KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

ÇOKGENSEL ALAN İÇERİSİNDE İKİ DEĞİŞKENLİ KEYFİ BİR DAĞILIMDAN RASTGELE SAYI ÜRETİLMESİ

YÜKSEK LİSANS TEZİ

Buğra Kaan TİRYAKİ

ARALIK 2014
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

ÇOKGENSEL ALAN İÇERİSİNDE İKİ DEĞİŞKENLİ KEYFİ BİR DAĞILIMDAN RASTGELE SAYI ÜRETİLMESİ

Buğra Kaan TİRYAKİ

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"YÜKSEK LİSANS (İSTATİSTİK)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 05.12.2014
Tezin Savunma Tarihi : 26.12.2014

Tez Danışmanı : Yrd. Doç. Dr. Orhan KESEMEN

Trabzon 2014

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü

İstatistik ve Bilgisayar Bilimleri Anabilim Dalında

Buğra Kaan TİRYAKİ tarafından hazırlanan

ÇOKGENSEL ALAN İÇERİSİNDE İKİ DEĞİŞKENLİ KEYFİ BİR DAĞILIMDAN RASTGELE SAYI ÜRETİLMESİ

başlıklı bu çalışma, Enstitü Yönetim Kurulunun 09 / 12 / 2014 gün ve 1580 sayılı kararıyla oluşturulan jüri tarafından yapılan sınavda

YÜKSEK LİSANS TEZİ

olarak kabul edilmiştir.

Jüri Üyeleri

Başkan	**: Prof. Dr. Mehmet AKBAŞ**	**…...…………………………**
Üye	**: Yrd. Doç. Dr. Orhan KESEMEN**	**…...…………………………**
Üye	**: Yrd. Doç. Dr. Uğur ŞEVİK**	**……...…………………………**

Prof. Dr. Sadettin KORKMAZ

Enstitü Müdürü

ÖNSÖZ

"Çokgensel Alan İçerisinde İki Değişkenli Keyfi Bir Dağılımdan Rastgele Sayı Üretilmesi" isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik ve Bilgisayar Bilimleri Anabilim Dalı, Yüksek Lisans Programı'nda hazırlanmıştır.

Başta tez çalışma sürecince bana bu konuyu öneren, sevdiren ve beni araştırmaya yönlendiren Karadeniz Teknik Üniversitesi, İstatistik ve Bilgisayar Bilimleri bölümü öğretim üyelerinden danışmam hocam Yrd. Doç. Dr. Orhan KESEMEN'e çok değerli bilgilerini benimle paylaştığı için ve bilim adamı olma yolundaki ilk adımlarımı atarken yolumu aydınlattığı için gönülden teşekkürü borç bilirim. Ayrıca yetişmemde katkısı olan tüm hocalarıma ve tez çalışması sürecinde hiçbir yardımdan kaçınmayan arkadaşım Özge TEZEL'e teşekkür ederim.

Son olarak, onlardan aldığım güç ile hayattaki tüm zorlukları yenebileceğime inandığım, beni hep seven ve destekleyen aileme en derin sevgi ve teşekkürlerimi sunarım.

Bu araştırmanın literatüre katkı sağlamasını dilerim.

Buğra Kaan TİRYAKİ

Trabzon 2014

TEZ BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum "Çokgensel Alan İçerisinde İki Değişkenli Keyfi Bir Dağılımdan Rastgele Sayı Üretilmesi" başlıklı bu çalışmayı baştan sona kadar danışmanım Yrd. Doç. Dr. Orhan KESEMEN'in sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 05/12/2014

Buğra Kaan TİRYAKİ

İÇİNDEKİLER

Yüksek Lisans

ÖZET

ÇOKGENSEL ALAN İÇERİSİNDE İKİ DEĞİŞKENLİ KEYFİ BİR DAĞILIMDAN RASTGELE SAYI ÜRETİLMESİ

Buğra Kaan TİRYAKİ

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik ve Bilgisayar Bilimleri Anabilim Dalı
Danışman: Yrd. Doç. Dr. Orhan KESEMEN

2014, 62 Sayfa

Bu çalışmada ret-kabul, ters dönüşüm ve karma yöntemleriyle iki değişkenli keyfi bir dağılımdan çokgensel bir alan içerisinde rastgele sayılar üretilmiştir. İlk olarak iki değişkenli bir uzayda rastgele sayı üretilirken doğal olarak neden dikdörtgensel bir bölge kullanmak gerektiği ortaya konulmuştur. Ancak pratik uygulamalarda her zaman dikdörtgensel bir alan olamayacağı için iki boyutlu bir alanda çokgensel bir yaklaşım kullanılarak istenen alan tanımlanmaya çalışılmıştır. Öncelikle istenen bölgeyi tanımlamak için çokgensel bir yaklaşım kullanılarak çokgeni tanımlayan düğüm noktaları bulunmuştur. Bulunan bu noktalara göre çokgensel bölge Delaunay üçgenleme algoritması ile üçgenlere ayrılmıştır. Elde edilen tüm üçgenlerin dağılım fonksiyonundaki olasılık değerleri hesaplanarak bu değerlerin ağırlıklarına göre rastgele bir üçgen seçilmiştir. Seçilen üçgen içerisinde verilen olasılık yoğunluk fonksiyonuna göre rastgele sayı üretilmiştir.

İki değişkenli dağılım fonksiyonlarından geliştirilen algoritmalar yardımıyla çeşitli geometrik alanlarda rastgele sayı üretimi için iki ayrı örnek belirlenmiştir. Bu örneklerden bir tanesinde üçgensel bir bölgede rastgele sayı üretimi ele alınırken, diğerinde Afrika ana kıtası şekliyle sınırlandırılmış bir bölge ele alınmıştır. Keyfi olarak belirlenen olasılık yoğunluk fonksiyonundan 100 ve 1000 tane olmak üzere iki farklı hacimde rastgele sayı, ret-kabul, ters dönüşüm ve karma yöntemle üretilmiştir.

Anahtar Kelimeler: İki değişkenli dağılım fonksiyonu, Çokgen tabanlı olasılık yoğunluk fonksiyonu, Çokgensel alanlar, Rastgele sayı üretimi.

Master Thesis

SUMMARY

GENERATE RANDOM NUMBERS FROM ARBITRARY BIVARIATE DISTRIBUTION IN POLYGONAL AREA

Buğra Kaan TİRYAKİ

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Statistical and Computer Sciences Graduate Program
Supervisor: Assist. Prof. Dr. Orhan KESEMEN
2014, 62 Pages

In this study, random numbers from bivariate distribution function are generated in polygonal region with accept-reject, inverse transformation and hybrid methods. Naturally why rectangular region is needed to use while generating random numbers in two dimension is revealed firstly. However, a rectangular area is not always used in real problems (not simulation or theory problems) for this reason the desired area is aimed to define using a polygonal approach. Firstly, all nodal points of defining polygon are found by using polygonal approach for defining the desired region. The polygonal area is divided into triangles by Delaunay triangulation algorithm using these points. A random triangle is selected by weight based on the value obtained by calculating all the triangles' probability values from distribution function. Then, random number is generated according to the probability density function in selected triangles.

Two samples are determined for random number generation in various polygonal areas by using algorithms which are developed from bivariate distribution functions. One of these samples, random number generation is discussed in triangular region and other, generating random number is discussed in a region that bounded by shape of African mainland. Lastly, random numbers are generated in these regions from arbitrary probability density function using accept-reject, inverse transformation and hybrid methods in simulation study.

Key Words: Bivariate distribution functions, Polygonal based probability density function, Polygonal areas, Random number generation.

ŞEKİLLER DİZİNİ

Sayfa No

TABLOLAR DİZİNİ

<u>Sayfa No</u>

SEMBOLLER DİZİNİ

PRN	: Pseudo Random Number
QRN	: Quasi Random Number
PRNG	: Pseudo Random Number Generation
QRNG	: Quasi Random Number Generation
p_{max}	: Olasılık kütle fonksiyonunun maksimum değeri
f_{max}	: Olasılık yoğunluk fonksiyonunun maksimum değeri
$U(a,b)$	: $[a,b]$ aralığında düzgün dağılım
$\overline{AB}$	: A ve B noktalarından geçen doğru denklemi
$\overline{AC}$	: A ve C noktalarından geçen doğru denklemi
$\overline{BC}$	: B ve C noktalarından geçen doğru denklemi
m_*	: (*) doğrusunun eğimi
c_*	: (*) doğrusunun sabiti
f_i	: i. gözlenen frekans
e_i	: i. beklenen frekans
c	: Sınıf sayısı

1. GENEL BİLGİLER

Rastgele sayılar ve simülasyon çok uzun süredir bilim dünyasında popüler bir konu olmuştur. Özellikle matematik ve istatistik alanlarında, üzerinde çok çalışılmış ve hala çalışılan konulardandır. Rastgele sayıların biyoloji, finans, sigorta, fizik, matematik, kriptografi, simülasyon, bilgisayar oyunları, istatistiksel örnekleme ve uygulamaları, istatistiksel hesaplama gibi çok geniş uygulama alanları vardır. Bu yüzden rastgele sayılar birçok alanda önemli bir yer tutmaktadır [1].

Eski zamanlarda, zar atma, çarkıfelek, rulet ve kart oyunları gibi yöntemler kullanılarak elle veya mekanik olarak rastgele sayılar üretilmiş ve sadece mekanik (ya da elektronik) aygıtların "gerçekten" rastgele sayılar elde edeceği düşünülmüştür. Bu eski geleneksel yöntemler, genel kullanım için çok yavaş bir şekilde çalışmaktadır, ayrıca dijital olmayan yöntemlerle oluşturulan diziler çoğaltılamamaktadır [2]. Bilgisayarın icadıyla birlikte rastgele sayı üretebilmek mümkün hale gelmiştir. Bir bilgisayarda dijital olarak rastgele sayılar üretmek, tablo hazırlanması ve bilgisayarın belleğinde depolanması aşamalarından oluşur [3]. 1955 yılında RAND şirketi [4] tarafından, bir milyon rastgele sayı içeren bir rastgele sayılar tablosu oluşturulmuştur. Bu yöntemin avantajı tekrar üretilebilir olması, dezavantajı ise hız eksikliği ve tablonun yorucu riski olmasıdır. Bu tür dezavatajlar yüzünden, bilgisayarlarda rastgeleliği elde etmenin tek yolu deterministik yöntemler kullanmaktır. Gerçek rastgelelik hariç, iki tür rastgele sayı üreteci vardır. Bunlar sözde ve yarı rastgele sayı üreteçleridir (PRNG-QRNG) [2,5–7].

1.1. Sözde Rastgele Sayılar

Doksanlı yılların başlarında sözde rastgele sayı üreteciyle ilgili genel kabul edilebilir bir algoritma bulunmamaktaydı. 1998 yılında, Japon matematikçiler Matsumoto ve Nishimura $2^{19937}-1$ periyotlu ilk sözde rastgele sayı üretme algoritması olan Mersenne-Twister algoritmasını geliştirmişlerdir [1].

Günümüzde ise modern yöntem bir bilgisayar yardımıyla başarılı sözde rastgele sayılar üretmektir. Bu sözde rastgele sayılar deterministik olarak üretilmesine rağmen hepsi $[0,1)$ aralığına indirgenen düzgün dağılıma sahip dizilerdir.

Sözde rastgele sayı üretmekte yaygın olarak kullanılan yöntem x_0 başlangıç değerini kullanarak başlar. Bu değer aynı zamanda tohum (çekirdek) olarak adlandırılır. Ve yinelemeli olarak x_n ‘e kadar n tane rastgele sayı bu tohum değerinden üretilir.

$$x_n = (ax_{n-1})\ mod(m) \tag{1}$$

$n \geq 1$ olmak üzere, a ve m pozitif tam sayılardır. ax_{n-1} değeri m’ye bölünür ve kalan x_n değeri olarak alınır. Böylece her bir üretilen x_n sözde rastgele sayı değerleri $[0, m)$ aralığında düzgün dağılıma sahip rastgele sayılar olur. Bu rastgele sayılar m sayısına bölünerek $[0,1)$ aralığında ondalıklı bir sayı biçimine dönüştürülür.

Eşitlik 1‘deki denklem çarpımsal eşleşik yöntem olarak adlandırılır. Sonlu sayıdaki m değeri için üretilen rastgele sayılar dizisi kendisini tekrar etmek zorunda kalacaktır ve tüm dizi tekrar etmeye başlayacaktır. Böylece farklı başlangıç değeri, a ve m katsayıları için dizinin kendisini tekrar etme sıklığının boyutu değişecektir. Bu nedenden dolayı, a ve m sabitlerinin seçimi için bazı kriterler belirlenmiştir.

Genellikle a ve m sabitleri üç kritere göre seçilmektedir:

1. Herhangi bir başlangıç tohumu için, üretilen rastgele sayılar dizisi, $[0, m)$ aralığında düzgün dağılmış bir rastgele sayı dizisi olmak zorundadır.
2. Herhangi bir başlangıç tohumu için, üretilen rastgele sayılar dizisinin kendisini tekrar etme sıklığı oldukça az olmalıdır. Çok büyük aralıklarla dizi başlangıç değerine dönmelidir.
3. Tüm değerler bilgisayar ortamında doğru bir şekilde üretilebilmelidir [8].

Araştırmalar, yukarıdaki üç şartın sağlanabilir olmasının, m'nin bilgisayar kelime işlemcisine uygun büyük bir asal sayı seçilmesine bağlı olduğunu göstermiştir. 32 bitlik kelime işlemcisinde ilk bit işaret biti olmak üzere, $m = 2^{31} - 1$ ve $a = 7^5$ katsayıları seçildiğinde en ideal sonuçlar alındığı gösterilmiştir. 36 bitlik kelime işlemcisinde ise, $m = 2^{35} - 31\ \ ve\ a = 5^5$ katsayı değerleri için sözde rastgele sayı üreteci en ideal şekilde çalışmaktadır [8].

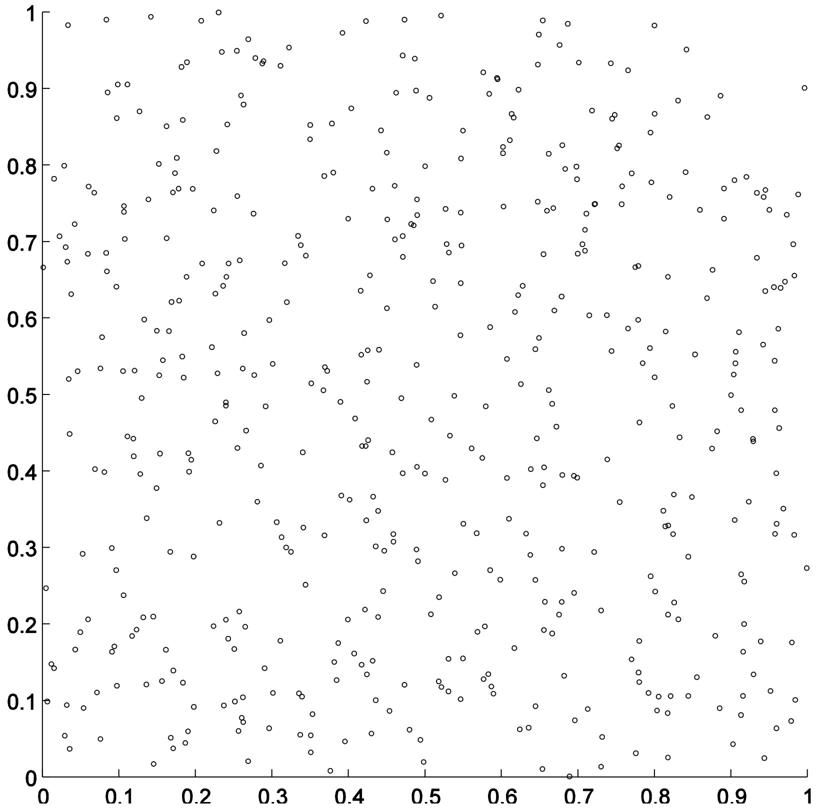
Bir diğer sözde rastgele sayı üreteci Eşitlik (2)’de gösterilmiştir.

$$x_n = (ax_{n-1} + c\ )\ mod(m) \tag{2}$$

Bu tür üreteçler karışık eşlenik yöntem (mixed congruential method) olarak adlandırılır. Çarpım ve toplam kısımlarından oluşmaktadır. Bu tür bir üreteç kullanıldığında, m katsayısı bilgisayar kelime işlemcisine eşitlendiğinde, oldukça etkili sonuçlar alınmaktadır.

Bilgisayar sistemleri ve simülasyonlarının başlangıç noktası olan sözde rastgele sayılar yukarıdaki gibi üreteçlerle üretilmektedir. Bununla beraber, bağımsız düzgün rastgele değişkenlerin oluşturmuş olduğu dizinin değerlerine yaklaşım olarak sözde (yalancı) rastgele sayı dizisi oluşturulmaktadır [8].

Şekil 1'de iki değişkenli sözde rastgele sayı üreteciyle üretilmiş rastgele sayıların grafiği çizilmiştir.



Şekil 1. Sözde rastgele sayı üreteciyle üretilmiş rastgele sayılar

1.2. Yarı Rastgele Sayılar

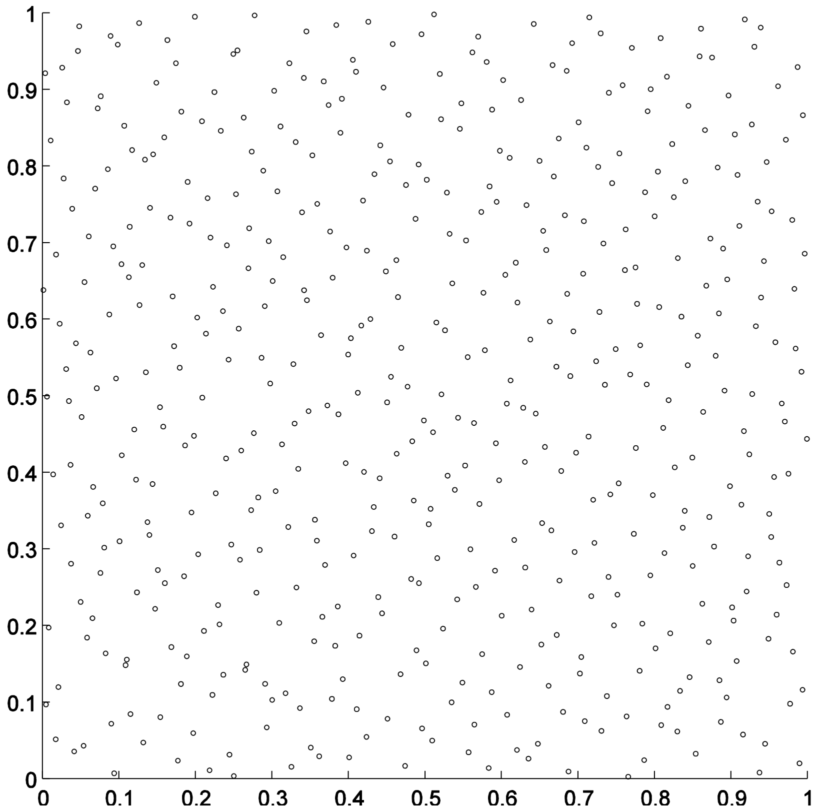
Yarı rastgele sayı üretecinin nasıl çalıştığını anlatabilmek için ilk olarak Monte-Carlo yönteminin tanımlanması gerekir. d boyutlu bir uzayda f integrallenebilir bir fonksiyon olmak üzere,

$I^d = [0,1]^d$ tanımlandığında, f fonksiyonunun I^d üzerinde integrali Monte Carlo yöntemi olarak tanımlanır ve eşitlik (3)'deki gibi gösterilir.

$$\int_{I^d} f(x)dx \approx \frac{1}{n}\sum_{i=1}^{n} f(X_i) \tag{3}$$

burada X_i'ler I^d de tanımlı bağımsız rastgele sayılardır.

Yarı-Monte Carlo yöntemi ile Monte Carlo yöntemi arasındaki fark X_i'lerin nasıl seçildiğidir. Yarı-Monte Carlo yöntemi, Halton, Sobel, Faure gibi düşük uyuşmazlık dizileri kullanırken, bilinen Monte Carlo yöntemi sözde rastgele dizisi kullanmaktadır. Ayrıca, düşük uyuşmazlık dizilerinin kullanmanın avantajı yakınsamanın daha hızlı olmasıdır [9]. Şekil 2'de Halton dizisi ile üretilmiş yarı rastgele sayıların grafiği çizilmiştir.



