)/p

Variance of the geometric distribution: var = $(1-p)/p^2$

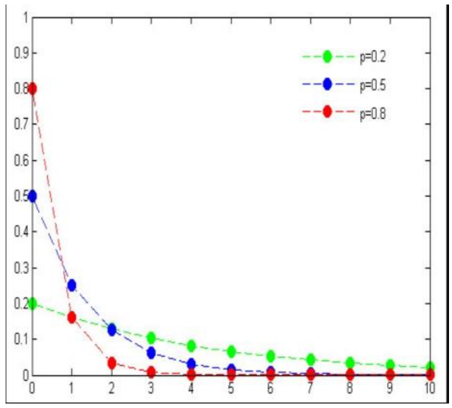

Figure 2.6 Geometric Distribution Probability Mass Function

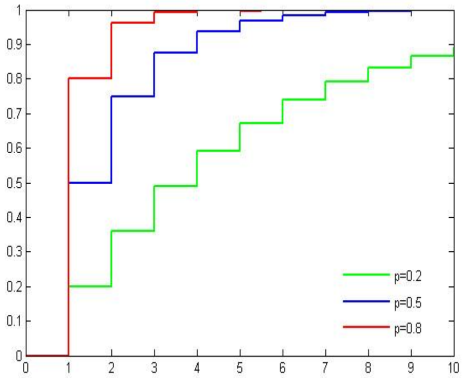

Figure 2.7 Geometric Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.3. Binomial distribution

The binomial distribution is a family of curves with two parameters. The binomial distribution is used to model the total number of successes in a fixed number of independent trials with the same probability of success (Probability Density Function, 2020). The binomial distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
N	Trial Number	Positive Integer
p	Probability of success in a single trial	$0 \leq p \leq 1$

Table 2.4 Binomial Distribution Parameters

Probability density function of the binomial distribution (pdf):

$$f(x|N,p) = \binom{N}{x} p^{x}(1-p)^{N-x}; x = 0,1,2,...., N,$$

Cumulative distribution function (cdf) of the binomial distribution:

$$F(x|N,p) = \sum_{i=0}^{x} \binom{N}{i} p^{i}(1-p)^{N-i}; x = 0,1,2,...., N,$$

Mean of binomial distribution: Np

Variance of the Binomial distribution: Np(1 – p)

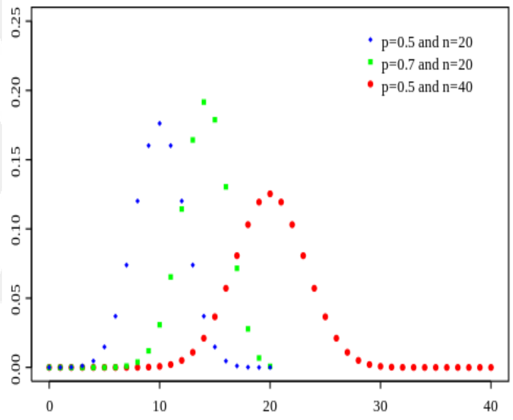

Figure 2.8 Binomial Distribution Probability Mass Function

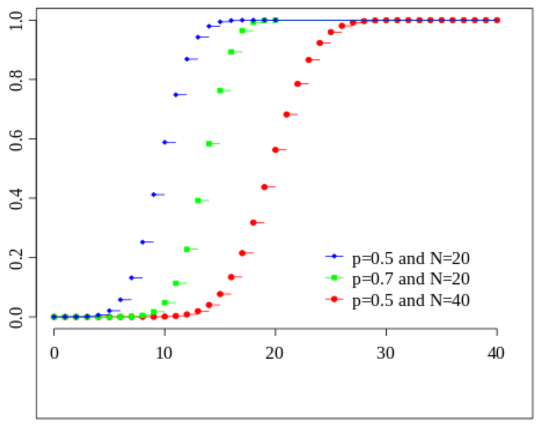

Figure 2.9 Binomial Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.4. Beta distribution

The beta distribution defines a curve family that is unique only in the zero range (0 1) because they are not zero. A more general version of the function assigns parameters to the endpoints of the range (*Probability Density Function*, 2020). The beta distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
a	First shape parameter	a>0
b	Second shape parameter	b>0

Table 2.5 Beta Distribution Parameters

Probability density function of the beta distribution (pdf):

$$y = f(x|a,b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} I_{[0,1]}(x)$$

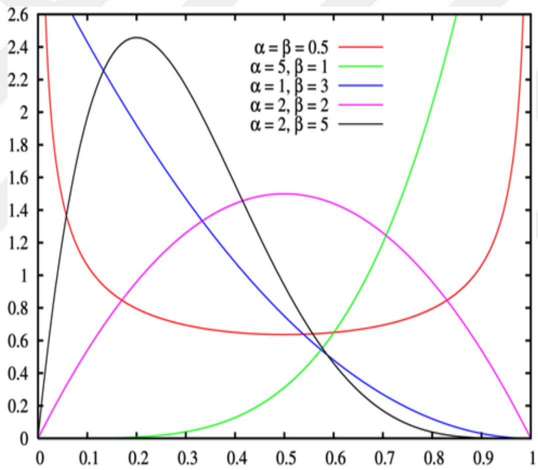

Figure 2.10 Beta Distribution Probability Mass Function

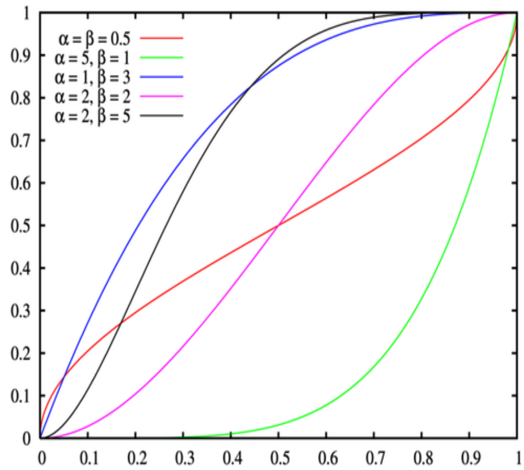

Figure 2.11 Beta Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.5. Weibull distribution

The Weibull distribution is a two-parameter curve family. This distribution is named for Waloddi Weibull, who presented it as a suitable analytical tool for modeling the breaking strength of materials. It includes current use, reliability and lifetime modeling (*Probability Density Function*, 2020). The Weibull distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
a	Scale	a>0
b	Shape	b>0

Table 2.6 Weibull Distribution Parameters

Probability density function of the Weibull distribution (pdf):

$$f(x|a,b) = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} e^{-(x/a)^b}$$

Cumulative distribution function (cdf) of the Weibull distribution:

$$p = F(x|a,b) = \int_0^x b a^{-b} t^{b-1} e^{-\left(\frac{t}{a}\right)^b} dt = 1- e^{-\left(\frac{x}{a}\right)^b}$$

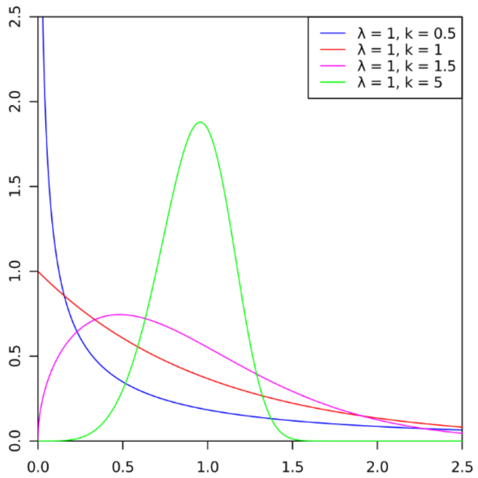

Figure 2.12 Weibull Distribution Probability Mass Function

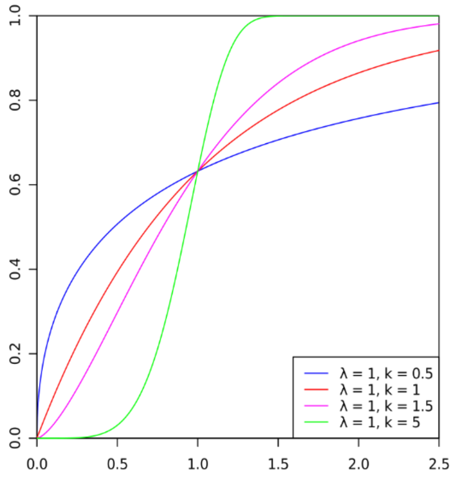

Figure 2.13 Weibull Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.6. Poisson distribution

The Poisson distribution is a single parameter curve family that models the number of times a random event occurs. This distribution is based on a specific time, distance, area, etc. It is suitable for applications involving counting the number of occurrences of a random event in quantities (Probability Density Function, 2020). The Poisson distribution uses the following parameters:

Parameter	Definition	Explanation
lambda (λ)	Mean	$\lambda \geq 0$

Table 2.7 Poisson Distribution Parameters

Probability density function of the Poisson distribution (pdf):

$$f = (x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}; x = 0,1,2,\ldots,\infty$$

Cumulative distribution function of Poisson distribution (cdf):

$$F(x|\lambda) = e^{-\lambda} \sum_{i=0}^{floor(x)} \frac{\lambda^i}{i!}$$

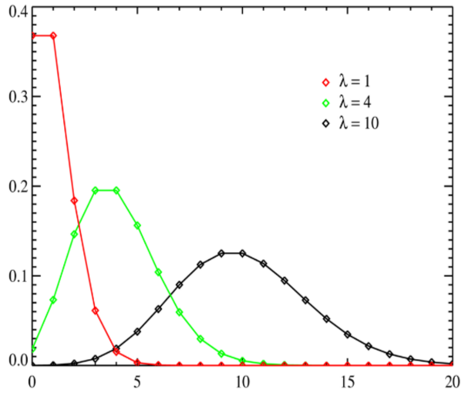

Figure 2.14 Poisson Distribution Probability Mass Function

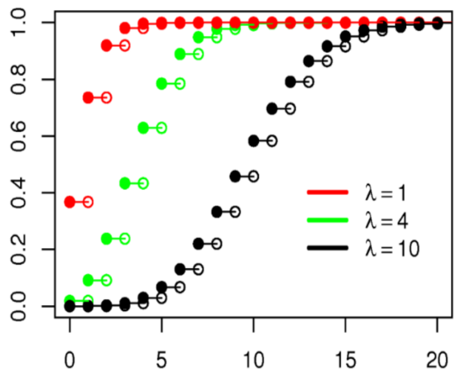

Figure 2.15 Poisson Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.7. Gamma distribution

The gamma distribution is a two-parameter curve family. Gamma distribution models compose the sum of exponentially distributed random variables and generalize both chi-square and exponential distributions (Probability Density Function, 2020). The gamma distribution uses the following parameters:

Parametre	Definition	Explanation
a	Shape	a>0
b	Scale	b>0

Table 2.8 Gama Distribution Parameters

Probability density function of the gamma distribution (pdf):

$$y = f(x|a,b) = \frac{1}{b^a\Gamma(a)} x^{a-1} e^{\frac{-x}{b}}$$

Cumulative distribution function (cdf) of the gamma distribution:

$$p = F(x|a,b) = \frac{1}{b^a\Gamma(a)} \int_0^x t^{a-1} e^{\frac{-t}{b}} dt$$

Mean of the gamma distribution: $\mu = ab$

Variance of the Gamma distribution: $\sigma = ab^2$

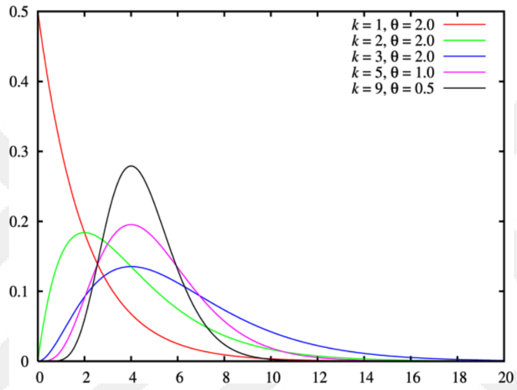

Figure 2.16 Gamma Distribution Probability Mass Function

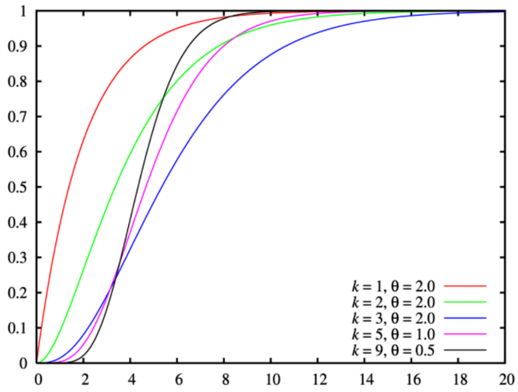

Figure 2.17 Gamma Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.8. Uniform distribution

Uniform distribution; each element of the probability can be found in the same size range for which the probability is supported, a family of probability distributions that show the same constant probability for each continuous value (Probability Density Function, 2020). Uniform distribution uses the following parameters.

Parameter	**Definition**	**Açıklama**
a	Lower endpoint	$-\infty < a < b$
b	Upper endpoint	$a < b < \infty$

Table 2.9 Uniform Definition Parameters

Probability density function of uniform distribution (pdf):

$$f(x|a,b) = \begin{cases} \left(\dfrac{1}{b-a}\right) & ;\ a \leq x \leq b \\ 0 & ;\ \text{otherwise} \end{cases}$$

Cumulative distribution function (cdf) of uniform distribution:

$$F(x|a,b) = \begin{cases} 0 & ;\quad x < a \\ \dfrac{x-a}{b-a} & ;\quad a \leq x < b \\ 1 & ;\quad x \geq b \end{cases}$$

Mean of uniform distribution: $\mu=1/2(a+b)$

Variance of uniform distribution: $\sigma^2 = 1/12(b-a)^2$

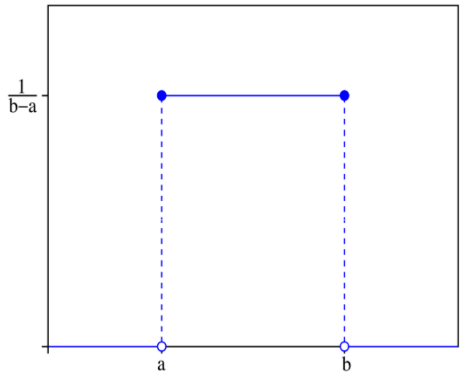

Figure 2.1 Uniform Distribution Probability Mass Function

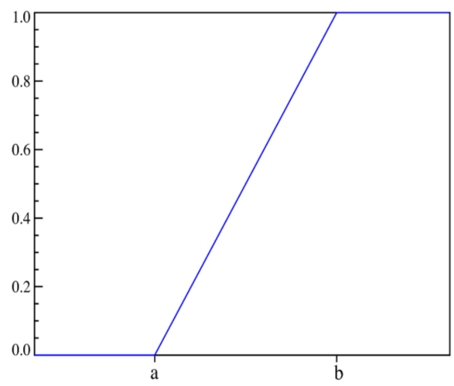

Figure 2.18 Uniform Distribution Cumulative Distribution Function (*Olasılık Dağılımı*, 2020)

2.3.9. Exponential distribution

The exponential distribution is a single parameter curve family. Exponential distribution models expect times when waiting for an additional amount of time is independent of how long you have waited (Probability Density Function, 2020). The exponential distribution uses the following parameters:

Parameter	Definition	Explanation
mu (μ)	Mean	μ > 0

Table 2.10 Exponential Distribution Parameters

Probability density function of the exponential distribution (pdf):

$$y = f(x|\mu) = \frac{1}{\mu} e^{\frac{-x}{\mu}}$$

Cumulative distribution function (cdf) of the exponential distribution:

$$p = F(x|\mu) = \int_0^x \frac{1}{\mu} e^{\frac{-t}{\mu}} dt = 1 - e^{\frac{-x}{\mu}}$$

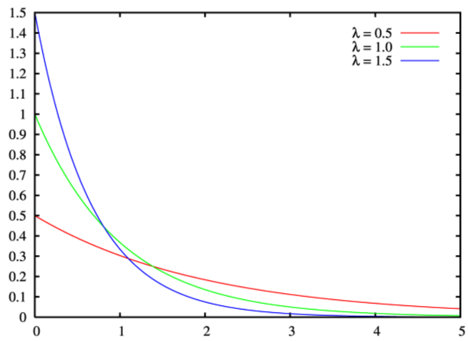

Figure 2.19 Exponential Distribution Probability Mass Function

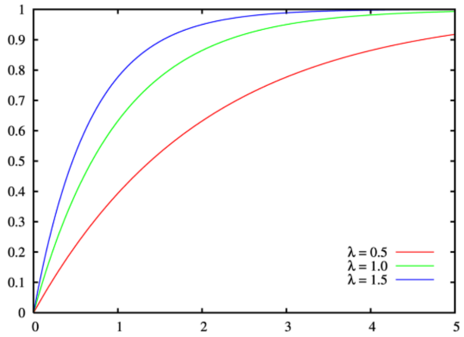

Figure 2.20 Exponential Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.10. Rayleigh distribution

The Rayleigh distribution is a special example of the Weibull distribution. If 'a' and 'b' are the parameters of the Weibull distribution, these parameters are equivalent when a = √2 and b = 2 are given. In communication theory, Nakagami distributions, Rician distributions and Rayleigh distributions are used to model diffuse signals reaching a receiver in multiple ways (Probability Density Function, 2020). Rayleigh distribution uses the following parameters;

