KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

BULANIK YERİNE KOYMA İLE RASTGELE ÖRNEKLEME

YÜKSEK LİSANS TEZİ

Ebru NAZ

Ocak 2018
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İSTATİSTİK VE BİLGİSAYAR BİLİMLERİ ANABİLİM DALI

BULANIK YERİNE KOYMA İLE RASTGELE ÖRNEKLEME

Ebru NAZ

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

Tezin Enstitüye Verildiği Tarih : 02 / 01 / 2018
Tezin Savunma Tarihi : 18 / 01 / 2018

Tez Danışmanı : Yrd. Doç. Dr. Orhan KESEMEN

Trabzon 2018

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İstatistik ve Bilgisayar Bilimleri Anabilim Dalında
Ebru NAZ Tarafından Hazırlanan

BULANIK YERİNE KOYMA İLE RASTGELE ÖRNEKLEME

başlıklı bu çalışma, Enstitü Yönetim Kurulunun 02 / 01 / 2018 gün ve 1734 sayılı
kararıyla oluşturulan jüri tarafından yapılan sınavda
YÜKSEK LİSANS TEZİ
olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Prof.Dr. İhsan ÜNVER ....

Üye : Yrd.Doç.Dr. Orhan KESEMEN ....

Üye : Yrd.Doç.Dr. Tolga BERBER ....

Prof. Dr. Sadettin KORKMAZ
Enstitü Müdürü

ÖNSÖZ

Bu tez, Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik ve Bilgisayar Bilimleri Anabilim Dalı, Yüksek Lisans Programı'nda hazırlanmıştır. Bu tez, "Bulanık Yerine Koyma ile Rastgele Örnekleme" için yapılmıştır. Tez çalışma süresinde değerli yardım ve katkılarıyla beni yönlendiren danışman hocam Yrd. Doç. Dr. Orhan KESEMEN'e teşekkürü bir borç bilirim.

Çalışmalarım boyunca maddi ve manevi desteklerini esirgemeyen anneme, babama ve eşime teşekkür ederim. Bu tezin, bundan sonraki çalışmalara katkı sağlamasını temenni ederim.

Ebru NAZ

Trabzon 2018

TEZ BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum "Bulanık Yerine Koyma ile Rastgele Örnekleme" 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. 02/01/2018

Ebru NAZ

İÇİNDEKİLER

Yüksek Lisans

ÖZET

BULANIK YERİNE KOYMA İLE RASTGELE ÖRNEKLEME

Ebru NAZ

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik ve Bilgisayar Bilimleri Anabilim Dalı
Danışman: Yrd. Doç. Dr. Orhan KESEMEN
2018, 63 Sayfa

Gerçek hayata birçok alanda kullanılan ve günlük yaşamın bir parçası haline gelen örnekleme yöntemleri istatistiksel çıkarımın ilk adımıdır. N büyüklüğündeki bir kitleden n büyüklüğünde alınan bir alt küme aracılığı ile kitle hakkında bilgi edinebilmek çıkarımsal istatistiğin ana amacıdır.

İstatistiksel örneklemede yerine koyma veya yerine koymama gibi Aristo'nun {doğru (1), yanlış (0)} mantık yaklaşımlıyla iki farklı teknik olarak kullanılmaktadır. Yerine koyma yaklaşımıyla kitleden çekilen her birimin süreç boyunca seçilme olasılığı korunurken, yerine koymama yaklaşımında çekilen her birimin seçilme olasılığı sıfırlanmakta ve dolayısıyla diğer birimlerin seçilme olasılıkları artmaktadır. Gerçek hayatta bilinçaltı etkisiyle yapılan birçok örnekleme Zadeh'in [0,1] yaklaşımı ile gerçekleştirilmektedir. Bu çalışmada, kitleden çekilen bir birimin çekilme olasılığı bulanıklaştırma parametresi yardımıyla düşürülerek yerine koyulmasıyla yerine koyma veya koymama aralığında bir tercih yapılmış olur. [0,1] aralığında değerler alan bulanıklaştırma parametresi 0 olduğunda yerine koymama, 1 olduğunda ise yerine koyma mantığı ile çalışmaktadır. Bulanık yerine koyma ile yapılan örnekleme yöntemlerinin başarımını ortaya koymak için kitle parametrelerin tahmini için benzetim teknikleri kullanılmıştır.

Anahtar Kelimeler: Örnekleme Yöntemleri, Rastgele Örnekleme, Yerine Koyarak Örnekleme, Yerine Koymadan Örnekleme, Bulanık Yerine Koyarak Örnekleme.

Master Thesis

SUMMARY

RANDOM SAMPLING WITH FUZZY REPLACEMENT

Ebru NAZ

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Statistical and Computer Science Graduate Program
Supervisor: Assist. Prof. Dr. Orhan KESEMEN
2018, 63 Pages

Sampling methods which is used in real life in many areas and become a part of daily life is the first step used in statistical inference. The size of the N received from a population through size of n a subset to learn about population inferential statistics is the main objective.

Statistical sampling is used as two different techniques approach rationale of Aristotle's {true (1), false (0)} like with replacement or without replacement. While the probability of selecting from population each unit is preserved approach with replacement throughout the process, the probability of selecting each unit in the approach without replacement is zero and therefore the probability of selecting other units is increasing. Many sampling that has performed subconscious effect in real life is carried out by approach with Zadeh's [0, 1]. In this study, the probability of selecting of a sample from the population is reduced through the help of the fuzzy parameter prefer is made between with replacement or without replacement. Values in the range [0, 1] which is the fuzzy parameter is work rational that zero is without replacement, one is with replacement. Simulation techniques have been used to estimate population parameters in order to demonstrate the success of the fuzzy replacement sampling methods.

Key Words: Sampling Methods, Random Sampling, Sampling with Replacement, Sampling without Replacement, Sampling with Fuzzy Replacement

ŞEKİLLER DİZİNİ

Sayfa No

TABLOLAR DİZİNİ

Sayfa No

ALGORİTMALAR DİZİNİ

SEMBOLLER DİZİNİ

N	Kitle hacmi (büyüklüğü)
μ	Kitle ortalaması
σ^2	Kitle varyansı
$\bar{x}$	Örneklem ortalaması
n	Örneklem hacmi (büyüklüğü)
S	Kitleyi temsil eden küme
S’	Karıştırılmış kitleyi temsil eden küme
s	Örneklemi temsil eden küme
p_i	Kitledeki i-nci birimin olasılığı
p_i'	Bulanık örneklemedeki i-nci ağırlıklı olasılık
$\hat{p}_i$	Bulanık örneklemedeki i-nci göreceli olasılık
x_i	Kitledeki i-nci birimin değeri
X_i	Örnekteki i-nci birimin değeri
U	Düzgün dağılımdan rastgele değişken
Uniform	Düzgün dağılım
m	Bulanıklaştırma parametresi
w_i	Bulanıklaştırma katsayısı

1. GENEL BİLGİLER

İstatistikte iki önemli ana dal vardır. Bunlar tanımlı (betimsel) ve çıkarımsal istatistik olarak bilinmektedir. Örnekleme kavramı, istatistiksel çıkarımların yapılması için kullanılan ilk tekniktir. Sonlu sayıda birime sahip kitleler hakkında bilgi sahibi olmak, bir araştırmacının en önemli amaçları arasında yer almaktadır. Bu bilgiyi elde etmek için yapılacak en sağlam yöntem tam sayımdır. Tam sayım, kitledeki her birime ulaşıp o birim hakkında istenilen bilginin temin edilmesi anlamına gelmektedir. Ancak gerçek hayatta bu her zaman mümkün olamamaktadır. Dolayısıyla bu sorunu çözmek için tüm kitle yerine kitleden küçük bir grup seçilerek tüm kitle hakkında genelleme yoluna gidilmektedir. Bu işleme örnekleme, bu işlem sonucunda elde edilen alt gruba ise örneklem denir. Diğer bir deyişle örnekleme, N birimden oluşan bir kitleden n birimden bir alt grubun (örneklemin) çeşitli teknikler yardımıyla seçilmesidir.

Günümüzde kullanılan gerçek anlamadaki örnekleme kavramı ve kestirim yöntemleri yirminci yüzyılda ortaya kondu. Yirminci yüzyılın ilk çeyreğinde basit rastgele örneklemenin temel yöntemleri ve formülleri çalışıldı. Neyman (1934) tarafından yapılan bir çalışmada olasılıksal örneklemenin matematiksel temeli ortaya konurken bir birimin bilinen bir dağılımdan rastgele seçildiği düşünülmektedir (Thompson, 2012).

Tabakalı örnekleme, sistematik örnekleme, küme örneklemesi ve çok aşamalı örnekleme gibi standart örnekleme yöntemleri 1930'ların sonuna doğru tanımlandı. Eşit olmayan olasılıktaki örnekleme yöntemleri ise 1940'lı ve 1950'li yıllarda çalışıldı.

Hansen ve Hurwitz (1943) yaptığı çalışmada yerine koyarak yapılan basit örnekleme yönteminde, değişken değerlerini kitle değerine oranlayarak oran tahmin edicilerini incelemiş ve yansız tahmin ediciler elde etmiştir.

Horvitz ve Thompson (1952) yaptığı çalışmada yerin koymadan yapılan herhangi bir örnekleme yöntemi ile yansız tahmin ediciler elde etmiştir. Sen (1953) ve Yates ile Grundy (1953) yerine koymadan örnekleme yöntemlerinde elde edilen tahmin ediciler ile kovaryans ve korelasyon katsayıları incelenerek varyansı minimum kılan bir varyans formülü bulmuşlardır. Bu varyans formülüne göre Horvitz & Thompson tahmin edicileri ile sıfır varyans elde edilebilen yerine koymadan örnekleme modelleri oluşturulmuştur.

Yerine koyularak ya da yerine koyulmadan yapılan örnekleme yöntemlerinde birçok bakış açısı geliştirilmiştir. Madow (1949) sistematik örnekleme üzerinde teoriler geliştirmiştir. Sınıflanabilen ve sınıflanamayan kitleler için sistematik örneklemenin uygulanabileceğini göstermiş ve tahmin edicilerini incelemiştir. Hartley ve Rao (1962), Fellegi (1963), Grundy (1954) de olasılıksal olan yaklaşımlar geliştirmeye çalışmıştır.

Günümüzde internet kullanımının artmasıyla birlikte büyük veri gruplarının elde edilmesi, depolanması ve çözümlenmesi kolaylaşmıştır. Büyük veri kitleleri hakkında bilgi çıkarımı yapmak bazen büyük maliyetlere ve zaman kaybına neden olmaktadır. Bu sorunları en aza indirebilmek için kitleyi temsil edebilecek örneklemin belirlenmesi gerekir. Ancak örneklemin belirlenmesinde ihtiyaca göre seçim yapmak gerekebilir. Bu seçimler göz önünde bulundurularak geliştirilen birçok örnekleme tekniği bulunmaktadır. Bu teknikler olasılıksal veya olasılıksal olmayan teknikler biçiminde iki grupta incelenebilir (Cochran, 2007). Bu gruplarda verilen tekniklerin dışında da birçok örnekleme tekniği bulunmaktadır. Ancak bu çalışmada en yaygın tekniklerden söz edilmiştir.

Örnekleme yöntemleri; birimlerin seçilme olasılıklarına göre, örneklemlerin içerdikleri birim sayısına göre ya da her bir birimin örnekleme alınmış adım sayısına göre sınıflandırılabilir. Örnekleme süreci çekim ve tahmin olmak üzere iki kesime ayrılır. Amaç araştırmalarda iyi bir örneklem ile yansız, tutarlı ve duyarlı tahminler yapabilmektir. İyi bir örneklem seçilebilmesi için doğru ve iyi bir örnekleme yöntemi seçmek gerekir (Çıngı, 1990).

İyi bir örneklemin genel karakteristiği, yeniden elde edilebilirliği, ulaşmadaki kolaylık ve düşük maliyetli olması sayılabilir.

1.1. Olasılıksal Örnekleme Yöntemleri

Olasılıksal örnekleme, kitledeki her birimin bir seçilme şansı olduğu seçme sürecidir. Bu yöntem ile çoğu araştırmanın sosyal önyargıların etkisinden kurtarılmasını sağlayan duyarlı bir yöntem olarak bilinmektedir. Sonuç olarak seçtiğimiz örnekleme tekniği, araştırmanın verilen soruna en iyi cevap verebilmesi beklenmektedir.

Öte yandan birimlerin çekilme olasılıkları her bir birim için eşit olabilir ya da olmayabilir. Her bir durumda farklı yöntemler kullanılır. Eşit olasılık ile seçim yapılır ise seçilen her bir birim bir sonraki seçim sürecinde kitleye geri konularak seçilir ya da yerine konmadan seçilir. Kitledeki her birimin seçilme olasılıklarının bilinmesi kitle dağılımı hakkında bilgi vermektedir. Olasılıksal örnekleme yöntemleri olarak Basit rastgele örnekleme, Sistematik örnekleme, Tabakalı örnekleme ve Kümeli örnekleme yöntemleri öne çıkmaktadır.

1.1.1. Basit Rastgele Örnekleme

Olasılıksal örnekleme yöntemleri içerisinde en çok kullanılan, kolay bir örnekleme yöntemidir. Kitledeki her birimin seçilme şansı eşit ve bilinmelidir (Olken ve Rotem, 1986; Tillé, 2006; Kılıç, 2013). Bu seçilme şansı (olasılığı) ile yerine konularak ya da konulmayarak seçilen birimlerden oluşturulan örnekleme yöntemidir. Bu yöntemin gerçekleştirilmesi kitle hacmi çok büyük olduğunda zor veya imkansız olabilmektedir. Basit rastgele örnekleme sürecini gerçekleştirirken bilgisayar ortamında sözde rastgele sayı üreteçleri kullanılarak yapılabildiği gibi müsabaka seçmelerinde sayışarak; yazı-tura atarak; zar atarak, torbadan isim çekerek gibi gerçek rastgele sayı üreteçleri yardımıyla yapılabilmektedir. Bu üreteçler dışında gerçek hayatta kendiliğinden gerçekleşen rastgele olaylar yardımıyla da seçilebilmektedir. Rastgele örnekleme ile ilgili geniş açıklama Yapılan Çalışmalar bölümünde verilmektedir.