Şekil 2. Halton dizisi ile üretilmiş yarı rastgele sayılar

Şekil 2'de iki değişkenli düzgün dağılımdan rastgele sayıların dağılımı gösterilmektedir. Bu dağılım dikkatlice incelendiğinde, Şekil 1'deki sözde rastgele sayıların dağılımına göre daha homojen bir dağılım olduğu görülmektedir. Şekil 2'deki rastgele üretilen sayılar, Şekil 1'deki gibi birbirine yakın değildir. Diğer yandan sözde rastgele sayı üreteci ile yarı rastgele sayı üreteci arasındaki ana fark sözde rastgele sayı üretecinin yarı rastgele sayı üretecinin aksine rastgele sayıları kullanmayıp deterministik olarak hesap yapmasıdır [1].

Büyük sayılar kanunu gereğince, n'nin sonsuza doğru gitmesi, Eşitlik 3'deki yaklaşım için neredeyse gerçeğe yakın sonuçlar sağlamaktadır, ayrıca beklenen integral hatası $\frac{1}{\sqrt{n}}$ ile sınırlıdır, d boyutuna bağlı değildir. Bu durumdan dolayı Monte Carlo yöntemlerinin geniş bir uygulama alanı vardır [9].

1.3. Tek Değişkenli Dağılımlar

Bir rastgele değişkenin aldığı değerler sonlu ya da sayılabilir sonsuzlukta ise kesikli rastgele değişken, bir aralıkta ya da birden çok aralıkta her değeri alabiliyorsa sürekli rastgele değişken denir [10].

Olasılık fonksiyonu, bir rastgele değişkenin ister kesikli ister sürekli olsun, alabileceği değerlerle, bu değerleri alma olasılıkları arasındaki ilişkiyi gösteren bir bağıntıdır. X rastgele değişkeninin x değerini alması olasılığı, $p(x) = P(X = x)$ şeklinde ifade edilmektedir.

Kesikli bir rastgele değişkenin verilen bir aralıktaki değerleri alması olasılığı, belirtilen aralıktaki bütün değerlere ait olasılıklarının toplamıyla gösterilir [11].

$$P(a \leq X \leq b) = \sum_{x=a}^{b} p(x) \tag{4}$$

X kesikli rastgele değişkeninin olasılık fonksiyonunun $p(x)$ olabilmesi için,

$$p(x) \geq 0, \quad \forall x \tag{5}$$

ve

$$\sum_{x} p(x) = 1 \tag{6}$$

koşullarını sağlamalıdır [10].

Sürekli bir rastgele değişkenin verilen bir aralıktaki değerlerini alması olasılığı, belirtilen aralıktaki olasılık fonksiyonunun integrali alınarak hesaplanmaktadır. Diğer bir deyişle, integral almakla yapılan iş, belirtilen aralıkta alan hesaplamakla aynı işleme denk gelmektedir [11].

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{7}$$

Ayrıca, sürekli bir değişkenin, herhangi bir değeri alması olasılığı sıfır olduğundan olasılık ifadesi Eşitlik 8'deki gibi iki şekilde de tanımlanabilir.

$$P(a \leq X \leq b) = P(a < X < b) \tag{8}$$

X sürekli rastgele değişkeninin olasılık fonksiyonunun $f(x)$ olabilmesi için,

$$f(x) \geq 0, -\infty < x < \infty \tag{9}$$

ve

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{10}$$

koşulları sağlanmalıdır [11].

1.3.1. Birikimli Dağılım Fonksiyonu

Dağılım fonksiyonu bir rastgele değişkenin herhangi bir değerden daha küçük değer alması olasılığını hesaplayan bir fonksiyondur ve $F(x)$ ile gösterilir. $F(x)$ dağılım fonksiyonu,

$$F(x) = P(X \leq x) \tag{11}$$

biçiminde ifade edilir.

X rastgele değişkeni kesikli ise,

$$F(x) = \sum_{x} p(x) \tag{12}$$

biçiminde gösterilir.

X rastgele değişkeni sürekli ise,

$$F(x) = \int_{-\infty}^{x} f(x)dx \tag{13}$$

biçiminde gösterilir.

X rastgele değişkeninin $[a, b]$ aralığında olasılık değeri Eşitlik 14'deki gibi hesaplanmaktadır [10].

$$P(a \leq X \leq b) = F(b) - F(a) \tag{14}$$

1.4. İki Değişkenli Dağılımlar

İki değişkenli dağılımlar iki boyutlu uzayda tanımlı fonksiyonlardır ve matematiksel olarak, ($f(x, y)$: $(x, y) \in \mathbb{R}$) biçiminde gösterilmektedir.

1.4.1. Ortak Olasılık Kütle Fonksiyonu

X ve Y kesikli rastgele değişkenler olmak üzere, X ve Y'nin ortak olasılık kütle fonksiyonu,

$$p(x, y) = P(X = x, Y = y) \; ya \; da \; P(X = x \Leftarrow Y = y) \tag{15}$$

biçiminde gösterilmekte ve

$$\sum_{x}\sum_{y} p(x,y) = 1 \tag{16}$$

koşulunu sağlamaktadır [12].

1.4.2. Marjinal Olasılık Kütle Fonksiyonu

X ve Y kesikli rastgele değişkenler olmak üzere, X'in olasılık kütle fonksiyonu $p_X(x) = P(X = x)$, Y'nin olasılık kütle fonksiyonu $p_Y(y) = P(Y = y)$ ile gösterilmektedir.

Her bir olasılık kütle fonksiyonunun belirtilen alanlarda toplam değeri 1'dir.

$$\sum_{x} p_X(x) = 1, \quad \sum_{y} p_Y(y) = 1 \tag{17}$$

X ve Y'nin ortak olasılık kütle fonksiyonundan, X ve Y'nin marjinal olasılık fonksiyonları Eşitlik 19-20'de verilen formüllerle hesaplanmaktadır [12].

Ortak olasılık kütle fonksiyonu,

$$p(x,y) = P(X = x, Y = y) \tag{18}$$

X'in marjinal olasılık kütle fonksiyonu,

$$p_X(x) = \sum_{y} p(x,y) \tag{19}$$

Y'in marjinal olasılık kütle fonksiyonu,

$$p_Y(y) = \sum_{x} p(x,y) \tag{20}$$

biçiminde gösterilmektedir.

1.4.3. Koşullu Olasılık Kütle Fonksiyonu

Y'in y değerini alma olasılığı verildiğinde, X'nin x değerini alma koşullu olasılığı;

$$p(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x,y)}{p_Y(y)} \tag{21}$$

X'in x değerini alma olasılığı verildiğinde, Y'nin y değerini alma koşullu olasılığı;

$$p(y|x) = P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{p(x,y)}{p_X(x)} \tag{22}$$

biçiminde hesaplanmaktadır [12].

1.4.4. Ortak Dağılım Fonksiyonu

X ve Y'nin ortak dağılım fonksiyonu Eşitlik 23'deki gibi,

$$F_{XY}(x,y) = \sum_{i=1}^{x} \sum_{j=1}^{y} p(x_i, y_j) \tag{23}$$

hesaplanmaktadır. Hesaplanan bu dağılım fonksiyonu,

$$0 \leq F_{XY}(x,y) \leq 1 \tag{24}$$

koşulunu sağlamalıdır.

1.4.5. X ve Y Değişkenlerinin Bağımsızlık Durumu

X ve Y rastgele değişkenleri birbirinden bağımsız olabilmesi için,

$$p(x, y) = p_X(x)\, p_Y(y) \tag{25}$$

ve

$$F_{XY}(x, y) = F_X(x)\, F_Y(y) \tag{26}$$

eşitliklerden herhangi birini sağlamalıdırlar [12].

1.4.6. Ortak Olasılık Yoğunluk Fonksiyonu

X ve Y sürekli rastgele değişkenler olmak üzere, X ve Y'nin ortak olasılık yoğunluk fonksiyonu üç boyutlu uzayda bir hacim oluşturmaktadır. Yüzey fonksiyonunun altında kalan hacim 1'e eşit olmalıdır [12]. Bu ifadeye göre,

$$f(x, y) \geq 0 \qquad \forall x, y \tag{27}$$

ve

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1 \tag{28}$$

koşulunu sağlamalıdır.

1.4.7. Marjinal Olasılık Yoğunluk Fonksiyonu

X ve Y sürekli rastgele değişkenler olmak üzere, X'in marjinal olasılık yoğunluk fonksiyonu $f_X(x)$, Y'nin marjinal olasılık yoğunluk fonksiyonu $f_Y(y)$ ile gösterilmektedir.

X ve Y'nin ortak olasılık yoğunluk fonksiyonundan, X ve Y'nin marjinal olasılık yoğunluk fonksiyonları 29-30'da verilen eşitliklerdeki formüllerle hesaplanmaktadır [12].

X'in marjinal olasılık yoğunluk fonksiyonu,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \tag{29}$$

ve Y'nin marjinal olasılık yoğunluk fonksiyonu,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \tag{30}$$

biçiminde verilmektedir.

X ve Y'nin olasılık yoğunluk fonksiyonlarının $(-\infty,\infty)$ aralığında integral değeri,

$$\int_{-\infty}^{+\infty} f_X(x) = 1\,, \qquad \int_{-\infty}^{+\infty} f_Y(y) = 1 \tag{31}$$

eşitlikte verildiği gibi 1'e eşit olmalıdır.

1.4.8. Koşullu Olasılık Yoğunluk Fonksiyonu

Y'in y değeri verildiğinde, X'nin koşullu olasılığı;

$$f(x|y) = f(x|Y=y) = \frac{f(x,y)}{f_Y(y)} \tag{32}$$

X'in x değeri verildiğinde, Y'nin koşullu olasılığı;

$$f(y|x) = f(y|X=x) = \frac{f(x,y)}{f_X(x)} \tag{33}$$

şeklinde hesaplanmaktadır [12].

1.4.9. Ortak Dağılım Fonksiyonu

X ve Y'nin ortak dağılım fonksiyonu 34'deki eşitlikte,

$$\mathrm{F_{XY}(x, y)} = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dy dx \tag{34}$$

biçiminde verilmiştir [12]. Bu eşitlik,

$$0 \leq F_{XY}(x, y) \leq 1 \tag{35}$$

koşulunu sağlamalıdır.

1.4.10. X ve Y Değişkenlerinin Bağımsızlık Durumu

X ve Y rastgele değişkenlerinin birbirinden bağımsız olması,

$$f_{XY}(x, y) = f_X(x)\ f_Y(y) \tag{36}$$

ve

$$F_{XY}(x, y) = F_X(x)\ F_Y(y) \tag{37}$$

koşullarına bağlıdır [12].

1.4.11. Dağılım Fonksiyonu ile Olasılık Yoğunluk Fonksiyonu Arasındaki İlişki

Dağılım fonksiyonunun x ve y değişkenlerine göre kısmi türevleri,

$$f(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \tag{38}$$

biçiminde verilen ortak olasılık yoğunluk fonksiyonuna eşittir [12].

Marjinal olasılık yoğunluk fonksiyonları,

$$f_X(x) = \frac{\partial F_X(x)}{\partial x} = F'_X(x) \tag{39}$$

ve

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = F'_Y(y) \tag{40}$$

biçiminde verilen marjinal dağılım fonksiyonun türevi ile elde edilmektedir.

1.5. Olasılık Kütle/Yoğunluk Fonksiyonundan Rastgele Değişken Üretilmesi

Tüm rastgele örnekleme yöntemleri, temelde standart düzgün dağılımlı bir rastgele değişkenin değerlerini, olasılık dağılımı verilen bir rastgele değişkenin değerlerine dönüştüren ve üreteç olarak adlandırılan birer algoritmadır. Rastgele örnekleme yapmanın ters dönüşüm ve ret-kabul yöntemi olmak üzere, genelde iki temel yaklaşımı vardır [13].

1.6. Ret-Kabul Yöntemiyle Rastgele Sayı Üretilmesi

Ret-kabul yöntemi, 1951 yılında von Neumann tarafından geliştirilmiştir. Yöntemde belirlenen dağılımdan rastgele örnek alınarak, seçilen örneklemin istenilen dağılım için uygun olup olmadığına karar verilir [14]. Ret-kabul yöntemi herhangi bir dağılımdan gözlemler oluşturmak için kullanılan temel bir yöntemdir. Bu yöntem n-boyutlu uzayda herhangi bir dağılım için kullanılabilir. Ret-kabul yöntemi, rasgele bir örnek değişkeninin yoğunluk/kütle fonksiyonunun altında kalan bölgede düzgün dağıldığı varsayımına dayanmaktadır [13,15,16].

1.6.1. Kesikli Dağılımdan Rastgele Sayı Üretilmesi

X rastgele değişkeninin olasılık kütle fonksiyonu p_j,

$$p_j = P\{X = j\} \tag{41}$$

biçiminde gösterilmektedir.

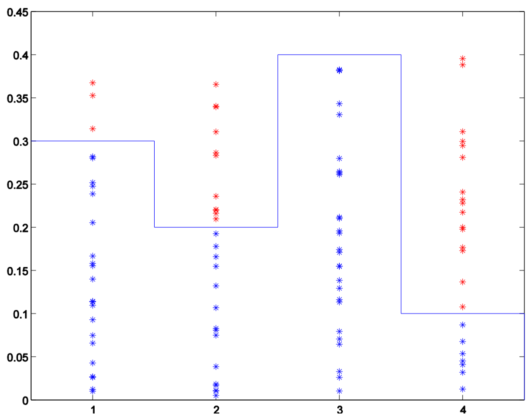
Algoritma 1 yardımıyla p_j olasılık kütle fonksiyonuna sahip X rastgele değişkeni üretilmektedir [8].

Algoritma 1. Tek değişkenli olasılık kütle fonksiyonundan rastgele değişken üretilmesi	
Adım 1:	p_j olasılık kütle fonksiyonun tanım aralığından eşit olasılıkla rastgele bir sayı üret (ξ).
Adım 2:	Düzgün dağılımdan bir rastgele sayı üret ($\eta \sim U(0, p_{max})$).
Adım 3:	Eğer $\eta < P(X = \xi)$ ise ξ kabul; aksi halde, ξ reddet ve Adım 1'e git.

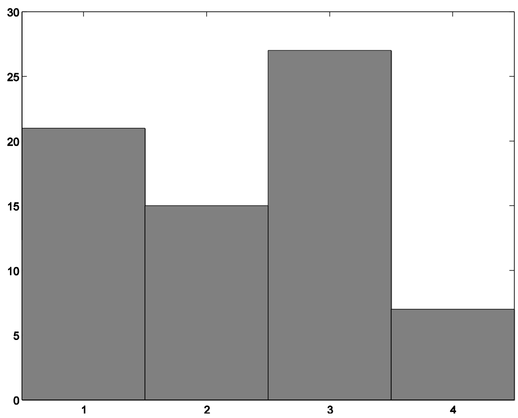
Algoritma 1'in, Tablo 1'deki olasılık kütle fonksiyonuna sahip bir rastgele değerler için 100 kez çalıştırılması durumunda, ret-kabul bölgeleri ve kabul bölgesine düşen rastgele sayıların histogram grafiği Şekil 3'de gösterilmiştir.

Tablo 1. Olasılık kütle fonksiyonu

$x_1 = 1$	$x_2 = 2$	$x_3 = 3$	$x_4 = 4$
0.3	0.2	0.4	0.1



(a)



(b)

Şekil 3. (a) 100 rastgele örneğin ret-kabul bölgelerine düşmesi (b) Kabul bölgesine düşen rastgele değişkenlerin histogram grafiği

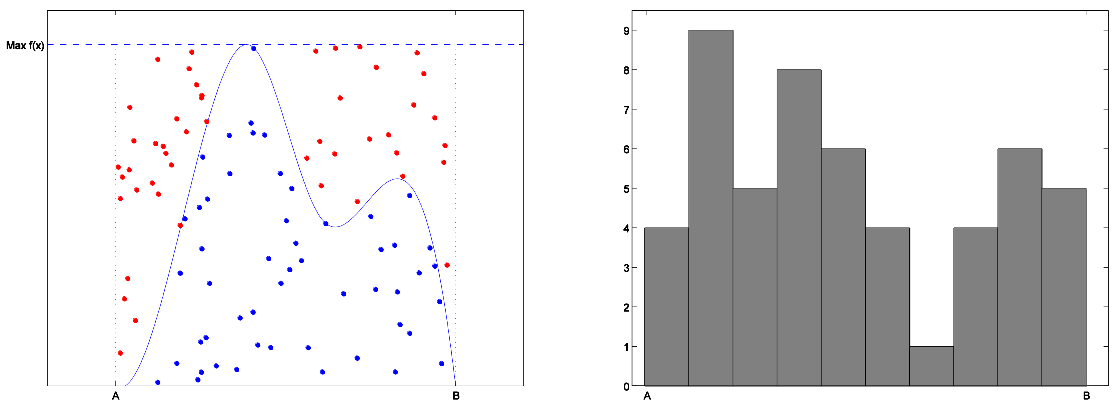
Şekil 3(a)'da, üretilen 100 tane rastgele örneğin, 70 tanesi mavi bölge ile gösterilen kabul bölgesine, geri kalan 30 tane rastgele sayı ise kırmızı bölge ile gösterilen ret bölgesinde yer almaktadır. Kabul edilen 70 rastgele örneğin histogram grafiği ise Şekil 3(b)'de gösterilmiştir.

1.6.2. Sürekli Dağılımdan Rastgele Sayı Üretilmesi

X rastgele değişkeninin olasılık yoğunluk fonksiyonu $f(x)$ ile ifade edilmek üzere, Algoritma 2 yardımıyla $f(x)$ olasılık yoğunluk fonksiyonuna sahip X rastgele değişkeni üretilmektedir [8].

Algoritma 2. Tek değişkenli olasılık yoğunluk fonksiyonundan rastgele sayı üretilmesi	
Adım 1:	$f(x)$ olasılık yoğunluk fonksiyonunu belirle.
Adım 2:	Fonksiyonun tanım aralığından eşit olasılıkla bir rastgele sayı üret (ξ).
Adım 3:	Düzgün dağılımdan bir rastgele sayı üret ($\eta \sim U(0, f_{max})$).
Adım 4:	Eğer $\eta < f(\xi)$ ise ξ kabul; aksi halde, ξ reddet ve Adım 2'e git.

Algoritma 2'nin 100 kez çalıştırılması sonucunda, iki modlu bir olasılık yoğunluk fonksiyonundan rastgele üretilmiş 100 örneklem için ret-kabul bölgeleri ve kabul bölgesine düşen rastgele sayıların histogram grafiği Şekil 4'de gösterilmiştir.



(a) (b)

Şekil 4. (a) 100 rastgele örneğin ret-kabul bölgelerine düşmesi (b) Kabul bölgesine düşen rastgele sayıların histogram grafiği

Şekil 4(a)'da, üretilen 100 tane rastgele örneğin, 52 tanesi mavi bölge ile gösterilen kabul bölgesine, geri kalan 48 tane rastgele sayı ise kırmızı bölge ile gösterilen ret bölgesinde yer almaktadır. Kabul edilen 52 rastgele örneğin histogram grafiği ise Şekil 4(b)'de gösterilmiştir.

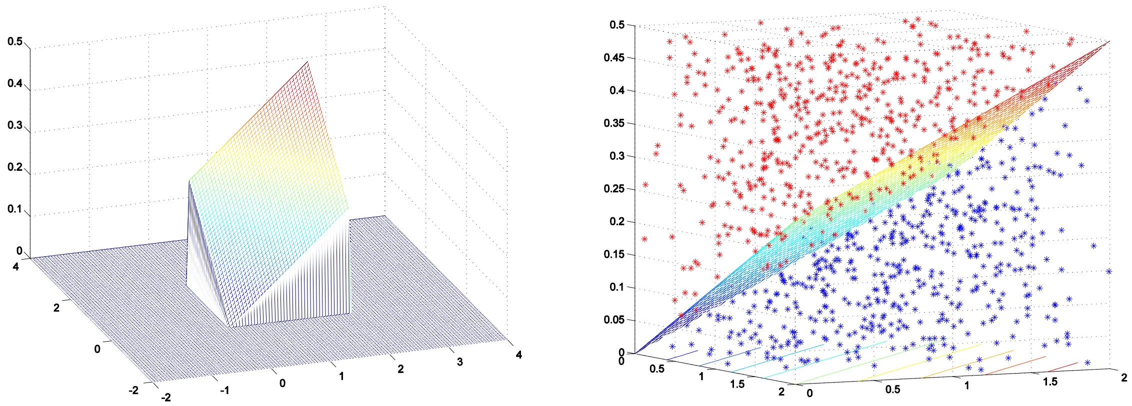
1.6.3. İki Değişkenli Sürekli Rastgele Sayı Üretilmesi

X, Y rastgele değişkeninin olasılık yoğunluk fonksiyonu $f(x, y)$ ile ifade edilmek üzere, Algoritma 3 yardımıyla $f(x, y)$ ortak olasılık yoğunluk fonksiyonuna sahip (X, Y) rastgele değişkeni üretilmektedir [8].