Parameter	Definition	Explanation
b	Scale	Positive Integer

Table 2.11 Rayleigh Distribution Parameters

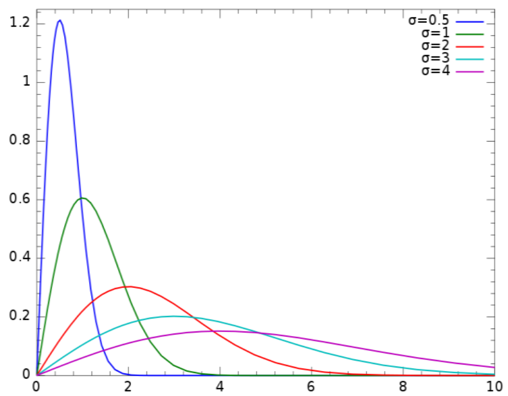

Figure 2.21 Rayleigh Distribution Probability Mass Function

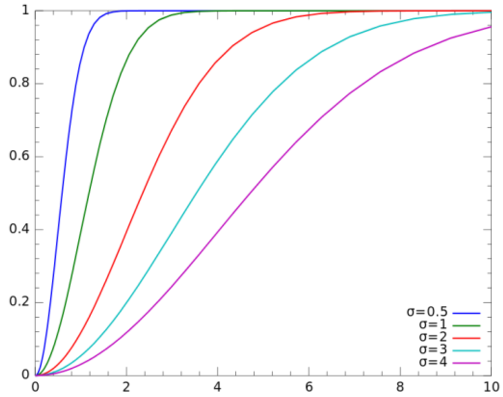

Figure 2.22 Rayleigh Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.11. Lognormal distribution

The lognormal distribution, sometimes called the Galton distribution, is a probability distribution whose logarithm is normally distributed. The lognormal distribution is valid when the amount of interest must be positive, since log (x) exists only when x is positive (Probability Density Function, 2020). The lognormal distribution uses the following parameters:

Parameter	Definition	Explanation
mu (μ)	Average of logarithmic values	$-\infty<\mu<\infty$
sigma (σ)	Standard deviation of logarithmic values	$\sigma\geq 0$

Table 2.12 Lognormal Distribution Parameters

Probability density function of the lognormal distribution (pdf):

$$y = f(x|\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{\frac{-(\log x-\mu)^2}{2\sigma^2}\right\}, \quad \text{for } x > 0$$

Cumulative distribution function (cdf) of the lognormal distribution:

$$p = F(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{1}{t} \exp\left\{\frac{-(\log t-\mu)^2}{2\sigma^2}\right\} dt, \quad \text{for } x > 0$$

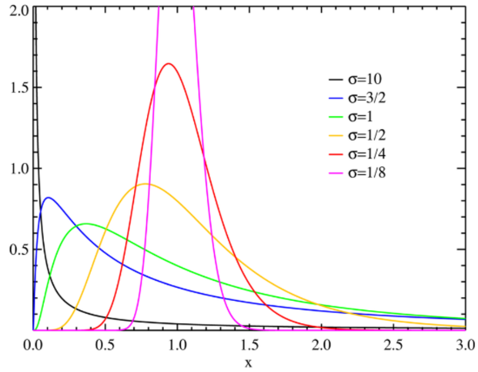

Figure 2.23 Lognormal Distribution Probability Mass Function

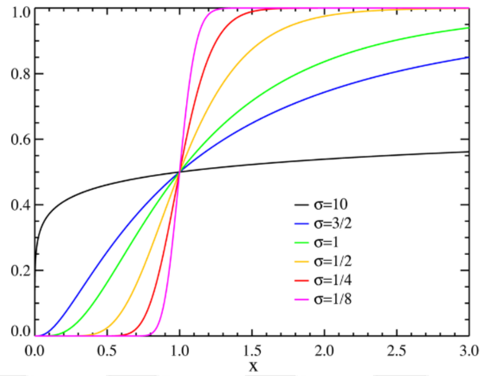

Figure 2.24 Lognormal Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.12. Logistics distribution

It is used for logistic distribution growth models and logistic regression. It has longer tails and higher flatness than normal distribution (Probability Density Function, 2020). Logistics distribution uses the following parameters:

Parameter	Definition	Explanation
mu (μ)	Mean	$-\infty<\mu<\infty$
sigma (σ)	Scale Parameter	$\sigma\geq 0$

Table 2.13 Logistics Distribution Parameters

Probability density function of logistics distribution (pdf):

$$y = f(x|\mu,\sigma) = \frac{\exp\left\{\frac{x-\mu}{\sigma}\right\}}{\sigma(1+\exp\left\{\frac{x-\mu}{\sigma}\right\})^2} \quad ; \quad -\infty < x < \infty$$

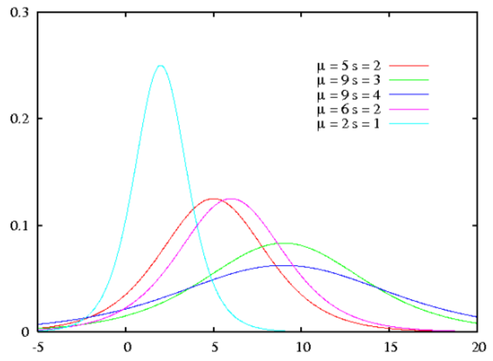

Figure 2.25 Logistics Distribution Probability Mass Function

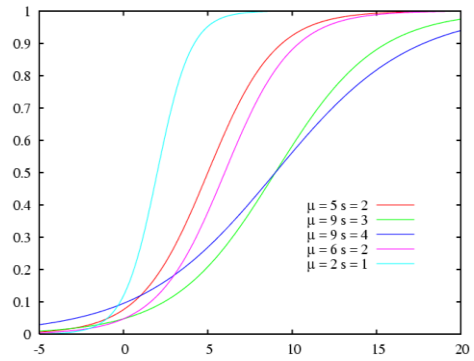

Figure 2.26 Logistics Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.13. Chi-Square distribution

The Chi-Square distribution is a single parameter curve family. The Chi-Square distribution is widely used in hypothesis testing, especially in the chi-square test for good fit (Probability Density Function, 2020). The Chi-Square distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
nu (ν)	Degree of Freedom	ν = 1, 2, 3,...

Table 2.14 Chi-Square Distribution Parameters

Probability density function of the chi-square distribution (pdf):

$$y = f(x|\upsilon) = \frac{x^{(\upsilon-2)/2}e^{-x/2}}{2^{\frac{\upsilon}{2}}\Gamma(\upsilon/2)}$$

Cumulative distribution function (cdf) of the chi-square distribution:

$$p = F(x|\upsilon) = \int_0^x \frac{t^{(\upsilon-2)/2}e^{-t/2}}{2^{\upsilon/2}\Gamma(\upsilon/2)}dt,$$

Mean of the chi-square distribution: ν

Variance of the Chi-square distribution: 2ν

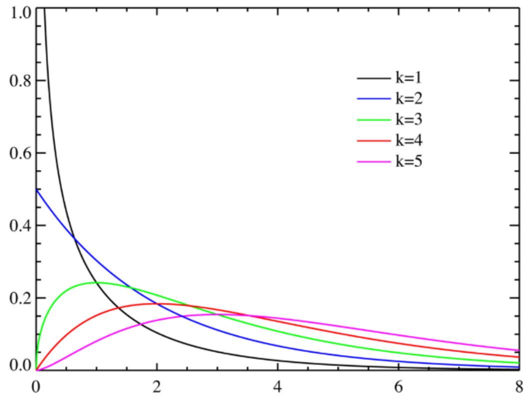

Figure 2.27 Chi-square Distribution Probability Mass Function

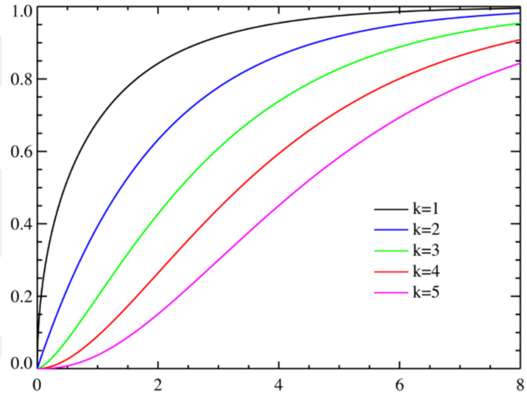

Figure 2.28 Chi-square Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.14. Loglogistic distribution

Loglogistic distribution is a probability distribution whose logarithm is the logistic distribution. This distribution is often used in survival analysis to model events that experienced an initial rate increase and then a rate decrease. It is also known as Fisk distribution in economics applications (Probability Density Function, 2020). The loglogistic distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
mu (μ)	Average of logarithmic values	μ>0
sigma (σ)	Scale parameter of logarithmic values	σ>0

Table 2.15 Chi-Square Distribution Parameters

Probability density function of logistics distribution (pdf):

$$f(x|\mu,\sigma) = \frac{1}{\sigma}\frac{1}{x}\frac{e^z}{(1+e^z)^2} \quad ; \qquad x \geq 0,$$

$$\text{where } z = \frac{\log(x)-\mu}{\sigma}$$

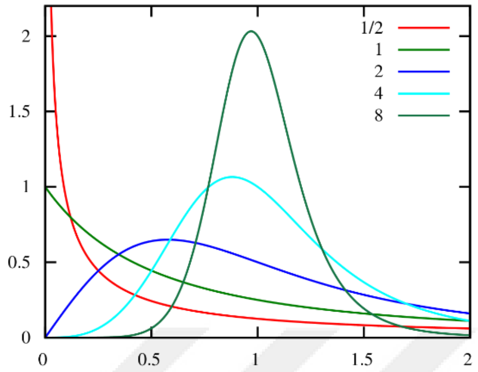

Figure 2.29 Loglogistic Distribution Probability Mass Function

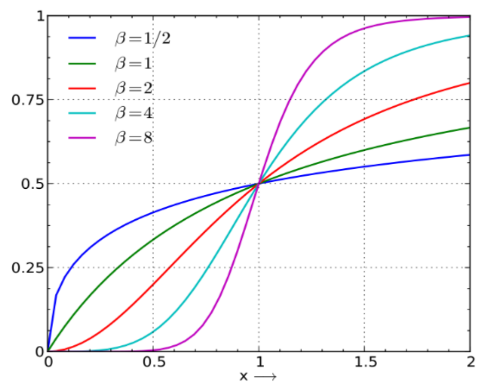

Figure 2.30 Loglogistic Distribution Cumulative Distribution Function(*Olasılık Dağılımı*, 2020)

2.3.15. Student's T distribution

The T distribution is a single parameter curve family. This distribution is typically used to test a hypothesis about the population mean when the population standard deviation is unknown (Probability Density Function, 2020). The T distribution uses the following parameters:

Parameter	**Definition**	**Explanation**
nu (ν)	Degree of Freedom	ν = 1, 2, 3,...

Table 2.16 Student's T Distribution Parameters

Probability density function of the T distribution (pdf):

$$y = f(x|\upsilon) = \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} \frac{1}{\sqrt{\upsilon\pi}} \frac{1}{\left(1+\frac{x^2}{\upsilon}\right)^{\frac{\upsilon+1}{2}}}$$

Cumulative distribution function of T distribution (cdf):

$$p = F(x|\upsilon) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)} \frac{1}{\sqrt{\upsilon\pi}} \frac{1}{\left(1+\frac{t^2}{\upsilon}\right)^{\frac{\upsilon+1}{2}}} dt$$

3. THE USAGE OF DISTRIBUTION FUNCTIONS IN ACQUISITION OF RANDOM NUMBERS

Random numbers are used in many fields. This topic is one of the most practical field of mathematics. Randomized processes take place in both theoretical and real world problems and these processes direct the flow of the problem. So at this point, random number generation becomes a very important issue because according to selection of the random number, the flow of the problem may go to wrong direction or right direction. Therefore, acquisition of random numbers is a subject which is worthy to be examined. At this point, distribution functions come into play. These functions show the probability distribution. Probability distribution as mentioned in section 2.1 defines the values and probabilities for a random event to occur. These values should include all possible outcomes for the event, and the sum of the probabilities must be exactly one or 100 percent. So the distribution functions specify the boundaries and the ruleset of the acquisition of the random numbers that will be used. Therefore, it can be said that the production of random numbers is not a totally random process. Because, probability distribution function selected for the random number generation specifies the characteristics of this process. So when the distribution function changes, the randomization also changes. In general, the uniform distribution function is used in random number generation. This approach reflects also into the programming languages either. So when it is wanted to get a random number, it generally means that this random number will be produced according to the specifications of the uniform distribution. Besides this, generating random number in programming languages also depends another factor called random seed. This seed is set as system clock on default. But this can be changed by the programmer on demand. So the programming language uses a seed as the basis of the random number generation and this generated number is not a real random but it is a pseudo-random number. Therefore, it can be said that random number generation is in real a deterministic process in programming languages. Because of this, always same number can be generated according to selected random seed. So changing seeds must be used in random number generation to not get same numbers. But as it is mentioned above the most important factor in random number generation is distribution function used, not the seed. Firstly, the distribution function specifies the interval where the random number will be selected. Secondly, the probability and cumulative distribution functions also specify the area where the random number is in according to specified parameters. So using only uniform distribution is not reasonable to use for all problems which need random numbers. There is a plenty of distribution functions which can be used. So using the suitable distribution function for a problem in randomness is a valuable issue to examine.

4. THE RANDOMNESS CHARACTERISTICS OF STOCHASTIC OPTIMIZATION METHODS

It is known that the numerical optimization methods can be deterministic or stochastics. The stochastics optimization methods are in our focus in this thesis. The stochastics optimization methods can be inspired from different fields but the main philosophy of these methods is based on random search mechanisms. These methods search the optimal solution according to the way specified by the algorithm. This searching moves ahead in a stochastics manner and this is specified by the randomization process used. This random behavior is a very powerful mechanism for stochastics methods. Because the search space is very wide and finding the optimal solution can be a very hard issue when it is used deterministic methods. But the random behavior in these problems directs the search process into areas where the optimal solution can be found more easily. So minimizing the search fields in each iteration narrows the search space. As it is mentioned above being a powerful mechanism, the randomness is also an uncontrolled notion. It is not based on an analytical background. Although it is known that the engineering wants to control all the aspects of a problem, the randomness is a blurred field and makes the problem solution hard. At this point it can be inferred that the making the randomization process with more analytical methods makes this issue more certain. Our thesis claims that specifying a more suitable distribution function for an algorithm makes randomization in an analytical way and the results become better. So a good mechanism should be made to try different distribution functions in a stochastics method while generating random numbers. This mechanism should also contain trying different parameters for the distribution function. Because the parameters of a distribution function specify the characteristics of this function. So a good combination of parameters with a suitable distribution function regulates the randomization process in an effective way and it is expected to make the optimization algorithm yield more effective results. If this hypothesis is proved with real world problem, its affect can be seen concretely. This hypothesis is put forwarded also for especially for real world problems. Because trying stochastic methods on real world problems in a search space step by step is not suitable for all real world problems. For example, this unsuitability can be sourced by hardware limitations. So there can be a limited set and try move right to reach solution. At this point a more rationalized randomization makes the usage of limited sources; i.e. time, hardware durability; more effective. So this method will diminish the ineffective ways of the stochastics mechanism which stems from the random steps. As a result, the powerful effect of the randomization with a controlled manner will resulted in better results which will not enforce the real world system.

5. USAGE EFFECT OF THE DISTRIBUTION FUNCTIONS IN STOCHASTICS MULTI-PARAMETERS DIVERGENCE OPTIMIZATION

In this section, the usage effect of the distribution functions in stochastics optimization methods is presented. SMDO method is an optimization method of set and trial mechanism. This algorithm is a metaheuristics algorithm which is identified as stochastics because this algorithm arranges its steps through the solution space by a randomized attitude. So the algorithm needs random numbers towards this optimal solution search phase. So it is valuable to examine the reaction of the algorithm against the change in random number acquisition mechanism for the random steps. So the results will give us evidence to infer strong conclusions. In this section firstly the SMDO method will be explained briefly, then the usage of distribution functions in SMDO method will be explained. At this point a practical application "SMDO Benchmark with Distribution Test Program" will be presented. By the help of this application the effectiveness of the different distribution functions in SMDO will be clearly examined by using mostly used benchmark functions. Then the results will be shown in tabular forms and the results will be analyzed.

5.1. The SMDO Method

SMDO is a stochastics method which is a very practical optimization algorithm especially used in real engineering problems. In this algorithm, the solution space is searched by set and trial logic around the last reached point. When the algorithm arrives a point, it specifies the points which can be selected. Then all the specified points are tried and the most suitable one is selected as the next point. At this point the most important factor is making the decision of where to go near around the present location. In the algorithm, this is specified by the randomization. The generated numbers are used for the possible forward or backward steps. This randomization could be done by deterministic steps in this algorithm but the random step method is preferred to not to fall into local optimal point. As it is known falling into local optimal point problem is a general problem for optimization algorithms. This problem prevents the algorithms to make closer to the solution. But this generally stems from the deterministic steps in the algorithms. SMDO uses stochastic method so that in all step decisions the step intervals become different and specified according to present situation. This provides a very flexible movement attitude so that risk of falling into local trap problem diminishes. This randomization is made with random numbers generated by the help of uniform distribution. So the general characteristics of the randomization is specified by this probability distribution. This yields very good results in benchmark functions and especially in real engineering problems. SMDO algorithm is essentially a very analytical algorithm except the randomization process. Because SMDO does all the process very cautiously with forward and backward trials. But in the randomization process all the work is transferred to uniform distribution which is an uncontrolled zone. The flow diagram of the SMDO algorithm is given in the Figure 5.1 below. As it is seen; firstly, the parameter initialization is done; then the error is calculated for the current location. If the error value calculated is below the minimum error specified before, the state is called as tuned state; otherwise the searching procedure goes forward. An assessment around the current location is done according to random step calculated in forward and backward direction and an error comparison is again done this time according to previous location's error value. The value below the error value of previous location specifies the direction of search; being as forward or backward. So as it can be seen from the algorithmic structure the random number used in calculating the step interval plays a crucial role in general flow. If this flow goes to wrong directions the time and accuracy of the algorithm can be effected.