1.1.2. Sistematik Örnekleme

Sistematik örnekleme sık sık basit rastgele örnekleme yöntemi yerine kullanılmaktadır. Bu yönteme k'ıncı yöntem de denilmektedir. Dizi şeklinde verilen bir kitlede belirlenen ilk k birim içerisinden herhangi bir birim başlangıç noktası olarak kabul edilir. Bundan sonra gelecek olan her k'ıncı birim seçilerek örnekleme dâhil edilir (Madow ve Madow, 1944; Madow, 1949). Bu şekilde yapılan örnekleme yöntemine sistematik örnekleme denir. Sistematik örnekleme yönteminde k değeri,

$$k = \left[\!\left[\frac{N}{n}\right]\!\right] \tag{1}$$

biçiminde bulunabilir. Burada N kitledeki örnek sayısı, n ise istenen örnek sayısını göstermektedir. İlk örneklem,

$$k_0 \sim [0, k] \tag{2}$$

aralığında düzgün dağılımdan veya keyfi seçilebilir. Sonraki örnek indisleri ise,

$$k_i = k_0 + i \cdot k \qquad (3)$$

biçiminde seçilir.

Sistematik örnekleme yöntemi basit rastgele örnekleme yöntemi kadar iyi sonuçlar vermektedir. Bu yöntemle bir nüfus listesinden, bir çalışanlar listesinden ve bilgisayardaki bir kayıt listesinden seçim yapılırken kullanılabilir.

1.1.3. Tabakalı Örnekleme

Tabakalı örnekleme yöntemi, rastgele örneklemeye göre örnekleme hatasını azalttığı için yaygın olarak tercih edilen bir yöntemdir (Imbens ve Lancaster, 1996; Hausman ve Wise, 1981). Her katman kitlenin en az bir özelliğini temsil eden bir altkümesidir. Örnek olarak erkekler ve kadınlar; öğretmenler ve öğrenciler; çalışanlar ve çalışmayanlar; yüksek gelirliler, orta gelirliler ve düşük gelirliler; eğitim seviyesi gibi verilebilir. Bu örneklemde araştırmacı öncelikle seçilen karakteristiğin kitledeki oranlarını belirler. Her tabaka için bu oranlara göre kitleden rastgele örneklem seçilerek örneklem yapılır. Seçilen örneklemin tabaklardaki oranı tam olarak yansıtabilmesi için yeteri büyüklükte örneklem yapılmalıdır.

1.1.4. Kümeli Örnekleme

Kümeli örnekleme, genelde tüm kitleye ulaşmanın imkansız olduğu durumlarda kullanılabilir (Kothari, 2004). Bu yöntem sık sık tabakalı örneklem ile karıştırılmaktadır. Oysa kümeli örnekleme tabakalı örneklemden oldukça farklıdır. Örneğin Türk halkı içerisinden anket yapmak için bir örneklem seçilmek istendiğinde, tüm illerden örneklem almak zor olabilir. Bunun yerine 81 ilden rastgele seçilen 15 il'e gidilerek rastgele örneklemle anket yapılabilir. Kümeli örneklemede kitle önceden kümelere ayrılmış olabildiği gibi sonradan kümelere ayrılabilmesi de mümkündür. Örneğin 16 sınıflı bir lisede aynı anda beden dersine çıkan 3 sınıf içerisinden örnekleme yapılabilir. Bu yönteme aşama açısından bakılırsa önce kümeler seçilir, daha sonra bu kümelerden örnekleme yapılır.

1.2. Olasılıksal Olmayan Örnekleme Yöntemleri

Örnekler seçilirken bir olasılık yaklaşımı kullanılmadan yapılan örnekleme teknikleridir. Seçim işlemi genelde yöneticiler veya araştırmacılar tarafından sübjektif yaklaşımla belirlenmektedir. Bu yöntemlerin seçilmesi önyargılı yaklaşımlara neden olmasına rağmen birçok uygulamada en iyi seçim olduğu gözlenmiştir.

1.2.1. Karar Örneklemesi

Karar örneklemesi bir ön bilgiye veya amaca göre ön yargılı yaklaşımla yapılan örneklemedir (Bowling, 2014). Örneğin Avrupa elemelerinde oynayacak Türk futbol takım oyuncularının belirlenmesi işlemi karar örneklemesine iyi bir örnek olarak verilebilir. Milli takıma seçilecek futbolcular teknik direktör tarafındaki liglerinde gösterdikleri performansa bakılarak seçilmektedir. Diğer bir örnek olarak, bir şirkete alınacak 20 yeni çalışan için verilen ilana 1200 kişi başvurduğunda bir ön eleme yapılması gerekir. Başvuran adaylardan özgeçmişlerinin yanı sıra yapılan mülakatlarda edinilen bilgilere göre seçim yapılabilir. Bu şekilde yapılan seçim rastgele olmayacaktır. Ancak şirketin çıkarları düşünüldüğünde en iyi seçimin karar örneklemesi olduğu görülmektedir.

1.2.2. Kota Örneklemesi

Kota örneklemesi, yapı bakımından tabakalı örneklemeye benzemesine rağmen, tabakalı örnekleme kitledeki oranlar korunurken, kota örneklemesinde oranlar araştırmacı tarafından belirlenmektedir (Moser, 1952). Böylelikle farklı özellikteki birimlerinde örneklemde temsili sağlanmasıdır. Örneğin seçimlerde bayanlara tanınan aday kotası verilebilir. Diğer bir örnek olarak lisans yerleştirmeleri için yapılan sınavlarda yabancıların yerli adaylarla yarışma şansı bulunmamaktadır. Oysa hükümetin stratejik planları doğrultusunda yabancı öğrencilere de lisans yerleştirmelerde şans vermek için kota uygulaması yapılmaktadır.

Kota örneklemesinde her grubun seçilme olasılıkları istenen oranlarda uygulanmasına rağmen, her grup kendi içerisinde eşit olasılıklara sahip olmaktadır. Bu olasılık çerçevesinde kota örneklemesi olasılıksal örnekleme grubuna dahil edilebilir.

1.2.3. Kartopu Örneklemesi

Kartopu örneklemesi, kitle birimlerine tam olarak nasıl ulaşılacağının bilinmediği durumlarda uygulanır (Goodman, 1961). Örnek olarak bir terör örgütü soruşturmasında suçluların kim olduğunun net olmadığı durumda savcının kimleri sorgulayacağı ve suçlayacağı başlangıçta bilinmemektedir. Elde edilen her ipucu sayesinde ulaşılan kişiler sorgulanarak verilen bilgilere göre diğer kişilere ulaşılmaktadır. Böylelikle örgütün diğer üyeleri bulunabilmektedir. Diğer bir örnek ise eskiden akademisyenlerin bir konu hakkında araştırma yaparken izledikleri yol verilebilir. Bir akademisyen bir makalede verilen konu ile ilgili araştırma yaparken o makalede verilen kaynaklara ulaşır. Daha bu kaynaklarda verilen ilgili diğer kaynaklara ulaşılır. Böylelikle konu ile ilgili kaynak araştırması yapmaktadır. Günümüzde bu işlemi arama motorları anahtar kelimeler yardımıyla yapmaktadırlar.

1.2.4. Kolay Örnekleme

Kolay örnekleme, genelde araştırmacıların iş gücü ve maliyeti düşürmek için ulaşılabilecek en uygun birimlerin seçilmesi işlemdir (Sedgwick, 2013). Örneğin bir araştırmacının bulunduğu şehirdeki insanların sokak hayvanları hakkındaki görüşlerini araştırdığını düşünelim. Bu araştırmacı tüm şehri dolaşmak yerine bulunduğu mahalledeki insanlar ile veya akrabaları ile görüşerek bir araştırma yapması buna örnek verilebilir. Örneğin bir öğretmenin kendi öğrencilerini kullanarak yaptığı araştırma gibi verilebilir.

1.3. Yerine Koyma veya Yerine Koymama Örneklemesi

Bir örnek yerine koyma veya koymama gibi iki yolla çekilir. Yerine koyma, kitledeki bir birimin çekilip örnekleme dahil edildikten sonra tekrar kitleye konmasıdır. Böylelikle aynı birimin tekrar seçilebilmesine imkan tanır. Yerin koymama ise kitleden seçilen birimin örnekleme eklendikten sonra tekrar geri kitleye konmamaktadır. Böylelikle aynı örnek tekrar seçilememektedir.

1.4. Parametre ve Örneklem Tahmini

Bir kitleyi temsil eden en önemli özelliklerden birisi kitlenin parametresidir. Genelde örneklemler alınan kitlenin birçok parametresi bilinmez. Amaç bu parametreleri tahmin etmek için alınan örneklemden yararlanmaktır. Dolayısıyla daha tutarlı bir tahmin yapılabilmesi için iyi bir örneklem seçilmesi araştırmacı için oldukça önemlidir. Parametre kitlenin sayısal olarak ölçülebilen herhangi bir özelliğidir. Parametre kitle hakkında önemli bilgiler verir. Her parametre tek bir değerdir ve genellikle Yunan harfleri ile gösterilir.

Örneklem tahmini, örneklemden faydalanarak kitle özelliklerini tahmin etmek amacıyla tanımlanan matematiksel eşitliktir. Doğrudan doğruya kitle özelliklerine ilişkin tahmin edicilere birinci derece tahmin ediciler denir. Örneğin, kitle ortalaması, kitle toplamı, belli özelliğe sahip birimler oranı, iki değişkenin birbirine oranı gibi tahmin ediciler birinci dereceden tahmin edicilerdir. Bu tahmin edicilerin varyansları, standart hataları tahmin edicilerine ise ikinci dereceden tahmin ediciler denilmektedir. İyi bir tahmin edicinin özellikleri tutarlı, yansız, duyarlı, değişim katsayısının küçük olması ve hata kareler ortalamasının küçük olmasıdır. Parametre ile tahmin edici arasındaki ilişkiyi, tahmin ediciler yardımı ile örneklemin parametre tahmini yapılabilir şeklinde söylenebilir.

En yaygın kitle parametreleri olarak kitle ortalaması, toplamı ve varyansı bu çalışmada incelenecektir.

N örnekten oluşan bir kitlenin ortalaması,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{4}$$

biçimindedir. Burada X_i kitledeki örnek değerlerini ve N ise kitledeki toplam örnek sayısını göstermektedir. Aynı kitlenin varyansı ise,

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (X_i - \mu)^2 \tag{5}$$

biçimindedir. N örnekten oluşan bir kitleden n sayıda rastgele seçilen bir örneklemin ortalaması,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

biçimindedir. Örneklem ortalaması kitle ortalamasının en iyi tahmin edicisidir (Yamane, 2001). Aynı örneklemin varyansı ise,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{7}$$

biçimindedir. Aynı şekilde örneklemden elde edilen örneklem varyansı aşağıdaki gösterdiği gibi kitlenin varyansının en iyi tahmin edicisidir (Yamane, 2001).

Örneklemden hesaplanan ortalama değerin beklenen değeri,

$$E\{\bar{x}\} = \mu \tag{8}$$

biçiminde gösterildiği gibi kitlenin ortalamasına eşittir. Bu ortalamanın varyansı ise yerine koyma ve koymama durumuna göre değişkenlik göstermektedir. Yerin koymama yöntemine göre örneklem ortalamasının varyansı,

$$\sigma_{\bar{x}}^2 = \frac{N-n}{N} \cdot \frac{S^2}{n} \tag{9}$$

biçiminde verilirken, yerine koyma yöntemine göre örneklem ortalamasının varyasnı ise,

$$\sigma_{\bar{x}}^2 = \frac{N-1}{N} \cdot \frac{S^2}{n} \tag{10}$$

biçiminde verilmektedir.

2. YAPILAN ÇALIŞMALAR

Örnekleme, gerçek hayatta çok sık kullanılan bir yöntemdir. Örnekleme sadece kitle hakkında bilgi almak için değil aynı zamanda değişik işlerde kullanılmak üzere kitleden bir grubun seçilmesi işlemi için de kullanılmaktadır. Keyfi olasılık değerlerine göre çekilen örnekler bilgisayar oyunlarından, eğitimde soru seçimine, sanayi benzetimlerine kadar birçok uygulama alanı bulunmaktadır. Dolayısıyla çoğu zaman değişen örnekleme yöntemi ön yargıların hakim olduğu bir seçim sürecine doğru kaymaktadır. Bu ise çoğu zaman haksızlıklara ve dolayısıyla tartışmalarla neden olmaktadır. Örnek olarak aşağıdaki sorunları verebilir:

- Bir birimde sabit ücretle çalışan kişilerin aynı miktarda iş yükü verilmesi gerekir. Ancak bazı işletmelerde verilen işler aynı ağırlıkta olmamaktadır. Bu durumda verilen her işe sırayla veya rastgele seçilen bir çalışan farklı ağırlıkta iş yapmakta ancak aynı ücret alması iş yerinde hoşnutsuzluklara neden olmaktadır. Dolayısıyla kolay (hafif) iş verilen bir çalışana başka bir hafif iş verilerek bu denge sağlanabilir. Diğer bir yaklaşım ise ağır iş yapan kişiye sırası geldiğinde sırası atlatılabilir. Öte yandan ağır iş yapan kişiye daha fazla ücret (performans ücreti) gibi bir yaklaşım izlenebilir. Bu durumda diğer çalışanların iş taleplerinde bir haksızlığı neden olabileceği gibi, işletmecinin bünyesinde bulunan atıl durumdaki iş gücü yerine ek masraf çıkarılması iş veren açısından istenmeyen bir durum oluşturabilir.
- Önemli spor karşılaşmalarının günümüzde televizyonlardan canlı olarak yayınlanması, karşılaşmaların farklı saatlere konmasına neden olmaktadır. Karşılaşmalar dar bir zaman dilimi yerine geniş bir zaman dilimine yayılarak televizyon izleyicisinin karşılaşmaları canlı izlemesine imkan sağlamaktadır. Eleme sistemine göre düzenlenen bir spor karşılaşmasında dinlenen bir yarışmacı, önceden yarışmış ve dinlenemeyen bir yarışmacı ile karılaştırılması karşılaşmanın adaletsizliğe neden olmaktadır.
- Sabah oturumunda sınava giren öğrenci, öğleden sonra başka bir oturumda sınava girdiğinde, sabah oturumunda sınava girmemiş öğrenci ile aynı kategoride değerlendirilmesi aynı şekilde adaletsizliğe neden olur.

