T.C.
KIRIKKALE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

GENETİK ALGORİTMA YAKLAŞIMIYLA KUMARASWAMY DAĞILIMI PARAMETRELERİNİN SIRALI KÜME ÖRNEKLEMESİ İLE TAHMİN EDİLMESİ

ADİL KILIÇ

ARALIK 2018

İstatistik Anabilim Dalında Adil KILIÇ tarafından hazırlanan GENETİK ALGORİTMA YAKLAŞIMIYLA KUMARSAWAMY DAĞILIMI PARAMETRELERİNİN SIRALI KÜME ÖRNEKLEMESİ İLE TAHMİN EDİLMESİ adlı Yüksek Lisans Tezinin Anabilim Dalı standartlarına uygun olduğunu onaylarım.

Prof. Dr. Sevgi YURT ÖNCEL
Anabilim Dalı Başkanı

Bu tezi okuduğumu ve tezin **Yüksek Lisans Tezi** olarak bütün gereklilikleri yerine getirdiğini onaylarım.

Doç. Dr. Güvenç ARSLAN
Danışman

Jüri Üyeleri

Başkan	: Prof. Dr. Sevgi YURT ÖNCEL	______________
Üye (Danışman)	: Doç. Dr. Güvenç ARSLAN	______________
Üye	: Dr. Öğr. Üyesi Sibel AÇIK KEMALOĞLU	______________

12/12/2018

Bu tez Kırıkkale Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu Yüksek Lisans derecesini onaylamıştır.

Prof. Dr. Mustafa YİĞİTOĞLU
Fen Bilimleri Enstitüsü Müdürü

ÖZET

GENETİK ALGORİTMA YAKLAŞIMIYLA KUMARSAWAMY DAĞILIMI PARAMETRELERİNİN SIRALI KÜME ÖRNEKLEMESİ İLE TAHMİN EDİLMESİ

KILIÇ, Adil
Kırıkkale Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı, Yüksek Lisans Tezi
Danışman: Doç. Dr. Güvenç Arslan
Aralık 2018, 60 sayfa

Bu tez çalışmasında, Kumaraswamy dağılımının parametrelerinin tahmin edilmesi için en çok olabilirlik yönteminde genetik algortimanın kullanılması araştırlmıştır. Ayrıca basit rasgele örneklemeye alternatif olarak sıralı küme örneklemesi de incelenmiştir.

Genetik algoritma, Kumaraswamy dağılımı parametrelerinin pozitif olma koşulunun hesaba katılması ve olabilirlik fonksiyonunun türev bilgisine ihtiyaç duymaması açısından kolaylık sağlamıştır. Bunun yanında sıralı küme örneklemesi tahmin edicileri basit rasgele örneklemeye kıyasla daha iyi sonuçlar vermiştir. Simülasyon çalışmasındaki hesaplamalar için R yazılımı kullanılmıştır.

Anahtar Kelimeler: Genetik Algoritma, Kumaraswamy Dağılımı, Basit Rasgele Örnekleme, Sıralı Küme Örneklemesi, Parametre Tahmini, Simülasyon

ABSTRACT

GENETIC ALGORITHM APPROACH TO PARAMETER ESTIMATION OF KUMARASWAMY DISTRIBUTION USING RANKED SET SAMPLING

KILIÇ, Adil
Kırıkkale University
Graduate School of Natural and Applied Sciences
Deparment of Statistics, M. Sc. Thesis
Supervisor: Doç. Dr. Güvenç ARSLAN

December 2018, 60 pages

In this thesis, the estimation of parameters of the Kumaraswamy distribution has been investigated by using maximum likelihood method with genetic algorithm. In addition, ranked set sampling is also investigated as an alternative for simple random sampling.

Genetic algorithm has two benefits for solving this problem. First benefit is that by using GA the pozitivity constraints for the parameters of the Kumaraswamy distribution are automatically satisfied. Second in GA use of derivatives is not needed. On the other hand ranked set sampling estimators give better results in comparison with simple random sampling estimators. R software was prefered for calculations in the simulation study.

Key Words: Genetic Algorithm, Kumaraswamy Distribution, Simple Random Samling, Ranked Set Sampling, Parameter Estimation, Simulation

TEŞEKKÜR

Bu tez çalışmasının her aşamasında engin bilgi birikimi ile katkılarını sunan ve hiçbir yardımını esirgemeyen değerli danışman hocam Sayın Doç. Dr. Güvenç ARSLAN’ a sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER DİZİNİ

Sayfa

ÖZET ... i
ABSTRACT ... ii
TEŞEKKÜR ... iii
İÇİNDEKİLER DİZİNİ ... iv
ÇİZELGELER DİZİNİ ... vi
ŞEKİLLER DİZİNİ ... vii
SİMGELER VE KISALTMALAR DİZİNİ ... viii
1. GİRİŞ ... 1
1.1. Literatür Özeti ... 1
1.2. Amaç ... 2
2. KUMARASWAMY DAĞILIMI ... 3
3. ÖRNEKLEME YÖNTEMİ İÇİN İKİ YAKLAŞIM ... 6
3.1. Basit Rasgele Örnekleme ... 7
3.2. Sıralı Küme Örneklemesi ... 11
3.2.1. Sıralı Küme Örneklemesi ile Örneklem Seçimi ... 12
4. PARAMETRE TAHMİNİ İÇİN GENETİK ALGORİTMA ... 18
4.1. Parametre Tahmini ... 19
4.1.2 Basit Rasgele Örnekleme ile Parametre Tahmini ... 19
4.2. Sıralı Küme Örneklemesi ile Parametre Tahmini ... 21
4.2. Genetik Algoritma ... 23
4.2.1. Genetik Algoritma Çalışma Biçimi ... 25
4.2.2. Genetik Algoritmadaki Temel Kavramlar ... 29
4.2.2.1. Popülasyon (Nüfus) ... 29
4.2.2.2. Kromozom ... 29
4.2.2.3. Gen ... 30
4.2.2.4. Kodlama ... 30
4.2.3. Genetik Algoritmadaki Operatörler ... 31

4.2.3.1. Seçim Operatörü 32
4.2.3.2. Çaprazlama Operatörü 34
4.2.3.3. Mutasyon Operatörü 37
4.2.4. Genetik Algoritma Uygulama Alanları 38
5. KUMARASWAMY DAĞILIMININ PARAMETRELERİNİN GENETİK ALGORİTMA İLE R PROGRAMINDATAHMİN EDİLMESİ 40
5.1. Basit Rasgele Örnekleme için Kumaraswamy Dağılımı Parametrelerinin Genetik Algoritma ile Tahmin Edilmesi 40
5.2. Sıralı Küme Örnekleme için Kumaraswamy Dağılımı Parametrelerinin Genetik Algoritma ile Tahmin Edilmesi 43
6. SİMÜLASYON ÇALIŞMASI 46
7. TARTIŞMA VE SONUÇ 56
KAYNAKLAR 58

ÇİZELGELER DİZİNİ

ÇİZELGE Sayfa

5.1. $a = 0.5$ ve $n = 20$ için Tahmin Sonuçları……….……….…… 48
5.2. $a = 0.5$ ve $n = 30$ için Tahmin Sonuçları……….……….…… 49
5.3. $a = 1$ ve $n = 20$ için Tahmin Sonuçları……….……….……… 50
5.4. $a = 1$ ve $n = 30$ için Tahmin Sonuçları……….……….………. 51
5.5. $a = 2$ ve $n = 20$ için Tahmin Sonuçları……….……….……… 52
5.6. $a = 2$ ve $n = 30$ için Tahmin Sonuçları……….……….………. 53
5.7. $a = 5$ ve $n = 20$ için Tahmin Sonuçları……….……….……….. 54
5.8. $a = 5$ ve $n = 30$ için Tahmin Sonuçları……….……….………. 55

ŞEKİLLER DİZİNİ

ŞEKİL Sayfa

2.1. Kumaraswamy Dağılımı Olasılık Yoğunluk Fonksiyonu…..……. 4
2.2. Kumaraswamy Birikimli Dağılım Fonksiyonu…..……..……… 5
3.1. SKÖ Seçim Şeması………………..………………..……………… 14
4.1. GA Akış Şeması………………..………………..………………….. 28
4.2. İkili Kodlamaya Örnek Kromozomlar………………...…………….. 31
4.3. Permütasyon Kodlamaya Örnek Kromozomlar………………...... 31
4.4. Tek Noktalı Çaprazlama………………..………………..…………. 35
4.5. Çift Noktalı Çaprazlama………………..………………..………….. 35
4.6. Düzgün Çaprazlama………………..………………..……………… 36
4.7. Aritmetik Çaprazlama………………..………………..…………….. 36
4.8. Ekleme Mutasyonu………………..………………..……………….. 37
4.9. Karşılıklı Değişim Mutasyonu………………..……………………...... 38
4.10. Kaydırma Mutasyonu………………..………………..…………… 38
4.11. Ters Çevirme Mutasyonu………………..………………..………. 38

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

a, b	Kumaraswamy dağılımı parametreleri
N	Kitle çapı
n	Örneklem çapı
m	Küme sayısı
r	Tekrar sayısı

KISALTMALAR

GA	Genetik Algoritma
BRÖ	Basit Rasgele Örnekleme
SKÖ	Sıralı Küme Örneklemesi
Var	Varyans
HKO	Hata Kareler Ortalaması
Etk	Etkinlik

1. GİRİŞ

1.1. Literatür Özeti

Genetik algoritmaya temel olan çalışmalar ilk olarak 1975'te John Holland tarafından ortaya çıkarılmıştır [1]. Genetik algoritmalar, büyük ve doğrusal olmayan arama uzaylarında klasik hesaplama yaklaşımlarının sonuç vermediği durumlarda tercih edilen, doğal seçilim ilkesini temel alan bir arama algoritmasıdır [2].

Genetik algoritmalar hakkında önemli bir gelişme ise, John Holland' ın doktora öğrencisi David E. Goldberg tarafından 1985 yılında hazırladığı "Gaz Boru hatlarının Genetik Algoritma Kullanılarak Denetlenmesi" konulu tez ile sağlanmıştır. Bu tez çalışmasından sonra Goldberg'in 1989 yılında yayımladığı "Makine Öğrenmesi, Arama ve Optimizasyon için Genetik Algoritma" adlı kitabı, hâlâ genetik algoritma konusundaki en kapsamlı çalışma olarak kabul görmektedir [3].

Bu algoritmanın her bir adımında bulunan çözüm içindeki her bireyin uygun çözüm olup olmadığını belirleyen bir uygunluk fonksiyonu vardır. Uygunluk fonksiyonundan gelen sonuca göre yüksek değeri olan bireyler, nüfustaki diğer bireyler ile çoğalmalarını mümkün kılar [4]. Bu bireylerden çaprazlama işleminden sonra çocuk denen yeni bireyler bulunur. Her çocuk kendisini meydana getiren ebeveynlerin özelliklerini taşır. Çocukların bulunmasında daha düşük uygunluk değeri olan bireylerin daha az seçilme şansı olduğundan bu bireylerin bir süre sonra nesli tükenir [5].

Sıralı küme örneklemesi 1952 yılında McInctrye tarafından bir meranın ürün hasılatını tahmin etmek için basit rasgele örnekleme yerine geliştirilmiştir [6]. Takahasi ve Wakimoto SKÖ ortalama tahmincisinin BRÖ kitle tahmincisine kıyasla daha küçük varyanslı kitle ortalaması için yansız olduğunu göstermiştir [7]. Dell ve Clutter sıralama hatası olsa da olmasa da SKÖ ortalamasını yine de yansız olduğunu göstermiştir [8]. Stokes SKÖ için sırlamada görsel değerlendirme veya uzman görüşü yerine birimleri sıralamak için yardımcı değişkenler kullanılmasını önermiştir [9]. Kitle ortalama tahmininin doğruluğunu arttırmak için Samawi tarafından Tabakalandırılmış Sıralı Küme Örneklemesi önerilmiştir [10]. Samawi ve arkadaşları dağılımın asimetrik olduğu durumda kitle tahmincisini tanımlamak için Extreme Ranked Set Sampling (ERSS) kullanmıştır [11]. Al-Saleh ve Al-Kadiri Double Ranked Set Sample (DRSS) geliştirmiştir [12]. Al-Saleh ve Al-Omari SKÖ' nün genelleştirilmiş hali olarak Çok Aşamalı SKÖ metodunu sunmuştur [13]. Muttlak asimetrik dağılımlarda daha iyi sonuç veren quartile SKÖ' yü önermiştir [14]. Al- Naseer and Mustafa Robust Extreme SKÖ metodunu önermiştir [15]. Jemain ve arkadaşları örneklemlerin çapını $m=3r$ (burada $r = 1, 2, 3, ...$) alarak Extreme Sıralı Küme Örneklemesi ile Çok Aşamalı Sıralı Küme Örneklemesini tek bir metotta bir araya getiren Dengelenmiş Gruplar Sıralı Küme Örneklemesini (Balanced Groups Ranked Set Sampling/BGRSS) ileri sürmüştür [16]. Arslan ve Öztürk konum ölçek dağılım ailesinde parametrik çıkarım için Partially Rank Ordered Set Samples (PROSS) tasarımı ile simülasyon çalışması yapmışlardır [17].

1.2. Amaç

Bu tez çalışmasında parametrelerinin pozitif olma koşulu bulunan Kumaraswamy dağılımının şekil parametrelerinin tahmin edilmesi problemi için parametrelere ait kısıtların da göz önüne alınarak genetik algoritma ile nasıl hesaplanabileceği araştırılmak istenmiştir. Ayrıca iki farklı örnekleme yönteminden elde edilen en çok olabilirlik tahmin edicilerinin performanslarının karşılaştırması hedeflenmiştir.

2. KUMARASWAMY DAĞILIMI

Kumaraswamy dağılımı 1980 yılında P. KUMARASWAMY tarafından hidroloji uygulamalarında alttan ve üstten sınırlı değişkenler için önerilmiştir [18]. $[0, 1]$ kapalı aralığında sürekli olasılık dağılımlar ailesinden iki sınırlı bir dağılımdır. Beta dağılımına benzerlik göstermesinin yanında olasılık yoğunluk fonksiyonu ve birikimli dağılım fonksiyonlarının kapalı formda yazılabilmesinden dolayı özellikle simülasyon çalışmalarında daha kolay bir kullanımı mevcuttur. Kumaraswamy dağılımı ile beta dağılımı arasındaki bir ilişki, $X_{(a,b)}$, a ve b parametreli Kumaraswamy dağılımlı rasgele değişken olmak üzere, $a=1$ ve $b=b$ olan beta dağılımlı $Y_{(1,b)}$ rasgele değişkeni ile arasındaki ilişkisi $X_{(a,b)}$ = $(Y_{(1,b)})^{1/b}$ şeklindedir [19].