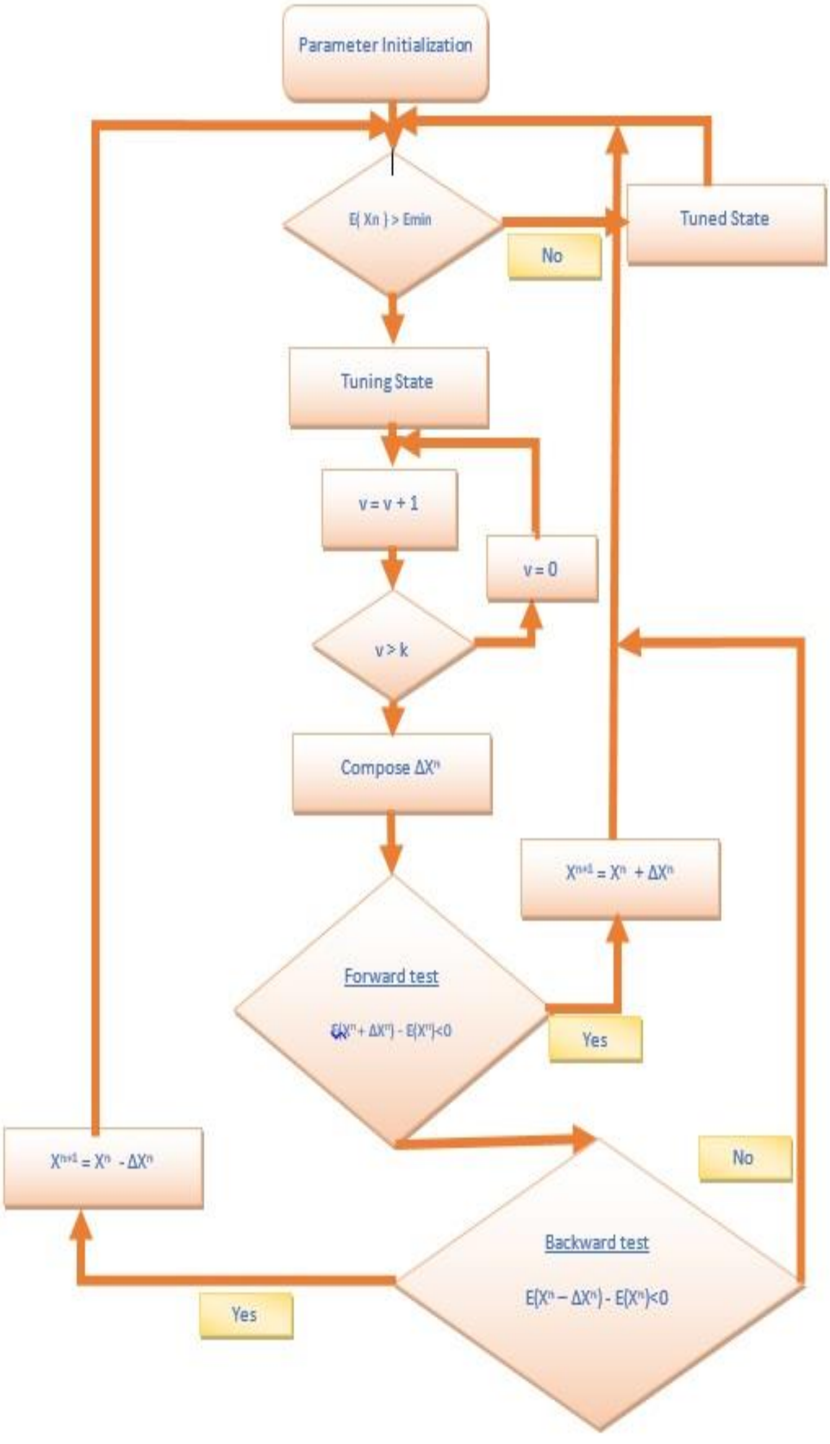

Figure 5.1 Flow Diagram of the SMDO Algorithm (Akpamukçu & Ateş, 2020)

5.2.The Usage of Distribution Functions in SMDO Method

In section 5.1 the importance of the random numbers in SMDO method has been emphasized. As it is mentioned, this randomization is done by the help of uniform distribution function. All engineering activities are based upon the ground of effectiveness. So inspiring from this general engineering discipline, it is claimed in this thesis that the randomization processes can be changed in a manner that the performance of the stochastic optimization algorithms can be enhanced. In this thesis, it is claimed that this aim can be realized by selecting the more appropriate distribution function in the randomization phase of the focused algorithm. This can be done by try and see method. For this aim, the procedure of the SMDO algorithm is slightly changed to try distribution functions in random steps. So measurable results can be taken and this will lead us to decide whether this claim is true. In the Figure 5.2 below the pseudocode of the modified SMDO algorithm is given. The basic change in this algorithm is giving the chance of selecting distribution function that will be used in the algorithm for random number generation. This is a small but a radical change. Because the characteristics of the randomization process is being changed. This selection opportunity is given at the beginning phase and it is used in the step size creation parts. So it is an absolute controlled experiment which holds all the other factors except distribution function used. After getting the results, they will be compared and the effect will be seen. And the reason of the change will be only the distribution function used.

```
Start
 Parameter Initialization
 Distribution Function Selection
 Starting Point
   if E(X^n) > E_Min
   Adaptation; Go to Adapted
        v= v+1;
        if v>parameterCount
          v=0;
        end
        Compose ΔX^n
        Forward test
        Create the step size according to selected distribution function
        if E(X^n +ΔX^n) - E(X^n )<0
           X^(n+1) = X^n + ΔX^n
           Go to Starting Point
        end
        Backward test:
        Create the step size according to selected distribution function
        if E(X^n +ΔX^n) - E(X^n )<0
           X^(n+1) = X^n - ΔX^n
           Go to Starting Point
        end
   Go to Starting Point
   Adopted
   end
End
```

Figure 5.2 Pseudocode of the Modified SMDO Algorithm (Akpamukçu & Ateş, 2020)

5.3. Benchmark Test Functions

In this study; while changing the distribution function for the performance evaluation, mostly used benchmark functions are used. These benchmark functions are firstly tried with the original SMDO algorithm and then with the modified versions of SMDO which only differs from the others from the used distribution function. At this point the benchmark functions play the role of objective function which is the one of the backbone notions of optimization algorithms. In this study, mostly used benchmark functions in optimization algorithms field are selected to get reliable results. These functions give opportunity to algorithm producers and evaluators to meet on a common base. The benchmark functions that is used in this study are; Ackley, Beale, Bohachevsky, Booth, Branin, DixonPrice, Easom, GoldsteinPrice, Griewank, Hump, Levy, Matyas, Perm, Powell, Rastrigin, Rosenbrock, Schwefel, Shubert, Sphere and Zakharov. These benchmark functions are commonly used in academic studies. So that it gives an opportunity to compare the changing results of an optimization algorithm with its modified version as in our case. These functions also give opportunity for us to compare with other algorithms. If we analyze these functions; they are comprised from mathematical formulas. But these formulas are special formulas which creates a curve on the space which takes different values increasing and decreasing. These ups and downs are several in the space in a manner that there are lots of local minimums and maximums. However, among these minimums and maximums, there is one global minima and one global maxima. The aim for an optimization algorithm is to find global minima and the global maxima among these local minimums and maximums. At this point there is a problem. The optimization algorithms generally fall into trap of assuming that an ordinary local minima or local maxima as the global one. This constitutes of a very big problem. So the optimization algorithms should be designed in a way so that they would not fall into these traps. At this point the benchmark functions help the algorithm producers to try their algorithms by providing a difficult medium. They play their role as being the objective function of optimization algorithms. The benchmark functions are comprised of formulas as it is stated above. But they are also defined with the component of dimension range and the optimal value. Dimension specifies the number of variables that a benchmark function uses. Range specifies the values that a benchmark function can take. The optimal value is the optimum value that benchmark function yields when it is in its global maxima or minima. The Table 5.1 below shows the benchmark functions used in this study.

Benchmark Function	Formula	Dim	Range	Optimal Value
Ackley	$f(x) = -a * \exp(\sqrt{\frac{1}{d}\sum_{1}^{d} x_i^2}) - \exp(\sqrt{\frac{1}{d}\sum_{1}^{d}\cos(cx_i)}) + a + \exp(1)$	*d*	[-32.768, 32.768]	0
Beale	$f(x) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2 + (2.625 - x_1 + x_1 x_2^3)^2$	2	[-4.5, 4.5]	0
Bohachecsky	$f_1(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) - 0.4\cos(4\pi x_2) + 0.7$ $f_2(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1)\cos(4\pi x_2) + 0.3$ $f_3(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1 + 4\pi x_2) + 0.3$	2	[-100, 100]	0
Booth	$f(x) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$	2	[-10, 10]	0
Branin	$f(x) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)\cos(x_1) + s$	2	[-5, 10]	0,3978
Dixon Price	$f(x) = (x_1 - 1)^2 + \sum_{i=2}^{d} i(2x_i^2 - x_{i-1})^2$	*d*	[-10, 10]	0
Easom	$f(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2)$	2	[-100, 100]	-1
Goldstein Price	$f(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times [30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	2	[-2, 2]	3
Griewank	$f(x) = \sum_{i=1}^{d}\frac{x_i^2}{4000} - \prod_{i=1}^{d}\cos(\frac{x_i}{\sqrt{i}}) + 1$	*d*	[-600, 600]	0
Hump	$f(x) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2$	2	[-5, 5]	0
Levy	$f(x) = \sin^2(\pi w_1) + \sum_{i=1}^{d-1}(w_i - 1)^2[1 + 10\sin^2(\pi w_i + 1)] + (w_d - 1)^2[1 + \sin^2(2\pi w_d)]$, where $w_i = 1 + \frac{x_i - 1}{4}$, for all i = 1, … , d	*d*	[-10, 10]	0
Matyas	$f(x) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$	2	[-10, 10]	0

Perm	f(x) = $\sum_{i=1}^{d}(\sum_{j=1}^{d}(j^{i}+\beta)((\frac{x_j}{j})^{i}-1))^2$	*d*	[-d, d]	0
Powell	f(x) = $\sum_{i=1}^{d/4}[(x_{4i-3}+10x_{4i-2})^2+5(x_{4i-1}+x_{4i})^2+(x_{4i-2}+2x_{4i-1})^4+10(x_{4i-3}+x_{4i})^4]$	*d*	[-4, 5]	0
Rastrigin	f(x) = $10d+\sum_{i=1}^{d}[x_i^2-10\cos(2\pi x_i)]$	*d*	[-5.12, 5.12]	0
Rosenbrock	f(x) = $\sum_{i=1}^{d-1}[100(x_{i+1}-x_i^2)^2+(x_i-1)^2]$	*d*	[-5, 10]	0
Schwefel	f(x) = $418.9829d-\sum_{i=1}^{d}x_i\sin(\sqrt{\lvert x_i\rvert})$	*d*	[-500, 500]	0
Shubert	f(x) = $(\sum_{i=1}^{5}i\cos((i+1)x_1+i))(\sum_{i=1}^{5}i\cos((i+1)x_2+i))$	*2*	[-5.12, 5.12]	-186,73
Sphere	f(x) = $\sum_{i=1}^{d}x_i^2$	*d*	[-5.12, 5.12]	0
Zakharov	f(x) = $\sum_{i=1}^{d}x_i^2+(\sum_{i=1}^{d}0.5ix_i)^2+(\sum_{i=1}^{d}0.5ix_i)^4$	*d*	[-5, 10]	0

Table 5.1 Properties of Benchmark Functions Used in This Study (Akpamukçu & Ateş, 2020)

5.4. SMDO Benchmark with Distribution Test Program

In order to test SMDO algorithm with distribution functions, a MATLAB toolbox called "SMDO Benchmark with Distribution Test Program" is developed and used. In this program parameter count of benchmark function, iteration number of the SMDO algorithm that will be run, error limit, divergence vector, initialization values, benchmark type and distribution type values are should be given as input. Parameter count is the number of parameters of a benchmark function that will be used. This is specified as the dimension of the benchmark function. Iteration number is the count of run that the SMDO algorithm will be run in an iterated way so as to get mean value of result. This iteration method will decrease the effect of extreme values that has been get. This method is used because of the stochastics characteristics of the SMDO method. Since the stochastic methods do not yield the same values for their each run and the results can fluctuate because of the random numbers generated. Taking the mean value of the results from specified iterated runs will give a fair assessment for an optimization algorithm. The error limit is the threshold value that is used in comparison with each error value calculation for each location reached in the solution space. So this value is used to halt the process. Taking this a value so small will result in a count of running of algorithm so less. Taking this value so big will result in count of running of algorithm so few. So this error limit value should be specified meticulously to reach the optimal value. The divergence vector values are the values used with random numbers to specify step length. So the divergence values for each parameter also play crucial role in stochastic steps. As it is known the random numbers can throw the steps in an unbalanced way. Here the divergence value for each parameter minimizes this effect and makes the solution search in a stable way. Finally, benchmark type information is the benchmark function that will be used and distribution type is the distribution function that will be used. Taking these inputs, the program outputs the mean error values by calculating by the help of SMDO algorithm. This program is a very effective tool to try the benchmark functions as objective functions of the SMDO algorithm with different iterations, different error limits, different initialization values for parameters and from the point view of this study with different distribution functions. So by the help of this tool, analytical and reliable results can be inferred about the usage effect of different distribution functions usages in stochastic optimization algorithms.

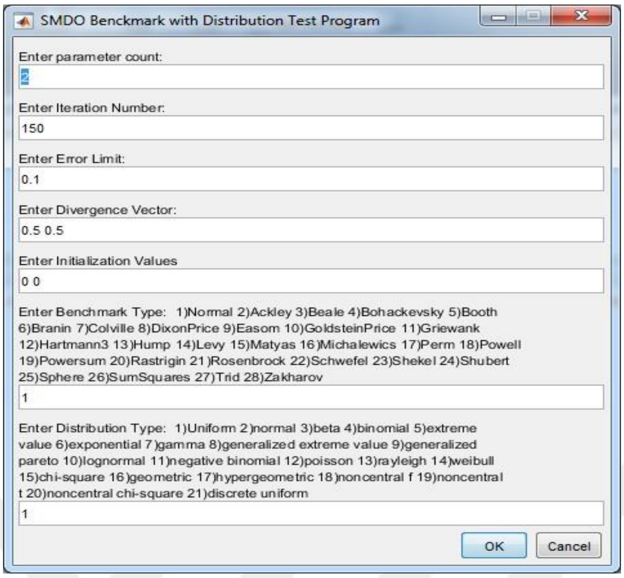

Figure 5.3 GUI of SMDO Toolbox (Akpamukçu & Ateş, 2020)

5.5. The Results of Usage of Distribution Functions in SMDO Method

Since there are a lot of distribution functions, it is adequate to select and see the results of a sample of them. This sampling method is used usually in statistics. In this study, the selected distribution functions are used instead of uniform distribution functions. Using these distribution functions the modified algorithm has been run with the count of iteration number specified. In each iteration a best value has been calculated. Finally using these best values, 1st statistical moment value (mean), 2nd statistical moment value (standard deviation), 3rd statistical moment (skewness) and 4th statistical moment (kurtosis) values are acquired. Besides getting the mean value form these best values; second, third and fourth moment values are calculated to make a deep statistical analysis about the behavior of the algorithm during these searching phase.

The aim of these study is to analyze the usage effect of different distribution functions instead of uniform distribution in SMDO. So the main aim is not measure the efficiency of SMDO. This study resembles a controlled experiment with changing only one component of the system in our case; distribution function used in random number generation. Although it is known that trying to specify some parameters in optimal values in the algorithm to make the results well, an extra effort has not been wasted for this. The main focus is directed to the randomization phase to see the effect of intended issue.

Firstly, the normal distribution function is used in acquiring random numbers parts in the SMDO. This modified version has been run 20 times with an iterated way on 20 different benchmark functions. Before this, the original SMDO algorithm has been run with these 20 benchmark functions and shown in the column of “Classical SMDO with Uniform (Mean)” in Table 5.2. Then the modified SMDO with normal distribution has been run and the results are shown in Table 5.2 in rows for each benchmark function. In the column of “Mean (First Moment)” the mean of the best values is shown. Then in the “Variance (Second Moment)” column the variance calculated from the best values, in the “Skewness (Third Moment)” column the skewness value and in the “Kurtosis (Fourth Moment)” column the kurtosis value calculated from the best values and shown in Table 5.2. As a result, in the benchmark functions; Beale, Booth, Branin, DixonPrice, Easom, GoldsteinPrice, Hump, Levy, Matyas, Perm, Rosenbrock, Schwefel, Shubert; the mean values calculated from the modified SMDO using the normal distribution function are better than the original SMDO using the uniform distribution. So 13 of 20 result shows a good change effect in the optimization algorithm. In 7 of 20 benchmark function which are; Ackley, Bohachecsky, Griewank, Powell, Rastrigin, Sphere, Zakharov; the mean values are worse or near the values of original SMDO. So it can be inferred from these values that there is an effect of using different distributions in randomization phase of an optimization algorithm and this effect is in the positive way. These validations can be also seen in the variance column too. These values are smaller than the first state. This values gives us a good evidence of using different distribution functions.

Secondly the beta distribution function is used in acquiring random numbers parts in the SMDO. This modified version again has been run 20 times with an iterated way on 20 different benchmark functions. Before this, the original SMDO algorithm has been run with these 20 benchmark functions and shown in the column of “Classical SMDO with Uniform (Mean)” in Table 5.3. Then the modified SMDO with beta distribution has been run and the results are shown in Table 5.3 in rows for each benchmark function. In the column of “Mean (First Moment)” the mean of the best values is shown. Then in the “Variance (Second Moment)” column the variance calculated from the best values, in the “Skewness (Third Moment)” column the skewness value and in the “Kurtosis (Fourth Moment)” column the kurtosis value calculated from the best values and shown in Table 5.3. As a result, in the benchmark functions; Ackley, Beale, Booth, Branin, DixonPrice, Easom, GoldsteinPrice, Hump, Levy, Perm, Powell, Rastrigin, Rosenbrock, Schwefel, Zakharov; the mean values calculated from the modified SMDO using the beta distribution function are better than the original SMDO using the uniform distribution. So 15 of 20 result shows a good change effect in the optimization algorithm. In 5 of 20 benchmark function which are; Bohachecsky, Griewank, Matyas, Shubert, Sphere; the mean values are worse or near the values of original SMDO. So it can be inferred from these values that there is an effect of using different distributions in randomization phase of an optimization algorithm and this effect is in the positive way. These validations can be also seen in the variance column too. These values are smaller than the first state. This values also gives us a good evidence of using different distribution functions.