Algoritma 3. İki değişkenli olasılık yoğunluk fonksiyonundan rastgele sayı üretilmesi	
Adım 1:	Olasılık yoğunluk fonksiyonunu belirle.
Adım 2:	Hem X'in tanımlandığı sınırlar içerisinden ve hem de Y'nin tanımlandığı sınırlar içerisinden düzgün dağılımdan iki rastgele sayı üret (r_1, r_2).
Adım 3:	Düzgün dağılımdan bir rastgele sayı üret $(\eta \sim U(0, f_{max}))$.
Adım 4:	Eğer $\eta < f(r_1, r_2)$ ise r_1 ve r_2 kabul et; aksi halde, r_1 ve r_2'yi reddet ve Adım 2'e git.

Eşitlik 42'deki gibi verilmiş bir ortak olasılık yoğunluk fonksiyonundan Algoritma 3 kullanılarak, 2000 tane rastgele sayı üretilmiştir. Üretilen 2000 tane rastgele sayıdan 1017 tanesi mavi ile gösterilen kabul bölgesine geri kalan rastgele sayılar ise kırmızı ile gösterilen ret bölgesine düşmüştür (Şekil 5(b)). Kabul bölgesi düşen 1017 tane rastgele sayının histogram grafiği ve beklenen frekansların histogram grafiği Şekil 6'da gösterilmiştir.

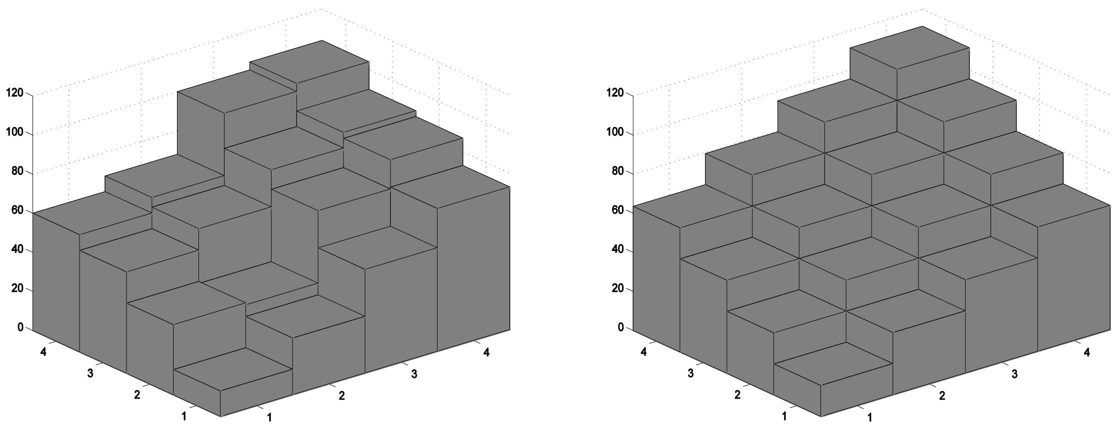
$$f(x, y) = \frac{x + y}{8}, \qquad 0 < x, y < 2 \tag{42}$$



(a) (b)

Şekil 5. (a) Ortak olasılık yoğunluk fonksiyonun gösterimi (b) Üretilen X, Y rastgele sayılarının ret ve kabul bölgelerine düşmesi

Ortak olasılık yoğunluk fonksiyonu tanımlandığı aralıkların dışında sıfır değerini almaktadır (Şekil 5(a)).



(a) (b)

Şekil 6. (a) Kabul edilen rastgele sayıların histogram grafiği (b) Beklenen frekansların sütun grafiği

H_0: Üretilen rastgele sayılar Eşitlik 42'de verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Eşitlik 42'de verilen ortak olasılık yoğunluk fonksiyonundan gelmemiştir.

Gözlenen ve beklenen frekanslar üzerinden ki-kare testi yapıldığında, $\chi_h^2 = 11.8588 < \chi_{c=15,\alpha=0.05}^2 = 24.9958$ olduğundan %95 güven düzeyinde H_0 hipotezi reddedilemez. Üretilen rastgele sayılar Eşitlik 42'de verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir ($c = SınıfSayısı - 1$).

1.7. Ters Dönüşüm Yöntemi ile Rastgele Sayı Üretilmesi

Ters dönüşüm yöntemi, rastgele örneklemeye tam ve doğrudan kuramsal bir yaklaşım olduğundan, en sık başvurulan yöntemdir. Bu yöntemin doğrudan uygulanması, ancak dağılım fonksiyonunun düzgün ve sürekli bir fonksiyon olması durumunda söz konusudur. Bir dizgenin girdi ve çıktılarının hepsi bu özellikte olmayabilir. Sanal deney modelindeki kimi değişkenlerin dağılımı deneysel, kesikli ve kesin sürekli artmayan nitelikte olabilir. Bu tür dağılımların ters dağılım işlevleri tanımlanabildiği ve bunlardan etkin olarak sanal gözlem üretilebildiği durumlarda da, ters dönüşüm yönteminin belli biçimleri uyarlanabilir [17,18].

Tersi alınabilen bir fonksiyon verildiğinde istenen bir bölgede etkili bir biçimde sözde rastgele sayılar oluşturulabilmektedir. Ancak tersi alınamayan dağılım fonksiyonları için ise sayısal yöntemler kullanılmaktadır.

1.7.1. Kesikli Dağılımdan Rastgele Sayı Üretilmesi

X rastgele değişkeninin olasılık kütle fonksiyonu

$$P\{X = x_j\} = p_j, \qquad j = 0, 1, \dots \tag{43}$$

eşitlik 43'deki gibi tanımlandığında, U standart düzgün dağılımdan bir rastgele sayı olmak üzere,

$$x = \begin{cases} x_0, & U < p_0 \\ x_1, & p_0 \le U < p_0 + p_1 \\ \dots, & \dots \\ x_j, & \sum_{i=0}^{j-1} p_i \le U < \sum_{i=0}^{j} p_i \\ \dots, & \dots \end{cases} \tag{44}$$

44 eşitliği sağlandığında, X rastgele değişkeninin olasılık kütle fonksiyonundan rastgele sayı üretilmektedir [8].

Algoritma 4. Tek değişkenli olasılık kütle fonksiyonundan rastgele değişken üretilmesi

Adım 1:	Olasılık kütle fonksiyonunu tanımla (p).
Adım 2:	Birikimli olasılık değerlerini bul (P_i).
Adım 3:	Standart düzgün dağılımdan bir rastgele sayı üret (U).
Adım 4:	Eğer $U < P_1$ ise $X = x_0$ ve döngüden çık.
	Eğer $U < P_2$ ise $X = x_1$ ve döngüden çık.
	Eğer $U < P_3$ ise $X = x_2$ ve döngüden çık.
	…
Adım 5:	Üretilmek istenilen rastgele sayı miktarına gelindiğinde dur; aksi halde, Adım 3'e git.

Tablo 2'deki gibi tanımlanan bir olasılık kütle fonksiyonundan 1000 tane rastgele sayı Algoritma 4 kullanılarak üretilmiştir.

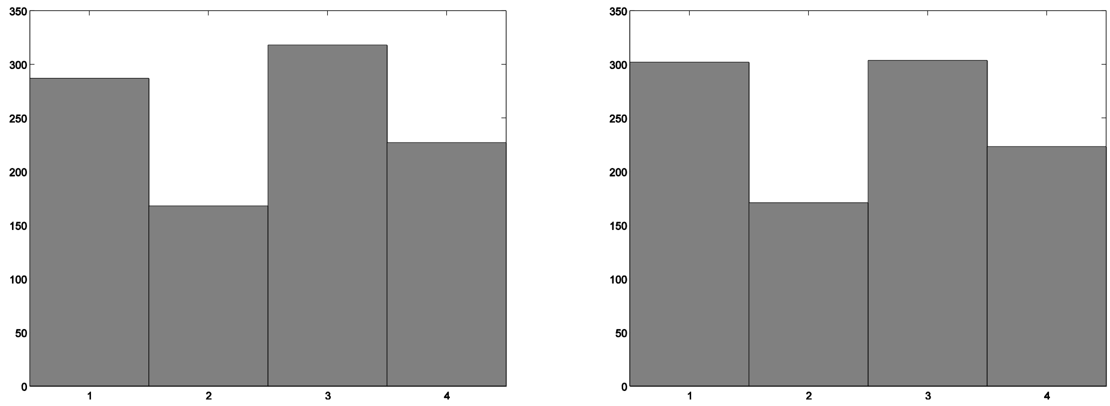
Tablo 2. Olasılık kütle fonksiyonu

$x_1 = 1$	$x_2 = 2$	$x_3 = 3$	$x_4 = 4$
0.302	0.1709	0.3038	0.2233

H_0: Üretilen rastgele sayılar Tablo 2'de verilen olasılık kütle fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Tablo 2'de verilen olasılık kütle fonksiyonundan gelmemiştir.

Gözlenen ve beklenen frekanslar üzerinden ki-kare testi yapıldığında, $\chi_h^2 = 1.5245 < \chi_{c=3,\alpha=0.05}^2 = 7.8147$ olduğundan %95 güven düzeyinde H_0 hipotezi reddedilemez. Üretilen rastgele sayılar Tablo 2'de verilen olasılık kütle fonksiyonundan gelmiştir $(c = SınıfSayısı - 1)$. Üretilen rastgele sayıların histogram grafiği Şekil 7(a)'da gösterilmiştir.



(a) (b)

Şekil 7. (a) Üretilen 1000 rastgele sayının histogram grafiği (b) Beklenen frekansların sütun grafiği

Diğer yandan x_j değerleri, $x_0 < x_1 < x_2 < \cdots < x_k$ şeklinde sıralandığında X'in dağılım fonksiyonu, $F(x_k) = \sum_{i=0}^{k} p_i$ olmaktadır $(x_i,\ i \geq 0)$. Bu durumda $F(x_{j-1}) \leq U \leq F(x_j)$ olduğunda $X = x_j$ olmaktadır. Diğer bir deyişle, U rastgele sayısı üretildikten sonra dağılım fonksiyonunda $F(x_{j-1}) - F(x_j)$ aralığı hesaplanarak X'in değeri bulunabilmektedir. Bu ifadeye eşdeğer olarak da $F(U)$'nun tersi de aynı işlemi yapmaktadır. Bu yüzden bu yöntem ayrık ters dönüşüm yöntemi olarak adlandırılmaktadır.

Ayrıca, bu yöntemle bir kesikli rastgele sayı oluşturmak için harcanan zaman, birikimli dağılım fonksiyonunda hesaplanan aralıkların sayısıyla doğru orantılıdır. Bu durumda, p_j'lerin azalan sırada kullanılması x_j'lerin olasılık değerlerinin bulunmasında daha kullanışlı olmaktadır [8].

1.7.2. İki Değişkenli Kesikli Rastgele Sayı Üretilmesi

X, Y rastgele değişkeninin olasılık kütle fonksiyonunu $p_{XY}(x, y)$ ile ifade edilirse bu fonksiyon

$$p_{XY}(x, y) = P\{X = x, Y = y\} \tag{45}$$

biçiminde gösterilmektedir. Algoritma 5 yardımıyla p_{XY} olasılık kütle fonksiyonuna sahip X, Y rastgele değişkeni üretilmektedir [19].

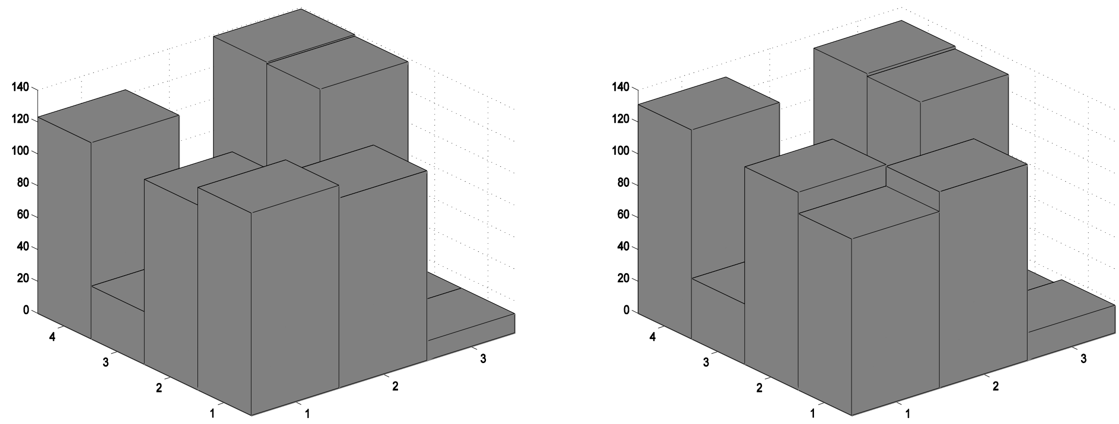
Algoritma 5. İki değişkenli olasılık kütle fonksiyonundan rastgele sayı üretilmesi.

Adım 1:	Ortak olasılık kütle fonksiyonunu belirle (p_{XY}).
Adım 2:	Ortak olasılık kütle fonksiyonundan X'in marjinal olasılık değerlerini bul (p_X).
Adım 3:	X'in marjinal olasılıklarından birikimli marjinal olasılıklarını hesapla (P_X).
Adım 4:	Y'nin koşullu olasılık değerlerini hesapla $h_Y(y\|X)$.
Adım 5:	Y'nin birikimli koşullu olasılık değerlerini hesapla $H_Y(y\|X)$.
Adım 6:	Düzgün dağılımdan iki tane rastgele sayı üret (r_X, r_Y).
Adım 7:	Eğer $r_X < P_X(1)$ ise $X = x_1$ ve döngüden çık. Eğer $r_X < P_X(2)$ ise $X = x_2$ ve döngüden çık. … Eğer $r_X < P_X(n)$ ise $X = x_n$ ve döngüden çık.
Adım 8:	Eğer $r_Y < H_Y(1)$ ise $Y = y_1$ ve döngüden çık. Eğer $r_Y < H_Y(2)$ ise $Y = y_2$ ve döngüden çık. … Eğer $r_Y < H_Y(n)$ ise $Y = y_n$ ve döngüden çık.
Adım 9:	Üretilmek istenilen rastgele sayı miktarına gelindiğinde dur, değilse Adım 6'ya git.

Algoritma 5, 1000 kez çalıştırılmıştır. Tablo 3'te verilmiş bir ortak olasılık kütle fonksiyonundan üretilmiş 1000 rastgele sayının histogram grafiği ve beklenen frekansların grafiği Şekil 8'de gösterilmiştir.

Tablo 3. Ortak olasılık kütle fonksiyonu

	$p_{XY}(X=1,Y)$	$p_{XY}(X=2,Y)$	$p_{XY}(X=3,Y)$	$p_{XY}(X=4,Y)$
$p_{XY}(X,Y=1)$	0.1106	0.1240	0.0378	0.1310
$p_{XY}(X,Y=2)$	0.1230	0.0858	0.0742	0.0214
$p_{XY}(X,Y=3)$	0.0172	0.0132	0.1300	0.1318



(a) (b)

Şekil 8. (a) Üretilen 1000 rastgele sayının histogram grafiği (b) Beklenen frekansların sütun grafiği

H_0: Üretilen rastgele sayılar Tablo 3'de verilen ortak olasılık kütle fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Tablo 3'de verilen ortak olasılık kütle fonksiyonundan gelmemiştir.

Gözlenen ve beklenen frekanslar üzerinden ki-kare testi yapıldığında, $\chi_h^2 = 8.186 < \chi_{c=11,\alpha=0.05}^2 = 19.6751$ olduğundan, %95 güven düzeyinde H_0 hipotezi reddedilemez. Üretilen rastgele sayılar Tablo 3'de verilen ortak olasılık kütle fonksiyonundan gelmiştir ($c = SınıfSayısı - 1$).

Diğer yandan, ortak olasılık kütle fonksiyonu, (46) eşitliğindeki gibi,

$$p_{XY}(x, y) = p_X(x)\, p_Y(y) \tag{46}$$

marjinal X ve marjinal Y'nin olasılıklarının çarpımıyla oluşturulabiliyorsa yani diğer bir deyişle X ve Y birbirinden bağımsız ise, koşullu olasılıklara gerek kalmadan da Algoritma 6 ile rastgele sayı üretilebilmektedir.

Algoritma 6. İki değişkenli olasılık kütle fonksiyonundan rastgele sayı üretilmesi

Adım 1: X'in ve Y'nin marjinal olasılıklarını belirle (p_X, p_Y).

Adım 2: X'in marjinal olasılıklarından birikimli marjinal olasılıklarını hesapla (P_X).

Adım 3: Y'in marjinal olasılıklarından birikimli marjinal olasılıklarını hesapla (P_Y).

Adım 4: Düzgün dağılımdan iki tane rastgele sayı üret (r_X, r_Y).

Adım 5: Eğer $r_X < P_X(1)$ ise $X = x_1$ ve döngüden çık.

Eğer $r_X < P_X(2)$ ise $X = x_2$ ve döngüden çık.

…

Eğer $r_X < P_X(n)$ ise $X = x_n$ ve döngüden çık.

Adım 6: Eğer $r_Y < P_Y(1)$ ise $Y = y_1$ ve döngüden çık.

Eğer $r_Y < P_Y(2)$ ise $Y = y_2$ ve döngüden çık.

…

Eğer $r_Y < P_Y(n)$ ise $Y = y_n$ ve döngüden çık.

Adım 7: Üretilmek istenilen rastgele sayı miktarına gelindiğinde dur, değilse Adım 4'e git.

Algoritma 6, 1000 defa çalıştırılmıştır. Tablo 4-5'deki gibi verilmiş marjinal olasılıklar ile oluşturulmuş ortak olasılık kütle fonksiyonundan üretilmiş rastgele sayıların histogram grafiği ve beklenen frekansların grafiği Şekil 9'da gösterilmiştir.

Tablo 4. X'in marjinal olasılıkları (p_X)

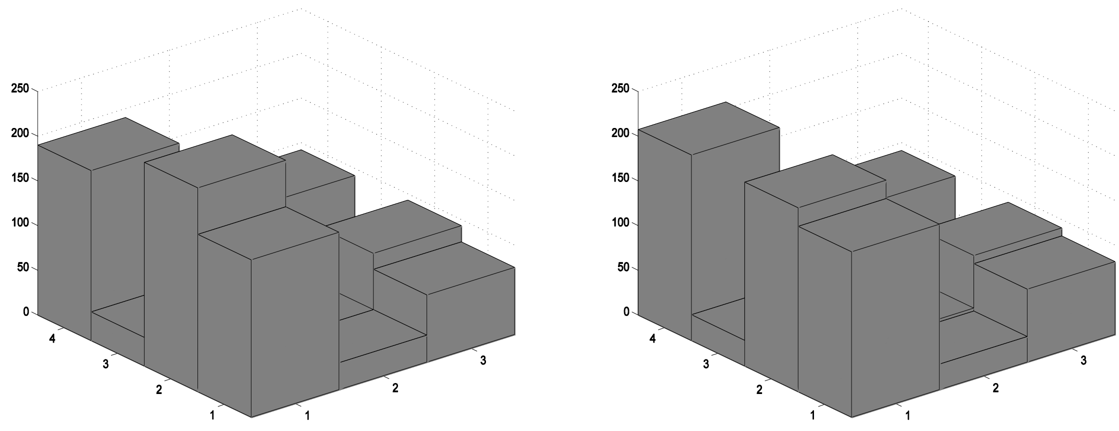
$p_X(x)$
$p_X(X = 1) = 0.2951$
$p_X(X = 2) = 0.3281$
$p_X(X = 3) = 0.0460$
$p_X(X = 4) = 0.3308$

Tablo 5. Y'nin marjinal olasılıkları (p_Y)

$p_Y(y)$
$p_Y(Y = 1) = 0.6271$
$p_Y(Y = 2) = 0.0967$
$p_Y(Y = 3) = 0.2762$

Tablo 6. Ortak olasılık kütle fonksiyonu (p_{XY})

$p_{XY}(x,y) = p_x \times p_y$			
0.185	0.206	0.029	0.207
0.029	0.032	0.004	0.032
0.081	0.091	0.013	0.091



(a) (b)

Şekil 9. (a) Üretilen 1000 rastgele sayının histogram grafiği (b) Beklenen frekansların sütun grafiği

H_0: Üretilen rastgele sayılar Tablo 6'da verilen ortak olasılık kütle fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Tablo 6'da verilen ortak olasılık kütle fonksiyonundan gelmemiştir.