Bu dağılıma ait sırasıyla birikimli olasılık yoğunluk fonksiyonu, dağılım fonksiyonu ve bazı dağılım karakteristikleri aşağıdaki verilen eşitliklerdeki gibidir.

Olasılık yoğunluk fonksiyonu,

$$f\left(x;a,b\right)=abx^{a-1}\left(1-x^{a}\right)^{b-1} \ , \ 0<x<1$$

Birikimli dağılım fonksiyonu,

$$F(x;a,b)=1-(1-x^{a})^{b} \ , \ 0<x<1$$

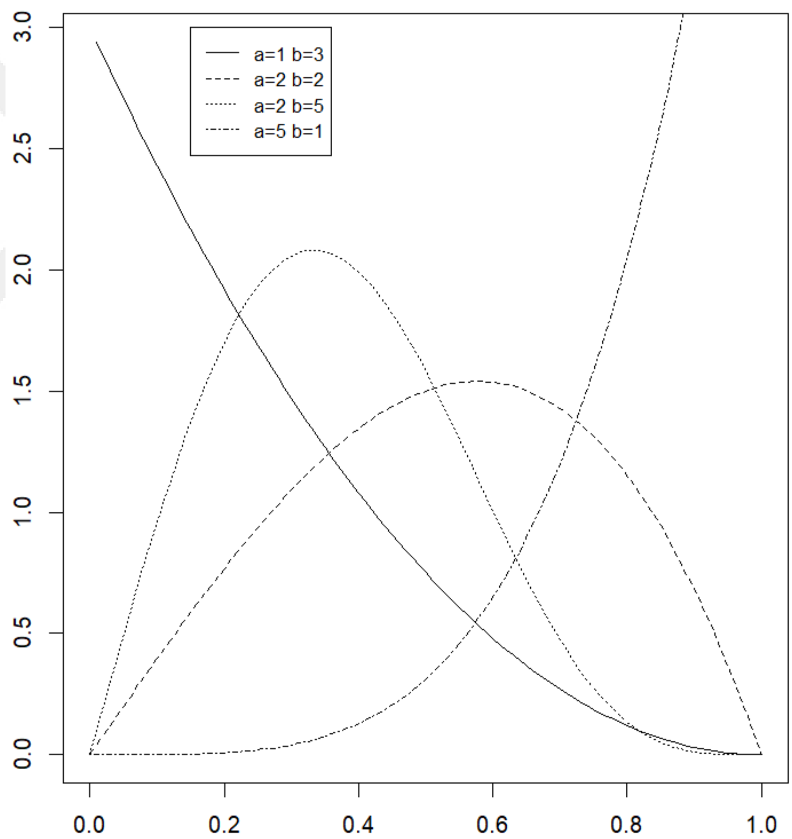
Beklenen değer,

$$E\left[X\right]=\frac{b\Gamma\left(1+\frac{1}{a}\right)\Gamma\left(b\right)}{\Gamma\left(1+\frac{1}{a}+b\right)}$$

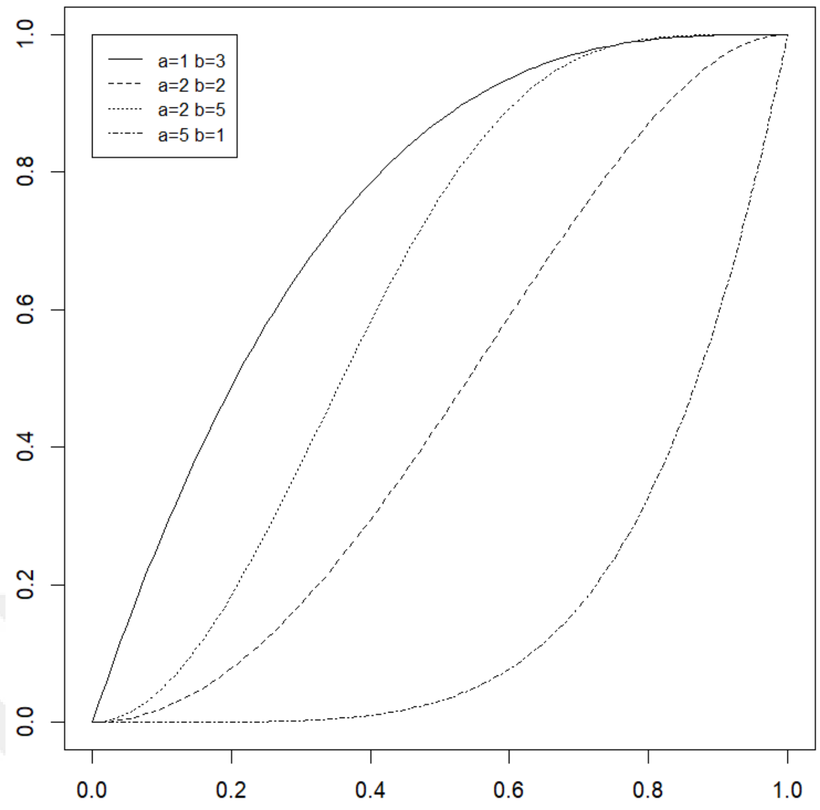
Kumaraswamy dağılımının ham momentleri,

$$E\left[X^n\right]=\frac{b\Gamma\left(1+\frac{n}{a}\right)\Gamma\left(b\right)}{\Gamma\left(1+b+\frac{n}{a}\right)}=b\mathrm{B}\left(1+\frac{n}{a},b\right)$$

eşitliği ile bulunur. Burada B, Beta fonksiyonunu ifade etmektedir. Bu ham momentler kullanılarak dağılımın varyansı, çarpıklığı ve basıklığı hesaplanabilir.



Şekil 2.1. Kumaraswamy Dağılımı Olasılık Yoğunluk Fonksiyonu



Şekil 2.2. Kumaraswamy Birikimli Dağılım Fonksiyonu

Kumaraswamy dağılımı şekil parametreleri a ve b parametreleri pozitif değerlidir. M. A. Hussian bu dağılımın hem basit rasgele hem de sıralı küme örneklemesi ile en çok olabilirlik ve bayes tahmin edicilerini MATHEMATICA yazılımını kullanarak karşılaştırmıştır [20].

Son yıllarda Kumaraswamy dağılımında x yerine başka bir dağılım yazılarak (Weibull, Pareto, Normal dağılım gibi) dağılımın genelleştirilmiş hali üzerine çalışmalar yapılmaktadır.

3. ÖRNEKLEME YÖNTEMİ İÇİN İKİ YAKLAŞIM

İstatistiğin en temel konularından birisi örnekleme kuramıdır. Bu konudaki bazı temel kavramlar kitle, örneklem ve örneklemedir. Kitle, hakkında bilgi edinmek istenen grubun bütününe denir. Örneklem, kitleden belirli yöntemler kullanılarak seçilmiş aynı nitelikteki birimlerden oluşan alt gruba denir. Örnekleme ise örneklem seçmek için kullanılan metotların tamamına verilen addır.

Tamsayım kitlenin her bir biriminin gözlenmesidir. Bu yöntem kitleye dair net sonuçlar ortaya koymasına karşın aslında kullanılması birçok araştırma alanı için pek kolay değildir ya da kullanımı yüksek maliyete sebep olur. Dolayısıyla tamsayımın uygulanmasının mümkün olmadığı veya zor olduğu yerlerde örnekleme yöntemleri önem kazanmaktadır. Örneklemenin tamsayıma göre artıları şu şekildedir;

- Örneklemenin toplam maliyeti tam sayıma kıyasla çok daha düşüktür.
- Örnekleme ile elde edilen gözlemler tamsayıma göre daha kısa zamanda elde edilir
- Örneklemeden gelen sonuçların doğruluk derecesi bazen tam sayıma göre daha yüksek olabilir.

Örnekleme kuramı iki temel aşamadan oluşur. Birinci aşama, örneklem seçimidir. Bu aşamada kitlenin karakteristiklerini iyi yansıtabilecek bir örneklem seçilmesi önemlidir. Araştırılan kitlenin genel yapısı göz önüne alınıp uygun olan örnekleme yöntemi kullanılarak örneklem elde edilir. Bu seçme işi yapılırken zaman, maliyet, emek vb. faktörler hesaba katılmalıdır. İkinci aşama ise kitle hakkında çıkarım yapma aşamasıdır. Parametreler örneklem seçilirken hangi örnekleme yöntemi kullanıldıysa ona göre tahmin edilir. Seçilen örneklemden gelen bu tahminlerin bazı özelliklere sahip olup olmadığına bakılır. Bu özellikler yansızlık, etkinlik, tutarlılık, yeterliliktir [21].

Örnekleme yöntemleri kitleden örneklem alınırken birimlerin seçiminin rasgele olup olmamasına göre olasılıklı olmayan ve olasılıklı örnekleme yöntemi olarak ikiye ayrılır.

Başlıca olasılıklı olmayan örnekleme yöntemleri şunlardır;

- Yargısal örnekleme
- Kota örneklemesi
- Mekanik örnekleme
- Kartopu örneklemesi

Bazı temel olasılıklı örnekleme yöntemleri de şunlardır;

- Basit rasgele örnekleme
- Tabakalı örnekleme
- Sistematik örnekleme
- Kademeli örnekleme
- Küme Örneklemesi
- Sıralı Küme Örneklemesi

3.1. Basit Rasgele Örnekleme

Basit rasgele örnekleme yöntemi diğer örnekleme yöntemlerine kıyasla en çok bilinen ve araştırmalarda yaygın kullanımı olan temel bir örnekleme yöntemidir. Bu yöntem aynı zamanda diğer pek çok örnekleme yönteminin de temelini oluşturması açısından önemli bir yeri mevcuttur. Basit rasgele örnekleme N çaplı bir kitleden n birimlik bir örneklem seçilmesinde her bir $\binom{N}{n}$ kombinasyonunun eşit seçilme olasılığının olduğu örnekleme yöntemidir. Kitlede bulunan her bir birimin örnekleme seçilme olasılıkları eşit ve $\frac{n}{N}$ ‘dir.

Basit rasgele ile elde edilmiş olasılık yoğunluk fonksiyonu $f(x)$ ve birikimli dağılım fonksiyonu $F(x)$ olan n birimlik $X_1, \dots, X_n$ rasgele değişken seti şu iki özelliği sağlar;

- Her bir X_i , $i=1,..,n$ aynı olasılık dağılımlıdır.
- Bu n adet rasgele değişkenler $X_1, \dots, X_n$ birbirinden bağımsızdır.

Kitleden birimlerin örnekleme alınmasının belirlenmesinde genel olarak rasgele sayılar tablosu ya da bilgisayar yazılımlarından faydalanılır. Homojen ve çok büyük olmayan bir kitlede uygulanması kolaydır. Eğer kitle homojen olmayan bir yapıya sahipse örneklem varyansı istenenden büyük bulunabilir. Kitle çapı çok fazla ya da büyük bir alana düzensiz bir şekilde yayılmışsa örneklem seçmek güçleşir. Bu yöntemde örneklem seçimi iadeli ve iadesiz olmak üzere iki şekilde yapılabilir. Ancak iadesiz basit rasgele örnekleme yöntemi iadeli basit rasgele örnekleme yönteminden daha yaygın kullanılır. Bunun sebebi iadesiz basit rasgele örnekleme yöntemi daha etkilidir. İstatistiksel çalışmalarda geniş yelpazede uygulanan basit rasgele örnekleme yöntemi temelde iki amaç için faydalanılır. İlk olarak seçilen örneklemden faydalanarak kitle ortalamasının tahmini ve diğeri ise kitle oranının tahmin edilmesidir [22].

Kitle dağılımının normal olduğu biliniyorsa bu durumda kitleden seçilecek n çaplı birbirinden farklı örneklemlerin ortalamaları da kitle ortalaması civarında normal dağılımlı olur. μ kitle ortalaması, $\bar{x}$ örneklem ortalaması ve s^2 örneklem varyansını ifade etmektedir. Kitle ortalaması μ örneklem ortalaması olan $\bar{x}$ ile aşağıdaki eşitlikteki gibi tahmin edilebilir.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\bar{x}$' nin varyansı şu eşitlik kullanılarak hesaplanabilir,

$$Var(\bar{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

Örneklem varyansı $s^2 = \frac{1}{(\text{n}-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$ eşitliği ifade edilir.

Örneklem ortalaması $\bar{x}$ kitle ortalaması μ için yansız bir tahmin edici olduğu aşağıdaki gibi gösterilir.

$$\begin{aligned}
E(\bar{X}) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mu \\
&= \frac{1}{n} n\mu = \mu
\end{aligned}$$

olur.

Ayrıca $\left(\frac{N-n}{N}\right)\frac{s^2}{n}$ tahmin edicisi de $Var(\bar{X})$ için yansız tahmin edici olduğu şu şekilde gösterilir,

$$\begin{aligned}
E(s^2) &= E\left[\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] \\
&= \left(\frac{1}{n-1}\right)E\left[\sum_{i=1}^{n}\left[\left(X_i - \mu\right) - (\bar{X} - \mu)\right]^2\right] \\
&= \left(\frac{1}{n-1}\right)E\left[\sum_{i=1}^{n}(X_i - \mu)^2 - n(\bar{X} - \mu)^2\right]
\end{aligned}$$

$$=\left(\frac{1}{n-1}\right)\left[\sum_{i=1}^{n}E(X_i-\mu)^2-nE(\bar{X}-\mu)^2\right]$$

$$=\left(\frac{1}{n-1}\right)\left[n\sigma^2-nVar(\bar{X})\right]$$

$$=\left(\frac{1}{n-1}\right)\left[n\sigma^2-n\left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n}\right]$$

$$=\frac{\sigma^2}{n-1}\left(n-\frac{N-n}{N-1}\right)$$

$$=\frac{N}{N-1}\sigma^2$$

olarak ifade edilir.