Thirdly the binomial distribution function is used in acquiring random numbers parts in the SMDO. This modified version again has been run 20 times with an iterated way on 20 different benchmark functions. Before this, the original SMDO algorithm has been run with these 20 benchmark functions and shown in the column of “Classical SMDO with Uniform (Mean)” in Table 5.4. Then the modified SMDO with beta distribution has been run and the results are shown in Table 5.4 in rows for each benchmark function. In the column of “Mean (First Moment)” the mean of the best values is shown. Then in the “Variance (Second Moment)” column the variance calculated from the best values, in the “Skewness (Third Moment)” column the skewness value and in the “Kurtosis (Fourth Moment)” column the kurtosis value calculated from the best values and shown in Table 5.4. As a result, in the benchmark functions; Easom, Matyas, Rastrigin, Rosenbrock, Shubert; the mean values calculated from the modified SMDO using the binomial distribution function are better than the original SMDO using the uniform distribution. So 5 of 20 result shows a good change effect in the optimization algorithm. In 15 of 20 benchmark function which are; Ackley, Beale, Bohachecsky, Booth, Branin, DixonPrice, GoldsteinPrice, Griewank, Hump, Levy, Perm, Powell, Schwefel, Sphere, Zakharov; the mean values are worse or near the values of original SMDO. In this distribution function there is not satisfying results but it can be inferred from 5 of 20 benchmark function there is an effect in the positive way. These validations can be also seen in the variance column too. These values are smaller than the first state. So it can be inferred that usage of suitable distribution function is also important. So getting the most suitable distribution function will help us in the improvement of results.

Fourthly the extreme value distribution function is used in acquiring random numbers parts in the SMDO. This modified version again has been run 20 times with an iterated way on 20 different benchmark functions. Before this, the original SMDO algorithm has been run with these 20 benchmark functions and shown in the column of “Classical SMDO with Uniform (Mean)” in Table 5.5. Then the modified SMDO with beta distribution has been run and the results are shown in Table 5.5 in rows for each benchmark function. In the column of “Mean (First Moment)” the mean of the best values is shown. Then in the “Variance (Second Moment)” column the variance calculated from the best values, in the “Skewness (Third Moment)” column the skewness value and in the “Kurtosis (Fourth Moment)” column the kurtosis value calculated from the best values and shown in Table 5.5. As a result, in the benchmark functions; Booth, Branin, DixonPrice, Easom, GoldsteinPrice, Hump, Levy, Perm, Powell, Rosenbrock, Schwefel, Shubert; the mean values calculated from the modified SMDO using the extreme value distribution function are better than the original SMDO using the uniform distribution. So 12 of 20 result shows a good change effect in the optimization algorithm. In 8 of 20 benchmark function which are; Ackley, Beale, Bohachecsky, Griewank, Matyas, Rastrigin, Sphere, Zakharov; the mean values are worse or near the values of original SMDO. So it can be inferred from these values that there is an effect of using different distributions in randomization phase of an optimization algorithm and this effect is in the positive way. These validations can be also seen in the variance column too. These values are smaller than the first state. This values also gives us a good evidence of using different distribution functions.

As we analyze the variance of the values, it can be seen that the uniform distribution does not show a stable performance in this issue. So by having lower variances, the selected distribution functions have stable performance in this subject. This is important for real life where the stable performance is important. So the second moment is a remarkable indicator in real engineering problems. The third and fourth moments can be also important is some aspects in real life but we cannot observe their effect in the benchmark functions.

SMDO with Normal Distribution	*Global Min Value*	*Mean(First Moment)*	*SMDO With Uniform (Mean)*	*Variance(Second Moment)*	*Skewness(Third Moment)*	*Kurtosis(Fourth Moment)*
SMDO with Ackley	0	0.0112	0.0079	0.01241	2.3824	10.2447
SMDO with Beale	0	0.0045	0.0065	0.0107	7.4792	74.6493
SMDO with Bohachecsky	0	0.0020	0.0018	0.0036	2.8920	13.0674
SMDO with Booth	0	0.0003	0.0010	0.0006	4.4293	29.6066
SMDO with Branin	0,3978	0.3910	0.4092	0.0554	-6.9216	48.9486
SMDO with DixonPrice	0	0.0039	0.0351	0.0053	3.3031	18.5576
SMDO with Easom	-1	-8.3330e-08	-2.9777e-08	1.16062e-07	-1.8169	5.3094
SMDO with GoldsteinPrice	3	3.2406	4.0397	0.6616	-0.9759	15.8988
SMDO with Griewank	0	3.1166e-05	2.8866e-05	2.8292e-05	0.6642	2.1408
SMDO with Hump	0	0.00253	0.0220	0.0027	2.3122	10.0579
SMDO with Levy	0	0.0002	0.0038	0.0004	5.1653	36.7470
SMDO with Matyas	0	3.24984e-05	3.3132e-05	3.0056e-05	0.6723	2.1384
SMDO with Perm	0	0.00756	0.0279	0.0093	2.8875	14.8627
SMDO with Powell	0	0.0005	0.0003	0.0013	4.7732	27.9739
SMDO with Rastrigin	0	0.0141	0.0067	0.02999	4.5009	30.5890
SMDO with Rosenbrock	0	0.1392	0.3961	0.1265	1.6300	7.3854
SMDO with Schwefel	0	813.7992	813.7996	115.3717	-6.9296	49.0199
SMDO with Shubert	-186,73	-22.31463	-18.3900	18.4898	-1.1027	2.8943
SMDO with Sphere	0	3.8099e-05	3.6085e-05	3.8897e-05	2.8275	20.0084
SMDO with Zakharov	0	4.5726e-05	4.0279e-05	6.4747e-05	4.3740	31.2018

Table 5.2 Results of SMDO Algorithm with Different Benchmark Functions Using Normal Distribution Function (Akpamukçu & Ateş, 2020)

SMDO with Beta Distribution	***Global Min. Val.***	***Mean(First Moment)***	***SMDO With Uniform (Mean)***	***Variance(Second Moment)***	***Skewness(Third Moment)***	***Kurtosis(Fourth Moment)***
SMDO with Ackley	0	0.0048	0.0079	0.00479	1.5709	5.3299
SMDO with Beale	0	0.0011	0.0065	0.003025	10.2732	128.1170
SMDO with Bohachecsky	0	0.0018	0.0018	0.002896	3.06244	14.8792
SMDO with Booth	0	9.4914e-05	0.0010	0.00013	3.8692	22.8552
SMDO with Branin	0,3978	0.3908	0.4092	0.05541	-6.9237	48.9667
SMDO with DixonPrice	0	0.00099	0.0351	0.001649	3.5398	17.58250
SMDO with Easom	-1	-1.4773e-08	-2.9777e-08	1.0938e-08	-0.8225	2.6445191
SMDO with GoldsteinPrice	3	3.0417	4.0397	0.4760	-4.7354	34.1134
SMDO with Griewank	0	3.3553e-05	2.8866e-05	3.1356e-05	0.5283	1.8592
SMDO with Hump	0	0.00186	0.0220	0.0032	3.1804	13.8229
SMDO with Levy	0	0.0001	0.0038	0.0002	4.35260	24.2882
SMDO with Matyas	0	3.555e-05	3.3132e-05	3.0747e-05	0.5315	1.9206
SMDO with Perm	0	0.00311	0.0279	0.0070	8.2227	86.2100
SMDO with Powell	0	0.0002	0.0003	0.0011	12.2108	163.6566
SMDO with Rastrigin	0	0.00128	0.0067	0.00345	6.15650	46.1952
SMDO with Rosenbrock	0	0.3233	0.3961	0.1038	-0.6704	4.05219
SMDO with Schwefel	0	813.7992	813.7996	115.3717	-6.9296	49.0199
SMDO with Shubert	-186,73	-11.54244	-18.3900	5.0931	0.6507	2.21161
SMDO with Sphere	0	3.3593	3.6085e-05	3.2951e-05	1.3627	7.1360
SMDO with Zakharov	0	2.7623e-05	4.0279e-05	2.9612e-05	2.2103	13.0656

Table 5.3 Results of SMDO Algorithm with Different Benchmark Functions Using Beta Distribution Function (Akpamukçu & Ateş, 2020)

SMDO with Binomial Distribution	***Global Min. Val.***	***Mean(First Moment)***	***SMDO With Uniform (Mean)***	***Variance(Second Moment)***	***Skewness(Third Moment)***	***Kurtosis(Fourth Moment)***
SMDO with Ackley	0	0.1056	0.0079	0.5248	4.7737	23.8915
SMDO with Beale	0	5.942	0.0065	7.0001	0.3370	1.11486
SMDO with Bohachecsky	0	0.02941	0.0018	0.1538	6.7805	58.9932
SMDO with Booth	0	0.1053	0.0010	0.3637	3.9940	19.5653
SMDO with Branin	0,3978	1.1269	0.4092	1.8028	4.5147	26.0063
SMDO with DixonPrice	0	0.4803	0.0351	0.2918	0.8972	4.72660
SMDO with Easom	-1	-2.1170e-08	-2.9777e-08	4.4813e-08	-4.34627	24.4713
SMDO with GoldsteinPrice	3	248.7368	4.0397	2854.2248	14.16455	201.75833
SMDO with Griewank	0	0.01319	2.8866e-05	0.0220	1.7918	6.0395
SMDO with Hump	0	0.2901	0.0220	0.1082	5.2853	38.8661
SMDO with Levy	0	0.1645	0.0038	0.3102	2.9967	15.5586
SMDO with Matyas	0	0.0019	0.0279	0.0110	5.5704	32.0303
SMDO with Perm	0	1.0699	0.0003	1.9015	6.0539	50.1328
SMDO with Powell	0	0.7457	0.0067	4.2443	5.5668	32.0037
SMDO with Rastrigin	0	0.04411	0.3961	0.2058	4.4399	20.7128
SMDO with Rosenbrock	0	64.0465	813.7996	114.9534	1.9502	5.8908
SMDO with Schwefel	0	789.9691	-18.3900	116.8345	-6.0896	41.3975
SMDO with Shubert	-186,73	-14.6498	3.6085e-05	10.8293	-2.6302	10.7316
SMDO with Sphere	0	0.01348	4.0279e-05	0.08305	9.0685	100.8007
SMDO with Zakharov	0	0.01573	0.0079	0.1058	9.7663	113.6100

Table 5.4 Results of SMDO Algorithm with Different Benchmark Functions Using Binomial Distribution Function (Akpamukçu & Ateş, 2020)

SMDO with Extreme Value Distribution	***Global Min. Val.***	***Mean(First Moment)***	***SMDO With Uniform (Mean)***	***Variance(Second Moment)***	***Skewness(Third Moment)***	***Kurtosis(Fourth Moment)***
SMDO with Ackley	0	0.0142	0.0079	0.01537	2.14544	9.2053
SMDO with Beale	0	0.0874	0.0065	0.9954	14.0801	200.12788
SMDO with Bohachecsky	0	0.0023	0.0018	0.0045	3.15584	13.8019
SMDO with Booth	0	0.0003	0.0010	0.00056	3.85149	26.6506
SMDO with Branin	0,3978	0.3915	0.4092	0.05555	-6.9114	48.8571
SMDO with DixonPrice	0	0.00376	0.0351	0.0045	2.5793	12.739
SMDO with Easom	-1	-7.8097e-08	-2.9777e-08	1.1694e-07	-2.0064	6.1050
SMDO with GoldsteinPrice	3	3.5185	4.0397	1.0694	2.41662	21.346
SMDO with Griewank	0	0.00017	2.8866e-05	0.00096	9.1354	90.4217
SMDO with Hump	0	0.003568	0.0220	0.00478	3.01209	14.6803
SMDO with Levy	0	0.0003	0.0038	0.000401	2.92719	15.7236
SMDO with Matyas	0	3.3554e-05	3.3132e-05	3.0966e-05	0.542472	1.93375
SMDO with Perm	0	0.0085	0.0279	0.01061	2.52858	11.58471
SMDO with Powell	0	0	0.0003	0	NaN	NaN
SMDO with Rastrigin	0	0.02889	0.0067	0.0856	6.8884	62.67449
SMDO with Rosenbrock	0	0.1388	0.3961	0.14363	2.27359	10.0489
SMDO with Schwefel	0	813.7992	813.7996	115.3717	-6.929	49.0199
SMDO with Shubert	-186,73	-22.0883	-18.3900	19.2324	-1.19970	2.9960
SMDO with Sphere	0	6.7565e-05	3.6085e-05	0.000192	11.5382	151.334
SMDO with Zakharov	0	0.0003	4.0279e-05	0.00123	7.0040	55.8365

Table 5.5 Results of SMDO Algorithm with Different Benchmark Functions Using Extreme Distribution Function (Akpamukçu & Ateş, 2020)

6. USAGE EFFECT OF THE DISTRIBUTION FUNCTIONS IN MONARCHY BUTTERFLY OPTIMIZATION METHOD

6.1.Monarchy Butterfly Optimization

Monarchy Butterfly Optimization (MBO) algorithm is a stochastics search based algorithm which tries to reach the optimal solution of a problem by imitating migration behavior of monarchy butterflies. This method struggles with the problems which have ambiguity in finding the optimal point in the solution space. Since migration of monarchy butterflies confronts with the difficulties of real life, this is a big motivation to apply this method used by these butterflies to engineering problems which also stem from real life. Especially the random characteristics of this behavior makes possible to create a stochastics numerical algorithm. As it is stated above, the algorithm tries a simulation of the migration behavior and to make this, it formulates this behavior with a specified set of parameters. All butterflies, each of which symbolizes a candidate solution for the problem, are comprised of a specified number of components which are named as variable. These variables represent the values in each dimension of our real problem. So the final aim is to find an excellent butterfly whose variable values will be used for the real problem and this will have resulted in the optimal point in the solution space. To reach this goal, the algorithm takes into account all the life cycle of monarch butterflies which are new births, death, migration between the lands where they in live interchangeably. The realization of this events is not done one-to-one but with a technique of an inspiration. For example, it is assumed that the population is always fixed numbered and a number of top best individuals will pass to next generation. So it is assumed that the individuals in latter generations will provide better variables which will be provide to approach the optimal point. The algorithm tries to reach this aim by applying some procedures and using the power of randomization. The algorithm is mainly based on the procedures of migration operator and butterfly adjusting operator which manage the change of the generations and individuals. The general flow of the algorithm starts with the initialization of the variables of individuals with predetermined values. These variables are the population size of the generation, the maximum generation count, the number of populations in the first land and the second land, the number of the variables in each butterfly, the ratio of monarch butterflies in land 1, maximum step size which shows the maximum walk step that a monarch butterfly can move in one step, butterfly adjusting rate which is used in butterfly adjusting operator, migration period and migration ratio which are used in the migration operator. These initializations direct the flow. So they should be set meticulously. Then a fitness value for each individual is calculated with the variables inside the individuals. A sorting is done according to this evaluation and then all population is divided into two subpopulations like in the real life. At this phase the migration operator is applied to all the individuals of Land 1 iteratively. Then a number called “r” is calculated as in the Eq. 6.1 via the random number “rand” generated with using uniform distribution function and “peri”, holding the value of the migration period, specified in the initialization phase.

$$r = rand * peri \tag{6.1}$$

At this point a comparison is made with the value of “p” called as the ratio of monarch butterflies. According to this comparison if the “r” is less than or equal to “p” an index value is generated to be used as an index to select a butterfly in Population 1. So the variable, which is the correspondent of current processed variable in the iteration, of this butterfly is used for the next generation. If the “r” is bigger than “p”, an index value is generated to be used as an index to select a butterfly in Population 2. So the variable, which is the correspondent of current processed variable in the iteration, of this butterfly is used for the next generation. So the critical parts in migration operator are; random numbers generated by the distribution function, “peri” which is the migration period and the “r” which is the ratio of monarch butterflies. After the end of application of migration operator, the butterfly adjusting operator is applied for all the butterflies in the Population 2. In this phase again a random number is generated with the uniform distribution function without using “peri” parameter. According to comparison of this random number with the “p” called as the ratio of monarch butterflies the correspondent variable value of the best individual is taken. If the generated random number is bigger than the ratio of monarch butterflies, then again a random number is acquired to use for index of the individual whose correspondent variable value will be used for the variable in recent iteration. At this point one more comparison is made with the predetermined value “BAR” and a new generated random number using uniform distribution function to decide whether use Levy flight algorithm to specify the value of variable current processed. If the random number is bigger than “BAR”, maximum step size parameter is used to calculate the weighting factor value which will be used with Levy flight to specify new variable value. So this procedure is repeatedly applied along with a count of specified generation number and finally a set of individuals is generated. Using this list; the top individual, which yields the best solution as being more close to optimal point with its variables, is specified as the optimal solution candidate.

6.2.The Usage of Distribution Functions in MBO Method

The original MBO algorithm, whose parts use random numbers, is based on uniform distribution. This provides the power of randomization to algorithm by being a stochastics algorithm. But here a question of whether using different distribution functions will provide better performance arises. So to try this idea some modifications should be made to algorithm. But these modifications should not change the basic flow of algorithm except changing the process of acquiring random numbers that are needed. For these purposes firstly the points of MBO algorithm where the random numbers used are specified as done in previous section. These critical points are located in migration operator and butterfly adjusting operator. Then a mechanism of selecting the distribution function that will be used in random number generation is provided. So using this mechanism firstly in the migration operator is modified as in algorithm 6.1. The random numbers that are generated are shown as "*randWithSelectedDistFunc*".