- Elektronik sınav sisteminde, soru bankasında bulunan sorular rastgele seçilerek sınav yapılmaktadır. Ancak bazı alanlarda soru üretme sayısı düşük olduğundan bu alanlarla ilgili sorular sorulmasına rağmen soru bankasından çıkarılmaz. Ancak bu soruların üst üste çıkmasının da engellenmesi gerekir.
- Çoktan seçmeli sınavlarda soruyla birlikte sorunun cevabı olabilecek seçenekler de verilir (http://oguzcetin.gen.tr/coktan-secmeli-dersler.html). Bu seçeneklerden bir tanesi doğru olmak zorundadır. Soru hazırlanırken doğru cevabın bulunacağı şıklar genelde soru sayısına göre eşit seçilmektedir. Örneğin, 5 seçenekli bir soru düzeninde 100 soru sorulacaksa cevapların 20 tanesinin A şıkkı olması istenir. Burada dikkat edilmesi gereken nokta 20 sorunun cevabının A şıkkı olmasının beklenmesi mi, yoksa tam olarak 20 sorunun cevabının A şıkkı olmalı mıdır? Öte yandan bir şıkkın arka arkaya fazla miktarda gelmesi öğrencinin cevaplar konusunda şüpheye düşmesine neden olmaktadır. Bu problemin çözülerek fazla miktarda aynı cevabın arka arkaya gelmesi önlenebilir.
- Bir basketbol maçında her takımın sahada 5 ve toplamda 12 oyuncusu bulunmaktadır. Bu oyuncular zamanla yer değiştirmektedirler. Bir teknik direktör başlangıçta 12 kişilik bir takımdan 5 oyuncuyu oynamak için saha sürer. Her takımın sınırsız oyuncu değiştirme hakkı vardır. Yedeğe alınan bir oyuncu yerine yedekte bekleyen diğer bir oyuncuyu saha sürer. Bu işlem yapılırken yeni yedeğe alınmış oyuncu dinlenmeden oyuna geri alınması mantıklı bir davranış sayılmaz dolasıyla yedeklerde bulunan her oyuncunun dinlenme, konum ve performansına göre tekrar oyuna seçilme olasılıkları ön plana çıkar. Bu olasılık duruma göre sürekli değişkenlik gösterebilir.
- İngilizce kelime ezberleten bir yazılım yazıldığını düşünelim. Her kelime tek tek ezberletilerek devam edilmektedir. Ezberlenen her kelime daha sonra tekrar gündeme getirilerek kelimenin pekişmesi sağlanır. Bu işlemde kelimler kitlesinde (sözlük) seçilen her kelime eşit olasılıkla seçilmesine rağmen tekrar sözcük havuzuna konur. Ancak hemen tekrar çıkması engellemek için seçilme olasılığı düşürülür. Bu olasılık değeri her seçim işlemiyle giderek artar ve daha sonraki seçimlerde tekrar çıkma olasılığı artarak kelimenin pekişmesi sağlanır.

2.1. Rastgele Örnekleme Algoritmaları

Birçok alanda yaygın bir şekilde kullanılan rastgele örnekleme iki farklı biçimde kullanılmaktadır. Bunlar yerine koyarak veya yerine koymayarak şeklinde sınıflandırılır. Rastgele örneklemenin basit rastgele örneklemden farklı olarak farklı olasılıklara sahip birimlerin oluşturduğu kitlelerden de örneklem çekebilmesidir.

2.1.1. Yerine Koyarak Örnekleme

Bir torbadan çekilen topun tekrar torbaya konarak yeniden bir top çekilmesi veya bir desteden çekilen kartın tekrar desteye konarak yeniden bir kart çekilmesi gibi yapılan her örnek çekimi ile genişleyen örneklem havuzundaki n birimin oluşturduğu kümeye yerine koyarak örnekleme denir. N birimden oluşan bir kitleden n birimden oluşan bir örneklem çekildiğinde N^n tane farklı örneklem oluşabilmektedir. Bir kitle $p = \{p_i, i = 1,2, \dots, N\}$ olasılıklarıyla $S = \{x_i, i = 1,2, \dots, N\}$ biçiminde birimlerden oluşsun. Verilen her bir olasılığa göre I indislerini temel alan bir rastgele sayı üretme algoritması istenen örneklemi oluşturacaktır. Eğer örneklemede her bir birim eşit olasılıkla çekiliyorsa Algoritma 1 yardımıyla yerine koyarak örnekleme yapılabilir.

Algoritma 1. Eşit olasılıkla yerine koyma tekniği ile rastgele örneklem seçimi

---

Adım 1. Kitle değerlerini $(x_i,\ i = 1,2, \dots, N)$

Adım 2. Örneklemin başlangıç indisini (j=1) belirle

Adım 3. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,

Adım 4. Çekilen indis değerini aşağıdaki eşitlik yardımıyla, $I = \lceil N \cdot U \rceil$ belirle

Adım 5. $X_j = x_I$ değerini seçilen birim değeri olarak al,

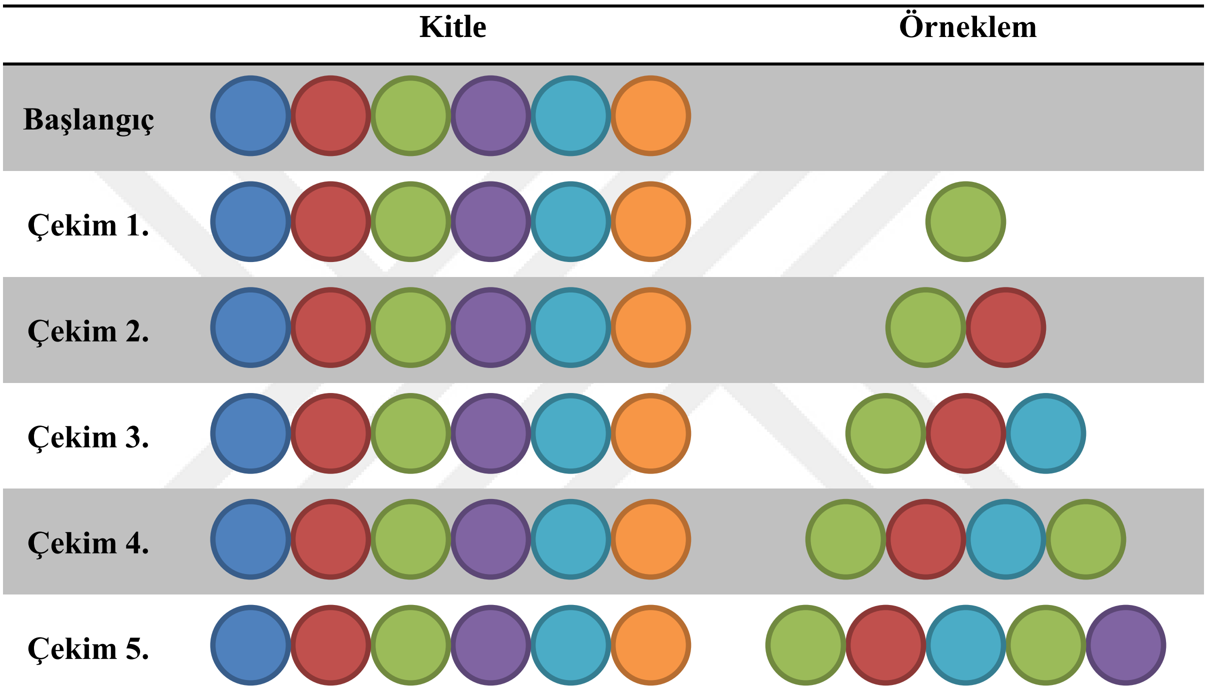
Adım 6. Eğer j=n ise işlemi sonlandır,

Adım 7. Örneklem indisini (j←j+1) biçiminde güncelleştir

Adım 8. Adım 3’e git.

Eşit olasılıkla yerine koyma tekniği ile rastgele örneklem seçimi için her topun sırasıyla çekildiği bir deneyin şematik gösterimi Tablo 1'de verilmiştir. Bu tabloya göre kitlede başlangıçta farklı renklerde 6 top bulunmaktadır. Kitleden rastgele çekilen her top örneklem sütununa eklenmiştir. 5 çekim sonunda örneklemde 5 tane top birikmiştir. Biriken bu toplardan ikisinin aynı renkli (yeşil) top 'tan oluştuğu görülmektedir. Bu ise yerine koyarak örnekleminin bir sonucudur.

Tablo 1. Yerine koyma tekniği ile örneklemenin şematik gösterimi

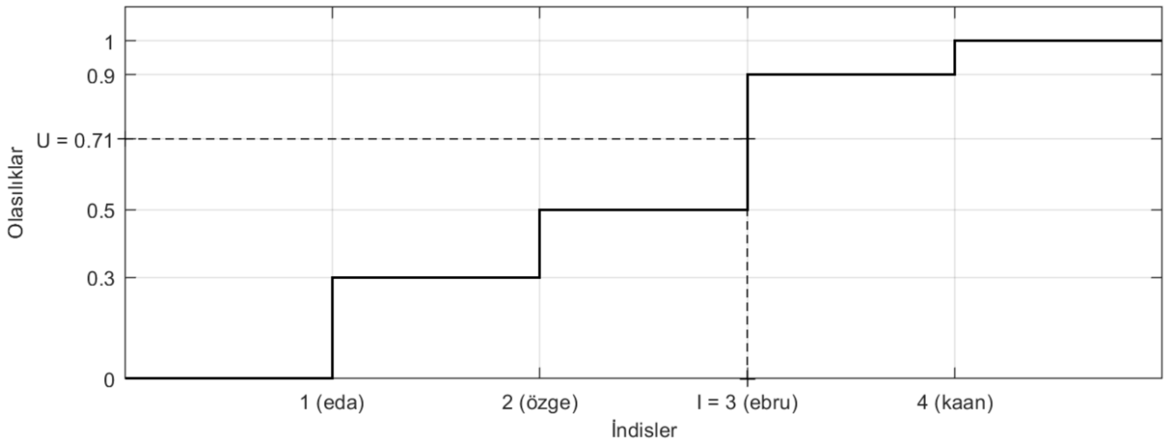


Eğer örneklemede her bir birim keyfi olasılıklarla çekiliyorsa bu durumda yapılan örnekleme sözde rastgele sayı üretme algoritmalarıyla yapılabilir. Algoritma 2'deki rastgele sayı üretme algoritması ters dönüşüm rastgele sayı üretme yönteminden (Devroye, 2013) esinlenerek geliştirilmiştir

Örnek olarak $p = \{0.3, 0.2, 0.4, 0.1\}$ olasılıklarıyla $S = \{eda, özge, ebru, kaan\}$ biçiminde birimlerden oluşan bir kümede düzgün dağılımdan rastgele çekilen değer $U = 0.71$ ise bunun karşılık geldiği indis 3'ü göstermektedir. 3 İndisindeki örnek değeri "ebru" olarak seçilmektedir (Şekil 1). Bu yaklaşımla istendiği kadar kitleden örnek çekilebilir. Bu ise örneklem büyüklüğünü çoğu zaman kitle büyüklüğünden büyük olmasına $(n > N)$ neden olmaktadır. Bu yöntemin diğer özellikleri ise eşit olmayan olasılıklara göre örnekleme yapmanın yanı sıra kitle olarak kesikli dağılım fonksiyonu alarak rastgele sayı üretimi gerçekleştirebilmesidir.

Algoritma 2. Keyfi olasılıklara göre yerine koyma tekniği ile rastgele örneklem seçimi

Adım 1. Kitle değerlerini $(x_i, i = 1,2, \dots, N)$ ve olasılıklarını $(p_i, i = 1,2, \dots, N)$ belirle,
Adım 2. Başlangıç örneklem indisini (j=1) ayarla,
Adım 3. Başlangıç kitle indisini $(I = 1)$ ayarla,
Adım 4. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,
Adım 5. $U < p_I$ ise Adım 9 git,
Adım 6. $U = U - p_I$ güncellemesini yap
Adım 7. İndis değerini $I = I + 1$ biçiminde güncelle,
Adım 8. Adım 4 git
Adım 9. $X_j = x_I$ değerini seçilen birim değeri olarak al.
Adım 10. Eğer j=n ise işlemi sonlandır,
Adım 11. Örneklem indisini (j←j+1) biçiminde güncelleştir,
Adım 12. Adım 3'e git.



Şekil 1. Yerine koyarak rastgele örnek çekmek.

2.1.2. Yerine Koymadan ve Eşit Olasılıklı Örnekleme

Bir torbadan çekilen top yerine konmadan başka bir topun çekilmesi veya bir desteden çekilen kartın tekrar desteye konmadan yeni bir kart çekilmesi gibi tanımlanan örnekleme işlemine yerine koymadan örnekleme denir. N birimden oluşan bir kitleden $n = N$ tane örnek yerine koymadan çekildiğinde $N!$ tane farklı örneklem oluşturulabilmektedir. Eğer $n < N$ biçiminde bir örnekleme gerekirse $N!/(N-n)!$ tane farklı örneklem oluşturulabilir.