Buradan,

$$E\left[\left(\frac{N-n}{N}\right)\left(\frac{S^2}{n}\right)\right]=\left(\frac{N-n}{N}\right)\left(\frac{1}{n}\right)\left(\frac{N}{N-1}\right)\sigma^2$$

$$=\left(\frac{N-n}{N-1}\right)\left(\frac{\sigma^2}{n}\right)=Var(\bar{X})$$

olarak elde edilir [22].

Böylece $\left(\frac{N-n}{N}\right)\frac{s^2}{n}$ tahmin edicisinin $Var(\bar{X})$ için yansız tahmin edici olduğu gösterilmiş olur.

$\left(\frac{N-n}{N}\right)$ N birimlik bir kitleden nispeten büyük n çaplı örneklem alındığında tahmin edicinin varyansını azaltan sonlu kitle faktörüdür [23].

3.2. Sıralı Küme Örneklemesi

Örnekleme seçilirken maliyet, zaman ve emek gibi etkenlerin düşük olmasına dikkat ederek kitle hakkında en iyi çıkarım yapılması istenir. Farklı örnekleme yöntemleriyle de küçük varyanslı tahmin ediciler bulunabilir. Basit rasgele örnekleme de sonlu bir kitleden n tane gözlem seçilerek bu gözlemlerden bir örneklemin oluşturulmasıdır. Fakat basit rasgele örnekleme yöntemi kullanılarak örneklem oluşturmak bazı araştırmalarda yüksek maliyete ve fazla zaman harcanmasına sebep olmaktadır. Bu yöntemde her birimin ayrı ayrı ölçülmesi gerekir ancak bu oldukça emek isteyen güç bir yoldur. Bu gibi sebeplerden ötürü basit rasgele örneklemeye alternatif bir yöntem olarak sıralı küme örneklemesi kullanılabilir.

Sıralı küme örneklemesi ilk olarak 1952 yılında McIntyre tarafından Avustralya' da mera verimini tahmin etmek için kullanılmıştır. Bu örnekleme yöntemi her bir gözlem için tam ölçüm yapılmadan örneklemeye seçilecek birimlerin belirlenmesinin sıralamaya dayalı olduğu bir yöntemdir. Bu örnekleme yönteminde örnekleme alınacak olan birimler ilgili değişkene göre sıralanır. Daha sonra bu örnekleme yönteminde sıralama gözleme dayalı olarak ya da yardımcı değişken kullanılarak yapılır. Sıralama işlemi yapıldıktan sonra bu sıralamaya göre tam ölçümü yapılacak olan gözlemler alınıp tam olarak ölçülür. Burada sıralama hatası söz konusu olabilir ancak yine de eşit örneklem çaplı basit rasgele örneklemeye göre daha iyi tahminler elde edilebilir. Tahminlerdeki duyarlılık örneklem çapına ve sıralamadaki hatalar gibi faktörlere bağlı olarak değişiklik gösterebilir. Bu yöntem özellikle bazı uygulamalarda geleneksel örnekleme yöntemlerinden daha kolay, daha ucuz ve uygulanabilirdir. Örneğin bir ormandaki ağaçların boy ortalamasını tahmin etmek istiyorsak tam ölçüm yapmak oldukça zordur. O halde tam ölçüm kullanmaktan ziyade görsel yolla sıralamaya dayalı olarak örneklem seçmek daha uygun olacaktır. Bu sebeple ilgili değişkenin ölçümü maliyetli veya fazlaca emek isteyen durumların olduğu tıp, tarım, ekoloji ve ormancılık gibi alanlarda sıralı küme örneklemesi yaygın olarak tercih edilmektedir. Yani bu yöntem bu gibi alanlarda minimum örneklem çapı ile kitle hakkında iyi bir çıkarım yapmaya olanak sağlar. Dolayısıyla bugün daha düşük maliyet ile etkin tahmin ediciler bulabilmek ve iyi parametre tahmini yapabilmek için sıralı küme örneklemesi tercih edilmektedir [24].

3.2.1. Sıralı Küme Örneklemesi ile Örneklem Seçimi

Sıralı küme örneklemesinde örneklem seçme işi iki aşamadan oluşur. İlk olarak kitleden m çaplı m adet basit rasgele örnekleme yöntemi ile seçilir. m birimden oluşan bu örneklere küme denir. Bu seçim kitleden m^2 tane birimden oluşan basit rasgele örnekleme yöntemi ile seçilmiş örneklemin m çaplı m tane kümeye tamamen rasgele olarak dağıtılmasıyla da gerçekleştirilebilir. Bu işlemden sonra her bir m birimlik küme ilgilenilen değişken bakımından kolay ve maliyeti düşük bir yolla tam ölçümü yapılmadan küçükten büyüğe doğru olacak şekilde sıralanır. Buradaki sıralama işlemi görsel bir sıralama, yardımcı değişken kullanımı ya da araştırmacının deneyimi olarak yapılabilir. İkinci aşama ise küçükten büyüğe doğru sıralan m tane kümeden birinci kümenin birinci sıradaki gözlemi, ikinci kümenin ikinci sıradaki gözlemi ve bu işlem aynı şekilde sürdürülerek $m.$ sıradaki kümenin $m.$ elemanı seçilir. m tane kümeden seçilen bu m çaplı örneklemdeki birimler istenen kesinliğe göre tam olarak ölçümü yapılır. Bu şekilde m çaplı sıralı küme örneklemi seçilmiş olur. Bu yöntemle m^2 tane birimden sadece m tanesinin tam ölçümü yapılmış olur. Ancak burada m birimlik kümelerin sıralanması yüksek hassasiyette olmadığından, literatürde sıralamadaki hatayı minimize etmek için m değerinin *2, 3, 4, 5* veya *6* olarak belirlenmesi tavsiye edilmektedir. Eğer bu yöntemle daha büyük bir örneklem seçmek istenirse yukarıda bahsedilen adım r defa tekrar edilir. Böylece kitleden $m^2.r$ gözlem seçilir ve bunların $m.r$ tanesi tam olarak ölçülür. Bu durumda her bir tekrar birbirinden bağımsızdır. Dolayısıyla sıra istatistikleri de aynı dağılımlı ve birbirinden bağımsız olur.

Adım adım sıralı küme örneklemesi ile örneklem seçme işlemi:

I. İlgilenilen kitleden rasgele olarak m^2 tane örneklem seçilir.

II. Seçilen m^2 çaplı bu örneklem rasgele olarak m adet m çaplı kümeye dağıtılır.

III. Bu m çaplı kümelerin her biri küçükten büyüğe doğru tam ölçmeden görsel yolla veya kolay ve pratik bir ölçümle sıralanır.

IV. Birinci kümeden en küçük birim, ikinci kümeden ikinci sıradaki birim ve aynı sıra takip edilerek $m.$ kümeden $m.$ sıradaki yani en büyük olan birim seçilir.

V. Bu prosedür ile seçilen m çaplı örneklem istenen hassasiyetle ölçümü yapılır ve daha büyük bir sıralı küme örneklemi için aynı adımlar tekrarlanır.

Bu adımlar Şekil 3.1 deki gibi şematize edilebilir. m, kümelerdeki birim sayısını ve r de döngü sayısını göstermektedir. $X_{(i:m)r}$: r inci döngüdeki m birimlik örneklemde i inci sıradaki değişkeni temsil etmektedir. Şemadaki her döngüde birbirinden bağımsız m adet küme vardır. Bu kümelerden seçilen birimler kalın font ile gösterilmiştir. Şekilde görüldüğü üzere her döngüden m tane birim tam ölçüm için alındığından sıralı küme örneklem çapı burada $m.r$ kadardır [24].

1.Döngü

$$
\begin{array}{ccccc}
\mathbf{X}_{\mathbf{(1:m)1}} & X_{(2:m)1} & X_{(3:m)1} & \cdots & X_{(m:m)1} \\
X_{(1:m)1} & \mathbf{X}_{\mathbf{(2:m)1}} & X_{(3:m)1} & \cdots & X_{(m:m)1} \\
 & & & \vdots & \\
X_{(1:m)1} & X_{(2:m)1} & X_{(3:m)1} & \cdots & \mathbf{X}_{\mathbf{(m:m)1}}
\end{array}
$$

2.Döngü

$$
\begin{array}{ccccc}
\mathbf{X}_{\mathbf{(1:m)2}} & X_{(2:m)2} & X_{(3:m)2} & \cdots & X_{(m:m)2} \\
X_{(1:m)2} & \mathbf{X}_{\mathbf{(2:m)2}} & X_{(3:m)2} & \cdots & X_{(m:m)2} \\
 & & & \vdots & \\
X_{(1:m)2} & X_{(1:m)2} & X_{(1:m)2} & \cdots & \mathbf{X}_{\mathbf{(1:m)2}}
\end{array}
$$

$$\vdots$$

r. Döngü

$$
\begin{array}{ccccc}
\mathbf{X}_{\mathbf{(1:m)r}} & X_{(2:m)r} & X_{(3:m)r} & \cdots & X_{(m:m)r} \\
X_{(1:m)r} & \mathbf{X}_{\mathbf{(2:m)r}} & X_{(3:m)r} & \cdots & X_{(m:m)r} \\
 & & & \vdots & \\
X_{(1:m)r} & X_{(1:m)r} & X_{(1:m)r} & \cdots & \mathbf{X}_{\mathbf{(1:m)r}}
\end{array}
$$

Şekil 3.1. SKÖ Seçim Şeması

Sıralı küme örneklemesinin örneklem ortalaması, kitle ortalaması tahmin edicisinin yansız bir tahmin edicisidir. Aynı zamanda bu tahmin edici eşit çaplı basit rasgele örneklem ortalamasının varyansından daha küçüktür [25]. Sıralı küme örneklemesinde kitle ortalama tahmin edicisinin etkinliği basit rasgele örneklemesi kitle ortalaması tahmin edicisinin etkinliğinden iyidir [8]. $X_{(1:m)r}, X_{(2:m)r}, \ldots, X_{(m:m)r}$ sıralı küme örneklemi olmak üzere, kitle ortalaması $\mu_{skö}$ ve kitle ortalaması tahmin edicisi $\overline{X}_{skö}$ olsun. O halde Sıralı küme örneklem ortalaması tahmin edicisi,

$$\overline{X}_{skö} = \frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} X_{(i:m)j} = \frac{1}{m}\sum_{i=1}^{m} X_{(i:m)}$$

şeklindedir.

Buradan $\overline{X}_{skö}$'nin yansızlığı,

$$E(\overline{X}_{skö}) = E\left(\frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} X_{(i:m)j}\right)$$

$$= \frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} E\left(X_{(i:m)j}\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m} E\left(X_{(i:m)}\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} x\, f_{X_{(i:m)}}(x)\,dx\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} x\frac{m!}{(i-1)!(m-i)!}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} x\, m\frac{(m-1)!}{(i-1)!((m-1)-(i-1))!}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} x\, m\binom{m-1}{i-1}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right)$$

$$= \int_{-\infty}^{\infty} x\, f(x) \left(\sum_{i=1}^{m} \binom{m-1}{i-1} \left[F(x)\right]^{i-1} \left[1-F(x)\right]^{m-i} \right) dx$$

Binom dağılımından bilindiği üzere,

$$\sum_{i=1}^{m} \binom{m-1}{i-1} \left[F(x)\right]^{i-1} \left[1-F(x)\right]^{m-i} = 1$$ dir.

Böylece,

$$E(\overline{X}_{sk\ddot{o}}) = \int_{-\infty}^{\infty} x\, f(x) dx = \mu_{sk\ddot{o}}$$

şeklinde gösterilir [26].

Sıralı küme örneklem ortalaması varyansı ise,

$$Var(\overline{X}_{sk\ddot{o}}) = Var\left(\frac{1}{m} \sum_{i=1}^{m} X_{(i:m)} \right)$$

$$= \frac{1}{m^2} Var\left(\sum_{i=1}^{m} X_{(i:m)} \right)$$

$$= \frac{1}{m^2} E\left[\left(X_{(i:m)} - \mu_{(i:m)} \right)^2 \right]$$

Ayrıca

$$E\left[\left(X_{(i:m)} - \mu \right)^2 \right] = E\left[\left(X_{(i:m)} - \mu_{(i:m)} + \mu_{(i:m)} - \mu \right)^2 \right]$$

$$= E\left[\left(X_{(i:m)} - \mu_{(i:m)} \right)^2 \right] + \left(\mu_{(i:m)} - \mu \right)^2$$

eşitliğinden faydalanarak,

$$Var(\overline{X}_{sk\ddot{o}}) = \frac{1}{m^2} E\left[\left(X_{(i:m)} - \mu_{(i:m)}\right)^2\right] - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

$$= \frac{1}{m^2}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} (x-\mu)^2 \frac{m!}{(i-1)!(m-i)!}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right) - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

$$= \frac{1}{m^2}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} m(x-\mu)^2 \frac{(m-1)!}{(i-1)!((m-1)-(i-1))!}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right) - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

$$= \frac{1}{m^2}\sum_{i=1}^{m}\left(\int_{-\infty}^{\infty} m(x-\mu)^2 \binom{m-1}{i-1}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} f(x)dx\right) - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

$$= \frac{1}{m}\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\left(\sum_{i=1}^{m}\binom{m-1}{i-1}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i}\right)dx - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

Binom dağılımından bilindiği üzere,

$$\sum_{i=1}^{m}\binom{m-1}{i-1}\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{m-i} = 1 \quad \text{idi.}$$

Böylece,

$$Var(\overline{X}_{sk\ddot{o}}) = \frac{1}{m}\int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

ve burada

$$\int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \sigma^2 \quad \text{dır.}$$

Sonuç olarak

$$Var(\overline{X}_{sk\ddot{o}}) = \frac{\sigma^2}{m} - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2 = Var(\overline{X}) - \frac{1}{m^2}\left(\mu_{(i:m)} - \mu\right)^2$$

şeklinde gösterilmiş olur [26].

4. PARAMETRE TAHMİNİ İÇİN GENETİK ALGORİTMA

Burada ele alınan parametre tahmin probleminde Kumaraswamy dağılımının hem basit rasgele örnekleme hem de sıralı küme örneklemesi ile bulunan olabilirlik fonksiyonlarını maksimum yapan değerlerin araştırılmasında *Genetik Algoritma* kullanılmıştır. Ayrıca Genetik Algoritma Kumaraswamy dağılımı parametrelerinin pozitif olması koşullarını da dikkate alma konusunda kolaylık sağlamıştır.