Gözlenen ve beklenen frekanslar üzerinden ki-kare testi yapıldığında, $\chi_h^2 = 6.9717 < \chi_{c=11,\alpha=0.05}^2 = 19.6751$ olduğundan, %95 güven düzeyinde H_0 hipotezi reddedilemez. Üretilen rastgele sayılar Tablo 6'da verilen ortak olasılık kütle fonksiyonundan gelmiştir ($c = SınıfSayısı - 1$).

1.7.3. Sürekli Dağılımdan Rastgele Sayı Üretilmesi

Rastgele sayı üretilmek istenen dağılımın olasılık yoğunluk fonksiyonu $f(x)$, dağılım fonksiyonu $F(x)$ olmak üzere, ters dönüşüm yöntemiyle rasgele sayı üretmek için Eşitlik 47 kullanılmaktadır [8]. U standart düzgün dağılımdan bir rastgele sayıdır.

$$\begin{aligned} F(x) &= U \\ X &= F^{-1}(U) \\ F_X(x) &= P\{X \le x\} \\ &= \{F^{-1}(U) \le x\} \end{aligned} \tag{47}$$

1.7.4. İki Değişkenli Sürekli Rastgele Sayı Üretilmesi

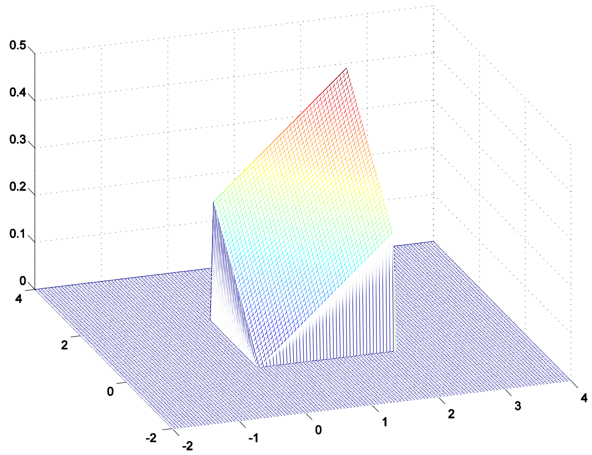
X, Y rastgele değişkeninin olasılık yoğunluk fonksiyonunu $f(x, y)$ ile ifade edilmek üzere, Algoritma 7 yardımıyla $f(x, y)$ ortak olasılık fonksiyonuna sahip (X, Y) rastgele değişkeni üretilebilmektedir [19].

Algoritma 7. İki değişkenli olasılık yoğunluk fonksiyonundan rastgele sayı üretilmesi

Adım 1:	Ortak olasılık yoğunluk fonksiyonunu belirle ($f(x, y)$).
Adım 2:	Ortak olasılık yoğunluk fonksiyonundan X'in marjinal olasılık yoğunluk fonksiyonunu bul ($f_X(x)$).
Adım 3:	X'in marjinal olasılık yoğunluk fonksiyonundan dağılım fonksiyonunu hesapla ($F_X(x)$).
Adım 4:	X'in dağılım fonksiyonunun tersini bul ($F_X^{-1}(x)$).
Adım 5:	Y'nin koşullu olasılık fonksiyonunu hesapla ($h_Y(y\|X)$).
Adım 6:	Y'nin koşullu dağılım fonksiyonunu hesapla ($H_Y(y\|X)$).
Adım 7:	Y'nin koşullu dağılım fonksiyonunun tersini hesapla ($H_Y^{-1}(y\|X)$).
Adım 8:	Düzgün dağılımdan iki tane rastgele sayı üret (r_X, r_Y).
Adım 9:	X'in dağılım fonksiyonun tersinde r_X yerine koy ve X'i üret ($X = F_X^{-1}(r_X)$).
Adım 10:	Adım 9'de üretilen X rastgele sayısını ve r_Y'yi, Adım 7'de hesaplanan koşullu dağılım fonksiyonun tersinde yerine yaz ve Y'yi üret ($Y = H_Y^{-1}(y = r_Y\|X)$).
Adım 11:	İstenilen sayıda rastgele sayı ikilisi (X, Y) üretildiyse dur, aksi halde Adım 8'ye git.

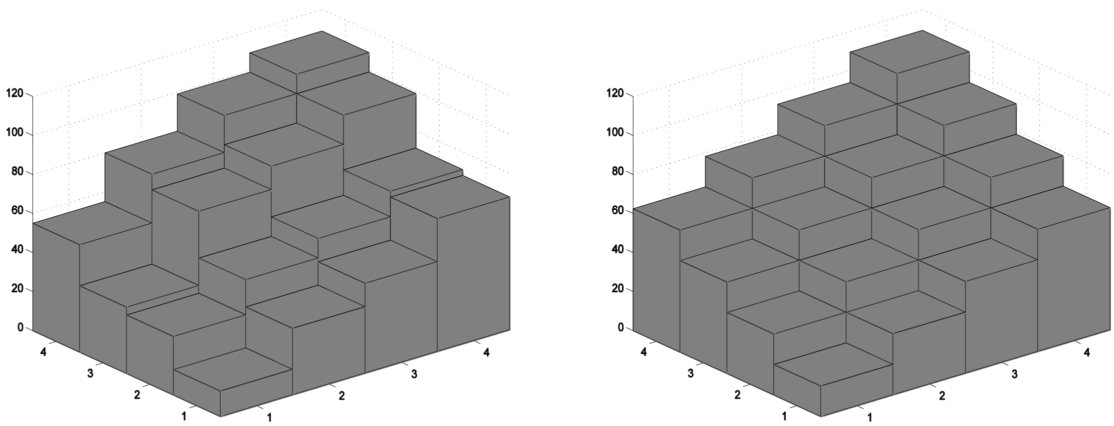
Eşitlik 48'deki gibi verilmiş bir ortak olasılık yoğunluk fonksiyonundan Algoritma 7 kullanılarak, 1000 tane rastgele sayı üretilmiştir. Üretilen 1000 tane rastgele sayının histogram grafiği ve beklenen frekansların histogram grafiği Şekil 11'de gösterilmiştir.

$$f(x,y) = \frac{x+y}{8}, \qquad 0 \leq x,y \leq 2 \tag{48}$$



Şekil 10. Ortak olasılık yoğunluk fonksiyonunun gösterimi

Ortak olasılık yoğunluk fonksiyonu tanımlandığı aralıkların dışında sıfır değerini almaktadır (Şekil 10).



(a) (b)

Şekil 11. (a) Üretilen 1000 rastgele sayının histogram grafiği (b) Beklenen frekansların sütun grafiği

H_0: Üretilen rastgele sayılar Eşitlik 48'de verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Eşitlik 48'de verilen ortak olasılık yoğunluk fonksiyonundan gelmemiştir.

Gözlenen ve beklenen frekanslar üzerinden ki-kare testi yapıldığında, $\chi_h^2 = 9.1906 < \chi_{c=15,\alpha=0.05}^2 = 24.9958$ olduğundan, %95 güven düzeyinde H_0 hipotezi reddedilemez. Üretilen rastgele sayılar Eşitlik 48'de verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir ($c = SınıfSayısı - 1$).

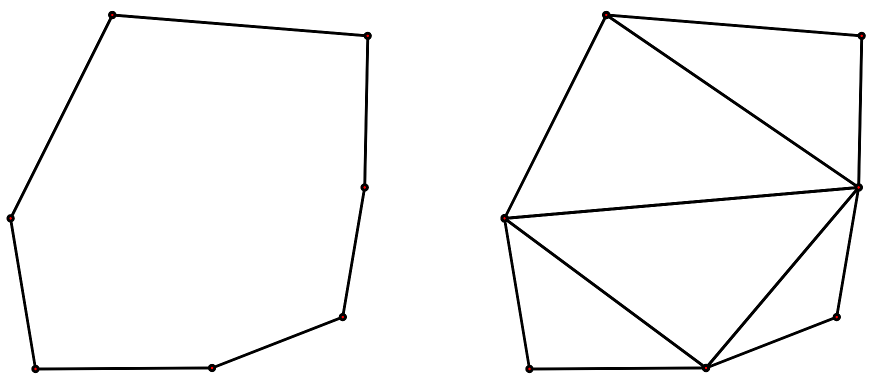
1.8. Üçgenleştirme (Delaunay) Algoritması

Delaunay algoritması, 1934 yılında Boris Delaunay [20] tarafından geliştirilmiştir. Her üç noktadan bir çember geçecek ve hiçbir çevrel çemberin içinde nokta kalmayacak şekilde, bir bölgeyi üçgen parçalara bölen bir algoritmadır [21].

Delaunay üçgenleştirme algoritması aşağıdaki belirtilen özellikleri sağlamaktadır.

1. Algoritmanın çalışmaya başladığı başlangıç noktası ne olursa olsun algoritma her seferinde aynı üçgenlemeyi yapmaktadır.
2. Oluşan üçgenlerin eş açılık özelliği vardır. Bu durum noktanın kendisinden uzakta bulunan noktalarla doğrudan üçgen kurabilmesinin önüne geçmektedir.
3. Algoritma sonucunda oluşan çevrel çemberlerin içerisinde nokta bulunmamaktadır.
4. Her bir noktayı birleştiren doğru parçaları oluşturulan üçgenlerin bir kenarını temsil etmektedir [21–24].

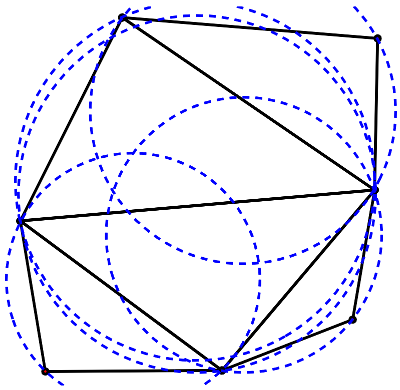
Şekil 12(a)'daki keyfi olarak çizilmiş çokgensel bir alan oluşturulmuştur. Bu çokgensel alanda Delaunay üçgenleştirme algoritması uygulandığında, Şekil 12(b)'deki gibi, üçgenlere bölünmüş bir alan elde edilmektedir.



(a) (b)

Şekil 12. (a) Keyfi olarak çizilmiş çokgensel alan (b) Delaunay üçgenleştirme algoritması uygulanmış çokgensel alan

Şekil 13’de ise her üç noktadan bir çevrel çember geçecek şekilde çokgenin parçalandığını ve çizilen tüm çevrel çemberlerin içinde nokta bulunmadığı gösterilmiştir.



Şekil 13. Çevrel çemberlerin gösterimi

1.9. Ki-Kare Uyum İyiliği Testi

İstatistik biliminde, örneklemin geldiği popülasyonun dağılımı örneklem dağılımını oluşturmaktadır. Örneklemin geldiği yığındaki birimlerin nasıl dağıldığını bilmek istatistikte hangi testin kullanılacağının belirlenmesinde oldukça önemlidir. Bu yüzden, örneklemin hangi dağılımdan geldiğinin test edilmesi gerekmektedir [25].

İstatistiksel bir test yapılacağı zaman, doğru bir sonuç için, kullanılacak olan testin belli varsayımları sağlanmış olması gerekmektedir. Gerekli varsayımları sağlamadan kullanılan testler, kullanım amacına uygun olmayarak hatalı sonuçlar verecektir. Her testin kendine özel varsayımları olacağı gibi çoğu test, sonuç çıkarılacak verinin normal dağılmış olması varsayımı altında çalışmaktadır. Bu tür testler parametrik testler olarak adlandırılır. Normal dağılmayan bir veri için parametrik olmayan test yöntemlerinin kullanılması gerekmektedir. Diğer yandan örneklem sayısının az olduğu durumlarda da hangi testin kullanılacağını belirlemek için, örneklemin nasıl dağıldığının test edilmesi gereken diğer bir durumdur. Örneklemin normal ya da başka herhangi bir dağılımdan gelip gelmediği test etmek için uyum iyiliği testleri kullanılmaktadır. Bu amaçla kullanılan uyum iyiliği testlerinden biri ki-kare uyum iyiliği testidir [26].

Ki–kare testi 1900 yılında Pearson [27] tarafından ortaya atılmıştır. Ki-kare istatistiği uyumun yeterliliği için karar vermek amacıyla kullanılır. “Uyum” terimi gözlenen örneklem dağılımlarının ya da deneysel sonuçlarla elde edilen dağılımların, beklenen ya da normal, binom, Poisson ve düzgün dağılım gibi kuramsal dağılımlarla karşılaştırılması için kullanılır. Beklenen frekansların eğrisi gözlenen frekansların eğrisi üzerine çizilir ve ki-kare istatistiği uyumun yeterli olup olmadığını belirler [10]. Bu testteki istatistik, sapmaların karelerinin toplamının beklenen frekansa bölünmesiyle elde edilmiştir [25].

$$\chi_h^2 = \sum_{i=1}^{c} \frac{(f_i - e_i)^2}{e_i} \tag{49}$$

Gözlenen örneklemin belirli sınıf aralıklarındaki frekans değerleri (f_i), kuramsal olarak hesaplanan o sınıf aralıklarındaki beklenen değerlerine (e_i), yakın olduğunda ki-kare istatistiği sıfıra yaklaşmaktadır. Bu durumda örneklemin dağılımıyla, kuramsal olarak verilen dağılım arasında fark yok denilmektedir. Ki-kare uyum iyiliği testi bu şekilde çalışmaktadır.

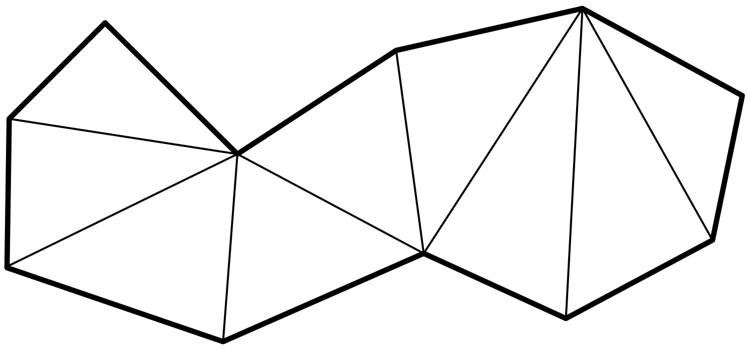
1.9.1. Çokgensel Alanlarda Ki-Kare Uyum İyiliği Testi

Ki-kare uyum iyiliği testi birçok uygulamada gözlenen değerlerin tek değişkenli dağılımlara uyup uymadığını kontrol etmek için kullanılmıştır. Ancak iki boyutlu bir alanda elde edilen gözlem değerleri test edilmesi için dörtgensel bir alanla çerçevelenmesi gerekmektedir. Ancak bilindiği gibi gerçek yaşamda elde edilen veriler her zaman dörtgen alan içinde olmazlar [28].

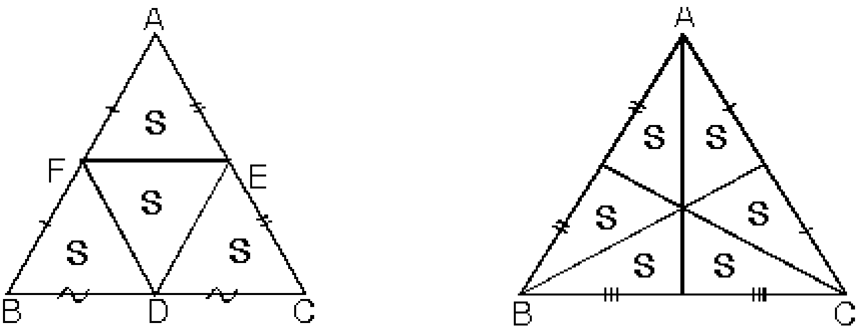
Keyfi bir alan içinde gözlenen verilerin bilindik bir geometrik şekil ile tanımlanmasında en uygun yöntem çokgensel tanımlamadır [29]. Çokgensel bir alan içerinde tanımlanan olasılık yoğunluk değerinden beklenen frekans sayısını bulmak için integralinin alınması gerekir. Bu ise oldukça zor olmaktadır.

Çokgensel alanda ki-kare uyum iyiliği testi yapabilmek için, çokgensel alan uygun üçgenlere bölünerek (Şekil 14) hem sınıf sayısı belirlenebilir hem de bu alandaki beklenen frekans sayısı hesaplanabilmektedir. Beklenen frekans sayısı, üçgen üzerinden integral alınmasında üçgen tepe noktasından iki parçaya bölünerek hesaplanabilmektedir. Aynı şekilde üçgen içerisine düşen gözlenen frekans sayısı da algoritmik olarak hesaplanabilmektedir [29].

Eğer üçgen sayısı ki-kare testi için yeterli değilse her üçgen kendi içerisinde kenarortayları yardımıyla dört parçaya veya altı parçaya bölünerek işlemler gerçekleştirilmektedir [28] (Şekil 15).



Şekil 14. Çokgensel alanın üçgenlere bölünmesi



(a) (b)

Şekil 15. Üçgenlerin kenarortaylar kullanılarak yeni üçgenlere bölünmesi

1.10. İkiye Yarılama Yöntemi

İkiye yarılama metodu matematik biliminde kök bulma problemi olarak adlandırılmaktadır. Bir aralığı tekrar tekrar ikiye böler ve kökün mutlaka içerisinde olduğu alt aralığı bir sonraki adım için işleme alır. Kolay uygulanabilir ve sağlam bir yöntemdir [30].

İkiye yarılama yönteminin özellikleri aşağıdaki gibidir [31].

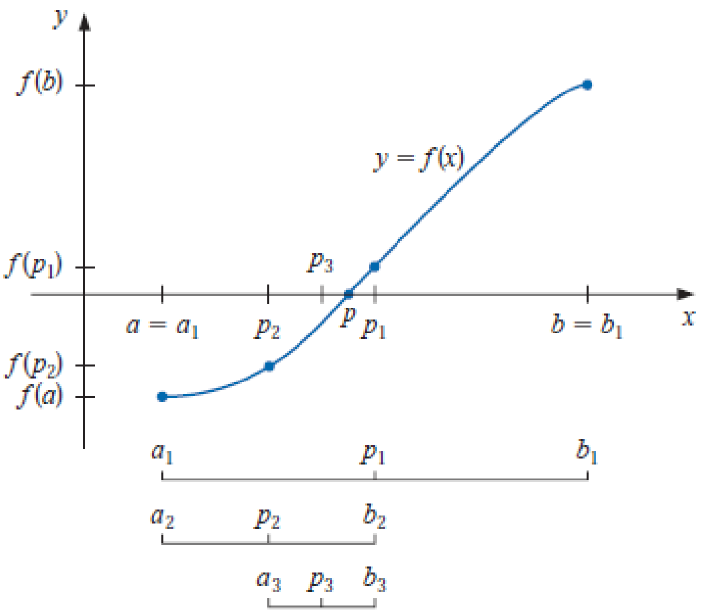
1. f, $[a, b]$ aralığında sürekli bir fonksiyon ve $f(a)$, $f(b)$ zıt işaretli olmalıdır.
2. Fonksiyonu sıfır yapan bir m noktası olmalıdır. ($f(p) = 0$)
3. Fonksiyonun kökü birden fazla olduğunda da çalışmaktadır.

4. $[a, b]$ aralığı her döngüde fonksiyonun kökünü içeren yarı alt aralıklara bölünür.

İkiye yarılama yönteminin algoritması [30] Algoritma 8'de verilmiştir.

Algoritma 8. İkiye yarılama yöntemi algoritması

Adım 1:	f fonksiyonunu, ε durma noktasını ve fonksiyonun tanımlandığı $[a, b]$ aralığını belirle.
Adım 2:	$a_i = a$, $b_i = b$, $p_i = \frac{a+b}{2}$ olarak belirle.
Adım 3:	Eğer $f(p_i) = 0$ ise $p = p_i$ ve dur, değilse Adım 4'e git.
Adım 4:	Eğer $f(p_i)$ ve $f(a_i)$ aynı işaretli ise $a_{i+1} = p_i$ ve $b_{i+1} = b_i$ olarak belirle, Değilse, $a_{i+1} = a_i$ ve $b_{i+1} = p_i$ olarak belirle.
Adım 5:	$\varepsilon > \lvert a_{i+1} - b_{i+1} \rvert$ dur aksi halde Adım 3'e git.