Algorithm 6.1: *Modified migration operator*

```
Begin
    for i= 1 to NP1 (for all monarch butterflies in Subpopulation 1) do
        for k=1 to D (all the elements in ith monarch butterfly) do
            Randomly generate a number randWithSelectedDistFunc by selected
            distribution function and its parameters;
            r= randWithSelectedDistFunc *peri;
            if r ≤ p then
                Randomly generate a number randWithSelectedDistFunc by
                selected distribution function and its parameters;
                Randomly select a monarch butterfly by using
                randWithSelectedDistFunc in Subpopulation 1 (say r1);
                Generate the kth element of the x_i^{t+1}
            else
                Randomly generate a number randWithSelectedDistFunc by
                selected distribution function and its parameters;
                Randomly select a monarch butterfly by using
                randWithSelectedDistFunc in Subpopulation 2 (say r2);
                Generate the kth element of the x_i^{t+1}
            end if
        end for k
    end for i
End.
```

Secondly in the butterfly adjusting operator is modified as in algorithm 6.2. The random numbers that are generated are also shown as *randWithSelectedDistFunc* as in migration operator.

Algorithm 6.2: *Butterfly adjusting operator*

Begin
    for j= 1 to NP2 (for all monarch butterflies in Subpopulation 2) do
        Calculate the walk step dx
        Calculate the weighting factor
        for k=1 to D (all the elements in jth monarch butterfly) do
            Randomly generate a number *randWithSelectedDistFunc* by selected distribution function and its parameters;
            if *randWithSelectedDistFunc* $\leq p$ **then**
                Generate the kth element of the x_j^{t+1}
            else
                Randomly generate a number *randWithSelectedDistFunc* by selected distribution function and its parameters;
                Generate the kth element of the x_j^{t+1}
                if *randWithSelectedDistFunc* $\leq BAR$ **then**
                    $x_{j,k}^{t+1} = x_{j,k}^{t+1} + \omega$ x $(dx_k - 0.5)$;
                end if
            end if
        end for ***k***
    end for ***j***
End.

So with these modifications the selected distribution function can be used. Besides this, some other mechanisms are added to MBO algorithm by providing the trial chance of benchmark functions and real engineering problems in a flexible way. The ability of changing the parameters of distribution functions is also added to get the more optimal usage combination of distribution function for the target objective function. So with the use of this method, the possibility of the best combination of distribution function and its parameters in the MBO algorithm that fit to selected objective function is increased. Finally, with the trials of this modified form of MBO, strong evidences will be acquired about the usage effects of using different distribution functions in stochastics algorithms which is the basic claim of this thesis.

6.3.Modified MBO with Benchmark Test Functions

After the necessary modifications have been made on MBO algorithm, 12 benchmark functions have been run with 17 different distribution functions to analyze the performance changes. The results of these runs are shown with some tables below. In these tables in each row the result of Modified MBO's with different distribution functions are shown. In these rows the original MBO result and the results of different optimization algorithms; which are ABC, ACO, BBO, DE, SGA; are also shown that are acquired from the original MBO study to make comparison (G. G. Wang et al., 2019) (A. Ateş & Akpamukçu, 2021). In these tables the results are shown in proportional values to each other. This is done by dividing all the values with the best value in the row. So the value of "1" shows the best value. This method gives simplicity as it is done in the original MBO study. (G. G. Wang et al., 2019)(A. Ateş & Akpamukçu, 2021).

In this study, benchmark functions had been run with M^2BO algorithm via using the distribution functions; Beta, Gamma, Chi-square, Binomial, F, Geometric, Exponential, Extreme Value, Generalized Extreme Value, Generalized Pareto, Student's T, Lognormal, Negative Binomial, Normal, Poisson, Rayleigh, Weibull. Below these runs the results of the basic MBO, ABC, ACO, BBO, DE, SGA algorithms are also given which are taken from the original MBO study. (G. G. Wang et al., 2019)(A. Ateş & Akpamukçu, 2021). The runs are done under the same conditions specified in the original MBO study.

In Table 6.1; Ackley benchmark function is used to analyze. When the results are analyzed it is seen that M^2BO-EXTREMEVALUE yields the best value. It can also be observed that most of the distribution functions apart from the uniform distribution make an enhancement in the result. So the clear effect of the distribution function change in an optimization function can be seen. It can be inferred that the using of proper distribution function can contribute positive effect on the randomization process of a stochastics optimization algorithm by reflecting positive effects on the results.

In Table 6.2; Bohachevsky benchmark function is used to analyze. In this benchmark function the performance of the basic MBO was below the literature as shown in table. The usage of different distribution functions especially SUDENT'S T yields the better value over the classical MBO algorithm. So there is a positive contribution of using a different distribution function over the classical algorithm. So again it can be inferred that the using of proper distribution function can contribute positive effect on the randomization process of a stochastics optimization algorithm by reflecting positive effects on the results.

In Table 6.3; Easom benchmark function is used to analyze. When the results are analyzed, it is seen that M^2BO-LOGNORMAL yields the best value. In this benchmark function the ABC algorithm was the best resulting algorithm. But when the modified MBO with lognormal function is applied to Easom benchmark function, it can be seen that the results are better. So the modified MBO with lognormal has passed not only the classical MBO but also the best performing algorithm. So this application can be also shown as a proof for the usage effect of different distribution functions in stochastics optimization algorithms.

In Table 6.4; Goldsteinprice benchmark function is used to analyze. For this benchmark function DE optimization algorithm was shown as the best performing algorithm in the original study. (A. Ateş & Akpamukçu, 2021). At this point modified MBO with different distribution functions couldn't pass that performance but M^2BO algorithm with student's distribution has passed the performance of classical MBO. So this also shows the enhancement that our thesis claims.

In Table 6.5; Hump benchmark function is used to analyze. In the original study the ABC algorithm yields best results. But as it can be seen from the table M^2BO with Lognormal and M^2BO with Normal algorithms have passed the performance of ABC. So changing the random number acquiring mechanism with using different distribution function made a drastic effect.

In Table 6.6; Matyas benchmark function is used to analyze. M^2BO with Normal yields better results according to basic MBO and other optimization algorithms.

In Table 6.7; Perm benchmark function is used to analyze. From the table it can be seen that M^2BO with Chi-square yields better results according to basic MBO but the results are not good according to DE algorithm. So there is again an improvement in the performance.

In Table 6.8; Rastrigin benchmark function is used to analyze. As it can be seen from the results M^2BO with Rayleigh gets better results according to classical MBO and existing literature. So there is again a very good performance via the distribution function change.

Like Rastrigin benchmark function, the benchmark functions Rosenbrock, Sphere and Zakharov as shown in Table 6.9, Table 6.10 and Table 6.11; M^2BO with Rayleigh produced best result according to classical MBO and existing literature.

In Table 6.12 Shubert benchmark function is used to analyze. At this study, none of MBO variations could get better results according to ABC algorithm but the M^2BO with Rayleigh got better results according to classical MBO.

In all test runs, conditions were kept same. So a fair assessment is tried to done to see clearly the effect of distribution functions. As it can be seen form the tables, this offered modification about the distribution functions in randomization processes of stochastics algorithms makes a great effect. First of all, it can be seen that all the results are better than the classical MBO algorithm. So the tuning of randomization with using proper distribution contribute an enhancement on basic algorithm. This is succeeded with not changing the backbone of algorithm. That is to say only random number acquisition process is changed. The migration process or butterfly adjusting operator mechanism is not changed. So a controlled experiment is realized. Besides this success, for some benchmark functions the modified algorithm overwhelmed not only classical MBO but also the other algorithms in the literature that are mentioned in the original article of MBO. (G. G. Wang et al., 2019)(A. Ateş & Akpamukçu, 2021). So the effect of usage of proper distribution functions in randomization process cannot be ignored. This method provides a good contribution to the search process of stochastics optimization algorithms. So the algorithms go closer to the optimal points. It can be inferred that this method increments the power of algorithms to solve problems.

	Mean	Best	Worst	Parameters
M^2BO-BETA	1,30	8,20	1,00	A=1 B=1
M^2BO-GAMMA	1,38	5,20	1,00	A=1 B=1
M^2BO-CHISQUARE	1,30	9,40	1,00	A=1
M^2BO-BINOMIAL	1,36	4,80	1,01	A = 2 B = 0,5
M^2BO-F	1,23	5,20	1,00	A=1 B=1
M^2BO-GEOMETRIC	1,45	2,60	1,00	A = 0,3
M^2BO-EXPONENTIAL	1,29	16,00	1,00	A=1
M^2BO-EXTREMEVALUE	1,00	19,60	1,00	A=0 B=1
M^2BO-GEN. EXT. VALUE	1,35	20,00	1,00	A=0 B=1 C=0
M^2BO-GEN. PARETO	2,24	6562,40	1,00	A=1 B=1 C=1
M^2BO-STUDENTST	1,24	57,80	1,00	A=1
M^2BO-LOGNORMAL	1,03	4,00	1,00	A=0 B=1
M^2BO-NEG. BINOMIAL	1,35	19,80	1,00	A=3 B=0,5
M^2BO-NORMAL	1,20	6,00	1,00	A=0 B=1
M^2BO-POISSON	1,50	1,00	1,00	A=1 B=0
M^2BO-RAYLEIGH	1,02	10,20	1,00	A=1
M^2BO-WEIBULL	1,24	11,20	1,00	A=1 B=1
MBO (Classical)	1,52	20,00	1,00	
ABC	3,27	2,4E6	1,00	
ACO	3,71	2,8E6	1,00	
BBO	2,02	1,24E6	1,00	
DE	2,95	2,4E6	1,00	
SGA	2,15	1,44E6	1,00	

Table 6.1 ACKLEY with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	62,67	20,22	260,99	A=1 B=1
M^2BO-GAMMA	52,83	16,53	200,69	A=0,9 B=0,9
M^2BO-CHISQUARE	71,58	1,00	335,05	A=1
M^2BO-BINOMIAL	2,46E2	2,63E3	252,29	A=1 B=0,5
M^2BO-F	53,50	1,04E3	294,35	A=2 B=2
M^2BO-GEOMETRIC	63,82	2,67E3	250,37	A=0,3
M^2BO-EXPONENTIAL	34,05	1,87E2	319,43	A=1
M^2BO-EXTREMEVALUE	59,99	7,04E4	338,21	A=4 B=30
M^2BO-GEN. EXT. VALUE	58,46	2,02E4	294,75	A=0 B=1 C=0
M^2BO-GEN. PARETO	2,38E3	3,5E10	248,83	A=1 B=1 C=1
M^2BO-STUDENTST	31,28	7,04E2	341,35	A=2
M^2BO-LOGNORMAL	57,54	1,92E2	203,61	A=0 B=2
M^2BO-NEG. BINOMIAL	57,70	16,97	249,23	A=2 B=0,45
M^2BO-NORMAL	47,22	8,96E3	294,93	A=0 B=5
M^2BO-POISSON	89,31	1,5E3	286,54	A=1
M^2BO-RAYLEIGH	4,12E2	2,86E4	222,89	A=1
M^2BO-WEIBULL	32,87	1,52E2	192,52	A=0,5 B=0,5
MBO (Classical)	69,49	1,62E3	293,62	
ABC	1,00	1,62E3	1,00	
ACO	2,27	1,83E3	1,63	
BBO	5,91	2,11E3	3,55	
DE	1,13	1,62E3	1,15	
SGA	2,01	1,62E3	2,05	

Table 6.2 BOHACHEVSKY with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	1,02	1,00	1,00	A=0,1 B=0,1
M^2BO-GAMMA	1,08	1,00	1,00	A=1,1 B=1,1
M^2BO-CHISQUARE	1,06	1,00	1,00	A=1
M^2BO-BINOMIAL	1,07	1,00	1,00	A=1 B=0,5
M^2BO-F	1,04	1,00	1,00	A=1 B=1
M^2BO-GEOMETRIC	1,04	1,00	1,00	A=0,4
M^2BO-EXPONENTIAL	1,04	1,00	1,00	A=1
M^2BO-EXTREMEVALUE	1,05	1,00	1,00	A=5 B=30
M2BO-GEN. EXT. VALUE	1,08	1,00	1,00	A=0 B=1 C=0
M^2BO-GEN. PARETO	1,53	0,46	1,00	A=1 B=1 C=1
M^2BO-STUDENTST	1,09	1,00	1,00	A=3
M^2BO-LOGNORMAL	1,00	1,00	1,00	A=0 B=10
M^2BO-NEG. BINOMIAL	1,05	1,00	1,00	A=1 B=0,5
M^2BO-NORMAL	1,05	1,00	1,00	A=0 B=5
M^2BO-POISSON	1,05	1,00	1,00	A=1,2
M^2BO-RAYLEIGH	1,08	1,00	1,00	A=0,4
M^2BO-WEIBULL	1,07	1,00	1,00	A=0,3 B=0,5
MBO (Classical)	1,09	1,00	1,00	
ABC	1,05	1.5E3	1,00	
ACO	1,12	2.0E3	1,00	
BBO	1,18	1.2E4	1,00	
DE	1,14	3.2E3	1,00	
SGA	1,10	91.68	1,00	

Table 6.3 EASOM with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	4,65	1,00	6,10	A=1 B=1
M^2BO-GAMMA	4,65	1,00	5,99	A=1 B=1
M^2BO-CHISQUARE	4,65	1,00	6,54	A=1
M^2BO-BINOMIAL	6,27	1,00	5,23	A=1 B=0,5
M^2BO-F	5,64	1,00	7,69	A=1 B=1
M^2BO-GEOMETRIC	14,81	1,03	5,33	A=0,01
M^2BO-EXPONENTIAL	4,46	1,00	6,40	A=1
M^2BO-EXTREMEVALUE	4,05	1,00	6,38	A=0 B=20
M^2BO-GEN.EXT.VALUE	3,89	1,00	5,72	A=0 B=1 C=0
M^2BO-GEN. PARETO	13,70	1,01	4,99	A=1 B=1 C=1
M^2BO-STUDENTST	3,40	1,00	6,68	A=1
M^2BO-LOGNORMAL	4,48	1,00	5,98	A=0 B=1
M^2BO-NEG. BINOMIAL	6,47	1,00	7,41	A=1 B=0,5
M^2BO-NORMAL	3,86	1,00	6,02	A=0 B=1
M^2BO-POISSON	6,49	1,00	6,08	A=1
M^2BO-RAYLEIGH	4,38	1,00	6,12	A=0,51
M^2BO-WEIBULL	4,03	1,00	6,20	A=1 B=1
MBO (Classical)	4,74	1,00	6,13	
ABC	1.43	1,00	1.00	
ACO	1.09	1,00	1.00	
BBO	1.58	1,00	1.97	
DE	1.00	1,00	1.02	
SGA	2.33	1,00	1.00	

Table 6.4 GOLDSTEINPRICE with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	1,28	1,00	277,77	A=1 B=1
M^2BO-GAMMA	1,06	1,00	287,40	A=1 B=1
M^2BO-CHISQUARE	1,43	1,00	155,16	A=1
M^2BO-BINOMIAL	2,80	1,00	440,37	A=1 B=0,5
M^2BO-F	1,40	1,00	223,63	A=2 B=2
M^2BO-GEOMETRIC	4,19	1,00	234,01	A=0,55
M^2BO-EXPONENTIAL	1,28	1,00	271,31	A=0,5
M^2BO-EXTREMEVALUE	1,39	1,00	97,80	A=4 B=30
M^2BO-GEN. EXT. VALUE	1,05	1,00	504,62	A=0 B=1 C=0
M^2BO-GEN. PARETO	84,54	591	594,86	A=1 B=1 C=1
M^2BO-STUDENTST	1,03	1,00	207,74	A=1
M^2BO-LOGNORMAL	1,00	1,00	89,62	A=0 B=10
M^2BO-NEG. BINOMIAL	1,45	1,00	343,25	A=0,6 B=0,125
M^2BO-NORMAL	1,00	1,00	154,02	A=0 B=3
M^2BO-POISSON	1,15	1,00	264,05	A=1
M^2BO-RAYLEIGH	1,33	1,00	302,87	A=0,4
M2BO-WEIBULL	1,36	1,00	927,43	A=0,5 B=0,5
MBO (Classical)	1,49	1,00	303,21	
ABC	1,25	1,00	1.00	
ACO	1,26	1,00	1.00	
BBO	1,28	1,00	1.00	
DE	1,25	1,00	1.00	
SGA	1,26	1,00	303,21	

Table 6.5 HUMP with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	1,42	1,92E3	2,13	A=1 B=1
M^2BO-GAMMA	1,12	1,17E3	1,45	A=0,9 B=0,9
M^2BO-CHISQUARE	1,17	5,31E3	1,65	A=1
M^2BO-BINOMIAL	2,52	28,37	1,71	A=1 B=0,5
M^2BO-F	1,44	5,39E3	1,35	A=1 B=1
M^2BO-GEOMETRIC	11,58	4,61E5	1,85	A=0,01
M^2BO-EXPONENTIAL	1,21	1,68E2	1,50	A=0,5
M^2BO-EXTREMEVALUE	1,26	6,55	1,34	A=4 B=30
M^2BO-GEN. EXT. VALUE	1,41	4,77E2	1,71	A=0 B=1 C=0
M^2BO-GEN. PARETO	15,88	6,11E7	1,31	A=1 B=1 C=1
M^2BO-STUDENTST	1,08	1,00	1,88	A=1
M^2BO-LOGNORMAL	1,39	87,99	1,27	A=0 B=5
M^2BO-NEG. BINOMIAL	2,02	9,61E3	1,74	A=1 B=0,5
M^2BO-NORMAL	1,00	2,95E2	1,53	A=0 B=1
M^2BO-POISSON	2,45	1,43E5	1,81	A=1
M^2BO-RAYLEIGH	1,33	32,10	1,36	A=0,3
M^2BO-WEIBULL	1,27	36,13	1,34	A=1 B=1
MBO (Classical)	1,55	319,80	1,08	
ABC	1,53	319,80	1.01	
ACO	1,53	319,80	1.00	
BBO	1,55	319,80	1.00	
DE	1,52	319,80	1.00	
SGA	1,53	319,80	2.05	