Kısaca en çok olabilirlik yönteminden bahsedersek, bu parametre tahmin yöntemi gözlemlere dayalı istatistiksel modelin parametrelerini tahmin etmek için sıklıkla tercih edilen bir yöntemdir. $X_1, X_2, ..., X_n$ bir rasgele örneklem olmak üzere $f(\,.\,;\theta)$ olasılık (yoğunluk) fonksiyonuna sahip olsun. Burada $\boldsymbol{\theta}$ bilinmeyen parametre değeridir. $f(x;\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i;\boldsymbol{\theta})$ olmak üzere $\mathbf{X} = \mathbf{x}$ olarak gözlendiğinde $\boldsymbol{\theta}$ nın bir fonksiyonu olan,

$$\mathrm{L}(\boldsymbol{\theta};\mathbf{x}) = f(\mathbf{x};\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta$$

olabilirlik fonksiyonudur. Θ parametre kümesi üzerinde eğer varsa $\mathrm{L}(\boldsymbol{\theta};\mathbf{x})$ ifadesini maksimum yapan $\boldsymbol{\theta}(\mathbf{x})$ değerine $\boldsymbol{\theta}$ nın en çok olabilirlik tahmini ve $\boldsymbol{\theta}(\mathbf{X})$ istatistiğine de $\boldsymbol{\theta}$ nın en çok olabilirlik tahmin edicisi denir.

Logaritma fonksiyonunun monotonluğu göz önüne alındığında,

$$\max_{\theta\in\Theta} \log L(\theta;x) = \max_{\theta\in\Theta} \prod_{i=1}^{n} f(x_i;\theta)$$

olarak yazılabilir. Bazen olabilirlik fonksiyonu karmaşık çarpımlı ifadeler içerebilir ve bu durumda logaritma almak fonksiyonun maksimum olduğu noktaların bulunmasında kolaylık sağlayabilir [27].

4.1. Parametre Tahmini

Burada seçilen iki farklı örnekleme yöntemine göre dağılım parametrelerinin tahmin edilmesinde farklılıklar mevcuttur. Her iki örnekleme için dağılım parametrelerinin olabilirlik fonksiyonları ve tahmin edicileri farklıdır. Bu iki farklı örnekleme yöntemine göre olabilirlik fonksiyonlarının ve tahmin edicilerin matematiksel olarak nasıl bulunduğu aşağıdaki iki alt başlıkta gösterilmiştir.

4.1.2 Basit Rasgele Örnekleme ile Parametre Tahmini

$X_1, X_2, ..., X_n$ n boyutlu a ve b parametreli Kumaraswamy dağılımından gelen rasgele örneklem olsun. a ve b ye ait olabilirlik fonksiyonu;

$$\begin{aligned} L_{brö}(a, b; x) &= \prod_{i=1}^{n} f_{brö}(x_i; a, b) \\ &= \prod_{i=1}^{n} ab x_i^{a-1}\left(1 - x_i^{a}\right)^{b-1} \\ &= a^n b^n \prod_{i=1}^{n} x_i^{a-1} \prod_{i=1}^{n}\left(1 - x_i^{a}\right)^{b-1} \end{aligned} \tag{1}$$

şeklinde yazılır.

Buradan logaritmik olabilirlik fonksiyonu;

$$\log L_{brö} = \log\left[a^n b^n \prod_{i=1}^{n} x_i^{a-1} \prod_{i=1}^{n}\left(1 - x_i^{a}\right)^{b-1}\right]$$

$$=\log a^n+\log b^n+\log\left(\prod_{i=1}^{n} x_i^{a-1}\right)+\log\left(\prod_{i=1}^{n}\left(1-x_i^{a}\right)^{b-1}\right)$$

$$=n\log a+n\log b+(a-1)\sum_{i=1}^{n}\log\left(x_i\right)+(b-1)\sum_{i=1}^{n}\log\left(1-x_i^{a}\right)$$

olarak elde edilir.

a ve b parametrelerinin basit rasgele örnekleme ile en çok olabilirlik tahmin edicilerini $a_{brö}$ ve $b_{brö}$ olarak gösterilsin. Yukarıda verilen olabilirlik fonksiyonunun her iki parametreye göre türevleri alınırsa,

$$\frac{n}{a}+\sum_{i=1}^{n}\log\left(x_i\right)-\left(b-1\right)\sum_{i=1}^{n}\frac{x_i^{a}\log a}{\left(1-x_i^{a}\right)}=0$$

$$\frac{n}{b}+\sum_{i=1}^{n}\log\left(1-x_i^{a}\right)=0$$

olur ve bu eşitliklerden,

$$b=\frac{-n}{\sum_{i=1}^{n}\log\left(1-x_i^{a}\right)}$$

ve a nin kapalı formu elde edilemeyen çözümü,

$$\frac{n}{a}+\sum_{i=1}^{n}\log\left(x_i\right)-\left(b-1\right)\sum_{i=1}^{n}\frac{x_i^{a}\log a}{\left(1-x_i^{a}\right)}=0$$

olarak bulunur.

Kumaraswamy dağılımı parametreleri için en çok olabilirlik tahmin edicileri olan $a_{brö}$ ve $b_{brö}$ nın doğrusal olmayan iki eşitlik olarak elde edildi. Dolayısıyla bu tahmin edicilerin hesaplanmasında başka alternatif yöntemlerle çözüm aranmalıdır [20].

4.2. Sıralı Küme Örneklemesi ile Parametre Tahmini

$X_1, X_2, \ldots, X_n \sim F(x)$ dağılımlı bir rasgele örneklem olsun. Buradan elde edilen sıralı küme örneklemesinde $X_{(i:m)j}$: sıralı küme örneklemesinde j inci döngüdeki m birimlik örneklemde i inci sıradaki değişkeni temsil etmektedir. Burada $i = 1, \ldots, m$ ve $j = 1, \ldots, \mathrm{r}$ olmak üzere örneklem büyüklüğü $n = m.r$ olan bir örneklem Kumaraswamy dağılımından sıralı küme örneklemesi ile elde edilsin. Gösterimi daha basit kılmak adına sıralı küme örneklemesi rasgele değişkeni $X_{(i:m)j}$ yerine X_{ij} olarak gösterilecektir. Gözlemler sıralı küme örneklemesi ile seçildiğinden dağılıma ait olasılık yoğunluk fonksiyonu yeniden ifade edilmelidir.

X_{ij} rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$f_{skö}(x_{ij}) = \frac{m!}{(i-1)!(m-i)!} f(x_{ij}) \left[F(x_{ij})\right]^{i-1} \left[1 - F(x_{ij})\right]^{m-i}$$

$$= \frac{m!}{(i-1)!(m-i)!} ab x_{ij}{}^{a-1} (1 - x_{ij}{}^{a})^{b(m-i+1)-1} \left[1 - \left(1 - x_{ij}{}^{a}\right)^{b}\right]^{i-1}$$

olur. Burada $x_{ij} \in (0,1)$ ve $a > 0,\ b > 0$ dır.

a ve b nin olabilirlik fonksiyonu,

$$L_{skö}(a, b; x) = \prod_{j=1}^{r} \prod_{i=1}^{m} f_{skö}(x_{ij}; a, b)$$

$$= \prod_{j=1}^{r} \prod_{i=1}^{m} \frac{m!}{(1-i)!(m-i)!} ab x_{ij}{}^{a-1} (1 - x_{ij}{}^{a})^{b(m-i+1)-1} \left[1 - \left(1 - x_{ij}{}^{a}\right)^{b}\right]^{i-1} \quad (2)$$

olarak yazılıabilir.

Buradan, logaritmik olabilirlik fonksiyonu,

$$\log L_{sk\ddot{o}}(a,b;x)=mr\log(a)+mr\log(b)+(a-1)\sum_{j=1}^{r}\sum_{i=1}^{m}\log(x_{ij})$$

$$+b\sum_{j=1}^{r}\sum_{i=1}^{m}(m-i+1)\log(1-x_{ij}{}^{a})-\sum_{j=1}^{r}\sum_{i=1}^{m}\log(1-x_{ij}{}^{a})$$

$$+\sum_{j=1}^{r}\sum_{i=1}^{m}(i-1)\log[1-(1-x_{ij}{}^{a})^{b}]+r\sum_{i=1}^{m}\log\left(\frac{m!}{(i-1)!(m-i)!}\right)$$

olarak bulunur.

a ve b parametrelerinin sıralı küme örneklemesi ile en çok olabilirlik tahmin edicilerini $a_{sk\ddot{o}}$ ve $b_{sk\ddot{o}}$ olarak gösterilsin. Yukarıda verilen olabilirlik fonksiyonunun her iki parametreye göre türevleri alınırsa;

a ’ya göre türev,

$$\frac{mr}{a}+\sum_{j=1}^{r}\sum_{i=1}^{m}\log(x_{ij})-b\sum_{j=1}^{r}\sum_{i=1}^{m}(m-i+1)\frac{x_{ij}{}^{a}\log(x_{ij})}{1-x_{ij}{}^{a}}$$

$$+\sum_{j=1}^{r}\sum_{i=1}^{m}\frac{x_{ij}{}^{a}\log(x_{ij})}{1-x_{ij}{}^{a}}+\sum_{j=1}^{r}\sum_{i=1}^{m}(i-1)\frac{bx_{ij}{}^{a}(1-x_{ij}{}^{a})^{b-1}\log(x_{ij})}{1-(1-x_{ij}{}^{a})^{b}}=0$$

b ’ye göre türev,

$$\frac{mr}{b}+\sum_{j=1}^{r}\sum_{i=1}^{m}(m-i+1)\log(1-x_{ij}{}^{a})-\sum_{j=1}^{r}\sum_{i=1}^{m}(i-1)\frac{(1-x_{ij}{}^{a})^{b}\log(1-x_{ij}{}^{a})}{1-(1-x_{ij}{}^{a})^{b}}=0$$

şeklinde elde edilir [20].

Aslında en çok olabilirlik yöntemi ile a ve b parametrelerinin tahmini problemi olabilirlik fonksiyonunu maksimum yapan a ve b değerlerinin bulunmasına dayanır. Bu çalışmada olabilirlik fonksiyonunu maksimum yapan a ve b değerleri araştırılırken olabilirlik fonksiyonunun türev bilgisini gerektirmediği aynı zamanda her iki parametrenin de pozitif olma kısıtlarının hesaplamaya katılmasında kolaylık sağladığından bu çalışmada *Genetik Algoritmaların* uygulanması araştırılmıştır.

4.2. Genetik Algoritma

Genetik algoritmanın temelleri ilk olarak John Holland tarafından Michigian Üniversitesi'nde ortaya konulmuştur. 1975 yılında John Holland bu konudaki çalışmalarını "Adaptation in Natural and Artificial Systems" adlı kitabında toplamıştır. Genetik algoritmalardaki dikkat çeken ilerleme John Holland'ın öğrencisi David E. Goldberg tarafından 1985 yılında tamamlanan "Gaz Boru hatlarının Genetik Algoritma Kullanılarak Denetlenmesi" başlıklı doktara tezi ile gerçekleşmiştir. Bu çalışmasının ardından David E. Goldberg'in 1989 yılında basılan "Makine Öğrenmesi, Arama ve Optimizasyon için Genetik Algoritma" adındaki kitabı genetik algoritmayı yeni bir seviyeye çıkarmıştır. Ayrıca bu kitap bugün dahi genetik algoritma konusunda en önemli referanslardan biridir.

Genetik algoritmadaki temel çalışma mantığı Charles Darwin'in 'Doğal Seleksiyon' ilkesidir. Darwin bu ilkeye ilişkin iki varsayım ortaya koymuştur. Bu varsayımlardan ilki, tüm canlılar fazla yavru doğurma kabiliyetine sahiptir. Bununla birlikte elenenler ile nüfus dengelenir. Diğer varsayım, belli bir tür içindeki bireyler kalıtsal özellik açısından birbirinden farlılık gösterir.

Genetik algoritmalar doğadaki evrimsel süreci temel fikir olarak kabul eden optimizasyon problemlerinin çözümü için geliştirilmiş sezgisel bir eniyileme yöntemidir. Çok boyutlu arama uzayında doğal seçilim ilkesine göre en güçlünün hayatta kaldığı bütünsel en iyi çözüm kümesini araştırır. Genetik algoritma rasgele olarak arama yöntemlerini esas alır parametre kodlama metotlarını kullanarak çalışır.

Genetik algoritmalar optimizasyon problemlerinin evrimsel süreci dijital ortamda taklit ederek çözümlerini araştırır. Diğer optimizasyon metotlarına paralel olarak çözüm için tek bir yapının geliştirilmesinden ziyade bu tipte yapıların oluşturduğu bir kümeyi geliştirir. Problemin mümkün olan çözümlerini ifade eden kümeye popülasyon adı verilir. Popülasyonlar kromozom ya da birey olarak adlandırılan sayı dizileriyle gösterilir. Kromozoma ait her birime gen adı verilir. Popülasyonu oluşturan bireyler evrimsel esaslara dayanan genetik algoritma işlemcileri vasıtasıyla saptanır.

Genetik algoritmayı diğer sezgisel optimizasyon yöntemlerinden ayıran bazı önemli özellikleri vardır. Bu özellikler aşağıdaki gibi sıralanabilir,

- Genetik algoritma parametre ile değil, direkt olarak parametre koduyla çalışır.
- Genetik algoritma aramaya tek bir noktadan değil seçilen noktalar kümesinde aramaya başlar. Dolayısıyla yerel en iyi çözümlerin yanında global en iyi çözümlere de ulaşır
- Genetik algoritmanın işleyişi içinde kullanılan operatörler rastlantısal metotları temel alır, kesin çözüm metotları kullanılmaz.
- Genetik algoritma çok sayıda değişkenlerle çalışabilir.
- Genetik algoritma problem hakkında detaylı bilgiye gerek duymaz [3].

Genetik algoritmanın sahip olduğu bazı avantajlar aşağıdaki gibi yazılabilir.

- Türev bilgisine ihtiyaç duymaz.
- Çok sayıda parametre ile çalışabilir.
- Sürekli ve ayrık parametreler için optimum çözümler bulur.
- Karmaşık amaç fonksiyonunu parametrelerini yerel optimum çözüme takılmadan hesaplar [4].