Şekil 16. İkiye Yarılama Yöntemi[30]

2. YAPILAN ÇALIŞMALAR

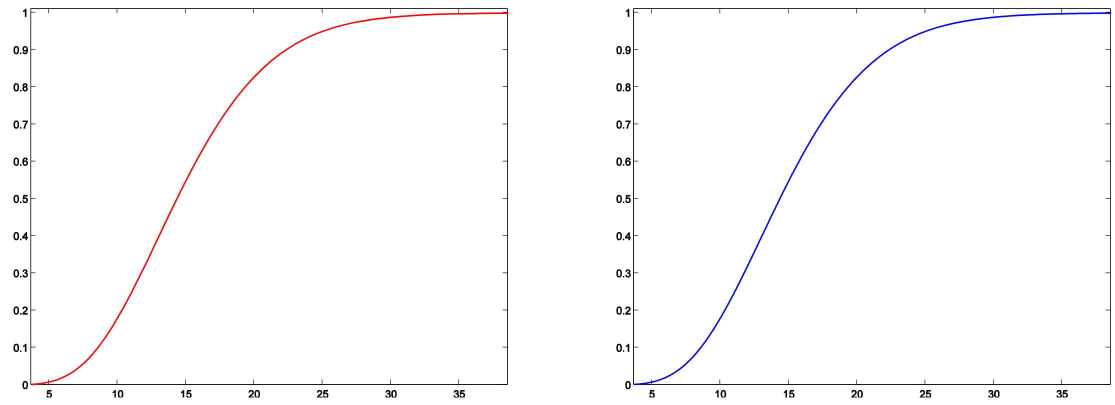
2.1. İki Değişkenli Dağılımlarda Ki-Kare Testi Simülasyonu

Dörtgensel bir alanda tanımlanmış iki değişkenli bir olasılık yoğunluk fonksiyonundan üretilen rastgele sayıların, belirtilen olasılık yoğunluk fonksiyonundan gelip gelmediği ki-kare ile test edilmişti. Aynı şekilde çokgensel alanlarda da üretilen rastgele sayıların seçilen olasılık yoğunluk fonksiyonundan gelip gelmediği de ki-kare ile test edilmişti. Bu kısımda, yukarıda bahsedilen ki-kare testinde hesaplanan ki-kare değerlerinin gerçekte ki-kare dağılıp dağılmadığı, simülasyonla gösterilmek istenmiştir. Ayrıca serbestlik derecesinin ne olduğu ya da ne olacağı ile ilgili belirsizliği ortadan kaldırmakta bu simülasyonun amacını oluşturmaktadır.

Eşitlik (47)'deki gibi tanımlanmış bir ortak olasılık yoğunluk fonksiyonundan 100 tane rastgele sayı üretilmiş ve bu sayıların (47) eşitliğindeki olasılık yoğunluk fonksiyonundan gelip gelmediği ki-kare testi ile test edilmiştir. Ki-kare testi için serbestlik derecesi sınıf sayısının bir eksiği olarak alınmıştır. Sınıf sayısı ise gözlenen frekansların ya da beklenen frekansların boyutu kadar olmaktadır.

Ki-kare testi 1000 kez tekrarlanmış ve hesaplanan 1000 tane ki-kare değerleri kullanılarak bu değerlerin 15 serbestlik dereceli ki kare dağılımından gelip gelmediği tek örneklem Kolmogorov-Smirnov testi [25,32] ile analiz edilmiştir. Serbestlik derecesi beklenen frekans sayısının veya gözlenen frekans sayısının bir eksiği olarak hesaplanmaktadır.

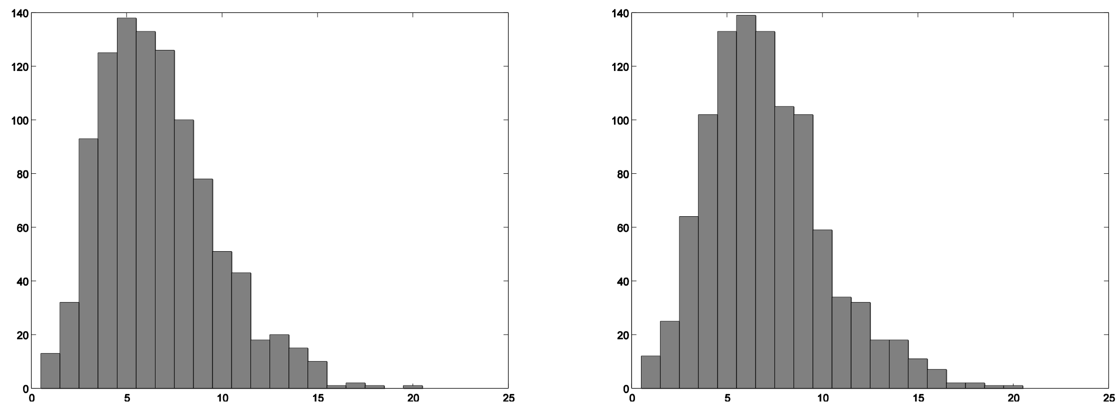
$$f(x,y) = \frac{1}{4} \qquad 0 < x, y < 2 \tag{50}$$



(a) (b)

Şekil 17. Birikimli ki-kare dağılım fonksiyonu (a) Hesaplanan ki-karelerin birikimli dağılımı (b)15 serbestlik dereceli birikimli ki-kare dağılımı

Şekil 17(a)'da her bir döngüde hesaplanan 1000 tane ki-kare değerlerinin birikimli dağılım grafiği çizilmiştir. Şekil 17(b)'de ise 15 serbestlik dereceli ki-kare dağılımı kuramsal olarak çizilmiştir. Elde edilen histogram grafikleri Şekil 18'de gösterilmiştir.



(a) (b)

Şekil 18. (a) Hesaplanan ki-karelerin histogram grafiği (b)15 serbestlik dereceli ki-kare dağılımının histogram grafiği

H_0: Örnekler 15 serbestlik dereceli ki-kare dağılımdan gelmiştir.

H_1: Örnekler 15 serbestlik dereceli ki-kare dağılımdan gelmemiştir.

Tek örneklem Kolmogorov-Smirnov testi için iki birikimli dağılım fonksiyonunun farklarından maksimum değeri alınarak test istatistiği $D_h = 0.0343$ olarak hesaplanmıştır. Tek örneklem Kolmogorov-Smirnov tablo değeri ise n=1000 için $D_{tablo} = 0.0430$,

$D_h < D_{tablo}$ olduğundan %95 güven düzeyinde H_0 reddedilemez. Örnekler 15 serbestlik dereceli ki-kare dağılımından gelmiştir.

$$D_{tablo} = \frac{1.36}{\sqrt{n}} = \frac{1.36}{\sqrt{1000}} = 0.0430 \tag{51}$$

2.2. Üçgensel Alanlarda Rastgele Sayı Üretilmesi

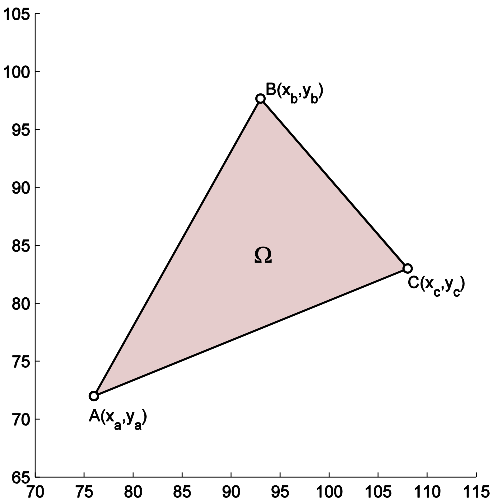
Üçgensel alan içerisinde rastgele sayı üretmek için kullanılacak tekniklerden ret-kabul tekniği Bölüm 2.3.1.'de karma teknik ise Bölüm 2.3.3'te genelleştirilmiş bir yapıda anlatılmıştır. Bu bölümde ters dönüşüm yöntemi ile üçgensel alanda rastgele sayı üretme üzerinde durulacaktır.

2.2.1. Ters Dönüşüm Yöntemi

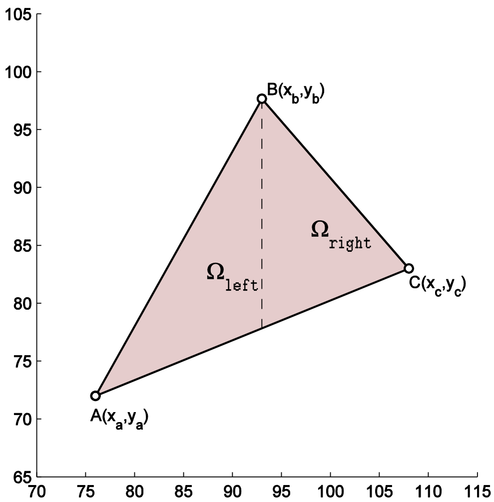
Dörtgensel alanlarda iki değişkenli olasılık yoğunluk fonksiyonlarının sabit sınırlara göre integral hesabı oldukça kolay olmasına rağmen üçgensel alanda integral alma işlemi biraz karmaşıktır. Bu karmaşıklığı ortadan kaldırmak için Şekil 19(a)'da Ω üçgen bölgesinin köşe koordinatlarından x bileşenlerinin $\{x_a, x_b, x_c\}$ ortancası bulunur. Bu ortancadan y eksenine paralel olarak çekilen bir doğru üçgeni iki parçaya bölerek Şekil 19(b) integral alınmasını kolaylaştırmaktadır. Ω bölgesi iki üçgene bölünerek (Ω_L, Ω_R) ayrı ayrı integral hesabı yapılır ve bunların toplamı,

$$\iint_{\Omega_L} f(x,y)dx\,dy + \iint_{\Omega_R} f(x,y)dx\,dy \tag{52}$$

biçiminde toplam integral bulunabilmektedir.



(a)



(b)

Şekil 19. Üçgensel bölgenin gösterimi; (a) Ω bölgesi; (b) Ω_L ve Ω_R üçgensel bölgelerinin belirlenmesi

Ω_L üçgeninin marjinal fonksiyonu,

$$g_{XL}(x) = \begin{cases} \displaystyle\int_{\overline{AC}}^{\overline{AB}} f(x,y)\,dy\,, & m_{\overline{AC}} x_b + c_{\overline{AC}} \leq y_b \\ \displaystyle\int_{\overline{AB}}^{\overline{AC}} f(x,y)\,dy\,, & m_{\overline{AC}} x_b + c_{\overline{AC}} > y_b \end{cases} \tag{53}$$

biçiminde verilmektedir. Burada $\overline{AB}$, A ve B noktalarından geçen doğru denklemi; $\overline{AC}$, A ve C noktalarından geçen doğru denklemidir. Elde edilen marjinal fonksiyonun dağılım fonksiyonu,

$$G_{XL}(x) = \int_{x_a}^{x} g_{XL}(x)dx\,, \qquad x \in [x_a, x_b] \tag{54}$$

biçiminde bulunur. Tüm bölgenin hacmi (olasılık değeri) ise,

$$I_L = G_{XL}(x_b) = \int_{x_a}^{x_b} g_{XL}(x)dx \tag{55}$$

biçiminde elde edilir. Aynı şekilde Ω_R üçgenin marjinal fonksiyonu,

$$g_{XR}(x) = \begin{cases} \displaystyle\int_{\overline{AC}}^{\overline{BC}} f(x,y)\,dy, & m_{\overline{AC}}x_b + c_{\overline{AC}} \leq y_b \\ \displaystyle\int_{\overline{BC}}^{\overline{AC}} f(x,y)\,dy, & m_{\overline{AC}}x_b + c_{\overline{AC}} > y_b \end{cases} \tag{56}$$

biçiminde verilmektedir. Burada $\overline{BC}$, B ve C noktalarından geçen doğru denklemi; $\overline{AC}$, A ve C noktalarından geçen doğru denklemidir. İkinci üçgenin marjinal fonksiyondan dağılım fonksiyonu,

$$G_{XR}(x) = I_L + \int_{x_b}^{x} g_{XR}(x)dx, \qquad x \in [x_b, x_c] \tag{57}$$

biçiminde elde edilir.

$X = x$ verildiğinde Ω_L bölgesinin Y'li koşullu olasılık yoğunluk fonksiyonu,

$$h_{YL}(y|x) = h_{YL}(y|X = x) = \frac{f(X,y)}{g_{XL}(X)} \tag{58}$$

biçiminde tanımlanır. Bu eşitlik $X < x_b$ ise geçerlidir. Aksi durumda, Ω_R bölgesinin Y'li koşullu olasılık yoğunluk fonksiyonu,

$$h_{YR}(y|x) = h_{YR}(y|X = x) = \frac{f(X,y)}{g_{XR}(X)} \tag{59}$$

biçiminde bulunur. Ω_L bölgesinin $X = x$ verildiğinde koşullu dağılım fonksiyonu,

$$H_{YL}(y; y_0|x) = \int_{y_0}^{y} h_{YL}(y|x)\,dy = \frac{1}{g_{XL}(X)} \int_{y_0}^{y} f(X,y)\,dy \tag{60}$$

biçiminde bulunur. Aynı biçimde Ω_R bölgesinin $X = x$ verildiğinde koşullu dağılım fonksiyonu,

$$H_{YR}(y; y_0|x) = \int_{y_0}^{y} h_{YR}(y|x)\, dy = \frac{1}{g_{XR}(X)} \int_{y_0}^{y} f(X, y)\, dy \tag{61}$$

biçiminde bulunur. Bulunan bu analitik eşitlikler Algoritma 9a'daki sırasıyla verilen ortak olasılık yoğunluk fonksiyonuna göre hesaplanır.

Algoritma 9a. Ters Dönüşüm yöntemi ile üçgensel alanda rastgele sayı üretme algoritmasının "analitik kısmı"

Adım 1: Ortak olasılık yoğunluk fonksiyonunu $f(x, y)$ ve Ω üçgensel bölgenin sınırlarını belirle.

Adım 2: Ω bölgesini ortanca apsise (x_b) göre iki üçgene (Ω_L, Ω_R) böl.

Adım 3: Her üçgenin marjinal fonksiyonlarını (53) ve (56) eşitliğinde olduğu gibi hesapla.

Adım 4: Sol üçgenin marjinal fonksiyonunun $[x_a, x_b]$ aralığında integralini al (I_L).

Adım 5: Adım 3'deki marjinal fonksiyonların dağılım fonksiyonlarını (54) ve (57) eşitliklerinde olduğu gibi hesapla.

Adım 6: $X = x$ verildiğinde Ω_L ve Ω_R bölgelerinin koşullu olasılık fonksiyonlarının dağılım fonksiyonlarını (60) ve (61) eşitliklerine göre hesapla. Bu dağılım fonksiyonlarının ters fonksiyonlarını belirle.

Algoritma 9a'da analitik olarak yapılan hesaplamalar Algoritma 9b'de kullanılarak üçgensel alanda rastgele sayı üretimi gerçekleştirilmiş olur.

Algortima 9b. Ters dönüşüm yöntemi ile üçgensel alanda rastgele sayı üretme algoritmasının "simülasyon kısmı"

Adım 1: $\xi_x \sim U(0,1)$ ve $\xi_y \sim U(0,1)$ biçiminde düzgün dağılımdan iki tane rastgele sayı üret.

Adım 2: ξ_x rastgele sayısını aşağıdaki eşitlikte olduğu gibi sınır dağılım fonksiyonun tersinde.

$$X = \begin{cases} G_{XL}^{-1}(\xi_x), & \xi_x \leq I_L \\ G_{XR}^{-1}(\xi_x), & \xi_x > I_L \end{cases} \tag{62}$$

yerine koyarak iki değişkenli rastgele sayının X bileşenini üret (Şekil 20(a)).

Adım 3: X noktasının Ω bölgesinin alt sınırını kestiği $D(X, y_X)$ noktasının y_x bileşenini aşağıdaki doğru denklemelerine göre hesapla.

$$y_x = \begin{cases} m_{\overline{AB}}X + c_{\overline{AB}}, & m_{\overline{AC}}x_b + c_{\overline{AC}} > y_b \; VE \; \xi_x \leq I_L \\ m_{\overline{BC}}X + c_{\overline{BC}}, & m_{\overline{AC}}x_b + c_{\overline{AC}} > y_b \; VE \; \xi_x > I_L \\ m_{\overline{AC}}X + c_{\overline{AC}}, & m_{\overline{AC}}x_b + c_{\overline{AC}} < y_b \end{cases} \tag{63}$$

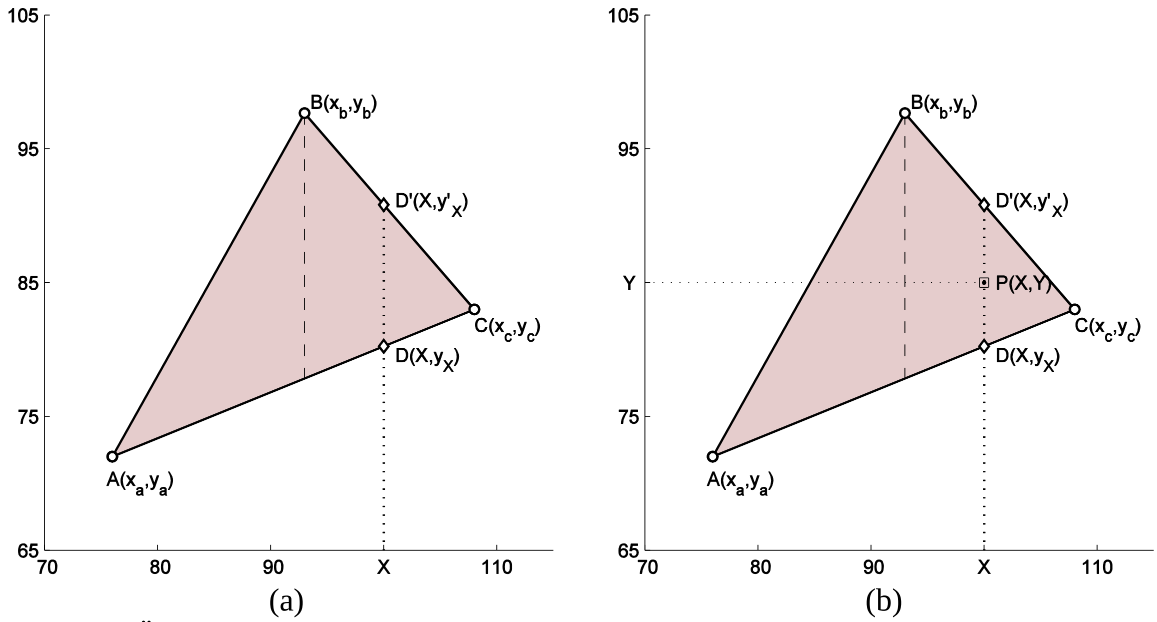
burada m_* (*) doğrusunun eğimini c_* ise (*) doğrusunun sabitini göstermektedir.

Adım 4: ξ_y rastgele sayısını, y_x ve X değerini aşağıdaki eşitlikte olduğu gibi koşullu dağılım fonksiyonun tersinde,

$$Y = \begin{cases} H_{YL}^{-1}\left(\xi_y; y_x|x\right), & \xi_x \leq I_L \\ H_{YR}^{-1}\left(\xi_y; y_x|x\right), & \xi_x > I_L \end{cases} \tag{64}$$

yerine koyarak iki değişkenli rastgele sayısının Y bileşenini üret (Şekil 20(b)).

Adım 5: Yeni bir sayı üretilecekse Adım 1’e giderek işlemi tekrarla.

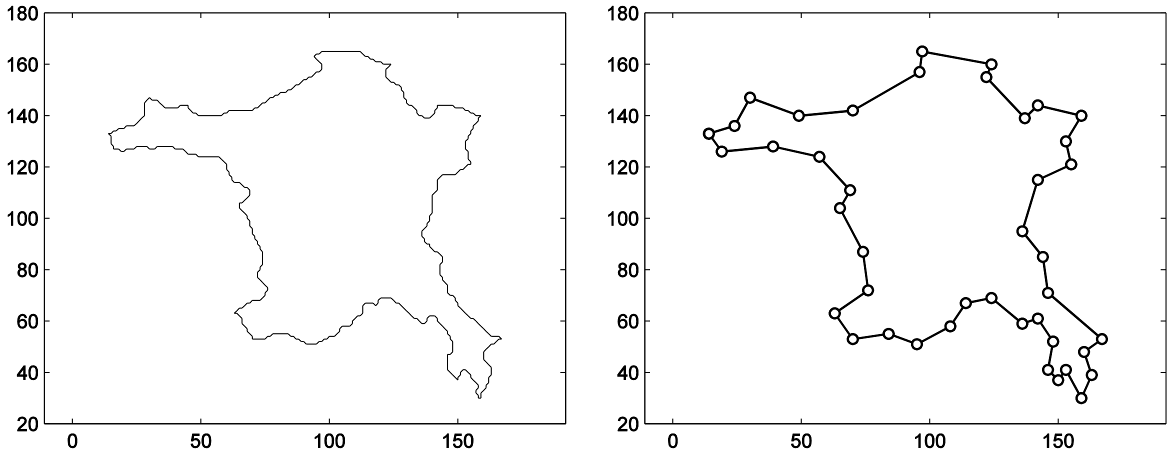


(a) (b)

Şekil 20. Üçgensel alanda iki değişkenli rastgele sayı üretimi; (a) iki değişkenli rastgele sayının X bileşenin üretilmesi; (b) iki değişkenli rastgele sayının Y bileşenin üretilmesi

2.3. Çokgensel Alanda Rastgele Sayı Üretilmesi

Genelde iki değişkenli istatistiksel hesaplamalar dikdörtgensel alanda yapılmaktadır. Ancak gerçek hayatta istatistiksel verilerin dörtgensel bir alanda olma olasılığı oldukça düşüktür. Bunun yerine keyfi bir alanda istatistiksel hesaplama yapma gereği ortaya çıkmaktadır. Bu çalışmada Ankara ilinin sınırları Şekil 21(a)'da verilmiştir. Sınırları keyfi olarak verilen bu alanda hesaplama yapabilmek için bu alanın sınırlarının çokgensel bir yapıda tanımlanması gerekir. Keyfi bir alanı çokgensel olarak tanımlayabilmek için sınırların baskın noktaları belirlenir. Bu baskın noktalar ister elle, ister bilgisayarda kullanılan baskın nokta algoritmalarından bölme-birleştirme, sıralı yaklaşım, çevre bölütleme, Teh-Chin algoritmaları vb. kullanılarak otomatik yapılabilmektedir [33–37] (Şekil 21(b)).