Table 6.6 MATYAS with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	1,17E5	3,42E3	4,72E5	A=1,2 B=1,2
M^2BO-GAMMA	1,16E5	1,12E4	4,31E5	A=1 B=1
M^2BO-CHISQUARE	1,02E5	9,66E3	2,69E5	A=1
M^2BO-BINOMIAL	2,57E5	6,51E3	4,73E5	A=1 B=0,5
M^2BO-F	1,20E5	1,22E4	4,39E5	A=1 B=1
M2BO-GEOMETRIC	9,56E5	8,90E4	3,40E5	A=0,01
M^2BO-EXPONENTIAL	1,29E5	6,97E4	6,33E5	A=1
M^2BO-EXTREMEVALUE	6,37E4	4,61E4	2,27E5	A=0 B=1
M^2BO-GEN. EXT. VALUE	1,35E5	1,31E4	4,40E5	A=0 B=1 C=0
M2BO-GEN. PARETO	2,76E5	1,05E5	4,01E5	A=1 B=1 C=1
M^2BO-STUDENTST	7,89E4	1,02E4	2,77E5	A=1
M^2BO-LOGNORMAL	1,17E5	9,76E3	2,64E5	A=0 B=1
M2BO-NEG. BINOMIAL	2,69E5	4,58E4	3,75E5	A=1 B=0,5
M^2BO-NORMAL	9,11E4	7,37E3	6,70E5	A=0 B=1
M2BO-POISSON	3,34E5	5,39E4	3,08E5	A=1
M2BO-RAYLEIGH	1,76E5	1,02E3	1,96E5	A=1
M^2BO-WEIBULL	1,09E5	1,26E5	2,61E5	A=0,6 B=0,6
MBO (Classical)	1,4E5	6.0E3	7.8E5	
ABC	7,3E4	2.8E4	1.6E6	
ACO	5,2E4	5.9E8	7.8E5	
BBO	4,1E5	5.9E8	1.0E8	
DE	1,00	1.00	1.00	
SGA	1,1E4	5.9E8	7.8E5	

Table 6.7 PERM with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	2,09	5,85	3,07	A= 0,1 B=0,1
M^2BO-GAMMA	1,99	9,71	3,13	A=1 B=1
M^2BO-CHISQUARE	2,19	46,76	3,08	A=1
M^2BO-BINOMIAL	3,11	5,27	3,09	A=1 B=0,5
M^2BO-F	1,85	8,06	3,11	A=1 B=1
M2BO-GEOMETRIC	1,86	8,76E2	3,15	A=0,01
M^2BO-EXPONENTIAL	2,01	2,63	3,14	A=1
M^2BO-EXTREMEVALUE	1,50	22,73	3,21	A=0 B=1
M^2BO-GEN. EXT. VALUE	2,20	17,92	3,22	A=0 B=1 C=0
M2BO-GEN. PARETO	5,48	7,72E4	3,06	A=1 B=1 C=1
M^2BO-STUDENTST	1,78	72,54	3,12	A=1
M^2BO-LOGNORMAL	1,30	24,40	3,05	A=0 B=1
M^2BO -NEG. BINOMIAL	3,56	1,00	3,05	A=1 B=0,5
M^2BO-NORMAL	2,10	13,79	3,13	A=0 B=1
M2BO-POISSON	1,30	9,10	3,02	A=2
M2BO-RAYLEIGH	1,00	34,94	3,09	A=1
M^2BO-WEIBULL	2,06	16,27	3,11	A=1 B=1
MBO (Classical)	2,63	52,25	3,10	
ABC	6,78	1,88E8	3,97	
ACO	12,52	5,17E8	7,89	
BBO	2,71	7,84E7	1,00	
DE	11,02	4,70E8	5,94	
SGA	3,60	1,1E8	1,77	

Table 6.8 RASTRIGIN with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	4,30	1,00	1,13	A=0,1 B=0,1
M^2BO-GAMMA	4,93	31,83	1,15	A=1 B=1
M^2BO-CHISQUARE	5,41	39,09	1,13	A=1
M^2BO-BINOMIAL	9,16	1,42	1,00	A=1 B=0,5
M^2BO-F	5,08	39,90	1,06	A=1 B=1
M^2BO-GEOMETRIC	4,65	1,57E3	1,16	A=0,01
M^2BO-EXPONENTIAL	4,95	86,72	1,11	A=1
M^2BO-EXTREMEVALUE	3,69	20,60	1,20	A=0 B=1
M^2BO-GEN. EXT. VALUE	6,62	39,39	1,09	A=0 B=1 C=0
M^2BO-GEN.PARETO	15,40	1,23E5	1,17	A=1 B=1 C=1
M^2BO-STUDENTST	6,31	18,98	1,05	A=1
M^2BO-LOGNORMAL	2,54	74,89	1,12	A=0 B=1
M^2BO-NEG. BINOMIAL	8,18	5,29	1,12	A=1 B=0,5
M^2BO-NORMAL	5,58	72,57	1,11	A=0 B=1
M^2BO-POISSON	1,83	45,89	1,09	A=2
M^2BO-RAYLEIGH	1,00	1,15E2	1,04	A=1
M^2BO-WEIBULL	4,17	4,05	1,11	A=1 B=1
MBO (Classical)	6,95	16,82	1,08	
ABC	43,36	3,87E6	27,89	
ACO	249,39	1,85E7	87,66	
BBO	14,87	1,19E6	10,18	
DE	38,63	4,04E6	22,95	
SGA	15,36	1,4E6	5,68	

Table 6.9 ROSENBROCK with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	9,32	26,14	18,76	A=1 B=1
M^2BO-GAMMA	8,07	1,40	19,32	A=1 B=1
M^2BO-CHISQUARE	7,57	4,71	19,12	A=1
M^2BO-BINOMIAL	16,85	4,31	20,25	A=1 B=0,5
M^2BO-F	6,45	1,00	20,08	A=1 B=1
M^2BO-GEOMETRIC	7,90	50,83	19,60	A=0,01
M^2BO-EXPONENTIAL	6,59	18,64	19,16	A=1
M^2BO-EXTREMEVALUE	4,28	31,36	19,72	A=0 B=1
M^2BO-GEN. EXT. VALUE	8,00	6,18	20,23	A=0 B=1 C=0
M^2BO-GEN. PARETO	23,49	3,83E5	19,74	A=1 B=1 C=1
M^2BO-STUDENTST	7,70	10,11	20,89	A=1
M^2BO-LOGNORMAL	3,07	1,25	20,31	A=0 B=1
M^2BO-NEG. BINOMIAL	16,17	4,65	20,32	A=1 B=0,5
M^2BO-NORMAL	5,86	42,28	19,71	A=0 B=1
M^2BO-POISSON	3,32	1,03	19,96	A=2
M^2BO-RAYLEIGH	1,00	18,00	20,08	A=1
M^2BO-WEIBULL	5,98	6,78	20,31	A=1 B=1
MBO (Classical)	11,29	2,47	21,21	
ABC	45,43	7,40E7	1,00	
ACO	117,52	1,73E8	5,44	
BBO	8,89	7,89E6	21,21	
DE	22,93	4,93E7	21,21	
SGA	9,60	1,01E7	21,21	

Table 6.10 SPHERE with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	2,06	44,24	1,65	A=1 B=1
M^2BO-GAMMA	1,93	2,47E2	2,01	A=1 B=1
M^2BO-CHISQUARE	1,90	31,54	2,26	A=1
M^2BO-BINOMIAL	3,05E2	5,84	2,34	A=1 B=0,5
M^2BO-F	1,83	23,87	1,22	A=1 B=1
M^2BO-GEOMETRIC	5,88	9,40E4	3,19	A=0,01
M^2BO-EXPONENTIAL	1,81	89,19	2,67	A=1
M^2BO-EXTREMEVALUE	1,61	9,44	1,42	A=0 B=1
M^2BO-GEN. EXT. VALUE	2,13	117,56	2,05	A=0 B=1 C=0
M^2BO-GEN. PARETO	2,06	8,99E4	2,14	A=1 B=1 C=1
M^2BO-STUDENTST	1,92	38,21	1,80	A=1
M^2BO-LOGNORMAL	1,50	1,75E2	3,78	A=0 B=1
M^2BO-NEG. BINOMIAL	2,49E2	1,00	2,13	A=1 B=0,5
M^2BO-NORMAL	2,07	8,42	1,62	A=0 B=1
M^2BO-POISSON	1,59	5,430E3	1,00	A=2
M^2BO-RAYLEIGH	1,00	1,97E2	3,74	A=1
M^2BO-WEIBULL	1,85	17,18	2,33	A=1 B=1
MBO (Classical)	2,23	55,64	1,77	
ABC	1,93	2,23E7	5,00	
ACO	1,46E2	1,72E7	2,04	
BBO	1,12	9,46E6	2,82	
DE	2,34	3,28E7	5,16	
SGA	1,76	1,45E7	30,38	

Table 6.11 ZAKHAROV with M^2BO (A. Ateş & Akpamukçu, 2021)

	Mean	Best	Worst	Parameters
M^2BO-BETA	63,69	29,00	1,99	A=1 B=1
M^2BO-GAMMA	80,60	1,00	1,77	A=1 B=1
M^2BO-CHISQUARE	50,25	1,00	1,42	A=1
M^2BO-BINOMIAL	244,56	39,00	1,78	A=2 B=0,7
M^2BO-F	74,69	9,00	1,54	A=1 B=1
M^2BO-GEOMETRIC	77,95	19,00	1,65	A=0,55
M^2BO-EXPONENTIAL	74,56	9,00	1,61	A=1,5
M^2BO-EXTREMEVALUE	72,14	9,00	1,19	A=0 B=10
M^2BO-GEN. EXT.VALUE	68,74	49,00	1,21	A=0 B=1 C=0
M^2BO-GEN. PARETO	701,42	5,42E4	1,62	A=1 B=1 C=1
M^2BO-STUDENTST	112,01	9,00	2,08	A=1
M^2BO-LOGNORMAL	64,25	9,00	1,28	A=0 B=2
M^2BO-NEG. BINOMIAL	115,86	1,00	1,00	A=1 B=0,5
M^2BO-NORMAL	79,74	1,00	1,84	A=0 B=5
M^2BO-POISSON	81,61	1,00	1,41	A=1
M^2BO-RAYLEIGH	196,77	9,00	1,31	A=1
M^2BO-WEIBULL	69,30	9,00	1,95	A=0,6 B=0,6
MBO (Classical)	82,54	29,00	1,62	
ABC	1,00	73,66	1,62	
ACO	63,93	1,31E4	35,41	
BBO	34,42	1,31E4	1,62	
DE	31,31	6,89E3	4,55	
SGA	2,76	1,31E4	1,62	

Table 6.12 SHUBERT with M^2BO (A. Ateş & Akpamukçu, 2021)

6.4.Tuning of 3 DOF Hover System Controller Parameters with Modified MBO

In this study, in order to apply the topic that is claimed in this thesis to a real engineering problem, the 3 Degree of Freedom Hover system is used. This system simulates the flight control experiment which is produced by Quanser. It can simulate the quadcopter flight system. This simulation is done by trying to act as if using the pitch, roll and yaw angles of the real system. To simulate these angles, the system uses multi input signals. At this point a K feedback gain matrix is used to control the simulation. This control is realized by using these multiple input signals in the simulation. So the main focus is on the K feedback gain matrix if the control performance would be enhanced. In the basic model LQR control structure is used to get the K matrix. In this model quadratic Riccati Equations, shown in Eq. 6.2, are solved according to Q and R symmetry matrices and the control signal is acquired shown in Eq. 6.3.(A. Ateş & Akpamukçu, 2021)

$$J = \frac{1}{2}\int_0^{\infty}\left[X^T Q X + U^T R U\right]dt \qquad (6.2)$$

$$U = KE \qquad (6.3)$$

K represents feedback gain matrix. E feedback errors from sensors.

After calculating the K matrix, the values are applied to system and the control is ensured. Since this calculation phase is complicated, in the literature the idea of using optimization algorithms came into prominence. In the literature for example the SMDO and DSO algorithms are used for this problem. In this study, as introduced in the above chapter M^2BO algorithm will be used by using different distribution functions to see the effect of the claims of this thesis on a real engineering problem. The results will be compared with classical MBO, SMDO and DSO.

6.4.1. K feedback gain matrix optimization with M^2BO for simulation model of 3 DOF hover:

In this part, K feedback gain matrix will be optimized with the M^2BO algorithm with using different distribution functions. The entities that will be optimized will be the four different parameters of K feedback gain matrix. The optimized parameters will be applied to system and the result changes will be observed. The flow chart of the proposed algorithm shown in Fig. 6.2 and the algorithm structure is shown in Fig. 6.1.

In this algorithmic structure the distribution function that will used should be selected first. In this study seventeen distribution function will be used to observe the change effect. After each run, the 4 parameters for the K feedback gain matrix are acquired to apply to 3 DOF Hover System.

It is known that all optimization problems need an objective function to minimize or maximize. In this problem since there are multi inputs and outputs, six error function values, that are calculated according to mean square error. A weighting coefficient is assigned with each error and a new formula of error, which is a combination of these six error functions, is formed as objective function. This function is called as multi-objective function (MOF). The formula is given in Eq. 6.4.(A. Ateş & Akpamukçu, 2021)

$$MOF = w1MSE1+w2MSE2+w3MSE3+w4MSE4+w5MSE5+w6MSE6 \quad (6.4)$$

In this formula w1, w2, w3, w4, w5, w6 are the weighting factors. MSE1, MSE2, MSE3, MSE4, MSE5, MSE6 are six mean square error functions that constitutes the MOF. In this study the weighting factors are set as “1”. So the main aim is to specify parameters that form the K matrix to minimize MOF. Table 34 shows the results that are acquired from the runs of M^2BO to gain that parameters.

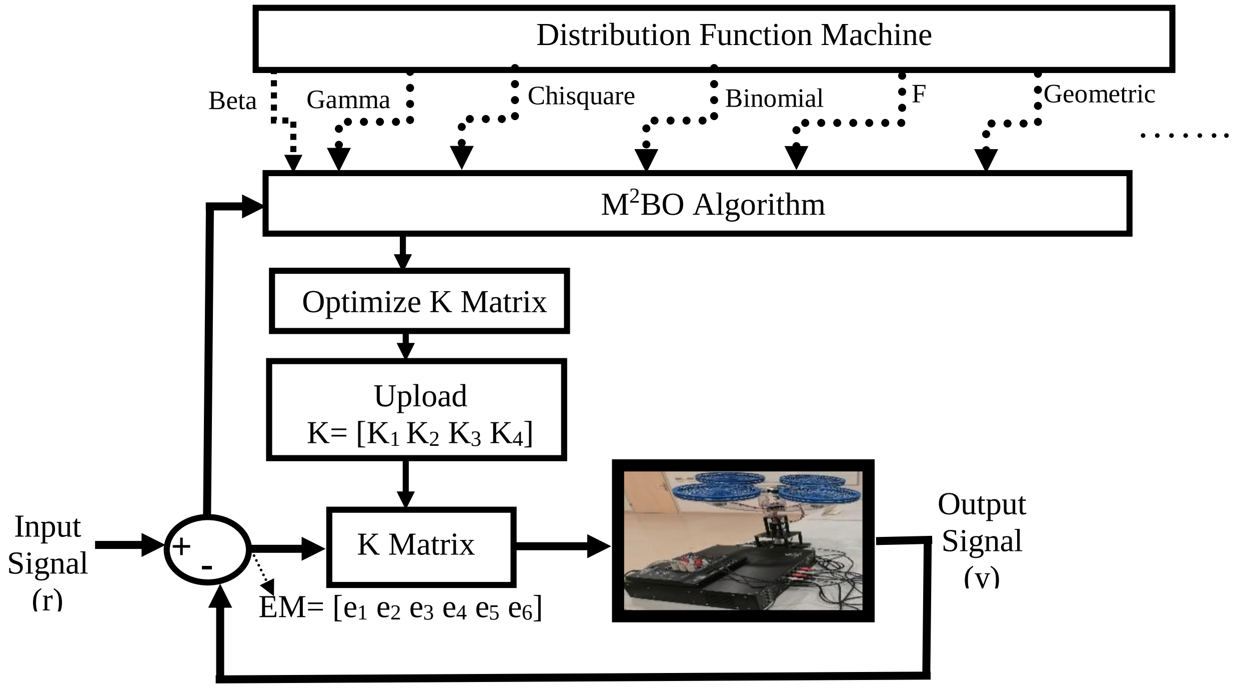

Figure 6.1 Optimization process of the 3 DOF Hover (A. Ateş & Akpamukçu, 2021)