4.2.1. Genetik Algoritma Çalışma Biçimi

Genetik algoritmadaki kavramalar evrim teorisinde kullanılan kavramlardır ve genetik algoritma da oradaki anlamına yakın olarak kullanılır. Popülasyon birçok bireyin bir araya gelmesiyle oluşur. Genetik algoritmada ise popülasyon olası pek çok çözümün olduğu kümedir. Olası çözümler probleme göre kodlanmış diziler olarak bulunur. Bu dizinin her bir elamanı birey olarak adlandırılır. Bu bireylerin her biri arama uzayında belirli bir bölgeyi ifade etmektedir.

Genetik algoritmada başlangıç popülasyonu genel olarak rasgele üretilir. Bu tezde ele alınan parametre tahmin problemindeki gibi dağılımın parametrelerine ait kısıtların dikkate alınmasını gerektiren kısıtlı optimizasyon problemlerinde kısıtlar göz önünde bulundurularak başlangıç popülasyonu oluşturulmalıdır. Popülasyonu oluşturan bireylerin uygunluk fonksiyonundaki değerlerine göre optimal çözüme yakın uygunluk değeri olan bireylerin tekrar seçilme şansı yüksektir. Düşük uygunluk değerine sahip bireyler bir sonraki adım için seçilme ihtimalinin az olmasından dolayı belli adımdan sonra tamamen nüfusun dışında kalırlar. Yeni popülasyon uygunluk değeri daha iyi olan bireylerden oluşur. Böylece her adımdaki popülasyon içi yüksek uygunluk değeri olan bireyler korunak arama uzayında optimal bir çözüm bulunabilir.

Genetik algoritmalar diğer optimizasyon metotlarına göre son derece büyük arama uzayı olan problemlerde başarılı sonuçlar vermektedir. Genetik algoritma kullanılmasındaki esas hedef diğer çözüm yollarının makul sonuçlar vermediği durumlarda optimizasyon problemlerine uygun çözümlerin aranmasıdır. Belli türdeki bazı problemlerin kendine özgü çözümleri vardır. Bu durumlarda genetik algoritma tercih edilmez.

Genetik algoritma işleyişinde üç adet evrim operatörü mevcuttur. Bu operatörler seçim, çaprazlama ve mutasyon operatörleri olarak ifade edilir. Bu operatörler genetik algoritma içindeki süreçte bir sonraki adımda oluşacak olan popülasyondaki her bir birey için uygulanır. Seçim operatörü, popülasyondaki her bireyin uygunluk değerine göre yeni popülasyonun oluşturulmasında hangi bireylerin seçileceğini belirler. Çaprazlama operatörü yüksek uygunluk değeri olan seçilmiş bireylerdeki kromozomların başlangıçta belirlen olasılığa göre karşılıklı olarak yer değiştirmesini sağlar. Böylece farklı özellik içeren bireyler oluşur. Mutasyon işlemi ise yine başlangıçta tercih edilen olasılığa göre yeni elde edilen bireylerin kromozomlarından rasgele birinin içindeki bir genin değişimidir.

Genetik algoritma da önemli bir başka etmen parametrelerdir. Çoğunlukla bu parametreler seçilecek popülasyon çapını, mutasyon ve çaprazlama olasılığını, nesil sayısı ve seçim operatörünün türünden oluşur. Literatürde belli tipteki problemler için bu parametrelerin uygun değerleri üzerine çalışmalar mevcuttur.

Genetik algoritma için birbirinden farklı durdurma kriteri vardır. Bunlar, algoritma çalışırken aranan çözüme ulaşıldığında, genetik algoritmanın başında belirlenen maksimum iterasyon sayısına ulaşıldığında ya da sürekli olarak uygunluk değerinin sabit kalması elde edilen en iyi bireyin ifade ettiği küme ilgili problemin optimum çözümü olarak kabul edilir.
Genetik algoritmanın çalışma şekli aşağıda ifade edilen adımlar ile özetlenebilir.

Genetik algoritma çalışma adımları,

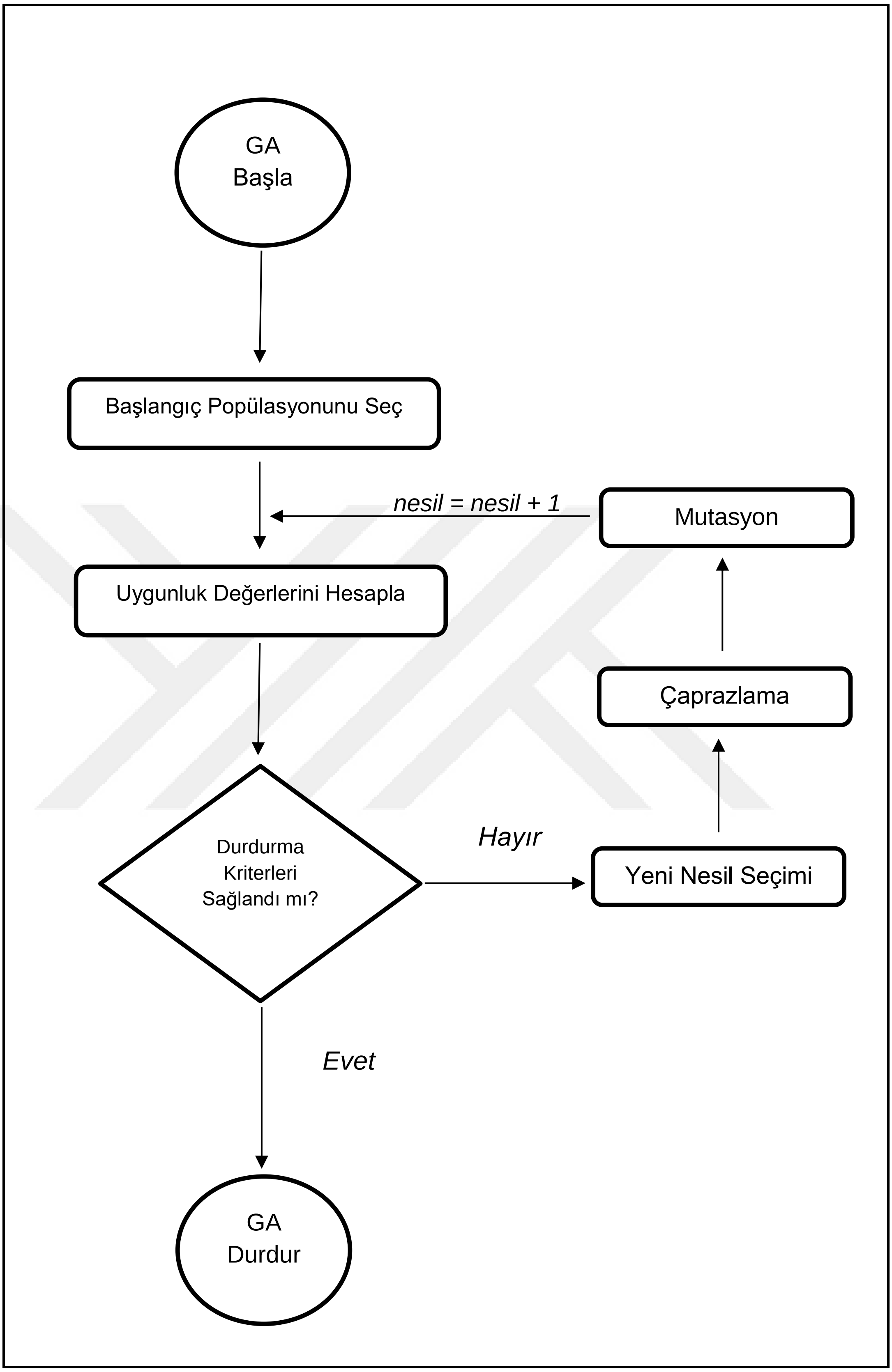
1) <u>Başlangıç:</u> Problemin çözümü için uygun olan n tane kromozomun olduğu rasgele popülasyon üretilir.
2) <u>Uygunluk:</u> Popülasyondaki her bir kromozom için uygunluk fonksiyonundaki değeri hesaplanır yüksek değerli olanlar saptanır.
3) <u>Yeni Popülasyon (Nesil):</u> Yeni bireylerden oluşan popülasyon bulunana kadar aşağıdaki adım takip edilir.
    i. *Seçim:* Daha iyi uygunluk değerine sahip iki adet ebeveyn seçilir.
    ii. *Çaprazlama:* Belirlenmiş olan çaprazlama olasılığına bağlı olarak yeni birey (çocuk) elde etmek için iki ebeveyn

kromozomları çaprazlanır. Çaprazlama işlemi yapılmadığında ise çocuk ebeveynlerin özelliklerini olduğu gibi alacaktır.

iii. *Mutasyon:* Belirlenmiş olan mutasyon olasılığına bağlı olarak çocuk birey bazı genlerinde farklılaşma olur.

iv. *Kabul etme:* Yeni birey yeni popülasyona dahil olur.

4) <u>Yenisiyle Değiştir:</u> Algoritmanın daha fazla tekrar etmesi için yeni popülasyon kullanılır.
5) <u>Test:</u> Durdurma kriteri sağlanıyorsa algoritma sonlandırılıp mevcut en son popülasyondaki en iyi çözüme dönülür.
6) <u>Döngü:</u> Eğer durdurma kriteri sağlanmıyorsa 2. adıma geri dönülür.



Şekil 4.1. GA Akış Şeması

4.2.2. Genetik Algoritmadaki Temel Kavramlar

Genetik algoritma kullanılarak ele alınan problemin kısıtlarına uygun çözümün elde edilmesinde algoritma yapısının inşası ve parametrelerin iyi bir şekilde belirlenmesi önemlidir. Bu kavramlar ve ihtiyaç duyulan parametrelerin tanımları genel olarak aşağıdaki alt başlıklarda ifade edilmiştir.

4.2.2.1. Popülasyon (Nüfus)

Probleme uygun olası çözüm kümelerinin ya da bireylerin bir araya gelerek oluşan başlangıç için genelde rasgele olarak belirlenen çözüm kümesinin adıdır. Popülasyonu oluşturacak birey sayısı problemin türü ve zorluğu dikkate alınarak araştırmacı tarafından sübjektif bir şekilde belirlenir.

Genetik algoritmanın işleyişinde elemine olabilecek bireylerin yerine yeni bireyler dahil edilmektedir. Popülasyonun çapı genetik algoritmanın çalışma süresine etki etmektedir. Büyük çaplı popülasyonlar algoritmanın çözüme ulaşma süresini uzatacağı gibi çok küçük popülasyonlar da ise en iyi çözüm bulunamadan algoritma sonlanabilir.

4.2.2.2. Kromozom

Çözüme ilişkin bilgiyi içeren ve bir ya da birden fazla genin bir araya gelerek oluşturduğu yapılara kromozom adı verilir. Kromozom oluşturan bireylere gen, popülasyonu oluşturan bireylere ise kromozom denir. Kromozomlar incelenen probleme ait verileri içerir. Bu sebeple genetik algoritmada ifade edilirken en dikkat edilmesi gereken unsurdur.

4.2.2.3. Gen

Genetik algoritmada tek başına bilgi barındıran en küçük birime gen adı verilir. Düşük düzeyde bilgi içeren bu yapıların yan yana dizilmesiyle olası çözüm kümesinin elemanları olan kromozomlar elde edilir. Genetik algoritma içerisindeki bu yapıların nasıl kodlanacağı araştırmacının elindedir. Herhangi bir gen ikili, onluk veya on altılık tabandaki sayılardan da oluşabilir. Bu sebeple probleme ve yazılan programa bağlı olarak genin bilgiyi nasıl taşıyacağı önem arz etmektedir.

4.2.2.4. Kodlama

Genetik algoritmanın performansına önemli ölçüde etki eden unsurlardan biri kodlama biçimidir. Ancak kodlama türü programa bağlı olduğu için her probleme uygun bir kodlama biçimi mevcut değildir. Bu sebeple birçok faklı kodlama biçimi geliştirilmiştir.

Kromozomlara bilginin kodlaması aşamasında üç hususa dikkat edilmelidir [28]. Bu hususlar,

- Kromozomların uygunluğu
- Kromozomun probleme ait kısıtları sağlayıp sağlamadığı
- Kodlamada kromozomun tek olması

Yaygın olarak kullanılan iki kodlama biçimi ikili ve permütasyon kodlamadır.

İkili Kodlama:

En sık tercih edilen kodlama biçimidir. Sayılar ikilik tabanda gösterildiği için ikili kodlama denmektedir. Bu kodlamada kromozomları oluşturan genler ya sıfır ya da bir değeri ile gösterilir. İkili tabanda kodlama sıklıkla kullanılsa da bazen negatif yönleri olabilir. Mesela çok değişkenli bir fonksiyonun optimal çözümü aranırken değişkenlerin sınırlarından kaynaklı yazılacak dizinin gereğinden fazla uzun olması söz konusu olabilir.

Kromozom 1	1011101100011011
Kromozom 2	1001011011101010

Şekil 4.2. İkili Kodlamaya Örnek Kromozomlar

Permütasyon Kodlama:

Bu kodlama türü çizelgeleme ve gezici satıcı gibi kombinatorik optimizasyon problemlerinde tercih edilmektedir. Yani sıralama problemlerinde permütasyonlu kodlama kullanılır. Bu kodlama biçiminde her bir kromozom birbiri ardına gelen sayı dizilerinin elemanı olan çok basamaklı bir sayı olarak gösterilir. Örnek olarak gezici satıcı problemi için araştırmaya konu olan şehirlerin arasındaki uzaklık biliniyorken bu şehirleri birbirine bağlayan en kısa yolun nasıl oluşturulacağı sorusunun çözümü için bu kodlama yardımıyla kromozomlar şehirlere nasıl bir sıra ile gidileceğini ifade eden permütasyonlar olarak yazılırlar.

Kromozom 1	153264798
Kromozom 2	856723149

Şekil 4.3. Permütasyon Kodlamaya Örnek Kromozomlar

4.2.3. Genetik Algoritmadaki Operatörler

Genetik operatörler genetik algoritmayı oluşturan ve algoritmanın çalışması sırasında popülasyonlara uygulanan işlemlerdir. Bu operatörlere ek olarak kısıtlı optimizasyon problemlerinde ihtiyaç duyulan ve problemin tipine uygun olarak ifade edilen bir diğer operatör tamir operatörüdür.