(a) (b)

Şekil 21. Çokgensel alan tanımlaması; (a) Keyfi bir alan; (b) Alan üzerinde baskın noktaların belirlenmesiyle elde edilen çokgensel alan

Çokgensel alanda rastgele sayı üretirken hem ret-kabul hem de ters dönüşüm yönteminde gerek duyulacak olan, bir noktanın çokgenin içerisinde olup olmadığının araştırılmasıdır. Bu problem çokgen içerisindeki nokta problemi olarak bilinmektedir [37]. Bu problemin çözümü için, rastgele seçilen (X, Y) noktasından $y = x_{min}$ doğrusuna çekilen dik bir doğru parçasının, çokgeni oluşturan tüm doğruları kaç noktadan kestiğinin belirlenmesi gerekir. Kesilen nokta sayısı tek ise nokta çokgenin içinde, eğer çift ise nokta çokgenin dışındadır. Bu işlem için MATLAB® yazılımında `inpolygon` komutu bulunmaktadır.

2.3.1. Ret Kabul Yöntemi

Dikdörtgensel bir alan içerisinde keyfi ortak dağılım fonksiyonundan rastgele bir sayı oluşturmak için ilk anlatılacak teknik ret-kabul tekniğidir. Dörtgensel ret-kabul tekniğine göre rastgele sayı üretme algoritması,

Algoritma 10. Çokgensel alanda ret-kabul tekniğine göre rastgele sayı üretme algoritması.

Adım 1: Ortak olasılık yoğunluk fonksiyonunu $f(x,y)$ ve çokgensel bölgenin düğüm koordinatlarını (x_k, y_k) belirle.

Adım 2: $X \sim U(x_{min}, x_{max})$ ve $Y \sim U(y_{min}, y_{max})$ biçiminde düzgün dağılımdan iki tane rastgele sayı üret.

Buradaki değişkenler çokgeni oluşturan noktalardan,

$$x_{min} = \min(x_k)$$
$$x_{max} = \max(x_k)$$
$$y_{min} = \min(x_k)$$
$$y_{max} = \max(y_k)$$

biçiminde belirlenmektedir.

Adım 3: Oluşturulan bu sayının (X, Y), çokgen içerisinde olup olmadığını kontrol et. Eğer sayı çokgen içerisinde değilse Adım 2'e giderek tekrar bir sayı (X, Y) üret.

Adım 4: $z \sim U(0, f_{max})$ biçiminde düzgün dağılımdan bir tane rastgele sayı üret. Burada f_{max} değeri, $f(x,y)$ fonksiyonun $f(x,y) \in \Omega$ bölgesindeki en büyük değeridir.

Adım 5: Eğer $z > f(X, Y)$ ise üretilen iki değişkenli rastgele sayıyı (X, Y) reddet ve Adım 2'ye git, değilse (X, Y) rastgele sayısını kabul et.

biçiminde verilmektedir.

2.3.2. Ters Dönüşüm Yöntemi

Ters dönüşüm yöntemine göre rastgele sayı üretmek için çokgensel alanın öncelikle üçgenlere bölünmesi gerekir. Bunun için Delaunay algoritması [20] kullanılır. Çokgensel alan birçok üçgene bölündükten sonra her bir üçgenin (Ω_j) olasılık değeri (üçgen içerisindeki olasılık değeri, P_j),

$$P_j = \int_{\Omega_j} f(x,y) d_x d_y \tag{65}$$

biçiminde tanımlanır. Üçgensel bir bölgede integral almak için Şekil 19(b)'de gösterildiği gibi üçgen iki parçaya bölünür. Daha sonra bunların integrali,

$$P_j = \left| \int_{x_a}^{x_b} \int_{\overline{AC}}^{\overline{AB}} f(x,y) dy dx \right| + \left| \int_{x_b}^{x_c} \int_{\overline{AC}}^{\overline{BC}} f(x,y) dy dx \right| \tag{66}$$

biçiminde ayrı ayrı hesaplanarak toplanır. Bu integrallerde doğruların integral sınırları yer değiştirebileceğinden integralin pozitif çıkması için mutlak değeri alınmaktadır. Tüm üçgenlerin olasılık değerlerinin toplamı,

$$\sum P_j = 1 \tag{67}$$

olasılık aksiyomları gereği 67 eşitliğinde olduğu gibi 1'e eşit olması gerekir.

Bu olasılık değerlerine göre üçgenler rastgele seçilir. Seçilen üçgenin içerisinde rastgele sayılar üretmek için Bölüm 2.2.'de verilen algoritma kullanılır. Ters dönüşüm yöntemi ile çokgensel alanda rastgele sayı üretme algoritması,

Algoritma 11. Ters dönüşüm yöntemi ile çokgensel alanda rastgele sayı üretme algoritması

Adım 1: Çokgensel alanı Delaunay üçgenleme yöntemi ile üçgenlere böl, (Şekil 22(a)). Eğer çokgensel alan dışbükey bir alan değilse Şekil 22(a)'da görüldüğü gibi üçgenleme algoritması çokgenin dışında da üçgen bulmaktadır. Bu sorunu çözmek için her üçgenin ağırlık merkezi bulunur ve bu merkezlerden çokgenin dışına düşenler silinir (Şekil 22(b)).

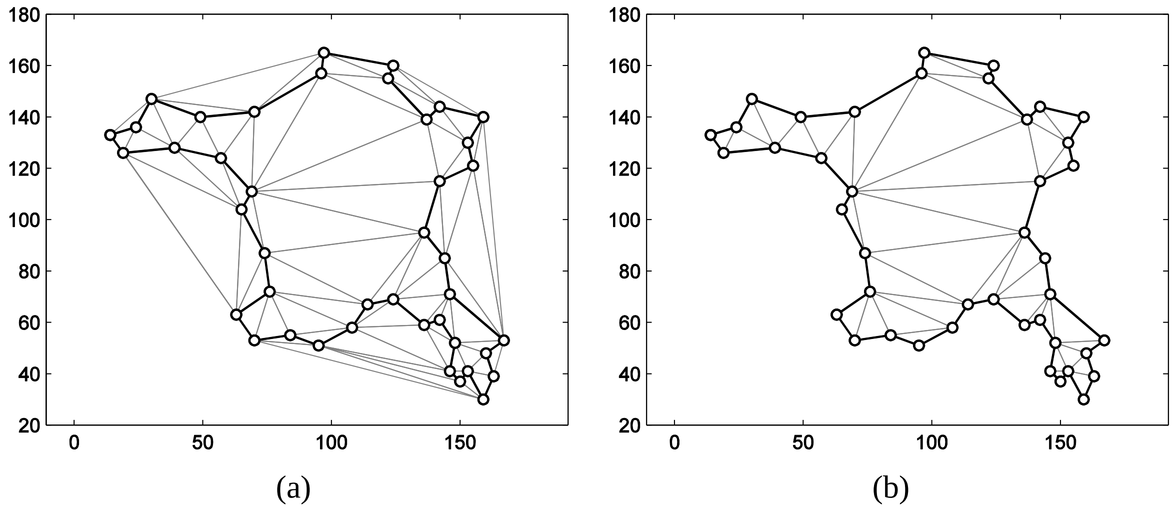
Adım 2: Oluşan her bir üçgenin Şekil 19(b)'de görüldüğü gibi iki üçgene bölünerek bunların seçilme olasılıklarını P_j, 66 eşitliği yardımıyla hesapla.

Adım 3: P_j olasılıklarına göre rastgele bir üçgen seç.

Adım 4: Seçilen üçgen içerisinde rastgele bir sayıyı (X, Y) Algoritma 9b yardımıyla belirle.

---

biçiminde verilmektedir.



Şekil 22. Çokgensel alana Delaunay üçgenleme yönteminin uygulanması; (a) çokgensel alanın üçgenlere bölünmesi; (b) çokgen dışında üçgenlenmiş bölgelerin atılması

Tersi alınabilen bir fonksiyon verildiğinde istenen bir bölgede etkili bir biçimde sözde rastgele sayılar oluşturulabilmektedir. Ancak tersi alınamayan dağılım fonksiyonları için ise sayısal yöntemler kullanılmaktadır.

2.3.3. Karma Yöntem

Çokgensel bir alan içerisinde iki değişkenli dağılımlardan rastgele sayı üretmek için diğer bir yaklaşım ise karma tekniktir. Bölüm 1.7.4'de verilen, dörtgensel alanda ters dönüşüm yöntemi ile rastgele sayı üretme algoritması kullanılarak iki değişkenli rastgele sayılar (X, Y) elde edilmektedir. Bu rastgele sayılardan çokgenin içine düşenler kabul edilip diğerleri reddedilerek istenen rastgele sayılar elde edilmektedir.

Algoritma 12. Karma yönteme göre çokgensel alan içerisinde iki değişkenli dağılım fonksiyonundan rastgele sayı üretme algoritması.

Adım 1: Ortak olasılık yoğunluk fonksiyonunu $f(x,y)$'i belirle.

Adım 2: Standart düzgün dağılımdan iki tane rastgele sayı üret (ξ_X, ξ_Y).

Adım 3: ξ_X'in değerini aşağıdaki ters marjinal dağılım fonksiyonunda yerine koyarak,

$$X = G_X^{-1}(\xi_X)$$

X rastgele sayı bileşeni üret.

Adım 4: X ve ξ_Y değerlerini aşağıdaki Y'nin ters koşullu dağılım fonksiyonunda yerine koyarak,

$$Y = H_Y^{-1}(\xi_Y | X = x)$$

Y rastgele sayı bileşenini üret.

Adım 5: Oluşturulan bu sayının (X, Y), çokgen içerisinde olup olmadığını kontrol et. Eğer sayı çokgen içerisinde değilse Adım 2'ye giderek tekrar bir sayı (X, Y) üret.

2.4. Simülasyon Çalışması

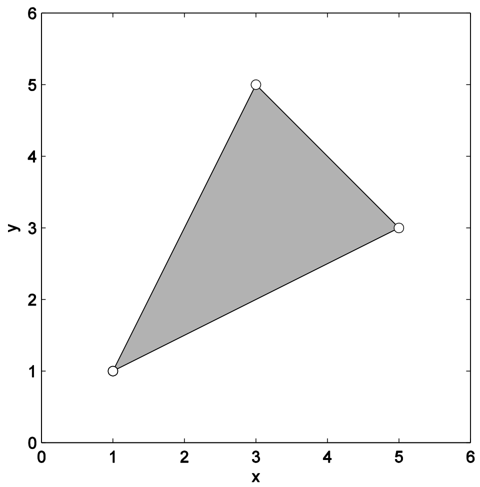
Bölüm 2.2 ve 2.3'te geliştirilen algoritmalar yardımıyla çeşitli geometrik alanlarda rastgele sayı üretimi için iki ayrı örnek belirlenmiştir. Bu örneklerden bir tanesinde üçgensel bir bölgede rastgele sayı üretimi ele alınırken, diğerinde Afrika ana kıtası şekli ele alınmıştır. Simülasyon ile ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayıların verilen ortak olasılık yoğunluk fonksiyonundan gelip gelmediği ve rastgele sayı üretme yöntemlerinden hangisinin daha hızlı çalıştığının bulunması amaçlanmıştır. Simülasyon Intel® Core(TM) i7-4770 CPU, 16 Gb Ram özellikli bir bilgisayarda yapılmıştır. Programlama dili olarak Matlab® yazılımı kullanılmıştır.

2.4.1. Üçgensel Bölgede Rastgele Sayı Simülasyonu

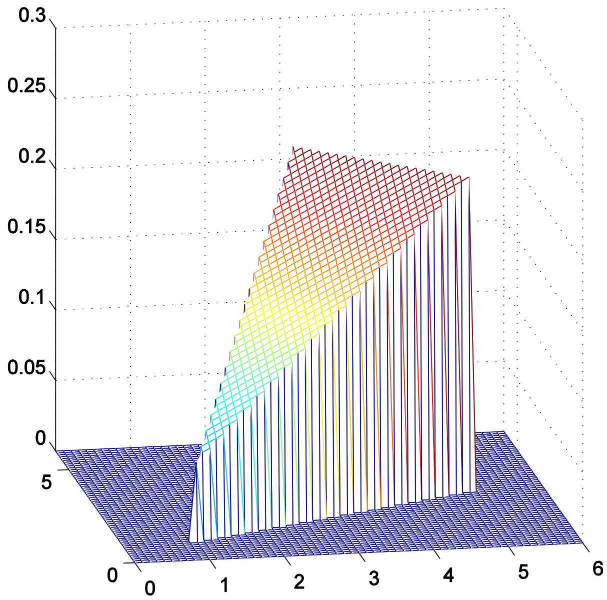
Birinci örnekte üçgensel bir alanda sınırlandırılmış bir olasılık yoğunluk fonksiyonu,

$$f(x,y) = \frac{1}{36}(x+y), \quad (x,y) \in \Omega \tag{68}$$

biçiminde tanımlanmıştır. Burada olasılık yoğunluk fonksiyonu 3 köşeden oluşan $\Omega = \{p_k: (x_k, y_k) = \{(1,1),(3,5),(5,3)\},\ k = 1,2,3\}$ bir bölge ile sınırlandırılmıştır. Bu bölge Şekil 23'de verilirken aynı bölge içerisindeki olasılık yoğunluk fonksiyonu Şekil 24'de gösterilmektedir.

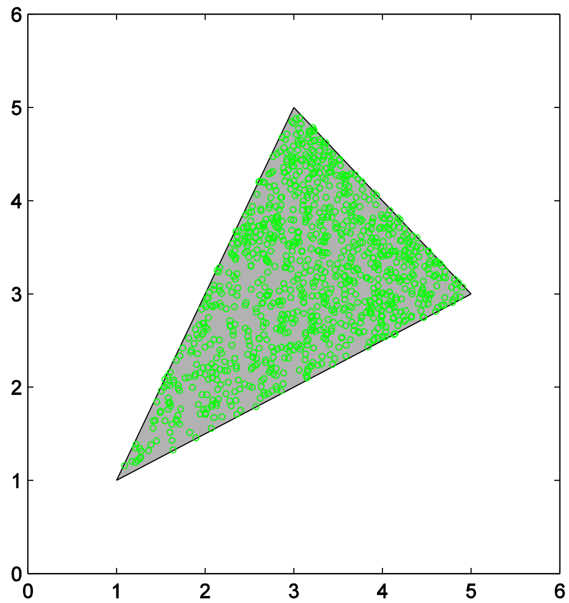


Şekil 23. Örnek 1'in tanımlanması: Üçgensel bölge

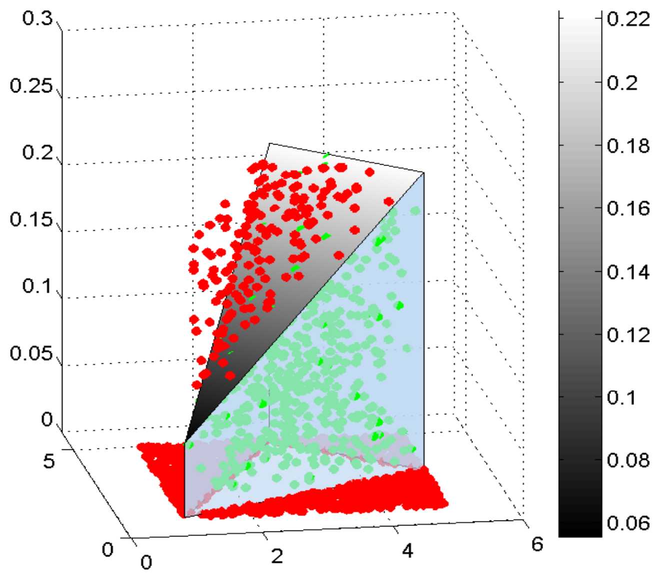


Şekil 24. Üçgensel bölge içerisindeki olasılık yoğunluk fonksiyonu

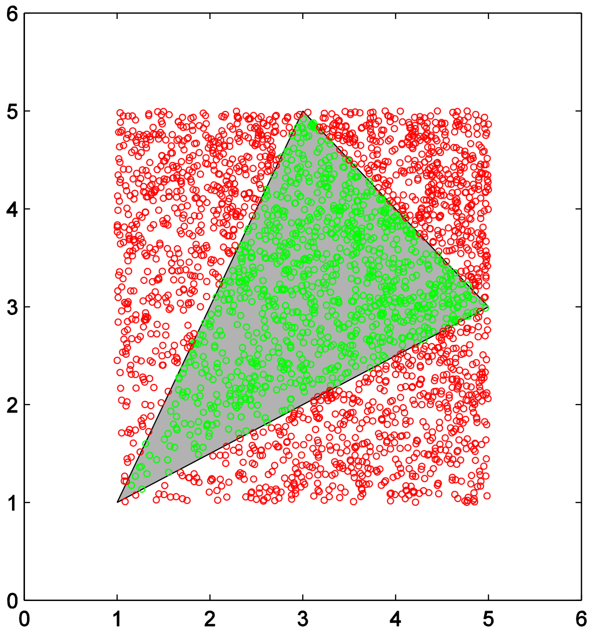
Verilen olasılık yoğunluk fonksiyonundan üç farklı yönteme göre rastgele sayılar üretilmiştir. Bu rastgele sayılar 100 ve 1000 olmak üzere iki farklı hacimde üretilmiştir. Yapılan bu işlem 10000 defa tekrarlanarak bir benzetim gerçekleştirilmiştir. Benzetimin bir aşamasında elde edilen 1000 rastgele sayı üç yöntem için de Şekil 25-27'de gösterilmiştir.