Bir kitle eşit olasılıklarıyla $S = \{x_i, i = 1,2, \ldots, N\}$ biçiminde birimlerde olsun. Bu kitleden yerine koymadan örnek çekilmek istenirse birkaç yaklaşım önerilebilir. Birinci yaklaşım veriler rastgele karıştırılır. İstenen sayıda örnek sırasıyla seçilerek örneklem yapılır (Algoritma 3). Bu yaklaşım kitledeki birim sayısı az ise iyi sonuçlar vermektedir. Kitledeki birim sayısı artıkça karıştırma işlem zamanı artacağından istenmeyen sonuçlara neden olur.

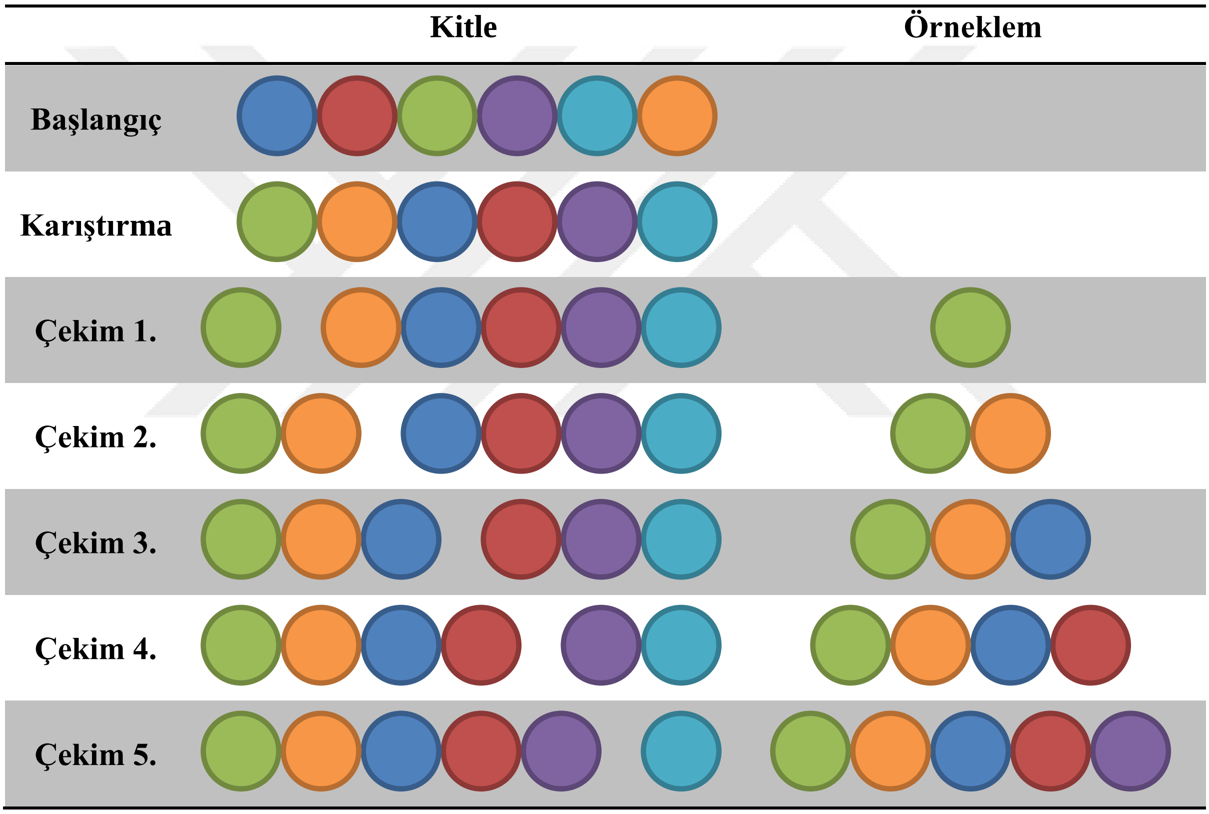
Örnek olarak eşit olasılıklarıyla $S = \{eda, özge, ebru, kaan\}$ biçiminde birimlerden oluşan bir küme Algoritma 3 yardımıyla rastgele karıştırılıyor. Elde edilen yeni dizi $S' = \{özge, ebru, kaan, eda\}$ biçiminde şekilleniyor. Bu diziden $n = 2$ örnekli bir örneklem seçildiğinde $s = \{özge, ebru\}$ örneklemi elde edilir.

Algoritma 3. Karıştırma yaklaşımlı eşit olasılıklı yerine koymama tekniği ile rastgele örneklem seçimi

Adım 1.	Kitle değerlerini ($x_i, i = 1,2, \ldots, N$) belirle
Adım 2.	Güncel indisi (I=1) ayarla,
Adım 3.	Yer değiştirilecek indisi (J~Uniform) [1,N] aralığında düzgün dağılımdan rastgele olarak belirle,
Adım 4.	I ve J-nci örnek değerlerini yer değiştir,
Adım 5.	İndis değeri I=N ise Adım 8'ye git,
Adım 6.	İndis değerini $I = I + 1$ biçiminde güncelle,
Adım 7.	Adım 3'e git,
Adım 8.	x_i kümesinden 1'den n'ye kadar örnekler seçilerek örneklemeyi gerçekleştir.

Eşit olasılıklı ve karıştırmalı yerine koymama tekniği ile rastgele örneklem seçimi için her topun sırasıyla çekildiği bir deneyin şematik gösterimi Tablo 2'de verilmiştir. Bu tabloya göre kitlede başlangıçta farklı renklerde 6 top bulunmaktadır. İlk olarak kitledeki toplar rastgele karıştırılır. Karıştırılan toplar sırasıyla kitleden çekilerek örneklem sütununa eklenmiştir. 5 çekim sonunda örneklemde 5 tane top birikmiştir. Biriken bu toplar karıştırılmış kitle topları ile aynı sırada oluştuğu görülmektedir. Bu ise karıştırmalı yerine koymama örnekleminin bir sonucudur.

Tablo 2. Karıştırmalı yaklaşımlı yerine koymama tekniği ile örneklemenin şematik gösterimi



Diğer bir yaklaşım ise rastgele indis değerleri üretilerek çıkan her örnek işaretlenerek seçilir. Bu yaklaşımla işaretlenen her örnek daha sonra seçilirse iptal edilir (Algoritma 4). Bu yaklaşım kitle birim sayısı az olduğunda gereksiz yere zaman kaybına neden olur. Kitle birim sayısı çok olduğu durumlarda çıkan bir birimin tekrar çıkma olasılığı düşük olacağından gereksiz zaman kaybı önlenmiş olur. Öte yandan tüm kitle yerine koymadan seçilecekse her birimin rastgele çekilme sonucunda tek tek çıkması ret-kabul yaklaşımından dolayı zaman alabilir.

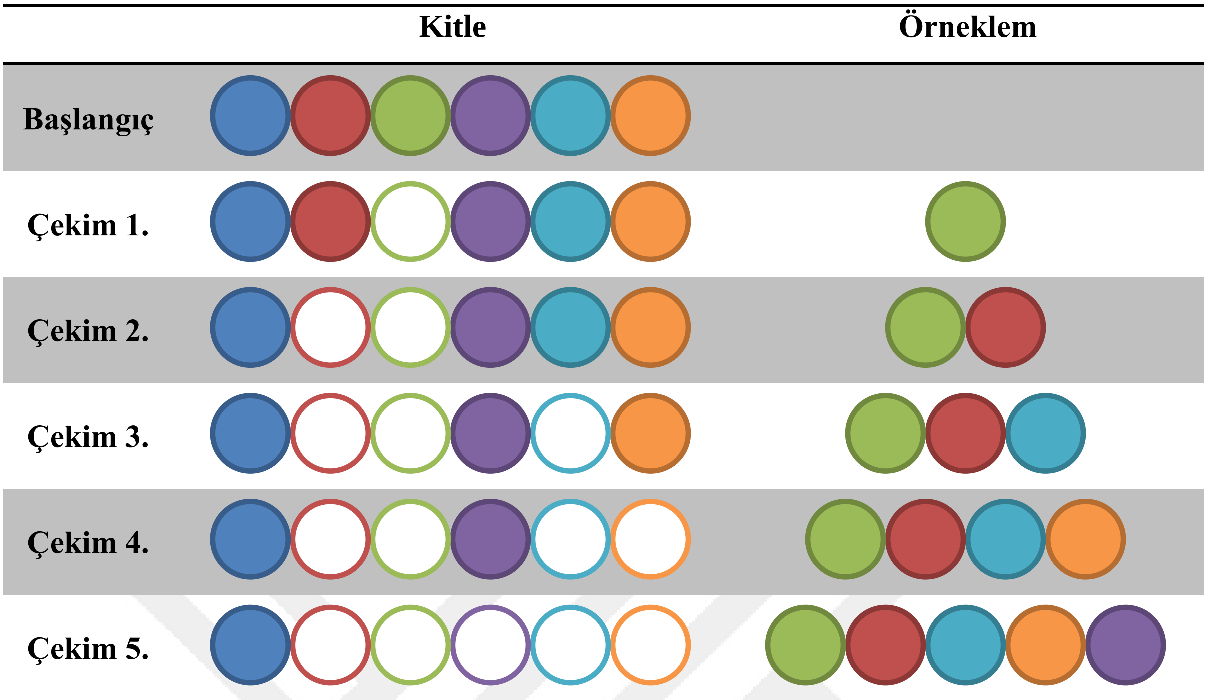
Örnek olarak eşit olasılıklarla $S = \{eda, özge, ebru, kaan\}$ biçiminde birimlerden oluşan bir kitleden Algoritma 4 yardımıyla $n = 2$ büyüklüğünde rastgele örneklem çekilmesi için kullanılması soncunda birinci örnek $L = \{özge\}$ seçilerek listeye eklenir ve ikinci örneğe sıra gelir. İkinci örnek olarak $özge$ tekrar çekilir ve L listesinde olduğundan yeniden çekme işlemi yapılır. Üçüncü çekilişte $ebru$ çekilir ve $ebru$ L listesinde olmadığı için listeye eklenir. Böylelikle $L = \{özge, ebru\}$ örneklem listesi elde edilir.

Algoritma 4. İşaretlemeli eşit olasılıklı yerine koymama tekniği ile rastgele örneklem seçimi

Adım 1. Kitle değerlerini (x_i, $i = 1,2,\ldots,N$) belirle,
Adım 2. İndis listesi L={}'yi boş küme olarak tanımla,
Adım 3. Başlangıç örneklem indisini ($j = 1$) ayarla,
Adım 4. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,
Adım 5. Çekilen indis değerini aşağıdaki eşitlik yardımıyla, $I = \lceil N \cdot U \rceil$ belirle,
Adım 6. Belirlenen I indisini L listesinde ara, eğer listede varsa Adım 4'e git,
Adım 7. I değeri L listesine ekle,
Adım 8. $X_j = x_I$ değerini seçilen birim değeri olarak al.
Adım 9. Eğer j=n ise işlemi sonlandır,
Adım 10. Örneklem indisini (j←j+1) biçiminde güncelleştir
Adım 11. Adım 4'e git.

İşaretlemeli eşit olasılıklı yerine koymama tekniği ile rastgele örneklem seçimi için her topun sırasıyla çekildiği bir deneyin şematik gösterimi Tablo 3'te verilmiştir. Bu tabloya göre kitlede başlangıçta farklı renklerde 6 top bulunmaktadır. Bu kitleden rastgele çekilen her top örneklem sütununa eklenirken aynı zamanda işaretlenerek kitledeki yerini korumaktadır. Böylelikle işaretlenen top tekrar çekilemeyeceği için yerine koymama tekniği kullanılmış olur. 5 çekim sonunda örneklemde 5 tane top birikmiştir. Biriken bu toplara karşılık gelen kitle toplarının işaretlendiği görülmektedir. Bu ise işaretlemeli yerine koymama örnekleminin bir sonucudur.

Tablo 3. İşaretlemeli yerine koymama tekniği ile örneklemenin şematik gösterimi



2.2. Yerine Koymadan ve Eşit Olmayan Olasılıklı Örnekleme

Yerine koymadan ve eşit olmayan olasılıkla örnekleme yapmak biraz kafa karmaşık bir durum oluşturmaktadır. Yerine koymalı sistemde her birim birden çok seçileceğinden çıkan örneklerle olasılıkları arasında bir ilişki olduğu Ki-kare testlerinden bilinmektedir. Oysa yerine koymadan örneklemde yüksek olasılık değerine sahip bir birim çekildikten sonra kitlede olmayacağı için elde edilen örneklem ile keyfi olasılıklar arasında bir ilişki bulmak imkansız gözükmektedir. Öte yandan keyfi olasılık değerlerine sahip birimlerin seçilme şansları yerine seçilme sıralarını öncelik veren bir yaklaşım değerlendirilebilir. Bu yaklaşım Algoritma 5'te verilmiştir.

Algoritma 5. Keyfi olasılıklara göre yerine koymama tekniği ile rastgele örneklem seçimi

Adım 1. Kitle değerlerini $(x_i, i = 1,2,\dots,N)$ ve olasılıklarını $(p_i, i = 1,2,\dots,N)$ belirle,

Adım 2. Başlangıç örneklem indisini (j=1) ayarla,

Adım 3. Başlangıç kitle indisini $(I = 1)$ ayarla,

Adım 4. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,

Adım 5. $U < p_I$ ise Adım 9 git,

Adım 6. $U = U - p_I$ güncellemesini yap

Adım 7. İndis değerini $I = I + 1$ biçiminde güncelle,

Adım 8. Adım 4 git

Adım 9. $X_j = x_I$ değerini seçilen birim değeri olarak al.