Bu operatörler algoritmadaki durdurma kriterlerini sağlamayan popülasyonlardan daha iyi bir popülasyon elde etmek için uygulanır. Bu operatörler durdurma kriterlerini sağlayan uygun popülasyona ulaşıncaya kadar çalışır. Yani genetik algoritmadaki operatörler daha iyi özellikteki bireylerden oluşan popülasyonlar üretmek için kullanılır.

4.2.3.1. Seçim Operatörü

Seçim operatörü en iyinin hayatta kalması ilkesine bağlı olarak çalışır. Popülasyondaki bazı kromozomlar yeni nesilde devam ederken bazıları da yok olur. Bu aşamada hangi kromozomun hayatta kalacağı yani hangisinin yeni nesilde varlığını sürdüreceğini belirleyen seçim operatörüdür. Seçme işlemi yüksek uygunluk değeri olan kromozomların düşük uygunluk değerine sahip kromozomların üzerine yeniden yazılmasıyla olur.

Yüksek uygunluk değerine sahip kromozomların hayatta kalma olasılığı her ne kadar yüksek olsa da düşük uygunluk değerine sahip kromozomlarında seçilmesinin mümkün olması önemlidir. Tersi durumda gelecek nesillerdeki kromozomların uygunluk değerlerinde olabilecek iyileşme engellenmiş olabilir. Seçme işlemi farklı bazı esaslara göre yapılabilir. Bu seçme işlemlerinden bazıları: rasgele seçim, rulet çarkı, turnuva seçim, sıralı seçim, kararlı durum seçimi, elitist seçim olarak yazılabilir.

Rasgele Seçim:

Bu seçimde mevcut popülasyondan rasgele olarak ebeveynler seçilir. Daha yüksek uygunluk değeri olan bireylere karşı bir ayrıcalık içermediği için genellikle bu seçim yönteminin kullanımı pek tercih edilmez.

Rulet Çarkı:

Rulet çarkı seçim yönteminde her kromozom uygunluk değerlerine göre bir çark üzerinde pasta dilimlerine benzer bir şekilde konumlanır. Çark üzerinde uygunluk değerinin toplam uygunluk değerine bölümü şeklinde oransal olarak sıralanan bu kromozomlardan değeri büyük olanın seçilme şansı daha yüksektir. Bu çark döndürülüp durduğu noktada karşılık gelen uygunluk oranı hangi kromozoma ait ise o kromozom seçilir. Buradaki sakınca sürekli olarak uygunluk değeri yüksek kromozomum seçilme durumu olabilir. Bu durum popülasyon içi çeşitliği etkileyebilir.

Turnuva Seçimi:

Turnuva seçim yönteminde kromozomlar rasgele olarak gruplara ayrılırlar ve bu gruptaki kromozomlar uygunluk değerine göre birbirleriyle rekabet eder. Bu gruplardan en yüksek uygunluk değerine sahip bireyler yeni nesil oluşumu için ebeveyn olarak seçilirler. Bu işleme toplam popülasyon sayısınca yeni birey oluşuncaya kadar devam edilir.

Bu seçim yönteminde turnuva grup çapı önemlidir ve bu seçim işleminin çalışmasını etkiler. Genel olarak turnuva grup sayısı iki seçilir Ancak bazı uygulamalarda ise daha fazla turnuva grubu belirlenir. Bu seçim küçük popülasyonlarla çalışılan durumlarda orantılı seçim yöntemlerine kıyasla daha iyi sonuç elde edilir.

Sıralı Seçim:

Sıralı seçim yöntemi rulet seçim yönteminde çok yüksek uygunluk değeri olan bir kromozomun bulunduğu durumda düşük uygunluk değerli kromozomların seçilme şansını önemli ölçüde azaltığı için önerilmiştir. Bu seçim yönteminde kromozomlar uygunluk değerlerine göre sıralanır. Seçime işlemi bu sıra dikkate alınarak yapılır. Bu durumda her kromozomun seçilme şansı makul ölçülerdedir. Ancak bu durum bu seçme yönteminde arama uzayının daha yavaş taranmasına sebep olacaktır.

Kararlı Durum Seçimi:

Kararlı durum seçiminin temel mantığı kromozomların büyük parçalarının yeni nesillere aktırılmasına dayanır. Her yeni nesilde oluşturulacak yeni kromozom yüksek uygunluk değerine sahip birkaç kromozomun seçilmesiyle oluşur. Bu şekilde oluşan yeni bireyler önceki popülasyondaki düşük uygunluk değerine sahip çıkarılmasıyla yeirne alınır. Böylece popülasyonun geri kalanı olduğu gibi yeni nesle taşınır.

Elitist Seçim:

Elitist seçim aslında bir seçme işleminden ziyade yüksek uygunluk değerine sahip bireylerin doğrudan yeni nesle aktarılmasını ifade eden bir yöntemdir. Buradaki hedef bir önceki jenerasyonun kazanımını korumaktır. Fakat çok sayıda yeni nesile taşınırsa egemen birey problemi olabilir. Bu durumda daha yüksek uygunluk değeri olan bireylerin oluşumu güçleşir.

4.2.3.2. Çaprazlama Operatörü

Bu operatör seçilen iki ebeveyn kromozomun belirli sayıdaki genlerinin karşılıklı olarak yer değiştirerek yeni bireylerin oluşmasını sağlamaya yarar. Yani bu operatör mevcut nesilden Genetik algoritmadan ebeveynlerin hangi noktadan çaprazlanacağı rasgele belirlenir. Genelde bu nokta bir tane olarak belirlenir ancak bazı durumlarda da birden fazla olması mümkündür. Bu operatör ebeveyn kromozomlarda işlem sonunda yaklaşık %50 ile %90'lık değişime sebep olabilir. Çaprazlama işlemi yapılacak kromozomlardan biri en yüksek uygunluk değerine sahip olan ile diğeri de rasgele olarak seçilir. Birçok çaprazlama şekli mevcuttur. Bunlardan bazıları tek noktalı, çift noktalı, düzgün ve aritmetik çaprazlamadır.

Tek Noktalı Çaprazlama:

Tek noktalı çaprazlamada öncelikle rasgele olarak bir nokta belirlenir. Her iki ebeveyn için rasgele seçilen noktadan yakın olan uca (sağ ya da sol) doğru olan genler karşılıklı olarak yer değiştirir.

Çaprazlanacak Ebeveyn Bireyler,

K1	101110110001**1010**
K2	100101101010**0110**

Çaprazlama Sonrası Çocuk Bireyler,

K1'	101110110001**0110**
K2'	100101101010**1010**

Şekil 4.4. Tek Noktalı Çaprazlama

Çift Noktalı Çaprazlama:

Çift noktalı çaprazlamada tek noktalı çaprazlamaya benzer şekilde seçilen iki ebeveyn kromozom üzerinde iki nokta belirlenir. Bu iki nokta arasında kalan gen parçaları karşılıklı olarak değiştirilir.

Çaprazlanacak Ebeveyn Bireyler,

K1	10111**0110**0011010
K2	10010**1101**0100110

Çaprazlama Sonrası Çocuk Bireyler,

K1	10111**1101**0011010
K2	10010**0110**0100110

Şekil 4.5. Çift Noktalı Çaprazlama

Düzgün Çaprazlama:

Düzgün çaprazlamada ebeveyn kromozomların uzunlukları kadar olmak üzere rasgele üretilmiş bir bit kalıbı kullanılır. Bu bit üzerinde 0 yazıyorsa çaprazlama için kullanılacak gen ilk ebeveynden, 1 yazıyorsa ikinci ebeveynden alınarak alınarak çaprazlama yapılır. İkinci yeni nesil kromozom için bu durumun tersi geçerlidir [29].

Örneğin, bit kalıbı: 10110010 olsun. Bu kalıba göre düzgün çaprazlama aşağıdaki gibidir.

Çaprazlanacak Ebeveyn Bireyler:

K1	10111011
K2	10010110

Çaprazlama Sonrası Çocuk Bireyler:

K1'	10111101
K2'	00000100

Şekil 4.6. Düzgün Çaprazlama

Aritmetik Çaprazlama:

Aritmetik çaprazlamada ele alınan iki ebeveynin kromozomları ve / veya gibi aritmetik işlemlere tabi tutularak yeni bireyler oluşturulur.

Örneğin, “ve” işlemi kullanılarak aritmetik çaprazlama aşağıdaki gibidir.

Çaprazlanacak Ebeveyn Bireyler,

K1	10011001
K2	11011010

Çaprazlama Sonrası Çocuk Birey,

K'	10011000

Şekil 4.7. Aritmetik Çaprazlama

4.2.3.3. Mutasyon Operatörü

Genetik algoritmada doğrudan rasgelelik prensibine dayanan işlem mutasyon operatörüdür. Genetik algoritmada bir yerden sonra popülasyondaki kromozomlarda benzerlik oluşabilir. Bu durumda çözüm uzayında daralma olabilir. Yani çaprazlama operatörü çeşitliliği korumak için yeterli olmayabilir. Bunun için mutasyon işlemi ile kromozomu oluşturan bir gen rasgele seçilip değeri değiştirilir. Bu şekilde popülasyon içi çeşitlilik sağlanır.

Mutasyon olasılığının yüksek olması problemin lokal optimum noktaya yakınmasını geciktirebilir. Bu sebeple mutasyon olasılığı küçük belirlenmeli. Böylelikle çaprazlama ile oluşan yüksek uygunluk değerine sahip bireylerin kaybolması engellenir. Bazı mutasyon işlem çeşitleri ekleme, karşılıklı değişim, kaydırma ve ters çevirme mutasyonlarıdır.

Ekleme Mutasyonu:

Ekleme mutasyonunda yeni nesle ait bir bireyden rasgele bir parça seçilip o parça rasgele bir yere konularak yapılır.

K1	2 7 3 4 9 1 **5** 8
K1'	2 7 3 **5** 4 9 1 8

Şekil 4.8. Ekleme Mutasyonu

Karşılıklı Değişim Mutasyonu:

Karşılıklı değişim mutasyonunda yeni nesle ait bir bireyden rasgele iki parça seçilip bu parçalar karşılıklı olarak yer değiştirir.

K1	2 **7** 3 4 9 **1** 5 8
K1'	2 **1** 3 4 9 **7** 5 8

Şekil 4.9. Karşılıklı Değişim Mutasyonu

Kaydırma Mutasyonu:

Kaydırma mutasyon yönteminde yeni nesle ait bir bireyden rasgele bir dizi gen seçilip o gen parçası rasgele bir konuma kaydırılarak yapılır.

K1	2 **7 3 4** 9 1 5 8
K1'	2 9 1 5 **7 3 4** 8

Şekil 4.10. Kaydırma Mutasyonu

Ters Çevirme Mutasyonu:

Ters çevirme mutasyonunda yeni nesle ait bir bireyden rasgele bir dizi gen seçilip o parçadaki genlerin ters sırada yazılmasıyla olur.

K1	2 7 **3 4 9 1** 5 8
K1'	2 7 **1 9 4 3** 5 8

Şekil 4.11. Ters Çevirme Mutasyonu

4.2.4. Genetik Algoritma Uygulama Alanları

Genetik algoritmalar karmaşık uzaylarda hızlı ve iyi sonuçlar verdiği görülmüştür. Genetik algoritmaların geniş bir uygulama alanı olmasının sebebi hesaplamalardaki kolaylığı ve problemin barındırdığı kısıtlara kolay bir şekilde uyum sağlayabilir olmasıdır. Genetik algoritmalar genellikle geleneksel metotların çözüm bulamadığı durumlarda tercih edilir ancak sezgisel bir metot olduğundan optimal çözümü tam olarak hesaplayamayabilir. Genetik algortiama İlk olarak doğrusal olmayan optimizasyon problemlerinde denenmiş daha sonra gezici satıcı, ders ve sınav programları gibi kombinatorik optimizasyon problemleri için de uygulanmış ve iyi sonuçlar verdiği görülmüştür [30]. Genetik algoritmanın optimizasyondaki uygulama alanları otomatik programlama ve bilgi sistemleri, mekanik öğrenme, ekonomik ve sosyal sistem modelleridir bunun yanında işletmedeki uygulama alanları ise finans, pazarlama, montaj hattı dengeleme problemi, çizelgeleme problemi, tesis yerleşim problemi, atama problemi, hücresel üretim problemi, sistem güvenilirliği problemi, taşıma problemi, gezgin satıcı problemi, araç rotalama problemi minimum yayılan ağaç problemleridir [31].

5. KUMARASWAMY DAĞILIMININ PARAMETRELERİNİN GENETİK ALGORİTMA İLE R PROGRAMINDATAHMİN EDİLMESİ

Kumaraswamy dağılımı parametrelerinin tahmin edilmesi probleminde a ve b' nin pozitif olma koşulu göz önüne alındığında bu tahmin problemi bir optimizasyon problemi olarak düşünülebilir. Bu optimizasyon problemi ise,

$$\begin{gathered} max\ L(a,b;x) \\ a > 0 \\ b > 0 \end{gathered}$$

olarak ifade edilebilir. Burada L , a ve b parametrelerinin bir fonksiyonu olan olabilirlik fonksiyonudur. Olabilirlik fonksiyonunu maksimum yapan pozitif a ve b değerlerinin bulunmasında klasik yaklaşım olan türev alarak analitik çözüm elde edilemediği için alternatif bir çözüm yolu olarak kıstlı optimizasyon problemi şeklinde ifade ederek çözüm aranabilir. Kumaraswamy dağılımının parametrelerinin tahmini için hem basit rasgele örnekleme hem de sıralı küme örneklemesi ile tahmin edilmesi problemleri ayrı ayrı kısıtlı birer optimizasyon problemleri olarak incelenebilir.