Şekil 25. Ters dönüşüm yöntemiyle üretilen rastgele sayılar



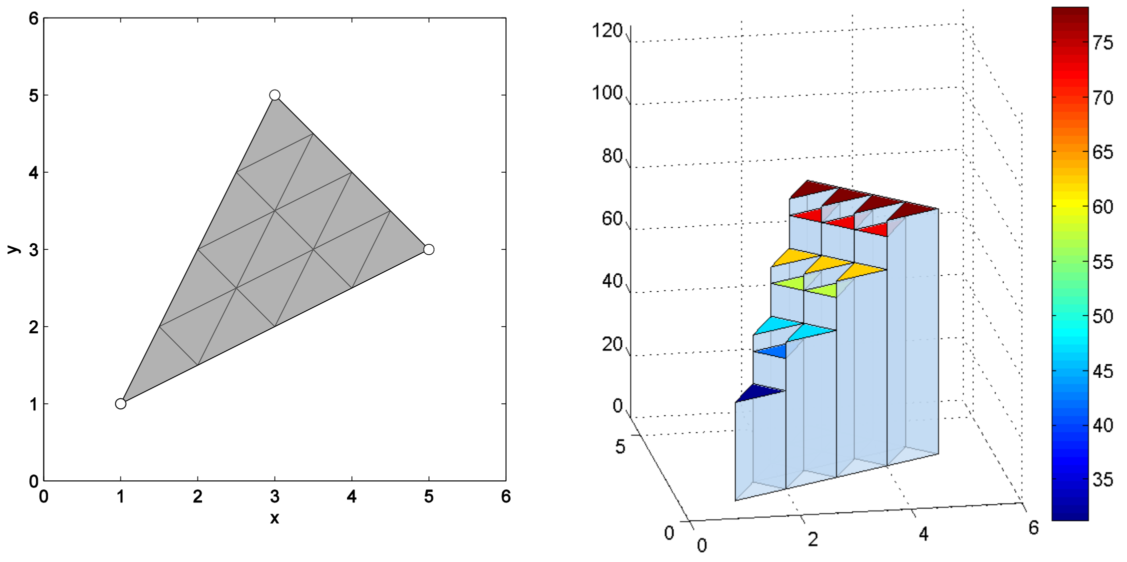
Şekil 26. Ret-kabul yöntemiyle üretilen rastgele sayılar



Şekil 27. Karma yöntemle üretilen rastgele sayılar

2.4.2. Üçgensel Bölgede Üretilen Rastgele Sayıların Ki-Kare Testi

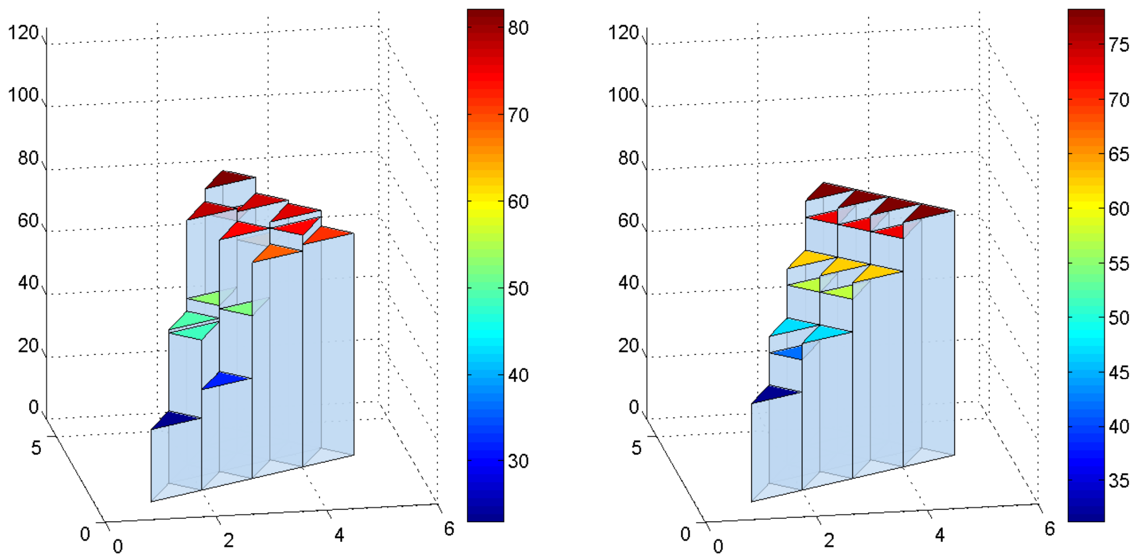
Üçgensel bölgede üretilen rastgele sayıların verilen olasılık yoğunluk fonksiyonuna uyup uymadığını göstermek için ki-kare uyum iyiliği testi kullanılmıştır. Ki-kare uyum iyiliği testini üçgensel bir bölgede kullanmak için öncelikle bölgenin sınıflara ayrılması gerekmektedir. Genelde iki değişkenli örnekler için dörtgensel aralıklarda sınıflama yapılmaktadır. Ancak üçgensel bir bölgede dörtgensel sınıflandırma yapmanın zorluğundan dolayı bölge üçgenlere bölünerek sınıflandırma yapılmıştır (Şekil 28(a)). Aynı küçük üçgenlerdeki beklenen frekanslarda önceki bölümlerde bahsedilen parçalı üçgenler üzerinden integral alınarak hesaplanabilmektedir (Şekil 28(b)).



(a) (b)

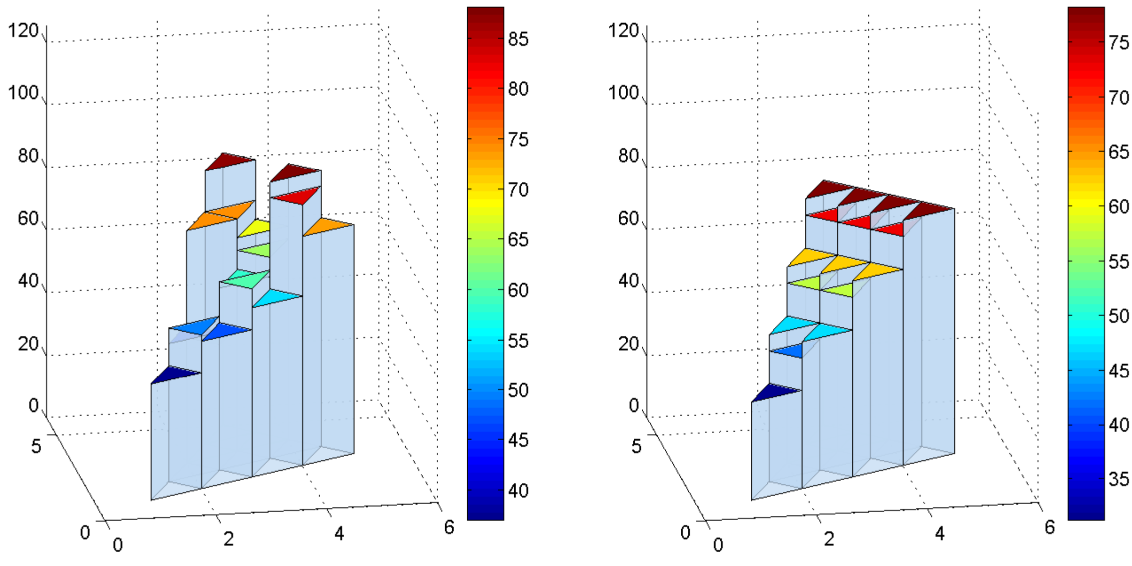
Şekil 28. Üçgensel bölgede beklenen frekans hesabı; (a) Üçgensel bölgenin küçük üçgenlere bölünmesi; (b) Her bir küçük üçgenin beklenen frekansları

Üç yönteme göre üretilen rastgele sayıları test etmek için üçgensel alan küçük üçgenlere bölünerek gözlenen frekansları bulunmaktadır (Şekil 29-31).



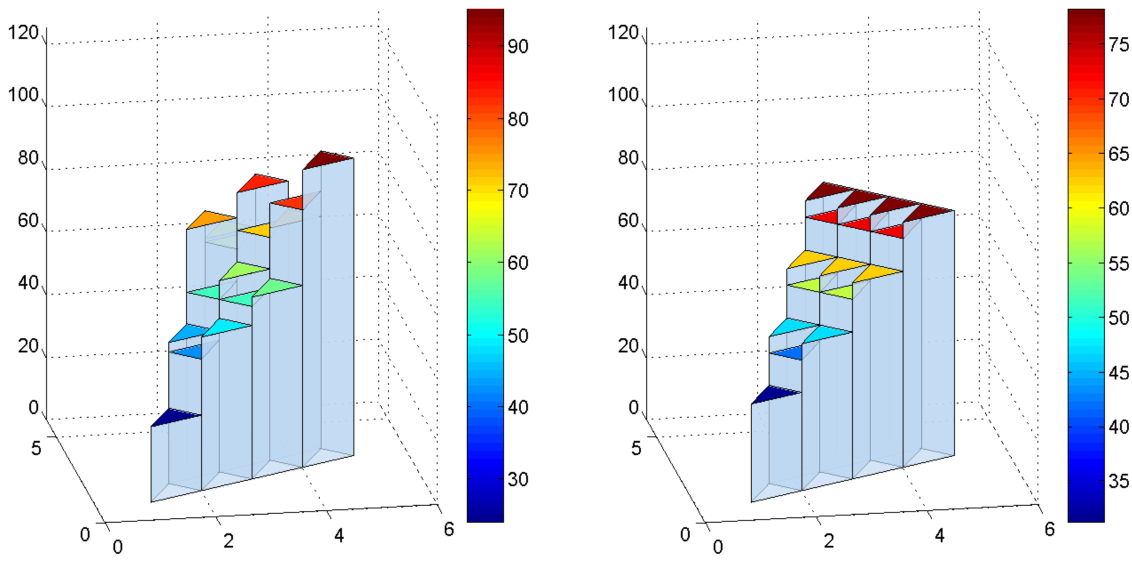
(a) (b)

Şekil 29. Gözlenen frekanslar; (a) Ret-kabul yöntemiyle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları



(a) (b)

Şekil 30. Gözlenen frekanslar; (a) Ters dönüşüm yöntemiyle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları



(a) (b)

Şekil 31. Gözlenen frekanslar; (a) Karma yöntemle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları

Her bir yöntem için gözlenen ve beklenen frekanslar kullanılarak 10000 defa ki-kare testi yapılmıştır. Denemeler sonucunda elde edilen sonuçlar özet halinde Tablo 7'de gösterilmiştir.

Tablo 7. N=10000 defa yapılan denemeler sonucunda elde edilen sonuçlar

Yöntem	n=100		n=1000	
	HKY(χ^2)	OHZ(sn)	HKY(χ^2)	OHZ (sn)
Ret-Kabul	0.9480	0.0373	0.9509	0.3651
Ters Dönüşüm	0.9493	0.0266	0.9487	0.2595
Karma	0.9512	0.0243	0.9499	0.2422

*(HKY: Hipotezin kabul yüzdesi; OHZ: Ortalama hesaplama zamanı)

H_0: Üretilen rastgele sayılar Eşitlik 68'de verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Eşitlik 68'de verilen ortak olasılık yoğunluk fonksiyonundan gelmemiştir.

Ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayılar %95 güven düzeyinde test edilmiştir. Bu üç yöntemde H_0 hipotezi iki örnek hacmi için de yaklaşık olarak %95 oranında kabul edilmiştir. Serbestlik derecesi sınıf sayısının bir eksiği 15 olarak alınmıştır. Diğer yandan, karma yöntemin, ortalama hesaplama zamanı bakımından diğer iki yöntemden daha hızlı çalıştığı görülmüştür.

2.4.3. Çokgensel Bölgede Rastgele Sayı Simülasyonu

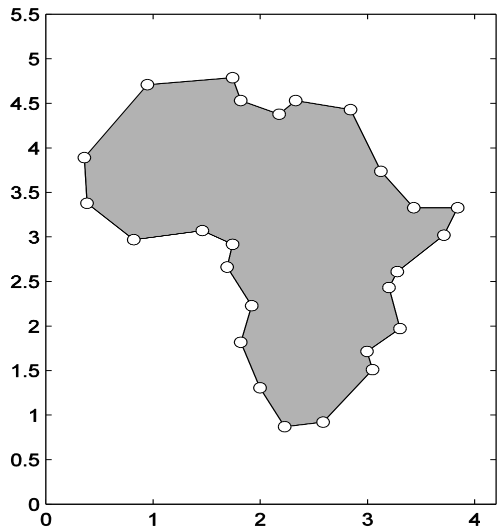
İkinci örnekte çokgensel alan olarak Afrika ana kıtasının şekli seçilmiştir. Afrika ana kıtasının tümü düzlemsel bir alan olarak kabul edilip ona göre rastgele sayı üretilmiştir.

$$f(x,y) = \frac{1}{36}(x+y), \quad (x,y) \in \Omega \tag{69}$$

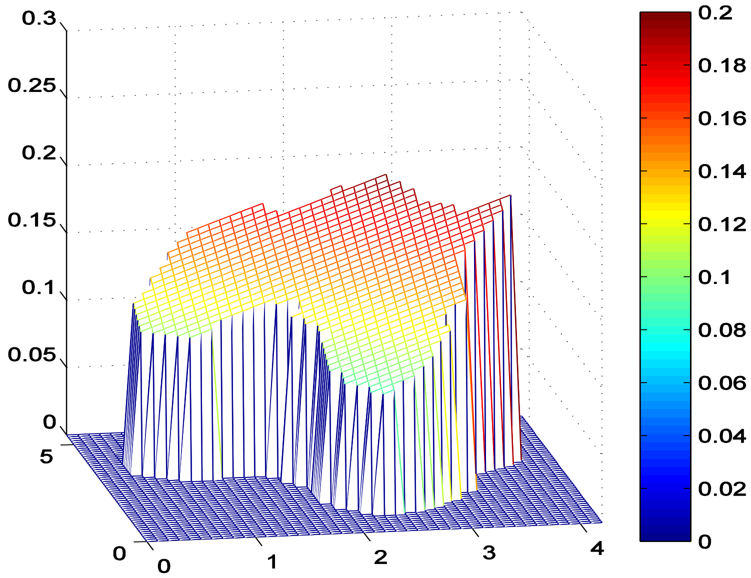
Burada olasılık yoğunluk fonksiyonu Afrika ana kıtası üzerinden oluşturulan bir bölge ile sınırlandırılmıştır. Burada olasılık yoğunluk fonksiyonu 26 köşeden oluşan

$$\Omega = \begin{Bmatrix} (0.95,4.71), & (1.74,4.79), & (1.82,4.53), & (2.18,4.38), & (2.33,4.53), \\ (2.84,4.43), & (3.12,3.74), & (3.43,3.33), & (3.84,3.33), & (3.71,3.02), \\ (3.28,2.61), & (3.2,2.43), & (3.3,1.97), & (3,1.72), & (3.05,1.51), \\ (2.59,0.92), & (2.23,0.87), & (2,1.31), & (1.82,1.82) & (1.92,2.23), \\ (1.69,2.66), & (1.74,2.92), & (1.46,3.07), & (0.82,2.97), & (0.38,3.38), \\ (0.36,3.89) & & & & \end{Bmatrix}$$

bir bölge ile sınırlandırılmıştır Bu bölge Şekil 32'de verilirken aynı bölge içerisindeki olasılık yoğunluk fonksiyonu Şekil 33'de gösterilmektedir.

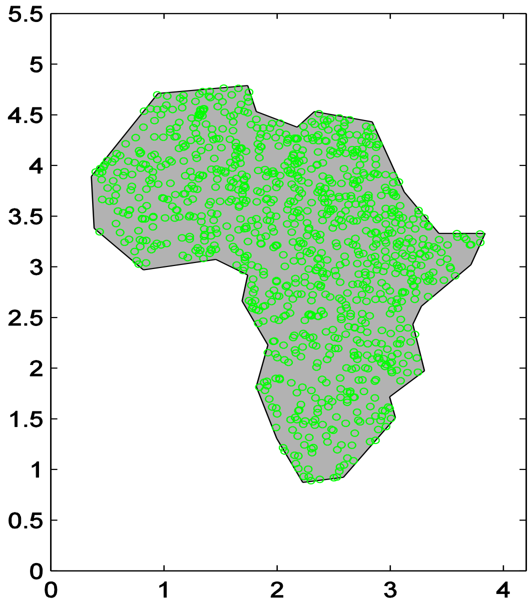


Şekil 32. Örnek 2'in tanımlanması: Çokgensel bölge

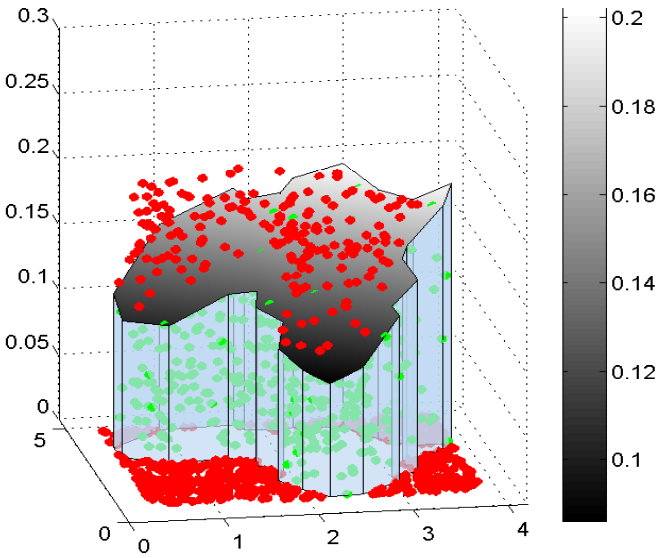


Şekil 33. Çokgensel bölge içerisindeki olasılık yoğunluk fonksiyonu

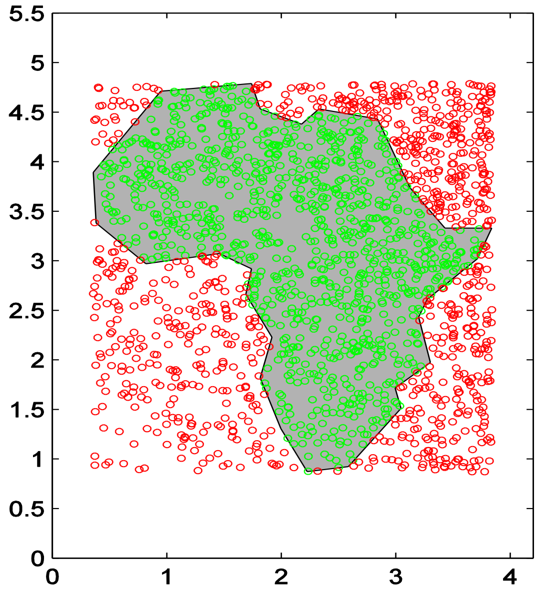
Verilen olasılık yoğunluk fonksiyonuna göre üç farklı tekniğe göre rastgele sayılar üretilmiştir. Bu rastgele sayılar 100 ve 1000 olmak üzere iki farklı hacimde üretilmiştir. Yapılan bu işlem 10000 defa tekrarlanarak bir benzetim gerçekleştirilmiştir. Benzetim bir aşamasında elde edilen 1000 rastgele sayı üç yöntem için de Şekil 34-36'da gösterilmiştir.



Şekil 34. Ters dönüşüm yöntemiyle üretilen rastgele sayılar



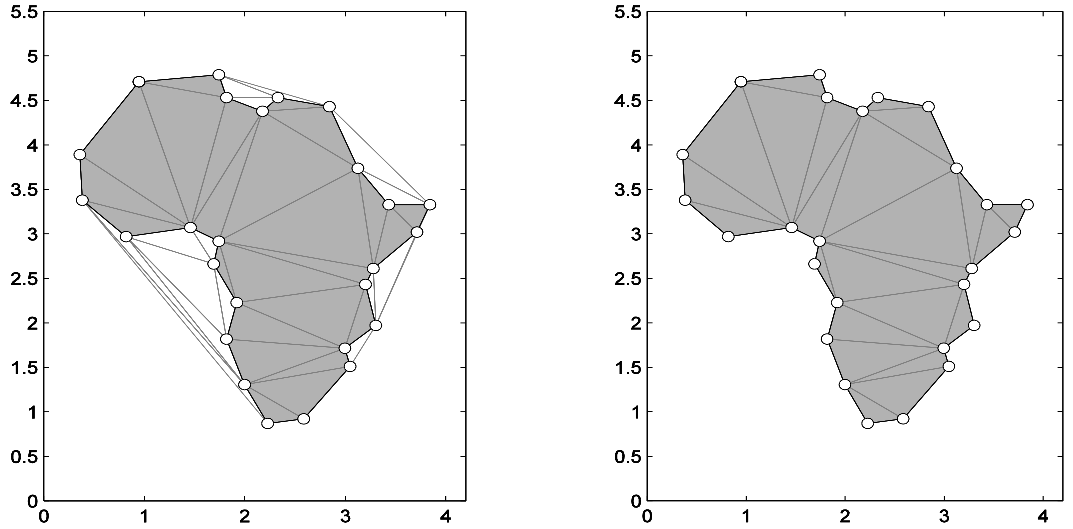
Şekil 35. Ret-kabul yöntemiyle üretilen rastgele sayılar



Şekil 36. Karma teknikle üretilen rastgele sayılar

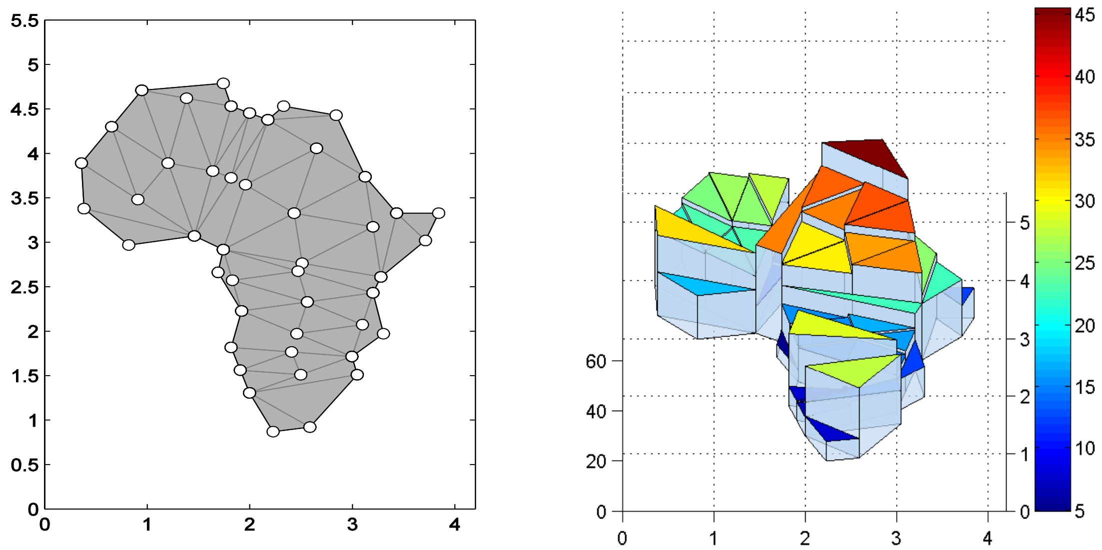
2.4.4. Çokgensel Bölgede Üretilen Rastgele Sayıların Ki-Kare Testi

Çokgensel bölgede üretilen rastgele sayıların verilen olasılık yoğunluk fonksiyonuna uyup uymadığını göstermek için ki-kare uyum iyiliği testi kullanılmıştır. Ki-kare uyum iyiliği testini çokgensel bir bölgede kullanmak için öncelikle bölgenin sınıflara ayrılması gerekmektedir. Genelde iki değişkenli örnekler için dörtgensel bir aralıklarda sınıflama yapılmaktadır. Ancak çokgensel bölgede dörtgensel sınıflandırma yapmanın zorluğundan dolayı bölge delaunay yöntemiyle üçgenlere bölünerek sınıflandırma yapılmıştır (Şekil 37). Aynı küçük üçgenlerdeki beklenen frekanslarda önceki bölümlerde bahsedilen parçalı üçgenler üzerinden integral alınarak hesaplanabilmektedir (Şekil 38).