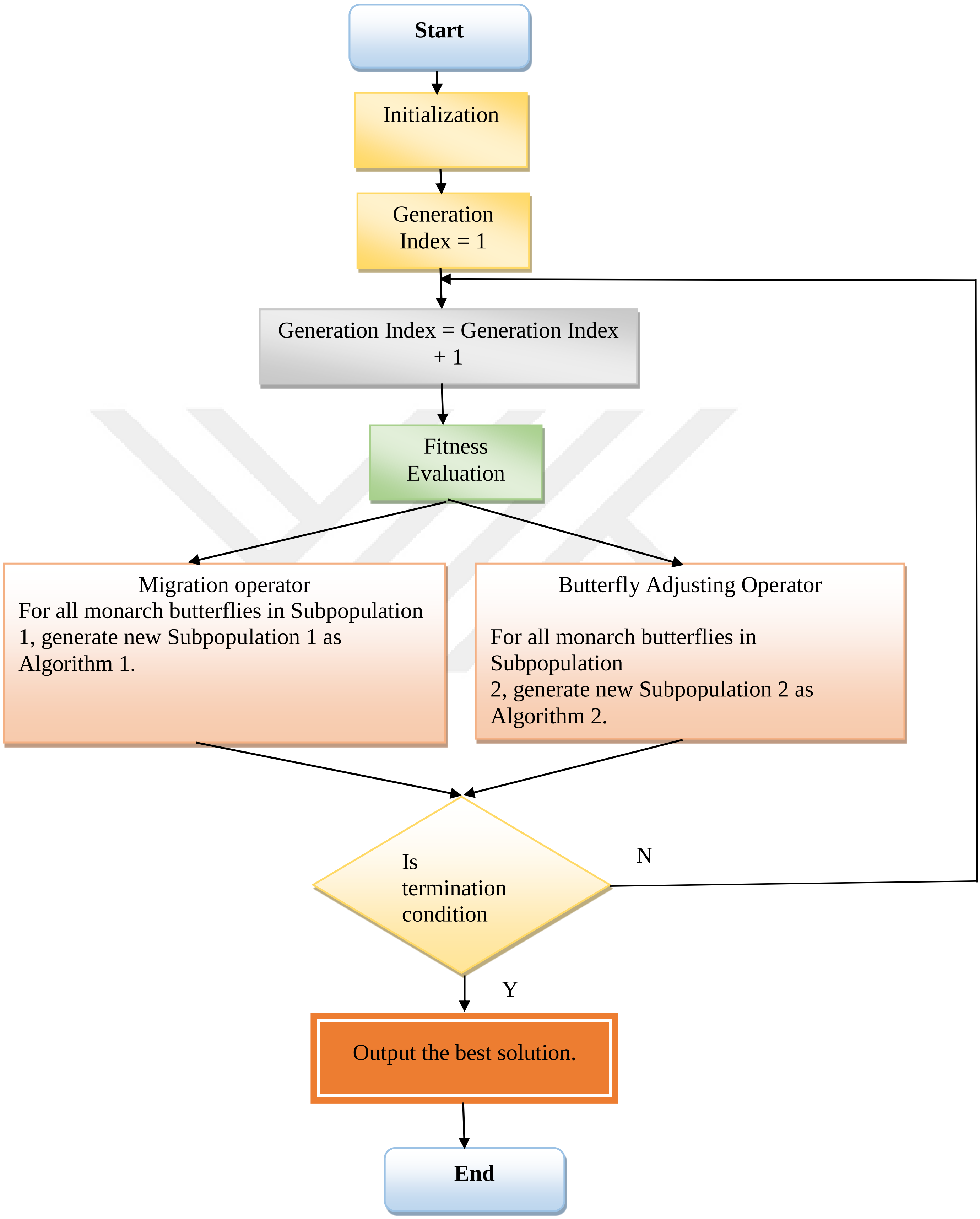

Figure 6.2 Flow chart of M²BO Algorithm (A. Ateş & Akpamukçu, 2021)

It is seen in the Table 6.13 that 17 different distribution functions are used with M^2BO. The K feedback gain matrix is formed with each parameter combination and applied to the simulation model of 3 DOF Hover system. As it is known, the results of SMDO and DSO for this system was obtained in (A. Ateş & Akpamukçu, 2021). By the way, the results for the classical MBO are also derived. Now in Figure 6.3, Figure 6.4 and Figure 6.5; the pitch, roll and yaw angles are depicted respectively.

In Figure 6.3 pitch angle responses are depicted. Firstly, it can be seen that classical MBO (red-dot) is very successful according to other optimization algorithms. It is known that all optimization algorithms are not suitable for a particular problem. But here it can be seen that MBO is very effective for 3 DOF Hover system controller problem. On the other hand, it can be seen from the Figure 6.3 that M^2BO algorithm with different distribution functions is better than the classical MBO. Especially M^2BO with Normal Distribution is the best performing algorithm. So it can be inferred that this concept of changing randomization via using different distribution functions is applicable not only for benchmark functions but also for the real engineering problems. This effect can also have seen from the settling and rising times too.

Likewise, the pitch angle, the roll angles also show the positive effect. As it is seen in Table 6.13 classical MBO overwhelmed the DSO and SMDO. It shows a very good performance as it is seen in Fig. 6.4. But M^2BO (red-dot), especially M^2BO with Normal Distribution, again shows a superior performance.

For the yaw angle the results are closer but M^2BO with Generated Pareto optimization has a better rise and settling time performance as it is seen in the Figure 6.5 and Table 6.13.

As a result, it can be said that MBO is a suitable algorithm for real engineering problems. Although this algorithm shows a good performance, better performances can be acquired by tuning the randomization process. In the result the net effect of this change is seen by the use of M^2BO algorithm with different distribution functions. This study is also important from the point of view that, this concept can enhance not only benchmarks but the real engineering problems.

	K1	K2	K3	K4	Dist. Parameters
M^2BO-BETA	57.8398	166.2913	41.4616	61.8107	A=1 B=1
M^2BO-GAMMA	59.7688	165.6915	42.2567	63.6452	A=1 B=1
M^2BO-CHISQUARE	58.0510	167.1579	41.1612	61.6889	A=1
M^2BO-BINOMIAL	58.0130	166.3435	41.4548	61.8323	A=1 B=0.5
M^2BO-F	58.7392	164.7325	42.0116	62.0901	A=1 B=1
M^2BO-GEOMETRIC	59.8624	165.8864	41.7143	65.6227	A=0.01
M^2BO-EXPONENTIAL	58.1030	165.1236	42.1255	64.5606	A=1
M^2BO-EXTREMEVALUE	59.0080	169.0899	41.2377	60.9626	A=0 B=1
M^2BO-GEN. EXT. VALUE	59.2014	166.8629	41.4716	61.8465	A=0 B=1 C=0
M^2BO-GEN. PARETO	59.4067	166.2895	40.5275	65.4693	A=1 B=1 C=1
M^2BO-STUDENTST	59.1608	167.9788	40.6408	61.6892	A=1
M^2BO-LOGNORMAL	57.8919	164.6466	43.0579	64.4928	A=0 B=1
M^2BO-NEG. BINOMIAL	59.0665	168.0757	41.3283	62.8102	A=1 B=0.5
M^2BO-NORMAL	58.2249	167.8959	40.4795	60.1212	A=0 B=1
M^2BO-POISSON	58.1361	166.4314	42.2236	64.7353	A=2
M^2BO-RAYLEIGH	57.9360	163.8931	42.9294	64.1748	A=1
M^2BO-WEIBULL	58.8751	165.1920	41.2942	62.6386	A=1 B=1
MBO (Classical)	64.4677	173.2275	52.6065	76.9007	
SMDO (Abdullah Ates et al., 2020)	61.2391	79.5224	48.1529	43.7932	
DSO (Abdullah Ates et al., 2020)	78.9259	68.7503	60.6965	68.1820	

Table 6.13 Controller Parameters for 3 DOF Hover System (A. Ateş & Akpamukçu, 2021)

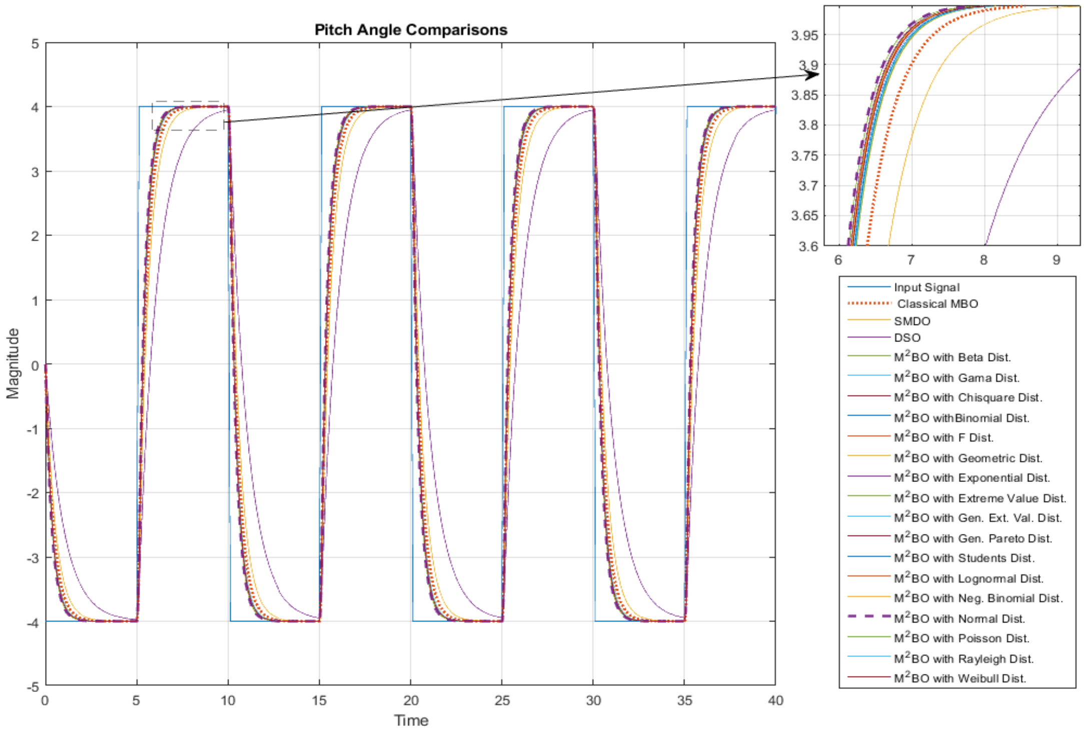

Figure 6.3 Pitch Angle Input Signals Comparisons for 3 DOF Hover Simulation Model (A. Ateş & Akpamukçu, 2021)

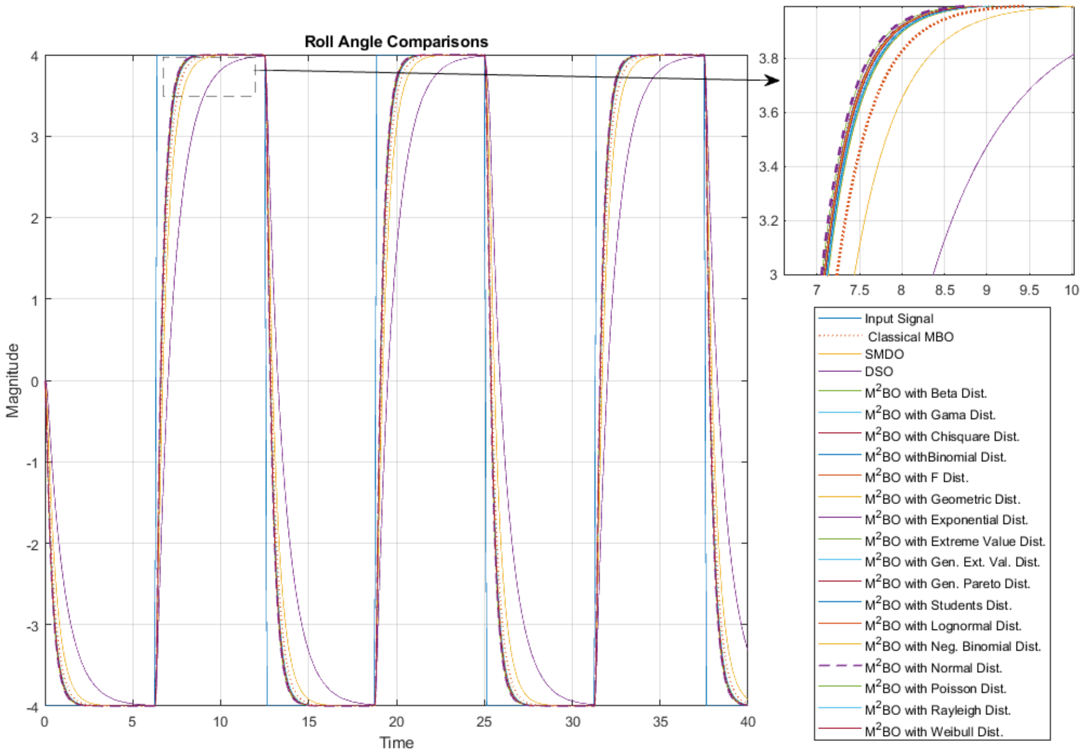

Figure 6.4 Roll Angle Input Signals Comparisons for 3 DOF Hover Simulation Model(A. Ateş & Akpamukçu, 2021)

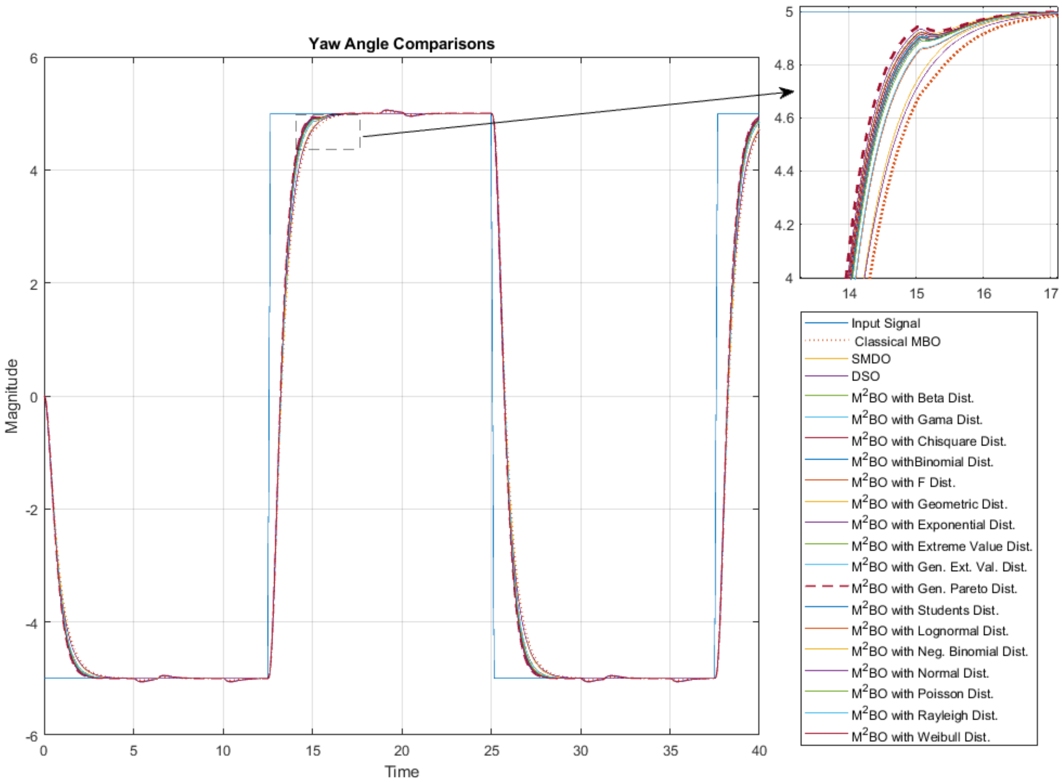

Figure 6.5 Yaw Angle Input Signals Comparisons for 3 DOF Hover Simulation Model (A. Ateş & Akpamukçu, 2021)

6.4.2. Experimental result:

In this study 3 DOF Hover system experiment set was used. This set was manufactured by Quanser. The system is comprised of four propellers and four DC motors. The set is depicted in Fig. 6.6. This system does its work by rotating the roll, pitch and yaw angles. Pitch and roll angles are used to produce lift force and the propeller motors are used to produce total torque. Two of the propellers are counter-rotating; so that the total torque in the system is balanced when the thrust of the four propellers is approximately equal (3 DOF Hover - Quanser,2020).

Figure 6.6 DOF Hover Experimental Setup (A. Ateş & Akpamukçu, 2021)

The values acquired by the M^2BO for K feedback gain matrix shown in the Table 6.13 were applied in the experiment set given in Figure 6.6. The pitch, roll and the yaw angles derived from the experiment set are comparatively.

In Figure 6.7 pitch angle responses are depicted in the experiment set. As it can be seen from the figure classical MBO (bold blue) has better rise and settlement time in comparison with the other optimization algorithms DSO and SMDO. On the other hand, it can be seen from the Figure 6.7 that M^2BO algorithm (bold green) with different distribution functions is better than the classical MBO. Especially M^2BO with Beta Distribution is the best performing algorithm. So it can be inferred that this concept of changing randomization via using different distribution functions is applicable to real engineering problems. In this case, this approach not only enhanced the original algorithm but also resulted in best result in the literature. This result gives us a good evidence of applicability of our approach.

Likewise, the pitch angle, the roll angles also show the positive effect. As it is seen in Figure 6.8 classical M^2BO with Beta Distribution shows a superior performance compared to other optimization algorithms.

Finally, for the yaw angle system, the positive effect can have seen in Figure 6.9. Especially M^2BO with Beta Distribution shows a superior performance compared to other optimization algorithms.

So from the above studies, it is obvious that the M^2BO algorithm is very successful in benchmark functions and 3 DOF Hover System which is a real engineering problem. The contribution of modifying the random number acquiring mechanism of stochastics algorithms is promising. As it can be seen above studies, the basic philosophy of the algorithm is not changed. The only thing that is done is tuning the randomization process. And this effort resulted in best optimal values for the algorithms and this made the target systems more robust. Evaluating this studies, an induction can be done as stochastics optimization algorithms can be enhanced by using the proper distribution functions in their random number acquirement steps.