Adım 10. Çıkan birimin olasılığını $(p_I = 0)$ sıfırla,

Adım 11. Tüm olasılıkları aşağıdaki eşitlik yardımıyla güncelle,

$$p_i \leftarrow \frac{p_i}{\sum_{k=1}^{N} p_k}, (\, i = 1,2,\dots,N)$$

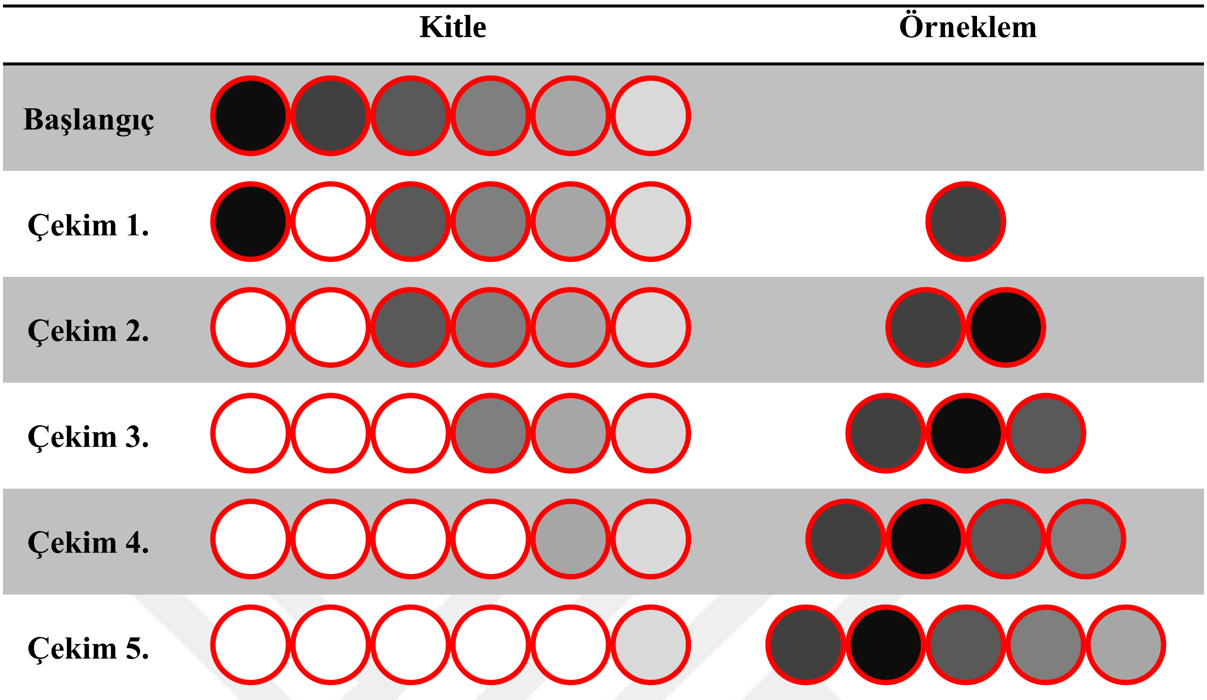
Adım 12. Eğer j=n ise işlemi sonlandır,

Adım 13. Örneklem indisini (j←j+1) biçiminde güncelleştir,

Adım 14. Adım 3’e git.

Keyfi olasılıklı yerine koymama tekniği ile rastgele örneklem seçimi için her topun sırasıyla çekildiği bir deneyin şematik gösterimi Tablo 4’te verilmiştir. Bu tabloya göre kitlede başlangıçta farklı renklerde 6 top bulunmaktadır. Bu topların olasılık değerleri topların koyuluklarına göre ayarlanmıştır. Koyu renkteki topların olasılığı yüksekken açık renkli topların olasılıkları düşüktür. Bu kitleden olasılıklarına göre rastgele çekilen her top örneklem sütununa eklenirken aynı zamanda olasılıkları sıfırlanarak kitledeki yerini konmaktadır. Böylelikle çekilen top tekrar çekilemeyeceği için yerine koymama tekniği kullanılmış olur. 5 çekim sonunda örneklemde 5 tane top birikmiştir. Biriken bu toplara karşılık gelen kitle toplarının olasılıklarının sıfırlandığı görülmektedir. Bu ise eşit olmayan olasılıklı yerine koymama örnekleminin bir sonucudur.

Tablo 4. Keyfi olasılıklı yerine koymama tekniği ile örneklemenin şematik gösterimi



Eşit olmayan olasılık örneklemesine diğer bir yaklaşım ise kitlede aynı birimden birden fazla olması durumunda geçerli olan yaklaşımdır. Bu durumda kitlede K tane farklı birim olsun. Kitle içindeki her birimin sayısı N_k ile gösterilirse, kitlede toplamda,

$$N = \sum_{k=1}^{K} N_k \tag{11}$$

tane birim bulunmaktadır. Bu durumda her birimin olasılık değeri,

$$p_k = \frac{N_k}{N}, \quad k = 1,2, \dots, K \tag{12}$$

olarak verilmektedir. Her çekilen birimde olasılıklar yeniden hesaplanır. Bu hesaplamada çekilen birimin olasılığı azalırken, diğer birimlerin olasılığı artmaktadır. Bu işlem adımları Algoritma 6’da verilmektedir.

Algoritma 6. Aynı örnekli birimlerden yerine koymadan tekniği ile rastgele örneklem seçimi

---

Adım 1. Kitle değerlerini $(x_i, i = 1,2, \ldots, K)$ ve her birimin sayısını $(N_k, k = 1,2, \ldots, K)$ belirle,

Adım 2. Başlangıç örneklem indisini (j=1) ayarla,

Adım 3. Olasılık değerlerini (12) eşitliğine göre belirle,

Adım 4. Başlangıç kitle indisini $(I = 1)$ ayarla,

Adım 5. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,

Adım 6. $U < p_I$ ise Adım 10 git,

Adım 7. $U = U - p_I$ güncellemesini yap,

Adım 8. İndis değerini $I = I + 1$ biçiminde güncelle,

Adım 9. Adım 5 git,

Adım 10. $X_j = x_I$ değerini seçilen birim değeri olarak al,

Adım 11. Seçilen birimin sayısı $N_I \leftarrow (N_I - 1)$ ve tüm birim sayısını $N \leftarrow (N - 1)$ güncelle,

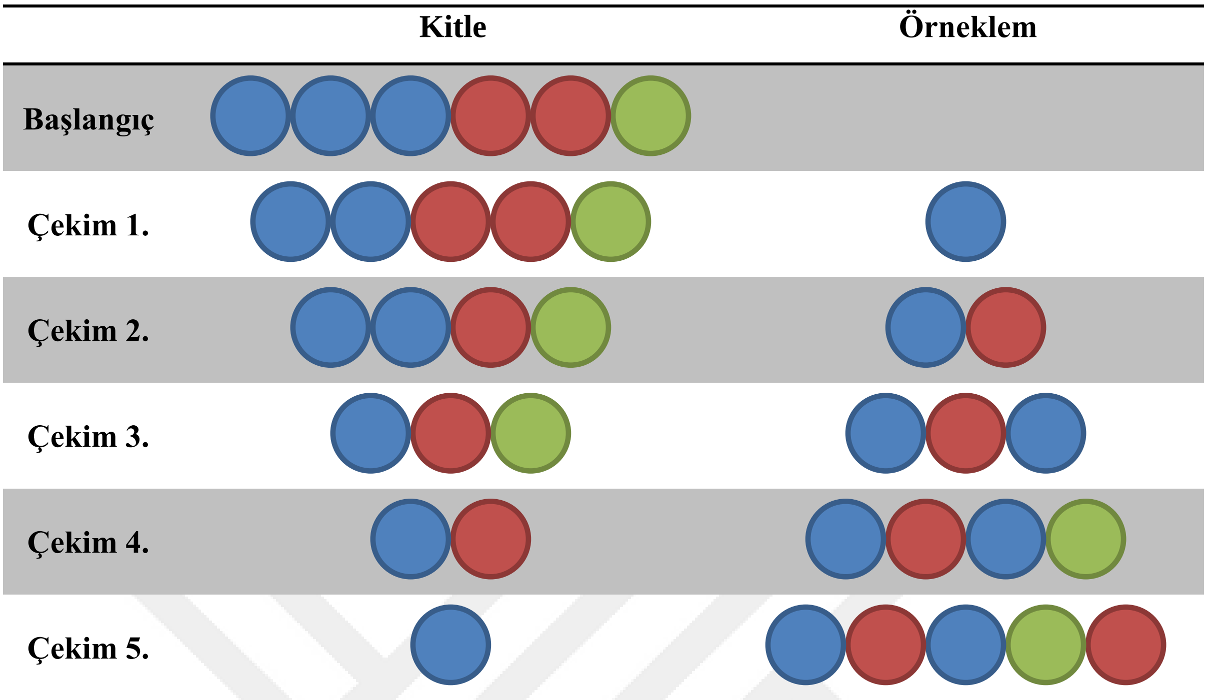
Adım 12. Eğer j=n ise işlemi sonlandır,

Adım 13. Örneklem indisini (j←j+1) biçiminde güncelleştir,

Adım 14. Adım 3'e git.

Aynı örnekli birimlerden oluşan bir kitleden yerine koymama tekniği ile rastgele örneklem seçiminde aynı örnek sayısı bilindiğinden çekilen her örnek türü bir azaltılarak yapılan bir deneyin şematik gösterimi Tablo 5'te verilmiştir. Bu tabloya göre kitlede başlangıçta 3 tane mavi 2 tane bordo ve 1 tane yeşil renklerde 6 top bulunmaktadır. Bu topların olasılık değerleri eşit olarak verilmektedir. Bu kitleden eşit olasılıklarına göre rastgele çekilen her top örneklem sütununa eklenirken aynı zamanda kitledeki yerini konmaktadır. Böylelikle çekilen top tekrar çekilemeyeceği için yerine koymama tekniği kullanılmış olur. 5 çekim sonunda örneklemde 5 tane top birikmiştir. Biriken bu toplara karşılık gelen kitle toplarının eksildiği görülmektedir. Bu ise eşit olasılıklı yerine koymama örnekleminin bir sonucudur.

Tablo 5. Keyfi olasılıklı yerine koymama tekniği ile örneklemenin şematik gösterimi



2.3. Bulanık Yerine Koyma ile Örnekleme

İstatistiksel örneklemede yerine koyma veya yerine koymama gibi Aristo'nun {doğru (1), yanlış (0)} mantık yaklaşımlıyla iki farklı teknik olarak kullanılmaktadır. Yerine koyma yaklaşımıyla kitleden çekilen her birimin süreç boyunca seçilme olasılığı korunurken, yerine koymama yaklaşımında çekilen her birimin seçilme olasılığı sıfırlanmakta ve dolayısıyla diğer birimlerin seçilme olasılıkları artmaktadır. Gerçek hayatta bilinçaltı etkisiyle yapılan birçok örnekleme Zadeh'in [0,1] yaklaşımı ile gerçekleştirilmektedir. Bu çalışmada, kitleden çekilen bir birimin çekilme olasılığı bulanıklaştırma parametresi yardımıyla düşürülerek yerine konulmasıyla yerine koyma veya koymama arasında bir tercih yapılmış olur. [0,1] aralığında değerler alan bulanıklaştırma parametresi $(m = 0)$ olduğunda yerine koymama, $(m = 1)$ olduğunda ise yerine koyma mantığı ile çalışmaktadır. Bulanıklaştırma parametresi $m \in [0,1]$ aralığında herhangi bir değeri alması ise bulanık yerine koyma anlamına gelmektedir.

Bir örneklemede istenen en temel yaklaşım gözlenen frekanslar ile beklenen frekanslar arasındaki farkın (hata) minimum hale getirilmesidir. Hatanın küçükleştirmesinde en sağlam yaklaşım kitleden seçilen birimin tekrar yerine konmaması olarak yapılabilir. Bu durumda N birimden oluşan bir kitleden $n > N$ şekilde bir örnekleme yapılacağı zaman istenen birim kitlede olmayabilir. Bu durumda seçilen birim tekrar kitleye konarak yeniden seçilme yoluna gidilebilir ve seçilen her birim daha önce seçilen örneklem kümesine eklenir. Ancak kitledeki birimlerin olasılıkları farklı olduğunda birebir seçim söz konusu olamaz. Bu durumda olasılığına göre her birimin beklenen frekans kadar aynı birimi kitleye konarak örneklem gerçekleştirilir. Ancak seçilecek birim sayısı n tane değer oluşacaksa bu örneklerin beklenen frekansları,

$$E_i = np_i \tag{13}$$

biçiminde bulunur. Bulunan beklenen değerler her zaman tam sayı çıkmayacağı için beklenen frekanslar ile gözlenen frekanslar eşit olmayacaktır. Küçük de olsa bir hata oluşacaktır. Öte yandan seçilen birim sayısı önceden bilinmiyorsa veya belli bir süreç boyunca örneklem yapılıyorsa yerine koymama yaklaşımı ile örnekleme yapmak özellikle eşit olasılıkta olmayan örnekleme imkansız hale gelmektedir.

Yerine koyma ve koymama yaklaşımının bir parametre yardımıyla birleştirilerek bu iki yaklaşımın arasında da bir örnekleme yapılabilir. Bu tarz örneklemeye bulanık yerine koyma ile örnekleme yöntemi olarak tanımlanabilir.