(a) (b)

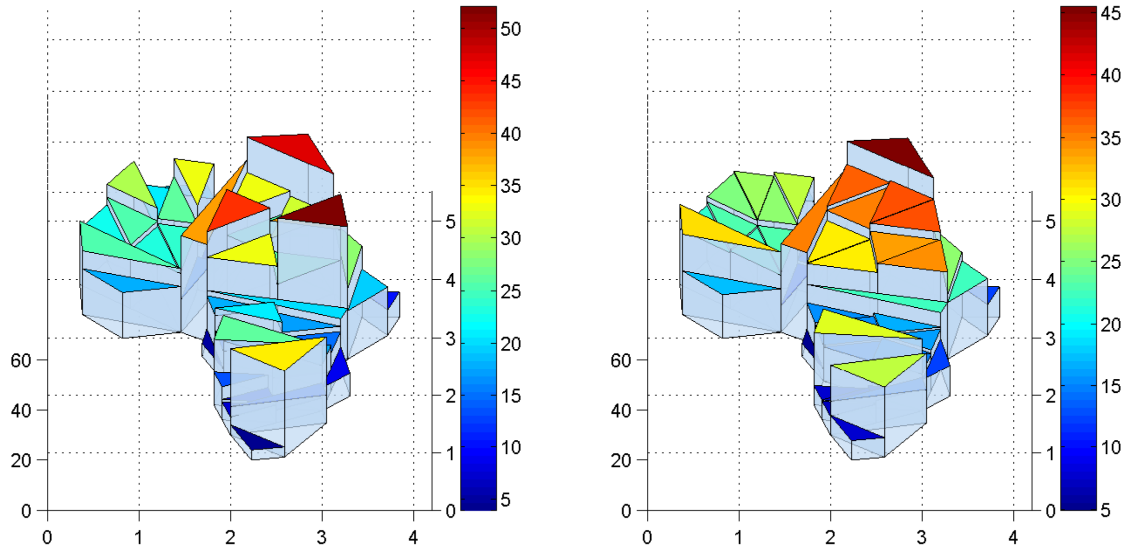
Şekil 37. Üçgenleme, (a) Çokgensel bölgenin üçgenlenmesi; (b) Çokgen dışındaki üçgenlerin silinmesi



(a) (b)

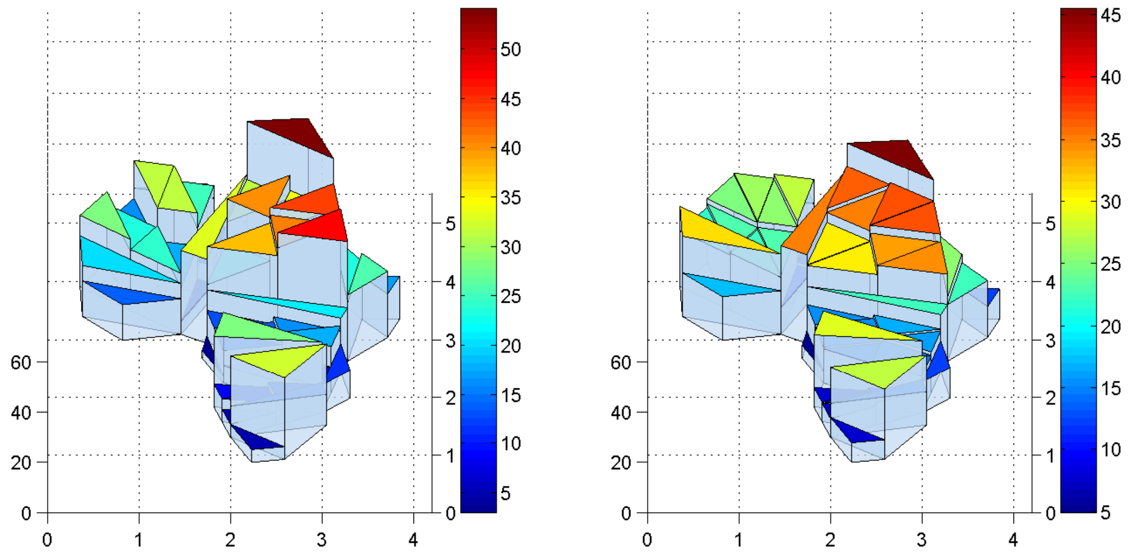
Şekil 38. Çokgensel bölgede beklenen frekanslar hesabı; (a) Çokgensel bölgenin küçük üçgenlere bölünmesi; (b) Her bir küçük üçgenin beklenen frekansları

Her tekniğe göre üretilen rastgele sayıları test etmek için küçük sınıflara bölünerek gözlenen frekansları bulunmaktadır (Şekil 39-41).



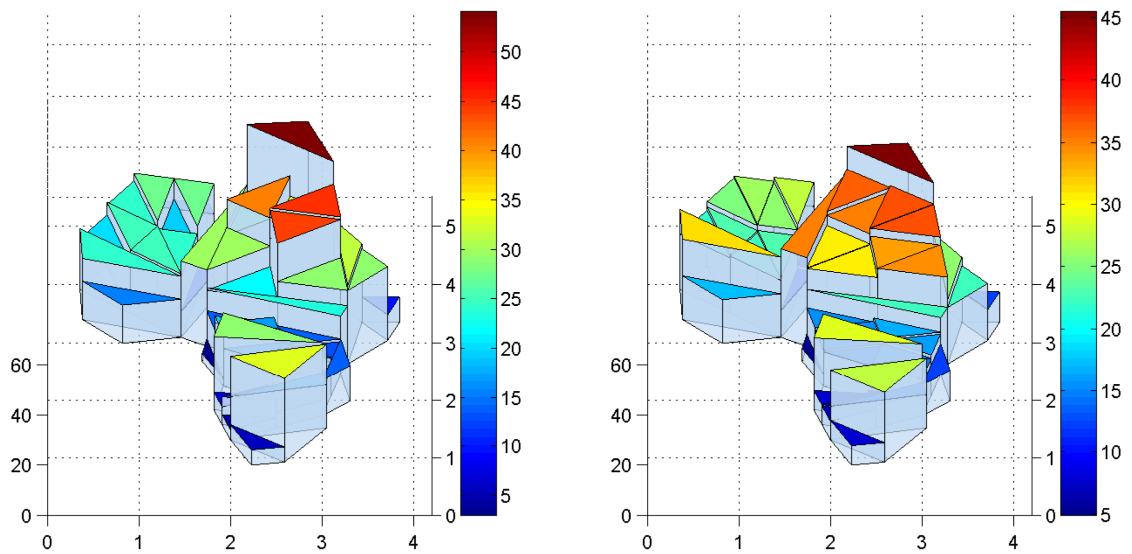
(a) (b)

Şekil 39. Gözlenen frekanslar; (a) Ret-kabul yöntemiyle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları



(a) (b)

Şekil 40. Gözlenen frekanslar; (a) Ters dönüşüm yöntemiyle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları



(a) (b)

Şekil 41. Gözlenen frekanslar; (a) Karma yöntemle üretilen rastgele sayıların gözlenen frekansları; (b) Her bir küçük üçgenin beklenen frekansları

Her bir yöntem için gözlenen ve beklenen frekanslar kullanılarak 10000 defa ki-kare testi yapılmıştır. Denemeler sonucunda elde edilen sonuçlar özet halinde Tablo 8'de gösterilmiştir.

Tablo 8. N=10000 defa yapılan denemeler sonucunda elde edilen sonuçlar

Yöntem	n=100		n=1000	
	HKY(χ^2)	OHZ(sn)	HKY(χ^2)	OHZ (sn)
Ret-Kabul	0.9518	0.3702	0.9490	0.7388
Ters Dönüşüm	0.9487	0.0624	0.9512	0.1235
Karma	0.9538	0.1969	0.9524	0.3936

*(HKY: Hipotezin kabul yüzdesi; OHZ: Ortalama hesaplama zamanı)

H_0: Üretilen rastgele sayılar Eşitlik 69'da verilen ortak olasılık yoğunluk fonksiyonundan gelmiştir.

H_1: Üretilen rastgele sayılar Eşitlik 69'da verilen ortak olasılık yoğunluk fonksiyonundan gelmemiştir.

Ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayılar %95 güven düzeyinde test edilmiştir. Bu üç yöntemde H_0 hipotezi iki örnek hacmi için de yaklaşık olarak %95 oranında kabul edilmiştir. Serbestlik derecesi sınıf sayısının bir eksiği 47 olarak alınmıştır. Diğer yandan, ters dönüşüm yönteminin, ortalama hesaplama zamanı bakımından diğer iki yöntemden daha hızlı çalıştığı görülmüştür.

3. BULGULAR VE SONUÇLAR

İki değişkenli dağılım fonksiyonlarından geliştirilen algoritmalar yardımıyla çeşitli geometrik alanlarda rastgele sayı üretimi için iki ayrı örnek belirlenmiştir. Bu örneklerden bir tanesinde üçgensel bir bölgede rastgele sayı üretimi ele alınırken, diğerinde Afrika ana kıtasının şekli ele alınmıştır. Her iki örnek için de tanım aralıkları, üçgen ve Afrika ana kıtası şekli ile sınırlandırılmış olan aynı olasılık yoğunluk fonksiyonu kullanılmıştır. Bu olasılık yoğunluk fonksiyonundan 100 ve 1000 tane olmak üzere iki farklı hacimde rastgele sayı, ret-kabul, ters dönüşüm ve karma yöntemler kullanılarak üretilmiştir.

Yapılan bu işlem 10000 kez tekrarlanarak bir simülasyon gerçekleştirilmiştir. Simülasyonda ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayıların istenilen dağılımdan gelip gelmediği ve rastgele sayı üretme yöntemlerinden hangisinin daha hızlı çalıştığının bulunması amaçlanmıştır. Üretilen rastgele sayıların verilen olasılık yoğunluk fonksiyonundan gelip gelmediğinin testi için ki-kare uyum iyiliği testi kullanılmıştır.

İlk örnekte tanım aralığı üçgen alanla sınırlandırılmış iki değişkenli keyfi bir olasılık yoğunluk fonksiyonundan, 100 ve 1000 tane olmak üzere iki farklı hacimde rastgele sayı 10000 kez üretilmiştir. Ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayılar %95 güven düzeyinde test edilmiştir. Bu üç yöntemde, H_0 hipotezi 10000 denemede iki örnek hacmi için de yaklaşık olarak %95 oranında kabul edilmiştir. Anlamlılık düzeyi $\alpha = 0.05$ olarak alındığından, elde edilen bu kabul yüzdesi üçgensel alanda rastgele sayı üretme algoritmasının her bir yöntem için doğru bir şekilde çalıştığını göstermektedir. Diğer yandan, karma yöntemin, ortalama hesaplama zamanı bakımından diğer iki yöntemden daha hızlı çalıştığı görülmüştür.

İkinci örnekte tanım aralığı Afrika ana kıtası şekli ile sınırlandırılmış iki değişkenli keyfi bir olasılık yoğunluk fonksiyonundan, 100 ve 1000 tane olmak üzere iki farklı hacimde rastgele sayı 10000 kez üretilmiştir. Ret-kabul, ters dönüşüm ve karma yöntemle üretilen rastgele sayılar %95 güven düzeyinde test edilmiştir. Bu üç yöntemde, H_0 hipotezi 10000 denemede iki örnek hacmi için de yaklaşık olarak %95 oranında kabul edilmiştir. Anlamlılık düzeyi $\alpha = 0.05$ olarak alındığından, elde edilen bu kabul yüzdesi çokgensel alanda rastgele sayı üretme algoritmasının her bir yöntem için doğru bir şekilde çalıştığını göstermektedir. Diğer yandan, ters dönüşüm yönteminin, ortalama hesaplama zamanı bakımından diğer iki yöntemden daha hızlı çalıştığı görülmüştür.

Sonuç olarak, bu çalışma ile ret-kabul, ters dönüşüm ve karma yöntemlerle tanım aralığı keyfi bir çokgensel alanla sınırlandırılmış iki değişkenli keyfi bir olasılık yoğunluk fonksiyonundan sözde rastgele sayılar, geliştirilen algoritmalar sonucunda basit, anlaşılır, tutarlı bir şekilde üretilmiştir.

4. ÖNERİLER

- Tanım aralığı keyfi bir çokgensel alanla sınırlandırılmış iki değişkenli bir olasılık yoğunluk fonksiyonundan istatistiksel simülasyonların gerçekleştirilmesi birçok pratik probleme çözüm oluşturabilir. Örneğin deprem görülme riski yüksek olan bölgelerden daha yüksek olasılıkla deprem üretmek.
- İki değişkenli olasılık fonksiyonlarının gerçek hayattaki uygulama alanları üzerine model kurularak istatistiksel olarak çıkarım ve analiz yapılabilir. Bu uygulama alanlarına deprem ve yangın bölgeleri simülasyonları, suç ve trafik yoğunluğu kestirimi, orman bölgeleri modellemesi örnek olarak verilebilir.
- Keyfi bir çokgensel alan içerisinde iki boyutlu konum tabanlı istatistiksel çalışmaların artmasını sağlanabilir.
- Keyfi bir çokgensel alan içerisinde veriyi en iyi temsil eden iki değişkenli olasılık yoğunluk fonksiyonunun belirlenmesi için bir çalışma yapılabilir.

5. KAYNAKLAR

1. Dutang, C. ve Wuertz, D., A Note on Random Number Generation, 2009.

2. Warnock, T., Random Number Generators,1987.

3. Rubinstein, R. Y., Simulation and the Monte Carlo method. John Wiley & Sons, 1981.

4. Page, E., Pseudo Random Elements for Computers Appl. Stat., 8, 2 (1959) 124–131.

5. Niederreiter, H., Quasi-Monte Carlo methods and pseudo-random numbers Bull. Am. Math. Soc., 84, 6 (1978) 957–1041.

6. Niederreiter, H., Random number generation and Quasi Monte Carlo methods. Philadelphia: Society for Industrial and Applied Mathematics, 1992.

7. http://www.mathworks.com/help/stats/common-generation-methods.html 29 Eylül 2014.

8. Sheldon M.Ross, Simulation. Elsevier, 2006.

9. Asmussen, S. ve W. Glynn, P., Stochastic Simulation: Algorithms and Analysis, 57. New York, NY: Springer New York, 2007.

10. Akdeniz, F., Olasılık ve İstatistik. Adana: Nobel Kitabevi, 2007.

11. Esin, A., Ekni, M. ve Gamgam, H., İstatistik. Ankara: Gazi Kitapevi, 2006.

12. http://www.efm.bris.ac.uk/ecdw/e12122/stats3.pdf, 4 Kasım 2014.

13. Robert, C. P. ve Casella, G., Monte Carlo Statistical Methods. New York, NY: Springer New York, 1999.

14. Neumann, J. Von, Various techniques used in connection with random digits Appl. Math Ser., (1951).

15. Neal, R. M., Slice sampling Ann. Stat., 31, 3 (2003) 705–767.

16. Bishop ve Cristopher, M., Pattern Recognition and Machine Learning. New York: Springer, 2006.

17. https://noppa.aalto.fi/noppa/kurssi/mat-1.3626/luennot/Mat-1_3626_lecture12.pdf, 4 Kasım 2014.

18. Devroye, L., Non-Uniform Random Variate Generation. New York, NY: Springer New York, 1986.

19. Walck, C., Hand Book on Statistical Distribution for experimentalists, Stockholm, 2007.

20. Delaunay, B., Sur la sphere vide Izv. Akad. Nauk SSSR, Otd. Mat. i …, (1934).

21. Yanalak, M., Sayısal Arazi Modellerinden Hacim Hesaplarında En Uygun Enterplasyon Yönteminin Araştırılması, İstanbul Teknik Üniversitesi, 1997.

22. Worboys, M. F. ve Duckham, M., GIS: a Computing Perspective. CRC press, 2004.

23. Gudmundsson, J., Haverkort, H. J. ve van Kreveld, M., Constrained higher order Delaunay triangulations Comput. Geom., 30, 3 (2005) 271–277.

24. http://www.hgk.msb.gov.tr/dergi/makaleler/126_6.pdf, 4 Kasım 2014.

25. Chakraborti, S. ve Gibbons, J. D., Nonparametric Statistical Inference, Fourth., 15, 2. Alabama, USA: Marcel Dekker Inc., 2003 421.

26. Ünver, Ö., Gamgam, H. ve Altunkaynak, B., Temel İstatistik Yöntemler. Ankara: Seçkin, 2011.

27. Pearson, K., On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling Philos. Mag. Ser. 5, 50, 302 (1900) 157–175.

28. Kesemen, O. ve Uluyurt, T., Bivariate Chi-Square Goodness Of Fit Test In Polygonal Areas, 8. International Statistics Congress, 2013.

29. Kesemen, O. ve Doğru, F. Z., Cumulative Distribution Functions of Two Variable In Polygonal Areas, 7. International Statistics Congress, 2011.

30. Burden, R. L. ve Faires, J. D., Numerical Analysis. Brooks/Cole, 2011 888.

31. https://www.math.ust.hk/~mamu/courses/231/Slides/ch02_1.pdf, 7 Kasım 2014.

32. Massey, F. J., The Kolmogorov-Smirnov Test for Goodness of Fit J. Am. Stat. Assoc., 46, 253 (1951) 68–78.

33. Wu, J.-S. ve Leou, J.-J., New polygonal approximation schemes for object shape representation Pattern Recognit., 26, 4 (1993) 471–484.

34. Ray, B. K. ve Ray, K. S., A new approach to polygonal approximation Pattern Recognit. Lett., 12, 4 (1991) 229–234.

35. Leu, J.-G. ve Chen, L., Polygonal approximation of 2-D shapes through boundary merging Pattern Recognit. Lett., 7, 4 (1988) 231–238.

36. Teh, C.-H. ve Chin, R. T., On the detection of dominant points on digital curves IEEE Trans. Pattern Anal. Mach. Intell., 11, 8 (1989) 859–872.

37. Güngör, Ç., İkil Görüntülerde Nesnelerin Baskın Noktalarının Belirlenmesinde Optimizasyon Algoritmaları, Karadeniz Teknik Üniversitesi, 2012.

ÖZGEÇMİŞ

Buğra Kaan Tiryaki, 3 Ocak 1991 tarihinde Seyhan‘da doğdu. Lisans öğrenimini Gazi Üniversitesinde, Fen Fakültesi, İstatistik Bölümünü 2012 yılında üçüncülükle tamamladı. 2013 yılında Karadeniz Teknik Üniversitesi, Fen Bilimler Enstitüsü, İstatistik ve Bilgisayar Anabilim dalında tezli yüksek lisans programına başladı. 2013 yılından beri Karadeniz Teknik Üniversitesi İstatistik ve Bilgisayar Bilimleri Bölümü’nde Araştırma Görevlisi olarak görev yapmaktadır.