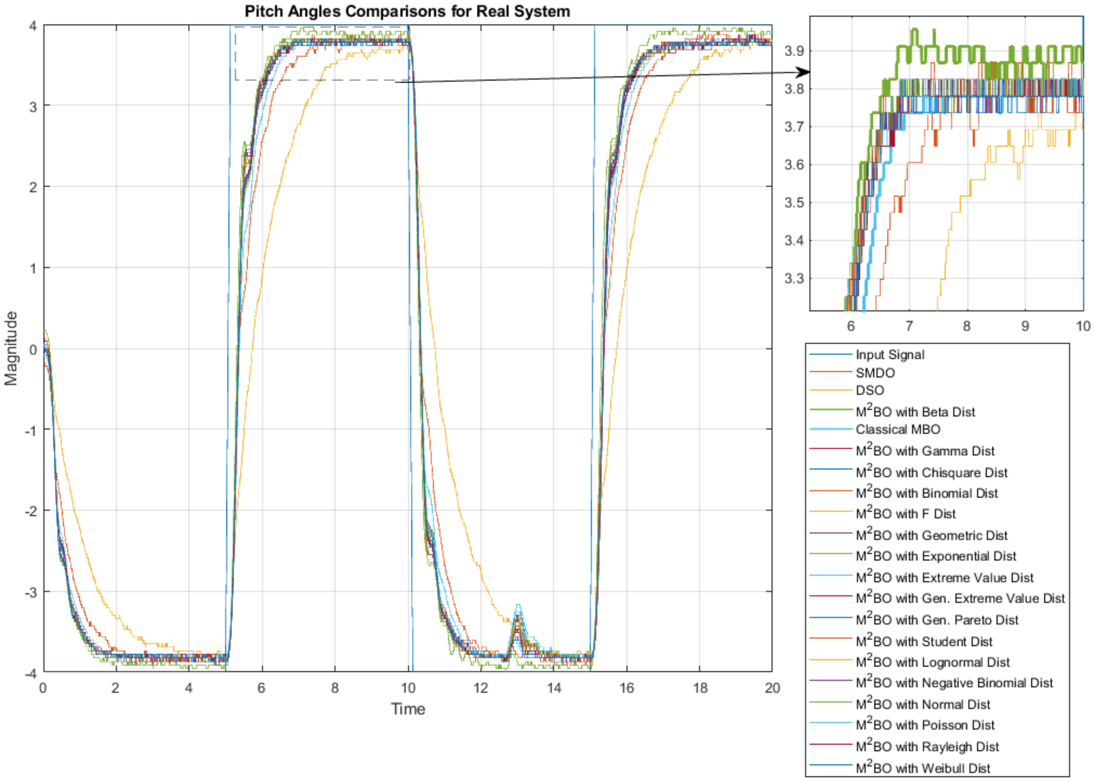

Figure 6.7 Pitch Angle Input Signals Comparisons for 3 DOF Hover Real System(A. Ateş & Akpamukçu, 2021)

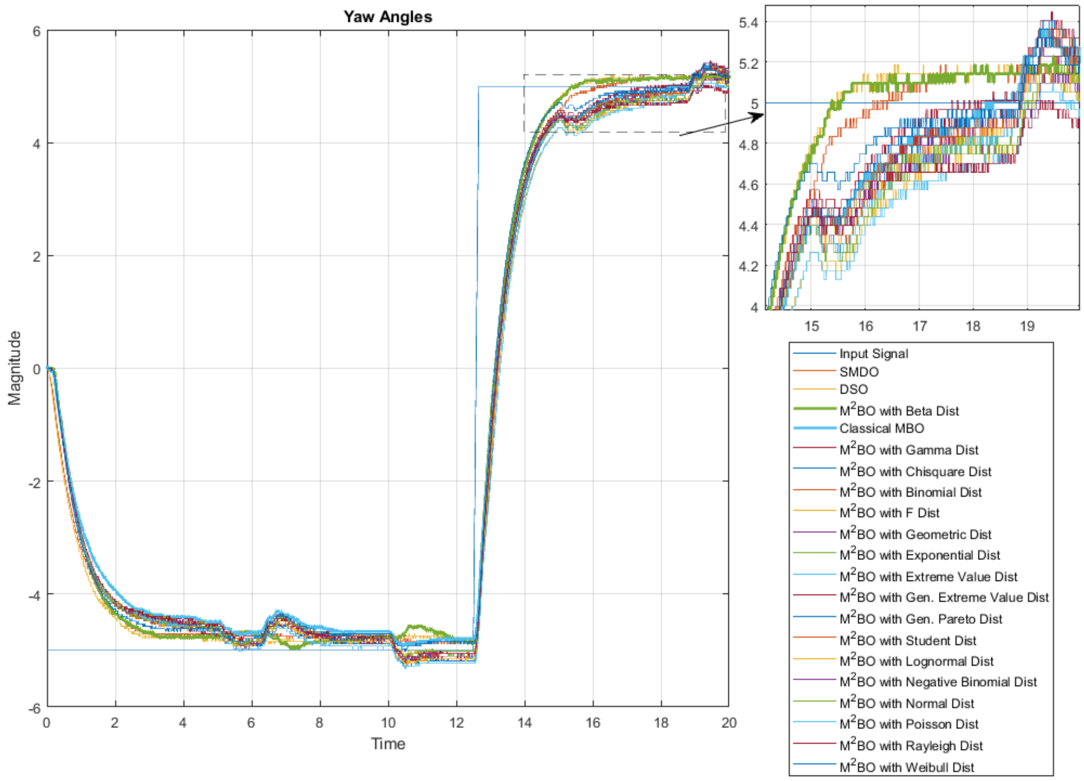

Figure 6.8 Yaw Angles Input Signals Comparisons for 3 DOF Hover Real System(A. Ateş & Akpamukçu, 2021)

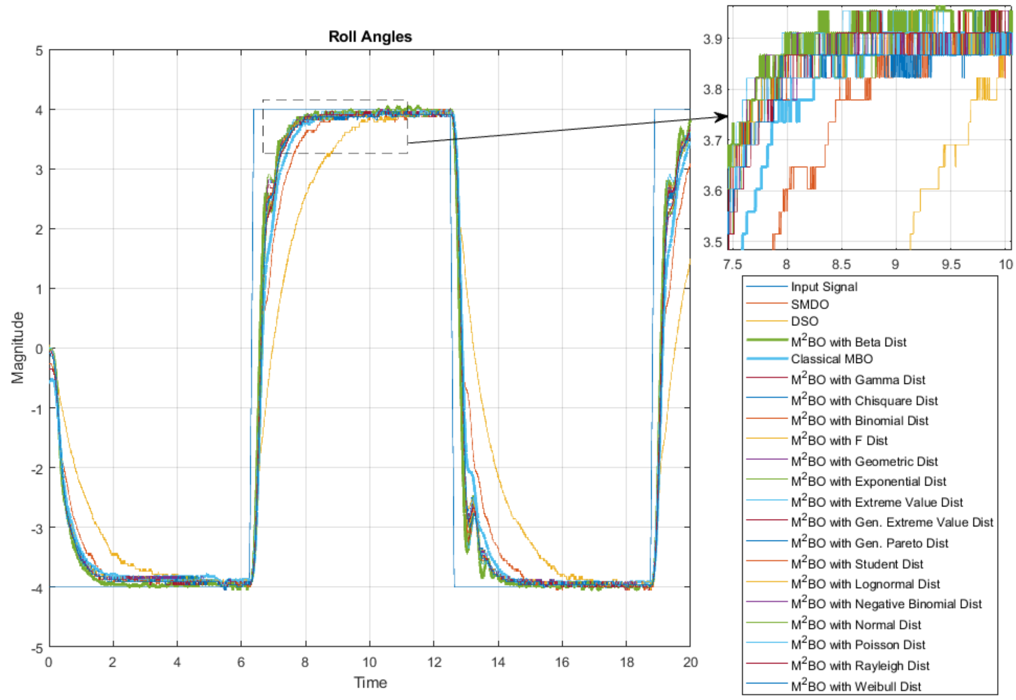

Figure 6.9 Roll Angle Input Signals Comparisons for 3 DOF Hover Real System (A. Ateş & Akpamukçu, 2021)

7. CONCLUSION

In this thesis the usage effect of the different distribution functions in stochastics optimization algorithms is analyzed. Firstly the characteristics of the distribution functions and statistical moments are presented. Probability distribution, statistical moments and finally 15 different distribution functions are analyzed mathematically. Secondly the usage of distribution functions in acquisition of random numbers is indicated so that the importance of distribution functions is shown in randomization processes and the main reason of selecting this topic in this thesis is emphasized. Then the randomness characteristics of stochastics methods is analyzed. Then the usage effect of the distribution functions in stochastic optimization methods is exemplified over the SMDO method. For this exemplification, an application is used so that all the results can be observed easily. So for these purposes; firstly, the original SMDO algorithm is explained. Then how the different distribution functions in the SMDO method will be used is presented. Then the benchmark test functions are explained and the specifications are shown. After that, the MATLAB toolbox “SMDO Benchmark Test Distribution Test Program” introduced. Then the results of using distribution functions in SMDO method is analyzed and compared with existing literature results and finally the original MBO method is explained. Then different distribution functions are used with MBO to change randomization process. Then this modified form of MBO is applied on most used benchmark functions. Furthermore modified MBO is applied on 3 DOF Hover system parameter tuning and the results are shown and compared.

So from all the analyzes taken part in this thesis the effect of using different distribution functions other than uniform distribution, which is classically used in literature, is shown. In most trials in this thesis this effect was positive. So this topic deserves to be evaluated in different areas. Especially in real engineering problems small improvements in processes are very valuable. As it is shown in this thesis, it is very probable by using suitable distribution functions in randomization phases of stochastics methods if they are used in the processes of real engineering problems. In this study very hard benchmark functions that are accepted in the literature are used. And the method offered gained success substantially which gives great hope for other studies. As it is known all problems have their own conditions that must be taken into account. But randomization is a general area which is used in lots of problem. So efforts to improve the performance of stochastic based optimization algorithms with reasonable analyzes, in our case with using suitable distribution functions, are promising studies.

8. REFERENCES

3 DOF Hover - Quanser. (n.d.).

Akdeniz, F. (2018). *Olasılık ve İstatistik*. Akademisyen Kitabevi A.Ş.

Akpamukçu, M., & Ateş, A. (2020). Analyzing of Usage Effect of the Distribution Functions for SMDO Algorithm via Benchmark Function with Matlab Toolbox. *DÜMF Mühendislik Dergisi*, *11*(3), 989–998. https://doi.org/10.24012/dumf.721670

Alagoz, B. B., Ates, A., & Yeroglu, C. (2013). Auto-tuning of PID controller according to fractional-order reference model approximation for DC rotor control. *Mechatronics*. https://doi.org/10.1016/j.mechatronics.2013.05.001

Ateş, A., & Akpamukçu, M. (2021). Modified Monarchy Butterfly Optimization with Distribution Functions and its Application for 3 DOF Hover Flight System. *(Under Review)Neural Computing and Applications*.

ATEŞ, A., & ALAGÖZ, B. (2019). *Farklı Olasılık Dağılım Fonksiyonlarının Stokastik Tabanlı Optimizasyon Algoritmalarına Etkilerinin Kontrolör Tasarım Uygulamalarında Değerlendirilmesi* . 21. Otomatik Kontrol Ulusal Toplantısı TOK 2019 (pp.1-6). Muğla, Turkey

Ates, A., Alagoz, B. B., Chen, Y. Q., Yeroglu, C., & Hosseinnia, S. H. (2019). Optimal Fractional Order PID Controller Design for Fractional Order Systems by Stochastic Multi Parameter Divergence Optimization Method with Different Random Distribution Functions. *2019 IEEE 7th International Conference on Control, Mechatronics and Automation, ICCMA 2019*, 9–14. https://doi.org/10.1109/ICCMA46720.2019.8988599

Ates, Abdullah, Alagoz, B. B., Kavuran, G., & Yeroglu, C. (2020). Fine-Tuning of Feedback Gain Control for Hover Quad Copter Rotors by Stochastic Optimization Methods. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, *7*. https://doi.org/10.1007/s40998-020-00323-7

Ateş, Abdullah, & Yeroglu, C. (2016). Optimal fractional order PID design via Tabu Search based algorithm. *ISA Transactions*, *60*, 109–118. https://doi.org/10.1016/j.isatra.2015.11.015

Ateş, Abdullah, & Yeroğlu, C. (2018). Modified Artificial Physics Optimization for Multi-parameter Functions. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, *42*(4), 465–478. https://doi.org/10.1007/s40998-018-0082-4

Bahadur, R. R. (1960). Some Approximations to the Binomial Distribution Function. *The Annals of Mathematical Statistics*, *31*(1), 43–54. https://doi.org/10.1214/aoms/1177705986

Chhikara, R. S., & Folks, J. L. (1974). Estimation of the inverse gaussian distribution function. *Journal of the American Statistical Association*, *69*(345), 250–254. https://doi.org/10.1080/01621459.1974.10480165

Chim, Y. C., Kassim, A. A., & Ibrahim, Y. (1999). Character recognition using statistical moments. *Image and Vision Computing*, *17*(3–4), 299–307. https://doi.org/10.1016/s0262-8856(98)00110-3

Coefficient of Kurtosis. (2020). Retrieved May 9, 2020, from http://grapherhelp.goldensoftware.com/WTOPICS/WKS_Kurtosis.htm

Glass, L., & Tobler, W. R. (1971). General: Uniform distribution of objects in a homogeneous field: Cities on a plain. *Nature*, *233*(5314), 67–68. https://doi.org/10.1038/233067a0

Krohling, R. A., & Dos Santos Coelho, L. (2006). Coevolutionary particle swarm optimization using gaussian distribution for solving constrained optimization problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, *36*(6), 1407–1416. https://doi.org/10.1109/TSMCB.2006.873185

Lee, C. Y., & Yao, X. (2004). Evolutionary Programming Using Mutations Based on the

Lévy Probability Distribution. *IEEE Transactions on Evolutionary Computation*, *8*(1), 1–13. https://doi.org/10.1109/TEVC.2003.816583
Lentini, A. G. (2020). *R ile Standart Normal Dağılım Olasılıklarının Grafiksel Gösterimi*. Retrieved May 9, 2020, from https://guisebbe.wordpress.com/2019/10/04/r-ile-standart-normal-dagilim-olasiliklarinin-grafiksel-gosterimi/
Marshall, A. W., & Olkin, I. (1967). A Multivariate Exponential Distribution. *Journal of the American Statistical Association*, *62*(317), 30–44. https://doi.org/10.1080/01621459.1967.10482885
Martin, H. R., & Honarvar, F. (1995). Application of statistical moments to bearing failure detection. *Applied Acoustics*, *44*(1), 67–77. https://doi.org/10.1016/0003-682X(94)P4420-B
Moment. (2020). Retrieved May 9, 2020, from https://tr.wikipedia.org/wiki/Moment_(matematik)
“Normal” Distribution Functions on Spheres and the Modified Bessel Functions on JSTOR. (2021). Retrieved March 1, 2021, from https://www.jstor.org/stable/2959410?seq=1#metadata_info_tab_contents
Olasılık Dağılımı. (2020). Retrieved May 9, 2020, from https://tr.wikipedia.org/wiki/Olasılık_dağılımı
On a Characterization of the Poisson Distribution on JSTOR. (2021). Retrieved March 1, 2021, from https://www.jstor.org/stable/25049332?seq=1#metadata_info_tab_contents
Probability Density Function. (2020). Retrieved May 9, 2020, from https://www.mathworks.com/help/stats/prob.normaldistribution.pdf.html
Rajabioun, R. (2011). Cuckoo optimization algorithm. *Applied Soft Computing Journal*, *11*(8), 5508–5518. https://doi.org/10.1016/j.asoc.2011.05.008
Shapiro, A. (2000). *Statistical Inference of Stochastic Optimization Problems* (pp. 282–307). Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-3150-7_16
Skewness. (2020). Retrieved May 9, 2020, from https://en.wikipedia.org/wiki/Skewness
Stacy, E. W. (1962). A Generalization of the Gamma Distribution. *The Annals of Mathematical Statistics*, *33*(3), 1187–1192. https://doi.org/10.1214/aoms/1177704481
Two Characterizations of the Geometric Distribution by Record Values on JSTOR. (2021). Retrieved March 1, 2021, from https://www.jstor.org/stable/25052122?seq=1#metadata_info_tab_contents
Wang, G. G., Deb, S., & Cui, Z. (2019). Monarch butterfly optimization. *Neural Computing and Applications*, *31*(7), 1995–2014. https://doi.org/10.1007/s00521-015-1923-y
Wang, Q. (2013). *Optimization with Discrete Simultaneous Perturbation Stochastic Approximation Using Noisy Loss Function Measurements*. http://arxiv.org/abs/1311.0042
Weibull, W. (1951) A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 18, 293-297. - References - Scientific Research Publishing. (2021). Retrieved March 1, 2021, from https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1292719
Yalçın, M. (2020). *Normal Dağılım*. Retrieved May 9, 2020, from https://medium.com/datarunner/normaldagilim-589846bb850a
Yalta, A. T. (2020). *No Title*. Retrieved May 9, 2020, from http://yalta.etu.edu.tr/files/ekonometri1-01-istatistiksel-kavramlarin-gozden-gecirilmesi-(s2,0).pdf
Yamaoka, K., Nakagawa, T., & Uno, T. (1978). Statistical moments in pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics*, *6*(6), 547–558.

https://doi.org/10.1007/BF01062109
Yang, D., Liu, Z., & Zhou, J. (2014). Chaos optimization algorithms based on chaotic maps with different probability distribution and search speed for global optimization. *Communications in Nonlinear Science and Numerical Simulation*, *19*(4), 1229–1246. https://doi.org/10.1016/j.cnsns.2013.08.017
Yeroğlu, C., & Ateş, A. (2014). A stochastic multi-parameters divergence method for online auto-tuning of fractional order PID controllers. *Journal of the Franklin Institute*, *351*(5), 2411–2429. https://doi.org/10.1016/j.jfranklin.2013.12.006

9. CURRICULUM VITAE

Name-Surname **:** Mehmet AKPAMUKÇU

Education:

- **Bachelor of Science** **:** 2006, Middle East Technical University, Engineering Faculty, Computer Engineering

Work Experience:

- 2014 - … T.C. Sağlık Bakanlığı , Computer Engineer
- 2011- 2014 Otokar A.Ş. , Software Engineer
- 2009- 2011 Havelsan A.Ş. , Software Engineer
- 2006- 2008 TSKB A.Ş. , Software Engineer

Publications:

1- Akpamukçu, M., & Ateş, A. (2020). Analyzing of Usage Effect of the Distribution Functions for SMDO Algorithm via Benchmark Function with Matlab Toolbox. *DÜMF Mühendislik Dergisi*, *11*(3), 989–998. https://doi.org/10.24012/dumf.721670

2- AKPAMUKÇU, M., ATEŞ, A., & ALAGÖZ, B. B. , (2020). Performance Analysis of SMDO Method with Benchmark Functions with Matlab Toolbox. Journal of the Institute of Science and Technology , vol.10, no.4, 2451-2460.