Yerine koymama yönteminde temel yaklaşım beklenen frekans ile gözlenen frekans arasındaki farkın sürekli kontrol altında tutmaktadır. Bu kontrolü gerçekleştirmek için beklenen frekans ile gözlenen frekansın oranı incelenmiştir. Bu oran,

$$r_i = \frac{f_i}{np_i}, \qquad i = 1,2,\dots,N \tag{14}$$

biçiminde hesaplanmıştır. Bu oranın birim sayısı önceden bilinmediğinde veya birer birer birimler seçildiğinde her aşamada bu oranın kontrol altında tutulması istenebilir bu durumda oran göreceli biçimde hesaplanabilir. Göreceli oran ise,

$$r_i = \frac{\hat{f}_i}{p_i}, \qquad i = 1,2,\dots,N \tag{15}$$

olarak verilebilir. Burada $\hat{f}_i$,

$$\hat{f}_i = \frac{f_i}{\sum_{j=1}^{n} f_j}, \quad i = 1,2,\dots,N \tag{16}$$

biçiminde göreceli frekanslar olarak verilmektedir. Bu oran değerinin (r_i) 1'den küçük olması karşılık gelen birimin az seçildiği, dolayısıyla o birimin olasılığı şişirilerek örnekleme denge getirilebilir. Bu oranın 1'den büyük olması ise karşılık gelen birimin gereğinden fazla seçilmesine neden olmuştur. Dolayısıyla bu birimin olasılığının bastırılması gerekir. Oran değerleri 1 olduğunda ise normal örneklemenin uygulanması gerekir. Bu oran değerlerindeki değişimin olasılık üzerindeki şişirme ve bastırma etkisi,

$$p_i' = p_i\, e^{-\left(\frac{\hat{f}_i}{p_i}\right)}, \quad i = 1,2,\dots,N \tag{17}$$

biçiminde tanımlanabilir. Bu tanımlamada p_i başlangıç (teorik) olasılık değerlerini gösterirken $\hat{p}_i$değerleri bir sonraki birimdeki olasılık değerlerini göstermektedir. Bu şekilde yapılan örneklemelerde örneklem hatası azalmaktadır. Bu durum ne yerine koyma ne de yerine koymama örneklemesine karşılık gelmektedir. Ancak tam anlamıyla bir kontrol sağlamak için bulanıklaştırma parametresi kullanılabilir. Bu parametre kullanımına göre örnekleme için değişken olasılıklar,

$$p_i' = p_i\, e^{-\left(\frac{\hat{f}_i}{p_i}\right)^{\frac{1-m^2}{m^2}}}, \quad i = 1,2,\dots,N \tag{18}$$

biçiminde verilebilir. Burada,

$$w_i = e^{-\left(\frac{\hat{f}_i}{p_i}\right)^{\frac{1-m^2}{m^2}}}, \quad i = 1,2,\dots,N \tag{19}$$

ifadesi olasılığın bulanıklaştırma katsayısı olarak tanımlanır. Bu katsayının bulanıklaştırma parametresine göre değişimi Tablo 6'de verilmiştir.

Tablo 6. Bulanıklaştırma parametresine göre bulanıklaştırma katsayısı.

$w_i = e^{-\left(\frac{f_i}{p_i}\right)^{\frac{1-m^2}{m^2}}}$		m		
		0	1/2	1
$\frac{f_i}{pi}$	0.5	1.00	0.61	0.37
	1	0.37	0.37	0.37
	2	0.00	0.14	0.37

Tablo 6'e göre bulanıklaştırma parametresi $(m = 0)$ olduğunda yüksek oranların olasılığını sıfırlarken düşük oranların olasılığı göreceli olarak artmaktadır. Böylelikle örneklemenin yerine koymama yaklaşımını desteklemektedir. Öte yandan bulanıklaştırma parametresi $(m = 1)$ olduğunda ise yerine koyma yaklaşımını desteklediği için olasılık değerlerine bir değişime gitmediği için istenen sonucu vermektedir.

Bulanıklaştırma katsayı ile başlangıç olasılık değerinin çarpımı,

$$p_i' = w_i p_i, \quad i = 1,2, \dots, N \tag{20}$$

biçiminde verilmektedir. Bu eşitlikte elde edilen olasılık değerlerinin toplamı 1'den büyük olacağından göreceli olasılıkların hesaplanması gerekir. Göreceli olasılıklar,

$$\hat{p}_i = \frac{p_i'}{\sum_{j=1}^{N} p_i'}, \quad i = 1,2, \dots, N \tag{21}$$

biçiminde bulunur. Bu göreceli olasılıklar çekilen her rastgele örnek öncesinde her seferinde tekrar tekrar hesaplanmalıdır.

Bulanık yerine koyma yönteminin algoritması Algoritma 7'de verilmiştir.

Algoritma 7. Bulanık yerine koyma yöntemi ile rastgele örnekleme algoritması.

Adım 1. Kitle değerlerini $(x_i, i = 1,2,\ldots,N)$ ve her birimin başlangıç olasılık değerlerini $(p_i, i = 1,2,\ldots,N)$ belirle,

Adım 2. Bulanıklaştırma parametresini ve çekilecek birim sayısını belirle,

Adım 3. Başlangıç örneklem indisini (j=1) ayarla,

Adım 4. Olasılık değerlerini (16)-(23) eşitliklerine göre belirle,

Adım 5. Başlangıç kitle indisini $(I = 1)$ ayarla,

Adım 6. [0,1) aralığında düzgün dağılımdan bir rastgele sayı (U) üret,

Adım 7. $U < \hat{p}_I$ ise Adım 11 git,

Adım 8. $U = U - \hat{p}_I$ güncellemesini yap,

Adım 9. İndis değerini $I = I + 1$ biçiminde güncelle,

Adım 10. Adım 6 git,

Adım 11. $X_j = x_I$ değerini seçilen birim değeri olarak al,

Adım 12. Seçilen birimin frekansını $f_I \leftarrow (f_I + 1)$ ve göreceli frekansını (16) eşitliğine göre güncelle,

Adım 13. Eğer j=n ise işlemi sonlandır,

Adım 14. Örneklem indisini (j←j+1) biçiminde güncelleştir,

Adım 15. Adım 4’e git.

3. BULGULAR

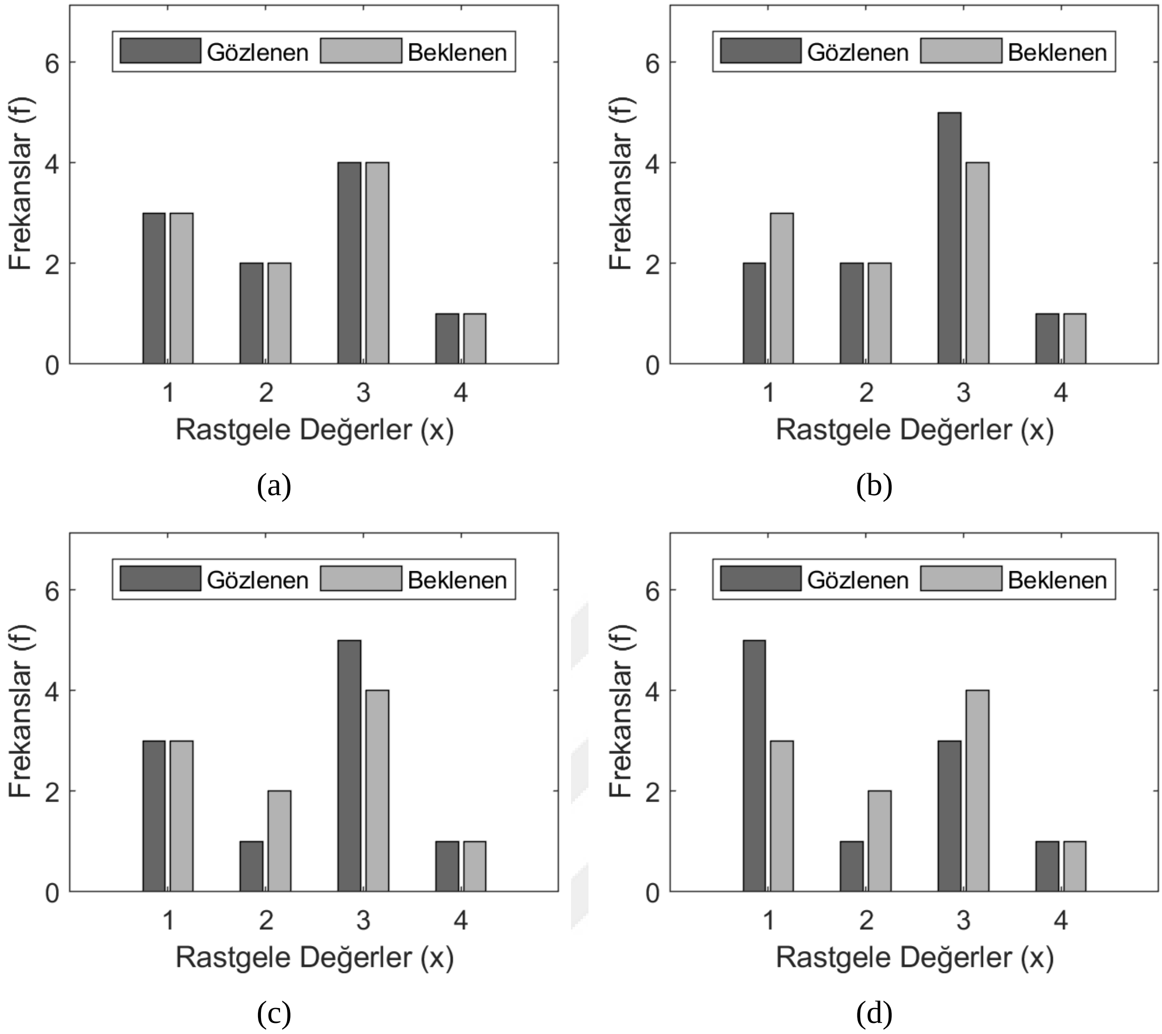
Önerilen yöntemin başarımını incelemek için keyfi ve teorik dağılımlara göre örneklemlerin benzetimleri gerçekleştirilmiştir. Yapılan örneklemlerin bulanıklaştırma parametresine göre başarımını ölçmek için ki-kare istatistiği kullanılmıştır.

3.1. Dağılımlarda Rastgele Örnekleme

Genelde örnekleme işlemi, kitle olarak verilen bir kümeden rastgele seçilen birimlerin oluşturduğu küme olarak tanımlanmaktadır. Eşit olmayan olasılıklara sahip birimlerin oluşturduğu bir kitleden örneklem oluşturmak için olasılık değerlerinin önceden bilinmesi gerekir. Bu olasılık değerleri teorik dağılım fonksiyonu olarak tanımlanabildiği gibi liste halinde de verilebilir. Bu dağılım yaklaşımı için 5 ayrı dağılım örneğine göre benzetimler yapılmıştır. Bu benzetimde farklı birim sayısına ve farklı bulanıklaştırma değerlerine göre örneklemler oluşturularak incelenmiştir.

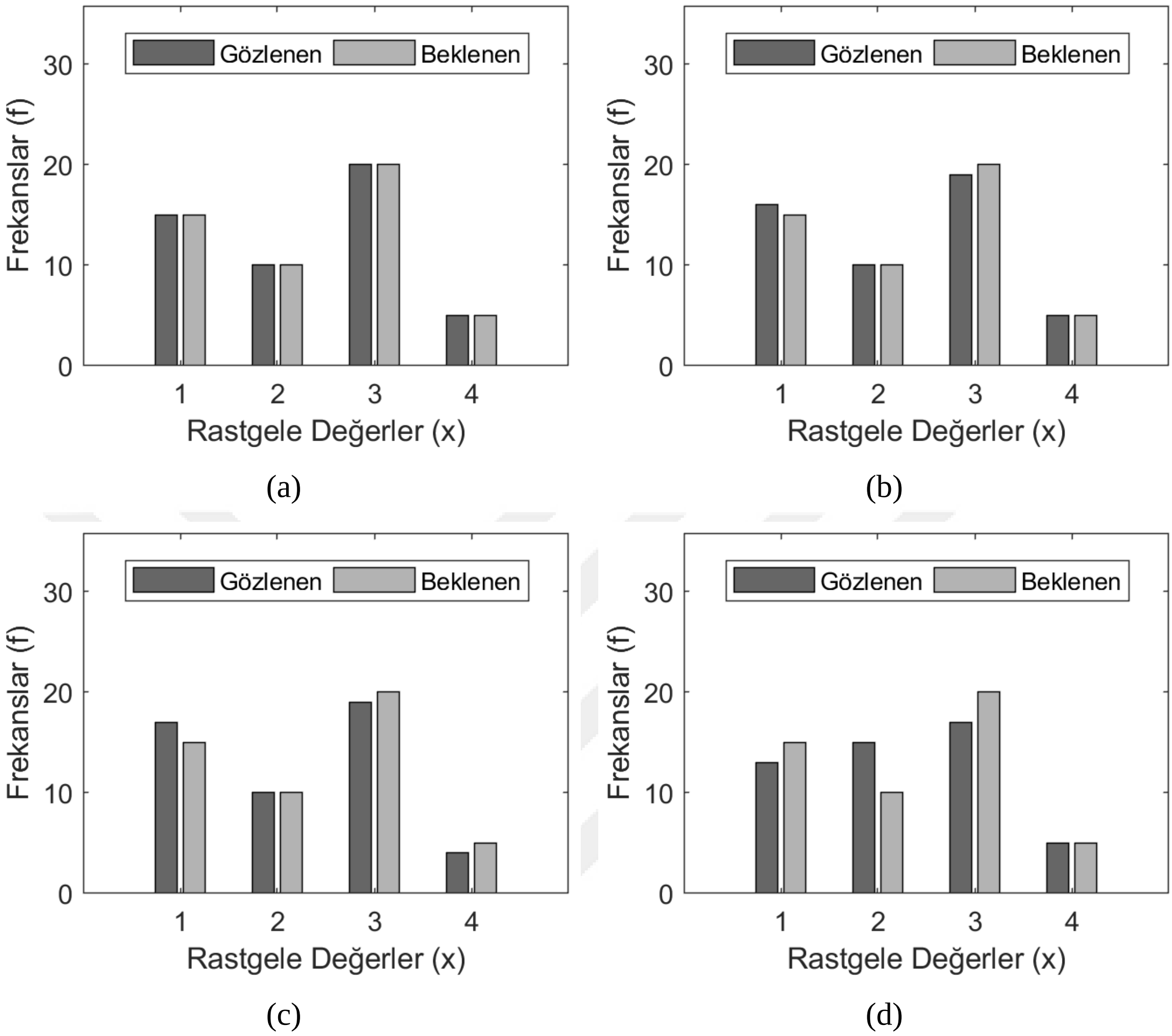
3.1.1. Keyfi Bir Dağılımdan Rastgele Örnekleme