5.1. Basit Rasgele Örnekleme için Kumaraswamy Dağılımı Parametrelerinin Genetik Algoritma ile Tahmin Edilmesi

Basit rasgele örnekleme için Kumaraswamy dağılımının olabilirlik fonksiyonu eşitlik (1) de ifade edildiği gibidir. Basit rasgele örnekleme için Kumaraswamy dağılımının parametrelerinin genetik algoritma ile tahmin edilmesinde R programındaki "GA" paketinden faydalanılmıştır [32]. Bu paketi kullanarak parametre tahmini yapmadan önce R programında yapılması gerekli adımlar aşağıda verilmiştir. Bunlardan birincisi Kumaraswamy dağılımından rasgele gözlem elde edilmesidir. Bu adım "Simülasyon Çalışması" başlığı altında ifade edilecektir. İkinci olarak ise maksimumu aranan olabilirlik fonksiyonunun komutları programa yazılmalıdır. Buradan hareketle basit rasgele örnekleme için Kumaraswamy dağılımı olabilirlik fonksiyonu $L_{brö}(a, b; x)$ ile gösterilmektedir. Bu fonksiyonun R programındaki kodları,

```
Likelihood.bro = function(ab, x) {
 a = ab[1]
 b = ab[2]
 n = length(x)

 carpim1 = 1
   for(i in 1:n){
     carpim1 = carpim1 * (x[i])^(a-1)
  }

 carpim2 = 1
   for(i in 1:n) {
     carpim2 = carpim2 * (1-(x[i]^a))^(b-1)
 }

 res  = a^n * b^n * carpim1 * carpim2

 return(res)
}
```

olarak yazılabilir. Burada, ab Kumaraswamy dağılımının parametre vektörü, x ise Kumaraswamy dağılımından elde edilen gözlemleri ifade etmektedir.

Bu iki adım tamamlandıktan sonra genetik algoritmanın üzerinde çalışacağı fitness fonksiyonu tanımlanır. Bu fonksiyon tanımlanırken $L_{brö}(a, b; x)$ fonksiyonu kullanılır.

```
fitness= function(ab, x) {
  res = Likelihood.brö(ab, x)

  return (res)
}
```

Son olarak "GA" paketi kullanılarak genetik algoritmayı çalıştıracak komut aşağıdaki gibidir.

```
GA.sol <- ga(type = "real-valued", fitness=Likelihood.srs, x=x, min = c(0.1, 0.1),
max = c(5, 5), popSize = 200, pcrossover = 0.7, pmutation = 0.3, maxiter = 100)
```

Burada min ve max argümanları çözümün hangi aralıkta aranacağını, popSize ise popülasyon(nüfus) sayısını, pcrossover ise kromozomların çaprazlanma olasılıklarını, pmutation ise çaprazlanan kromozomların mutasyona uğrama olasılıklarını ve maxiter maksimum iterasyon sayısını ifade etmektedir. Bu argümanlara atanacak değerler optimizasyon problemine uygun olarak seçilmelidir.

$L_{brö}(a, b; x)$ fonksiyonu için çalıştırılan genetik algoritma sonuçları ise,

```
teta = GA.sol@solution
```

komutu ile ekrana yazdırılabilir. Bu adımların tümü bir denemeyi ifade etmektedir. Bu adımlar istenen sayıda tekrar edilerek daha hassas sonuçlar elde edilebilir.

5.2. Sıralı Küme Örnekleme için Kumaraswamy Dağılımı Parametrelerinin Genetik Algoritma ile Tahmin Edilmesi

Sıralı küme örneklemesi için Kumaraswamy dağılımının olabilirlik fonksiyonu eşitlik (2) de ifade edildiği gibidir. Burada da bu örnekleme yöntemi ile parametre tahmini yapılırken R programındaki “GA” paketinden faydalanılmıştır. Yine paketteki ilgili komutu uygulamadan önce bölüm 3.2.1. de anlatıldığı gibi sıralı küme örneklemesinden gözlem üretilip gerekli fonksiyonların yazılması gerekmektedir.

Sıralı küme örneklemesi için Kumaraswamy dağılımının olabilirlik fonksiyonu $L_{skö}(a, b; x)$ ile gösterilmektedir. Bu fonksiyonun R programındaki kodları,

```
ab = c(a,b)
Likelihood.sko = function(ab, m, r, x) {
 a = ab[1]
 b = ab[2]
 X = matrix(x, nrow=m, ncol=r)

 carpim1 = 1
 for(j in 1:r) {
  for(i in 1:m)
    carpim1 = carpim1 * (X[i,j])^(a-1)
 }

 carpim2 = 1
 for(j in 1:r) {
  for(i in 1:m)
    carpim2 = carpim2 * (1 - X[i,j]^a)^(b*(m-i+1)-1)
 }

 carpim3 = 1
```

```
 for(j in 1:r) {
  for(i in 1:m)
    carpim3 = carpim3 * (1 - (1 - X[i,j]^a)^b)^(i-1)
 }

  c = 1
  for(i in 1:m) {
    c = c * (factorial(m)/(factorial(i-1)*factorial(m-i)))
 }

 res = c^r * a^(m*r) * b^(m*r) * carpim1 * carpim2 * carpim3

 return(res)
}
```

olarak yazılabilir. Burada, ab Kumaraswamy dağılımının parametre vektörünü, x Kumaraswamy dağılımından sıralı küme örneklemesi ile elde edilen gözlemleri, m ve r ise küme çapı ile döngü sayısını ifade etmektedir.

Bu iki adım tamamlandıktan sonra genetik algoritmanın üzerinde çalışacağı fitness fonksiyonu tanımlanır. Bu fonksiyon tanımlanırken $L_{skö}(a, b; x)$ fonksiyonu kullanılır.

```
fitness = function(ab, m, r, x) {
 res = Likelihood.sko(ab, m, r, x)

 return (res)
}
```

Son adımda “GA” paketi kullanılarak genetik algoritmayı çalıştıracak komut aşağıdaki gibi yazılır.

```
GA.sol <- ga(type = "real-valued", fitness=Likelihood.sko, m=m, r=r, x=x,
  min = c(0.1, 0.1), max = c(5, 5), popSize = 200, pcrossover = 0.7,
  pmutation = 0.3, maxiter = 100)
```

$L_{skö}(a, b; x)$ fonksiyonu için çalıştırılan genetik algoritma sonuçları ise,

```
teta = GA.sol@solution
```

komutu ile ekrana yazdırılabilir. Bu adımların tümü bir denemeyi ifade etmektedir. Bu adımlar istenen sayıda tekrar edilerek daha hassas sonuçlar elde edilebilir.

6. SİMÜLASYON ÇALIŞMASI

Burada ele alınan parametre tahmin problemi için en çok olabilirlik fonksiyonlarının genetik algoritma kullanılarak çözülmesinde R yazılımı kullanılmıştır. R yazılımında kullanıcılara büyük kolaylık sağlayan başka araştırmacılar tarafından hazırlanmış birçok konuya ilişkin paketler bulunmaktadır [33]. Bu simülasyon çalışmasında genetik algoritmadaki hesaplamalar için “*GA*” paketi kullanılmıştır [32]. Ayrıca Kumaraswmay dağılımından sayı üretmek için “*VGAM*” paketi kullanılmıştır [34]. Bu paket yardımıyla sayı üretmek için ilgili komut,

$$\mathrm{rkumar}(\mathrm{n};\ \mathrm{shape1};\ \mathrm{shape2})$$

şeklinde ifade edilir. Burada n gözlem sayısı, $shape1$ ve $shape2$ ise gözlem üretmek istediğimiz Kumaraswamy dağılımının parametre değerleridir.

Kumaraswamy dağılımının bilinmeyen parametrelerine ait tahmin edicilerin performansını değerlendirebilmek için yan, hata kareler ortalaması ve etkinlikleri hem basit rasgele hem de sıralı küme örneklemesi için örneklem çapının 20 ve 30 olduğu durumlardaki R sonuçları Çizelge 5.1.-6. daki tablolarda verilmiştir. R yazılımı yardımıyla yapılan bu simülasyonda her bir deney için 1000 tekrar yapılmıştır.

Parametre tahmininde yan, hata kareler ortalaması ve etkinlik aşağıdaki eşitliklerdeki gibi ifade edilmektedir.

a tahmin edilmek istenen parametre ve $\hat{a}$ da tahmin edicisi olmak üzere;

Tahmin edici için yanlılık,

$$Yan(a) = E[a] - a \tag{3}$$

Hata kareler ortalaması,

$$HKO(a) = E_a[(a - a)^2]$$
$$= Var(a) + Yan(a)^2 \quad (4)$$

$a_{brö}$, a parametresi için basit rasgele örnekleme ile elde edilen tahmin edici ve $a_{skö}$, sıralı küme örneklemesi ile elde edilen tahmin ediciyi göstermektedir. Buna göre etkinlik,

$$Etk\left(a_{skö}\right) = \frac{HKO\left(a_{brö}\right)}{HKO\left(a_{skö}\right)} \quad (5)$$

olarak ifade edilir.

Buna göre, eşitlik (3), (4) ve (5) kullanılarak tahmin edicilerin performansları hesaplanmıştır. Hem basit rasgele örnekleme hem de sıralı küme örneklemesi için bulanan tahmin edicilere ait sonuçlar Çizelge 5.1. – 5.8. de verilmiştir. Bu sonuçlara göre sıralı küme örneklemesi tahmin edicilerinin basit rasgele örneklemeye göre daha küçük yanlılık ve hata kareler ortalamasına sahip olduğu görülmüştür. Her iki örnekleme yönteminde de örneklem çapı arttıkça Kumaraswamy dağılımının b şekil parametresinin hem yanlılığı hem de hata kareler ortalaması değerinin arttığı görülmüştür. Ayrıca sıralı küme örneklemesinde küme çapı artıp döngü sayısı azaldığında ise sıralı küme örneklemesi tahmin edicilerinin yanlılıklarının ve hata kareler ortalama değerlerinin azaldığı görülmektedir.

Çizelge 5.1. $a=0.5$ ve $n=20$ için Tahmin Sonuçları

Tablo 1 *a=0.5* ve *n=20* için Yan, HKO ve Etkinlik												
n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
20	0,5	0,088374	0,060193	0,05735	0,031771	2, 10	0.07534	0,054308	1,108361	0,048454	0.022385	1,419275
						4, 5	0.059749	0.032536	1,850055	0.042348	0.016695	1,903007
						5, 4	0.051246	0.029633	2,031296	0.036108	0.016203	1,960761
20	1	0,06536	0,032461	0,167268	0,17643	2, 10	0.050801	0.025337	1,281168	0.123049	0.137414	1,283930
						4, 5	0.045840	0.020459	1,586597	0.112418	0.111933	1,576213
						5, 4	0.041864	0.017754	1,828357	0.097812	0.086123	2,048584
20	3	0,025548	0,012906	0,436129	1,256876	2, 10	0,025408	0,011296	1,142457	0,401267	1,175696	1,069048
						4, 5	0.024611	0,0080508	1,603098	0.346893	0.952953	1,318927
						5, 4	0.025173	0.008002	1,612878	0.325377	0.897973	1,399680

Çizelge 5.2. $a = 0.5$ ve $n = 30$ için Tahmin Sonuçları

Tablo 2 *a=0.5* ve *n=30* için Yan, HKO ve Etkinlik													
n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
30	0,5	0,061985	0,035985	0,037967	0,016953	2, 15	0.054098	0.027070	1,329356	0.037366	0.014222	1,192040
						3, 10	0.037109	0.021082	1,706897	0.023994	0.010783	1,572155
						5, 6	0.039403	0.017302	2,079833	0.025053	0.0083764	2,023973
30	1	0,036555	0,017926	0,090156	0,093921	2, 15	0.031813	0.016286	1,100730	0.087642	0.084995	1,105014
						3, 10	0.030848	0.012745	1,406522	0.081264	0.069973	1,342249
						5, 6	0.022565	0.010672	1,679726	0.055357	0.048851	1,922581
30	3	0,024608	0,009472	0,325066	0,949518	2, 15	0.020251	0.007743	1,223223	0.276112	0.808685	1,174151
						3, 10	0.016001	0.006633	1,427887	0.251781	0.755933	1,256087
						5, 6	0.016482	0.005129	1,846512	0.252278	0.640009	1,483602

Çizelge 5.3. $a=1$ ve $n=20$ için Tahmin Sonuçları

Tablo 3 *a=1* ve *n=20* için Yan, HKO ve Etkinlik													
n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
20	0,5	0.203177	0.275269	0.065377	0.034602	2, 10	0.140301	0.169071	1,628123	0.046883	0.023575	1,467771
						4, 5	0.132686	0.141853	1,940519	0.041436	0.017903	1,932779
						5, 4	0.112759	0.122256	2,251575	0.034852	0.014289	2,421516
20	1	0.119265	0.132142	0.141339	0.196927	2, 10	0.086105	0.101852	1,297387	0.118082	0.158831	1,239852
						4, 5	0.072270	0.076830	1,719929	0.099678	0.102097	1,92881
						5, 4	0.068969	0.068492	1,929282	0.089662	0.086843	2,267623
20	3	0.061031	0.055177	0.377469	1,229051	2, 10	0.054179	0.042201	1,307477	0.388056	1,13269	1,085073
						4, 5	0.046867	0.035579	1,55083	0.341375	1,022173	1,20239
						5, 4	0.043480	0.030719	1,796142	0.307069	0.906718	1,355494

Çizelge 5.4. $a=1$ ve $n=30$ için Tahmin Sonuçları

Tablo 4 *a=1* ve *n=30* için Yan, HKO ve Etkinlik

n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
30	0,5	0.1330166	0.1540919	0.03901874	0.0206853	2, 15	0.1089445	0.1089727	1,414041	0.03577562	0.01448028	1,428515
						3, 10	0.101119	0.1032881	1,491865	0.029383	0.01140256	1,814093
						5, 6	0.05909006	0.06540928	2,355811	0.02210925	0.007980109	2,592107
30	1	0.08230964	0.07547317	0.09773945	0.09284563	2, 15	0.07567052	0.05965786	1,2651	0.08997415	0.0785448	1,182072
						3, 10	0.06616342	0.05217908	1,446426	0.07720715	0.06071519	1,529199
						5, 6	0.0449612	0.04104466	1,838806	0.06010113	0.04929051	1,883641
30	3	0.04889742	0.0387163	0.3556983	0.999691	2, 15	0.04490134	0.03091828	1,252214	0.3130421	0.8605425	1,161699
						3, 10	0.03389245	0.02554992	1,51532	0.2415101	0.7561437	1,322091
						5, 6	0.03078124	0.02158976	1,793271	0.2612763	0.6714346	1,488888