Olasılık değerleri $p = [0.3\ 0.2\ 0.4\ 0.1]$ olan bir kitledeki değerler $x = \{1,2,3,4\}$ olarak verilsin. Verilen olasılıkların toplamı 1 olduğundan keyfi bir dağılım olarak kabul edilebilir. Bu keyfi dağılımın oluşturduğu kitleden seçilen 10 birim için beklenen ve gözlenen frekansların histogramı Şekil 2'de verilmektedir. Şekil 2 (a) ve (b)'de bulanıklaştırma değeri $m = 0, 1/3$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 2 (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



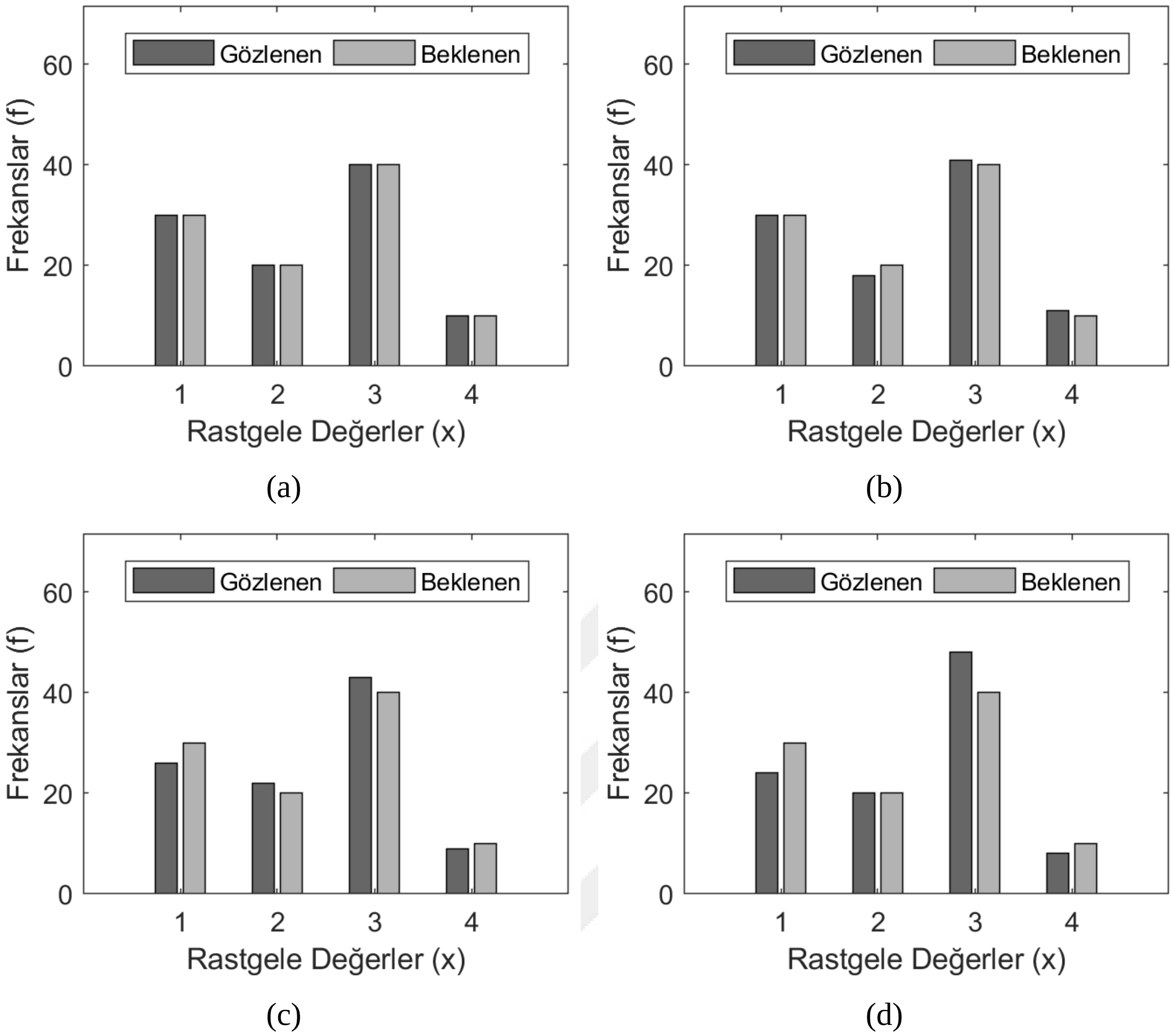
Şekil 2. Keyfi dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 10); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık değerleri $p = [0.3\ 0.2\ 0.4\ 0.1]$ olan bir kitledeki değerler $x = \{1,2,3,4\}$ olarak verilsin. Verilen olasılıkların toplamı 1 olduğundan keyfi bir dağılım olarak kabul edilebilir. Bu keyfi dağılımın oluşturduğu kitleden seçilen 50 birim için beklenen ve gözlenen frekansların histogramı Şekil 3'te verilmektedir. Şekil 3 (a)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 3 (b), (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



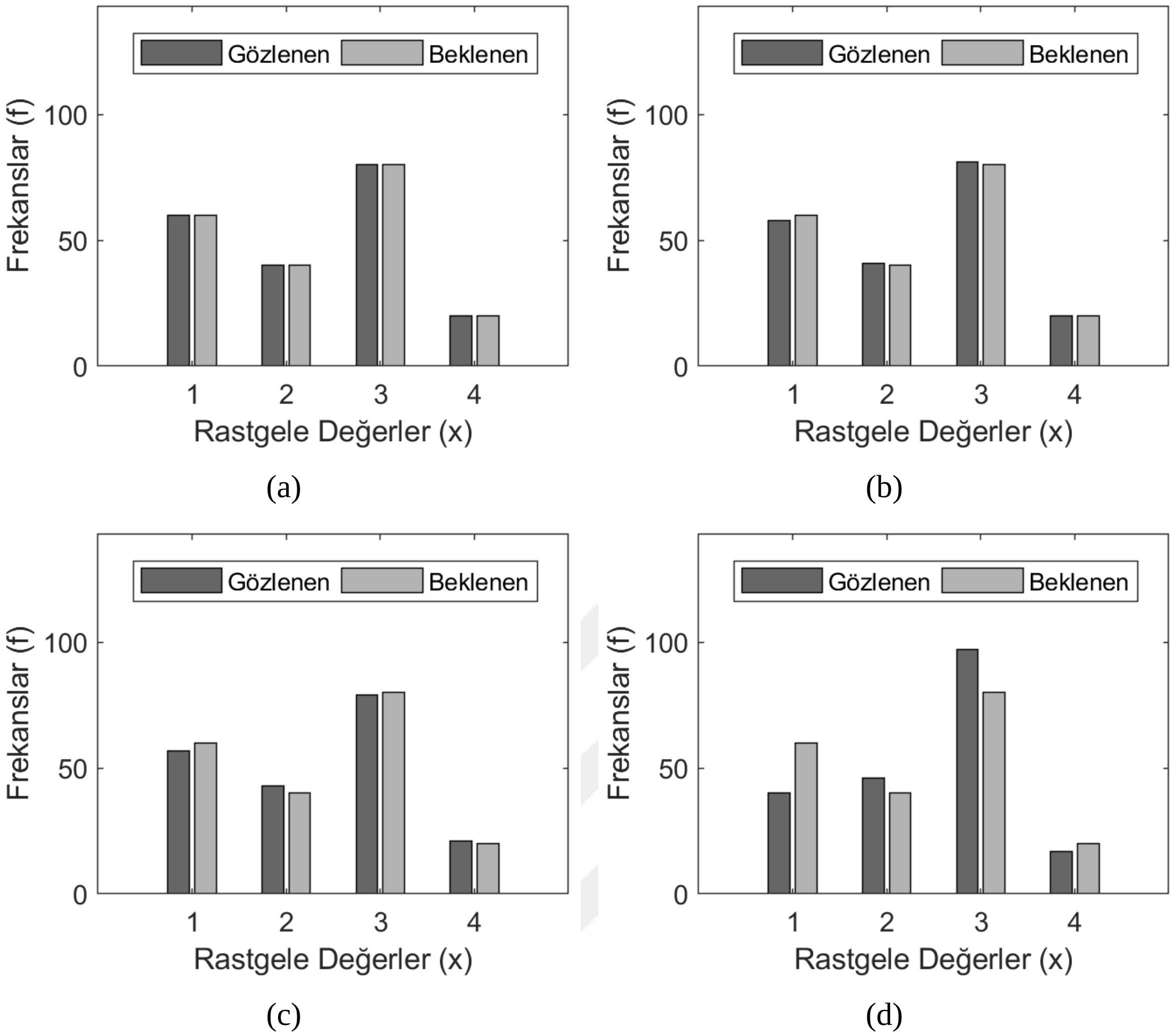
Şekil 3. Keyfi dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 50); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık değerleri $p = [0.3\ 0.2\ 0.4\ 0.1]$ olan bir kitledeki değerler $x = \{1,2,3,4\}$ olarak verilsin. Verilen olasılıkların toplamı 1 olduğundan keyfi bir dağılım olarak kabul edilebilir. Bu keyfi dağılımın oluşturduğu kitleden seçilen 100 birim için beklenen ve gözlenen frekansların histogramı Şekil 4'te verilmektedir. Şekil 4 (a)'da bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 4 (b), (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 4. Keyfi dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 100); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

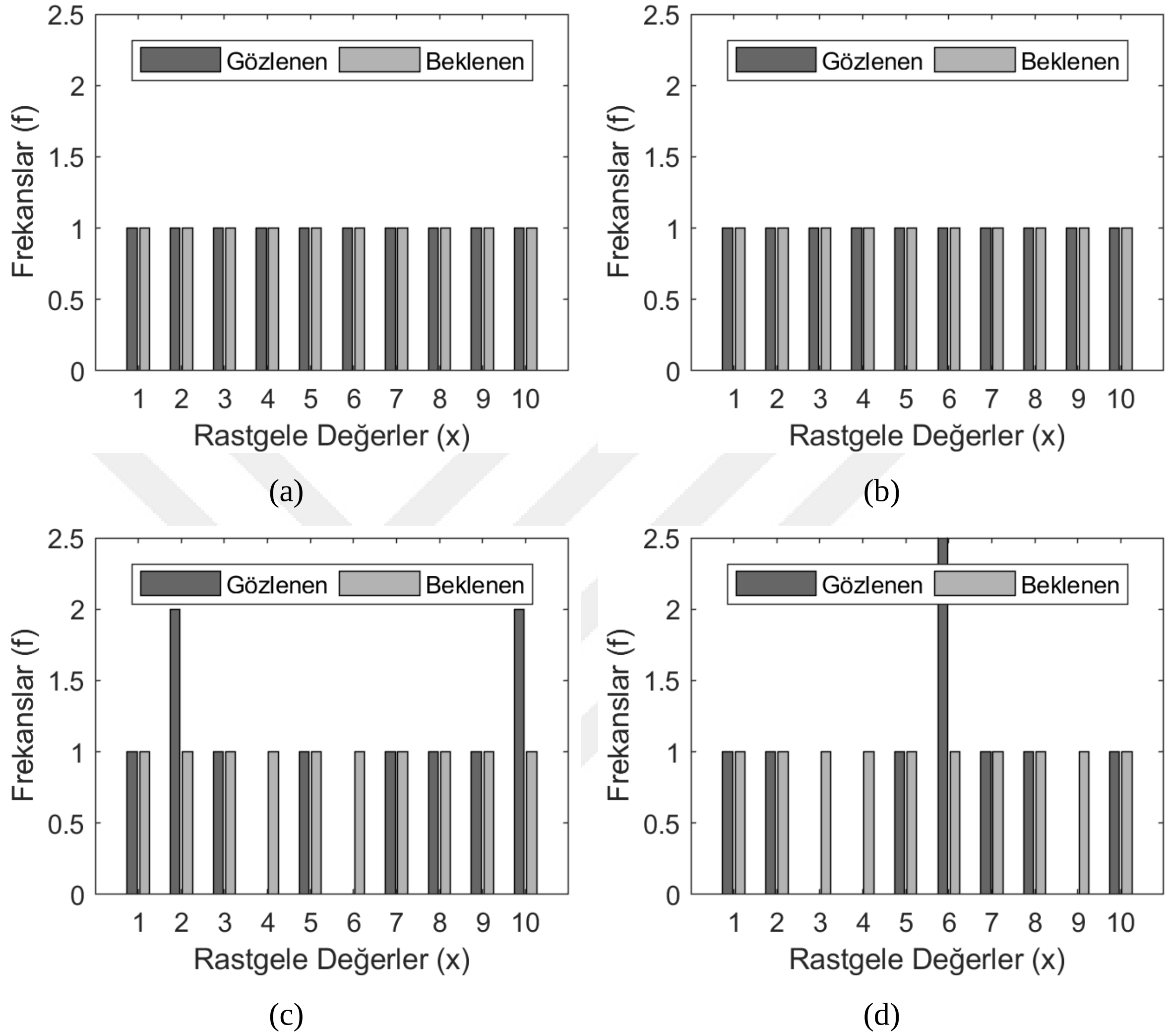
Olasılık değerleri $p = [0.3\ 0.2\ 0.4\ 0.1]$ olan bir kitledeki değerler $x = \{1,2,3,4\}$ olarak verilsin. Verilen olasılıkların toplamı 1 olduğundan keyfi bir dağılım olarak kabul edilebilir. Bu keyfi dağılımın oluşturduğu kitleden seçilen 200 birim için beklenen ve gözlenen frekansların histogramı Şekil 5’te verilmektedir. Şekil 5 (a)’de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 5 (c) ve (d)’de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 5. Keyfi dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 200); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