Çizelge 5.5. $a=2$ ve $n=20$ için Tahmin Sonuçları

Tablo 5 *a=2* ve *n=20* için Yan, HKO ve Etkinlik													
n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
20	0,5	0,382023	0,941333	0,060942	0,029546	2, 10	0.290714	0.743901	1,265401	0.048928	0.023084	1,279941
						4, 5	0.245506	0.552013	1,705271	0.038185	0.016930	1,745178
						5, 4	0.228650	0.523412	1,798455	0.034047	0.014731	2,005656
20	1	0,246559	0,536713	0,166546	0,192373	2, 10	0.227287	0.414813	1,293867	0.135167	0.137564	1,398428
						4, 5	0.177227	0.335844	1,598103	0.122047	0.103851	1,852388
						5, 4	0.174240	0.303449	1,768706	0.109885	0.095698	2,010213
20	3	0,138711	0,228225	0,432455	1,341007	2, 10	0.111385	0.174348	1,309016	0.436769	1,215651	1,103118
						4, 5	0.075750	0.136543	1,671446	0.324192	0.979115	1,369611
						5, 4	0.106859	0.136831	1,667926	0.372746	0.989681	1,354989

Çizelge 5.6. $a=2$ ve $n=30$ için Tahmin Sonuçları

Tablo 6 *a=2* ve *n=30* için Yan, HKO ve Etkinlik													
n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
30	0,5	0,261381	0,516551	0,046363	0,017242	2, 15	0.190457	0.450267	1,147209	0.035338	0.013733	1,255468
						3, 10	0.160694	0.351617	1,469070	0.026096	0.010813	1,594450
						5, 6	0.128157	0.269591	1,916050	0.022614	0.0081098	2,126082
30	1	0,124311	0,288139	0,087494	0,089463	2, 15	0.116544	0.229367	1,256236	0.08098	0.073754	1,212987
						3, 10	0.142315	0.227652	1,265698	0.081721	0.067241	1,330475
						5, 6	0.119027	0.158514	1,817756	0.073107	0.049286	1,815167
30	3	0,107734	0,143861	0,32609	0,881651	2, 15	0.089293	0.122493	1,174439	0.306599	0.825095	1,068545
						3, 10	0.074883	0.111372	1,291716	0.264314	0.761836	1,157272
						5, 6	0.065650	0.083537	1,722128	0.252783	0.623547	1,413929

Çizelge 5.7. $a = 5$ ve $n = 20$ için Tahmin Sonuçları

Tablo 7 *a=5* ve *n=20* için Yan, HKO ve Etkinlik

n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
20	0,5	-0,41633	0,621401	0,006843	0,014403	2, 10	-0,446375	0,617766	1,005884	-0.003083827	0.009962	1,445686
						4, 5	-0,3462043	0,41925	1,482174	-0.008382018	0.006195	2,324850
						5, 4	-0,3532828	0,426078	1,45842	-0.008343514	0.005507	2,615274
20	1	-0,35753	0,456462	0,012417	0,062108	2, 10	-0.3504003	0.429697	1,062288	-0.02032202	0.043735	1,420087
						4, 5	-0.3184731	0.366049	1,246998	-0.03211455	0.027561	2,253416
						5, 4	-0.3287275	0.358925	1,271746	-0.03129469	0.025262	2,458538
20	3	-0,30483	0,328324	-0,13472	0,546238	2, 10	-0.281435	0.298837	1,098672	-0,1297896	0.408731	1,336423
						4, 5	-0.2435803	0.218021	1,505930	-0.1098621	0.300562	1,817388
						5, 4	-0.2555532	0.221061	1,485219	-0,1049975	0.265820	2,054913

Çizelge 5.8. $a=5$ ve $n=30$ için Tahmin Sonuçları

Tablo 8 *a=5* ve *n=30* için Yan, HKO ve Etkinlik

n	*b*	$Yan(a_{brö})$	$HKO(a_{brö})$	$Yan(\hat{b}_{brö})$	$HKO(\hat{b}_{brö})$	*m, r*	$Yan(a_{skö})$	$HKO(a_{skö})$	$etk(a_{skö})$	$Yan(\hat{b}_{skö})$	$HKO(\hat{b}_{skö})$	$etk(\hat{b}_{skö})$
30	0,5	-0,37602	0,465254	0,001317	0,009231	2, 15	-0.3653908	0.459669	1,01215	-0.0127483	0.0066456	1,388916
						3, 10	-0.3252529	0.3584408	1,297996	-0.007982075	0.0050668	1,821720
						5, 6	-0.3085028	0.328711	1,415388	-0.009368466	0.0039365	2,344738
30	1	-0,30126	0,326815	0,000318	0,043268	2, 15	-0.3130707	0,319501	1,022892	-0.01927488	0.028803	1,502219
						3, 10	-0.2954372	0.284515	1,148673	-0.0219347	0.0228527	1,893358
						5, 6	-0.2768795	0.247779	1,318976	-0.03424835	0.0174273	2,482787
30	3	-0,2809	0,267421	-0,11048	0,393714	2, 15	-0,2336411	0,183719	1,455597	-0.0831274	0.2851602	1,380679
						3, 10	-0,233713	0,178674	1,496692	-0,117977	0.2438374	1,614661
						5, 6	-0,2244707	0,174645	1,531223	-0,1215469	0.1902251	2,069731

7. TARTIŞMA VE SONUÇ

Bu tez çalışmasında parametre tahmin problemlerinde kolay bir şekilde türevi alınamayan karmaşık olabilirlik fonksiyonların çözümünde ve burada olduğu gibi parametrelere ilişkin kısıtların olduğu durumlarda çözüm için genetik algoritmanın değerlendirilebileceği gösterilmiştir. Ayrıca iki farklı örnekleme yöntemi kullanılarak bu yöntemlerden elde edilen tahmin ediciler karşılaştırılmıştır.

Simülasyon sonuçlarında görüldüğü üzere sıralı küme örneklemesi ile elde edilen tahmin edicilerin basit rasgele örnekleme ile elde edilen tahmin edicilere kıyasla daha küçük yan ve hata kareler ortalamasına sahip olduğu ve hata kareler ortalaması oranlarıyla bulunan etkinlikler açısından da daha iyi olduğu gösterilmiştir. Ayrıca sıralı küme örneklemesinde küpe çapı *m*’ nin arttığı ve döngü sayısı *r*’ nin azaldığı durumlarda da sıralı küme örneklemesi tahmin edicilerinin performansında iyileşme olduğu görülmektedir. Bunun yanında hem basit rasgele örnekleme hem de sıralı küme örneklemesi için örneklem çapı *n* arttığında her iki örnekleme yöntemine ait tahmin edicilerin hata kareler ortalaması’ nın (HKO) azaldığı görülmektedir.

Çizelge 5.3 ve 5.4’ deki sonuçlar ile bu tezde referans alınan çalışma olan M. A. Hussian’ nın çalışmasındaki sonuçlar [20] ile karşılaştırıldığında genetik algoritma Kumaraswamy dağılımının b parametresini her iki örnekleme yöntemi için de önemli ölçüde daha küçük hata kareler ortalaması (HKO) ile tahmin ettiği görülmüştür. Bunun yanında yine her iki örnekleme için de genetik algoritma Kuamraswamy dağılımının a ve b parametrelerini daha küçük yanlılıkla tahmin etmiştir. Bu sonuçlar doğrultusunda genetik algoritmanın Kumaraswamy dağılımı parametrelerini iki farklı örnekleme yönteminde de daha iyi tahmin ettiği söylenebilir.

Sıralı küme örneklemesinde genel olarak en çok olabilirlik fonksiyonu daha karmaşık bir yapıdadır. Bu nedenle bu tahmin problemi kısıtlı bir optimizasyon problemi olarak ifade edilerek problemin çözümü için Newton-Raphson veya Karush-Kuhn-Tucker koşullarını temel alan Lagrange algoritması düşünülebilir. Problemin çözümü için de R gibi yazılımlarda bazı optimizasyon fonksiyonları kullanılabilir. Ancak bu tür problemlerin çözümünde yakınsama ve başlangıç değerlerinin seçimine bağlı hesaplama zorlukları olduğu saptanmıştır. Bu nedenle de bu çalışmada olduğu gibi bu tür durumlarda genetik algoritma yaklaşımının iyi bir alternatif olabileceği görülmüştür.

Sonuç olarak parametrelere ait kısıtların olduğu tahmin problemleri kısıtlı optimizasyon problemi olarak düşünülüp farklı metotlar denenebilir. Ayrıca kısıtlı optimizasyon problemi olarak ifade edilen tahmin problemlerinde çözüm için kullanılacak yöntem için başlangıç değerlerinin belirlenmesinde genetik algoritmadan elde edilen sonuçlar da kullanılabilir.

KAYNAKLAR

[1] Holland, J.H., Adaptation in Natural and Artifcial Systems, *MIT Press,* 1975.

[2] Louis, S., Li, G., Case injected genetic algorithms for travelling salesman problems, *Information Sciences,* no. 12, pp. 201-225, 2000.

[3] Goldberg, D., Genetic Algorithms in Search, Optimization and Machine Learning, *Addison-Wesley,* 1989.

[4] Beasley, D., Bull, D. R., and Martin, R. R., An overview of genetic algorithms: Part 1, fundamentals, *University computing,* vol 2, no. 15, pp. 56-69, 1993.

[5] Beasley, D., Bull, D. R., ve Martin, R. R., An overview of genetic algorithms: Part 2, research topics, *University computing,* vol 4, no. 15, pp. 170-181, 1993.

[6] McIntyre, G. A., A method for unbiased selective sampling, using ranked sets, *Australian Journal of Agricultural Research,* pp. 385-390, 1952.

[7] Takahasi, K., and Wakimoto, K., On unbiased estimates of the population mean based on the sample stratifed by means of ordering, *Annals of the Institute of Statistical Mathematics,* pp. 1-31, 1968.

[8] Dell, T. R., and Clutter, J. L., Ranked set sampling theory with order statistics background, *Biometrics,* pp. 545-555, 1972.

[9] Stokes, S. L., Ranked set sampling with concomitant variables, *Communications in Statistics-Theory and Methods,* pp. 1207-1211, 1977.

[10] Samawi, H. M., Stratified ranked set sample, Pakıstan Journal of Statistics-All Series, no. 12, pp. 9-16, 1996.

[11] Samawi, H. M., Ahmed, M. S., and Abu-Dayyeh, W., Estimating the population mean using extreme ranked set sampling, *Biometrical Journal,* cilt 5, no. 38, pp. 577-586, 1996.

[12] Al-Saleh, M. F., and Al-Kadiri M. A., Double-ranked set sampling, *Statistics & Probability Letters,* cilt 2, no. 48, pp. 205-212, 2000.

[13] Al- Saleh, M. F., and Al-Omari A. I., Multistage ranked set sampling, *Journal of Statistical planning and Inference,* cilt 2, no. 102, pp. 273-286, 2002.

[14] Muttlak, H. A., Investigating the use of quartile ranked set samples for estimating the population mean, *Applied Mathematics and Computation,* vol 2, no. 146, pp. 437-443, 2003.

[15] Al-Nasser, A. D., and Mustafa A. B., Robust extreme ranked set sampling, *Journal of Statistical Computation and Simulation,* vol 7, no. 79, pp. 859-867, 2009.

[16] Jemain, A. A., Al-Omari, A., and Ibrahim K., Balanced groups ranked set samples for estimating the population median, *In: New Developments in Applied Statistics. Nova Science Publishers, Inc.,* 2012.

[17] Arslan, G., and Ozturk O., Parametric inference based on partially rank ordered set samples, *Journal of the Indian Statistical Association,* no. 51, pp. 1-24, 2013.

[18] Kumaraswamy, P., A generalized probability density function for double-bounded random processes., *Journal of Hydrology,* vol 1-2, no. 46, pp. 79-88, 1980.

[19] Jones, M. C., Kumaraswamy's distribution: A beta-type distribution with some tractability advantages, *Statistical Methodology,* vol 1, no. 6, pp. 70-81, 2009.

[20] Hussian, M. A., Bayesian and maximum likelihood estimation for Kumaraswamy distribution based on ranked set sampling, *American Journal of Mathematics and Statistics,* vol 1, no. 4, pp. 30-37, 2014.

[21] Cingi, H., Örnekleme Kuramı 3. Baskı, Ankara: Bizim Büro Basımevi, 2009.

[22] Cochran, W. G., Sampling Techniques 3d Ed., Wiley, 1977.

[23] Scheaffer, R. L., Mendenhall, W. and Ott, L., Elementary survey sampling, California: Duxbury Press, 1990.

[24] Patil, G. P., Surucu, B. and Egemen D., Ranked set sampling, *Wiley StatsRef: Statistics Reference Online,* 2002.

[25] Wolfe, D. A., Ranked set sampling: its relevance and impact on statistical inference, *ISRN Probability and Statistics,* 2012.

[26] Chen, Z., Bai Z. and Sinha B., Ranked set sampling: theory and applications, Springer Science & Business Media, 2003.

[27] Öztürk, F., Akdi, Y., Aydoğdu, H. ve Karabulut, İ., Parametre Tahmini ve Hipotez Testi, Ankara: Seçkin Yayıncılık, 2006.

[28] Cheng, R., Gen, M. and Tsujimura, Y., A tutorial survey of job-shop scheduling problems using genetic algorithms—I. Representation., *Computers & industrial engineering,* vol 4, no. 30, pp. 983-997, 1996.

[29] Syswerda, G., Uniform crossover in genetic algorithms, *In: Proceedings of the third international conference on Genetic algorithms. Morgan Kaufmann Publishers,* pp. 2-9, 1989.

[30] Liepins, G. E., Hilliard, M. R., Richardson, J. and Palmer, M., Genetic algorithms applications to set covering and traveling salesman problems, *In: Operations research and Artificial Intelligence: The integration of problem-solving strategies. Springer, Dordrecht,* pp. 29-57, 1990.

[31] Emel, G. G., ve Taşkın, Ç., Genetik Algoritmalar ve Uygulama Alanlari, *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,* cilt 1, no. 21, pp. 129-152, 2002.

[32] Scrucca, L., GA: A Package for Genetic Algorithms in R., Journal of Statistical Software, 53(4), 1-37., 2013. [Çevrimiçi]. Available: http://www.jstatsoft.org/v53/i04/.

[33] R Core Team (2017). R: A language and environment for statistical computing., R Foundation for Statistical Computing, Vienna, Austria., [Çevrimiçi]. Available: https://www.R-project.org/.

[34] Yee, T. W., The VGAM Package for Categorical Data Analysis., Journal of Statistical Software, 32(10), 1-34., 2010. [Çevrimiçi]. Available: http://www.jstatsoft.org/v32/i10/.