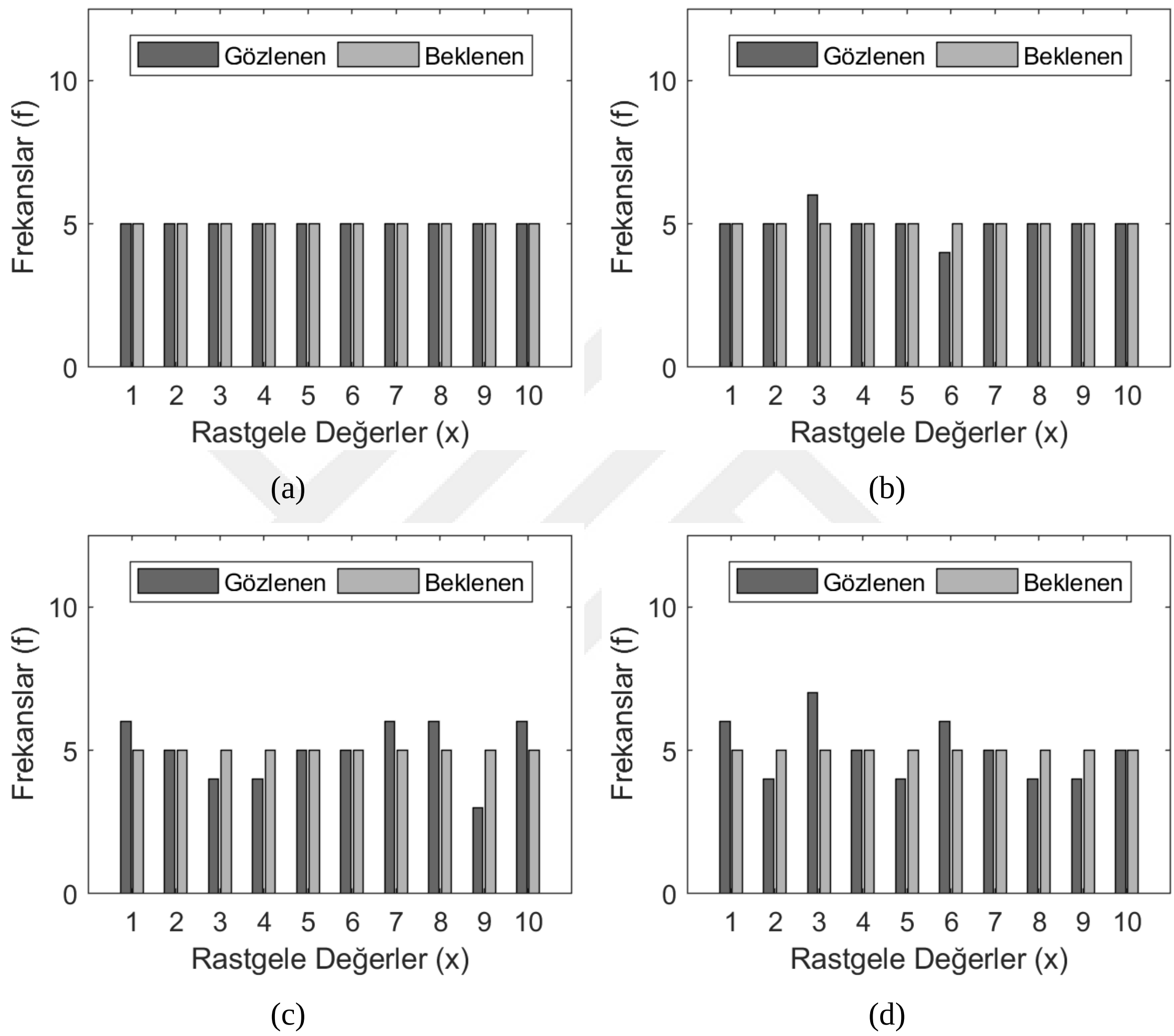
3.1.2. Düzgün Dağılımdan Rastgele Örnekleme

Olasılık yoğunluk fonksiyonu $f(x) = \frac{1}{N}$; $x = 1,2, \dots, N,$ $N = 10$ olan bir kitledeki değerler $[1,10]$ aralığında tanımlansın. Bu düzgün dağılımın oluşturduğu kitleden seçilen 10 birim için beklenen ve gözlenen frekansların histogramı Şekil 6'da verilmektedir. Şekil 6 (a) ve (b)'de bulanıklaştırma değeri $m = 0, 1/3$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 6 (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



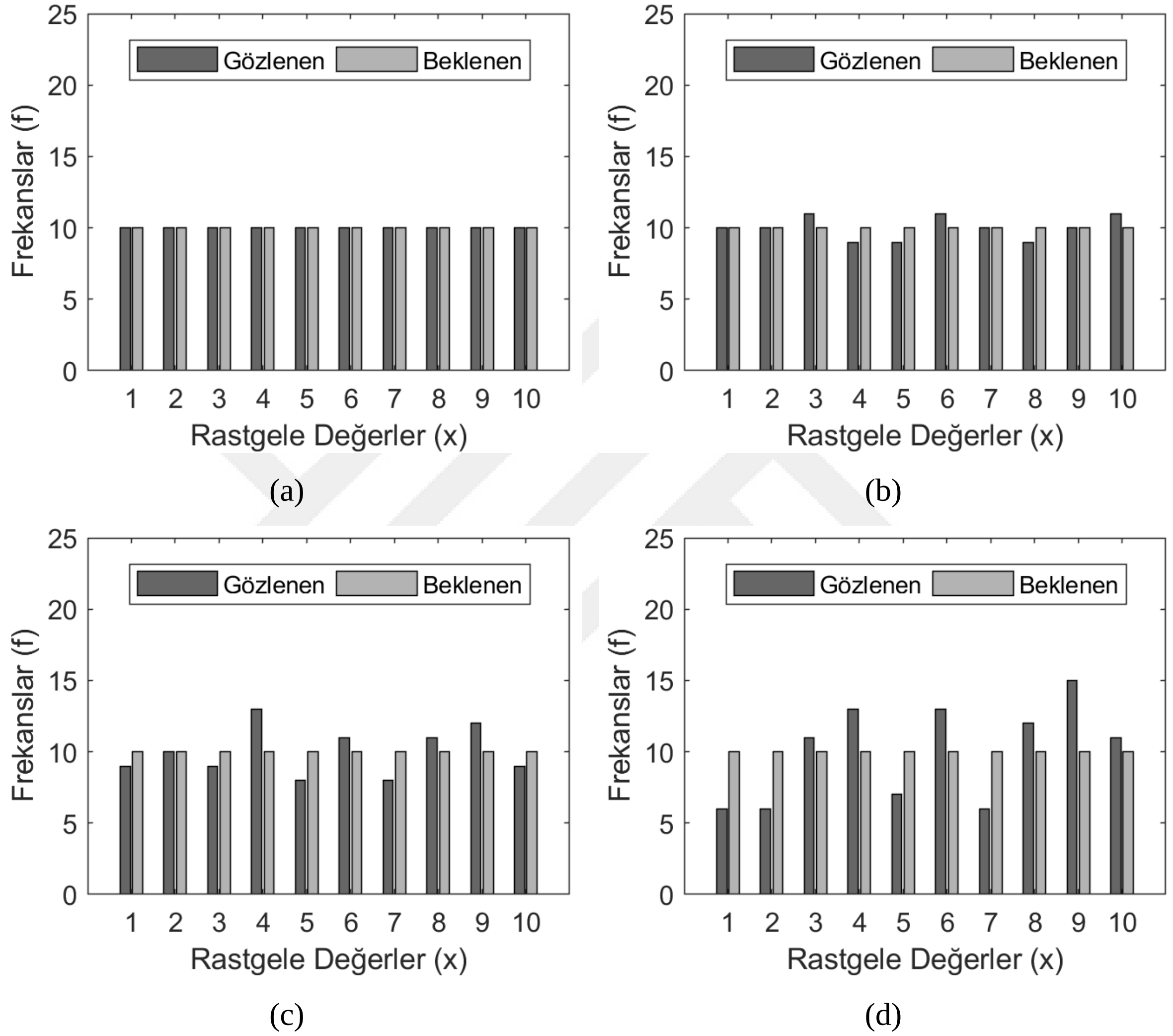
Şekil 6. Düzgün dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 10); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \frac{1}{N}$; $x = 1,2, \ldots, N,\ \ N = 10$ olan bir kitledeki değerler [1,10] aralığında tanımlansın. Verilen olasılıkların toplamı 1 olduğundan ve her bir değerin olasılığı aynı olduğundan düzgün bir dağılım olarak kabul edilebilir. Bu düzgün dağılımın oluşturduğu kitleden seçilen 50 birim için beklenen ve gözlenen frekansların histogramı Şekil 7'de verilmektedir. Şekil 7 (a)'da bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 7 (b), (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



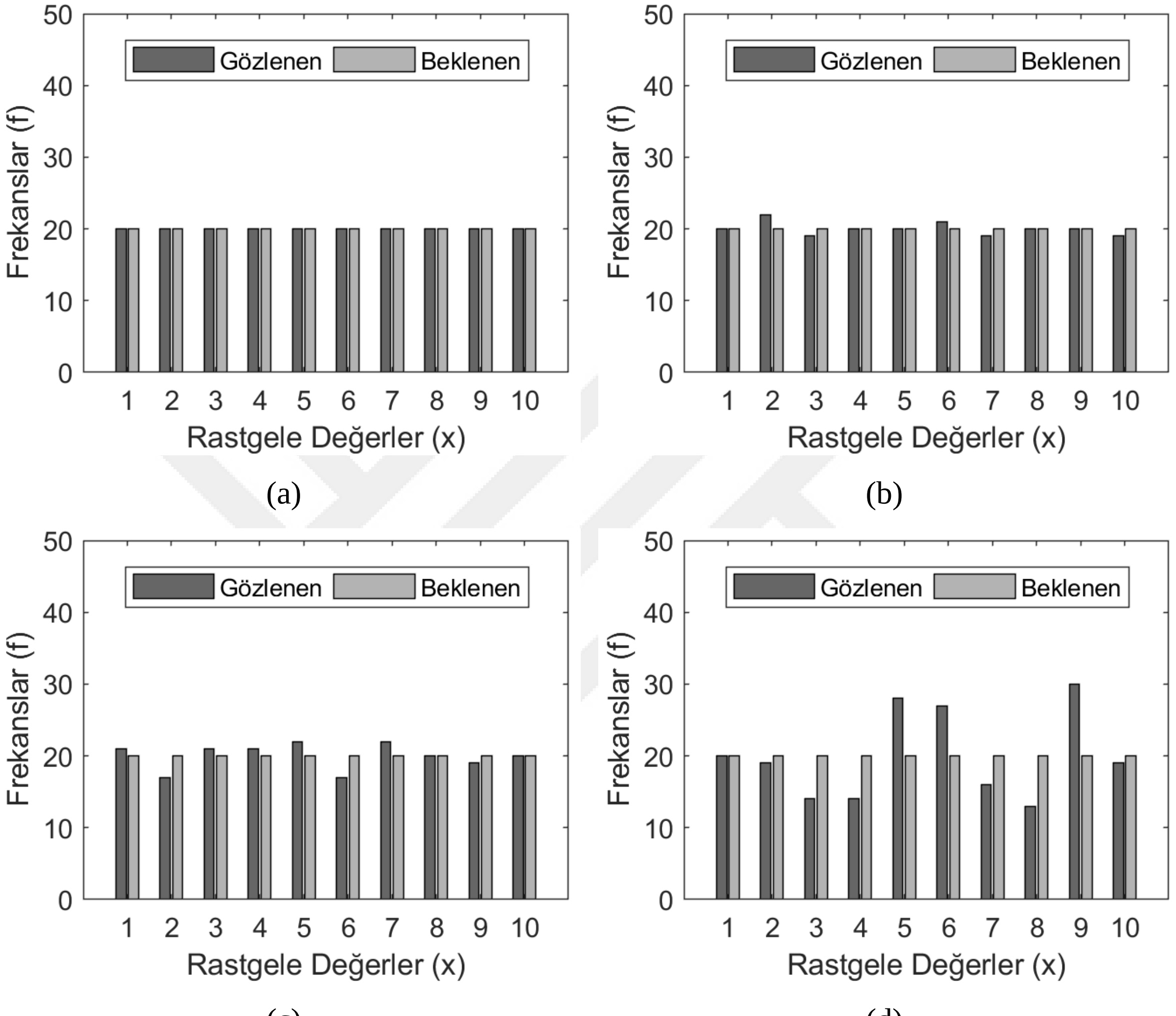
Şekil 7. Düzgün dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 50); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \frac{1}{N}$; $x = 1,2, \dots, N,\ N = 10$ olan bir kitledeki değerler [1,10] aralığında tanımlansın. Bu düzgün dağılımın oluşturduğu kitleden seçilen 100 birim için beklenen ve gözlenen frekansların histogramı Şekil 8'de verilmektedir. Şekil 8 (a)'da bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 8 (b), (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 8. Düzgün dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 100); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \frac{1}{N}$; $x = 1,2,\dots,N,\ \ N = 10$ olan bir kitledeki değerler $[1,10]$ aralığında tanımlansın. Bu düzgün dağılımın oluşturduğu kitleden seçilen 200 birim için beklenen ve gözlenen frekansların histogramı Şekil 9'da verilmektedir. Şekil 9 (a)'da bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 9 (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.

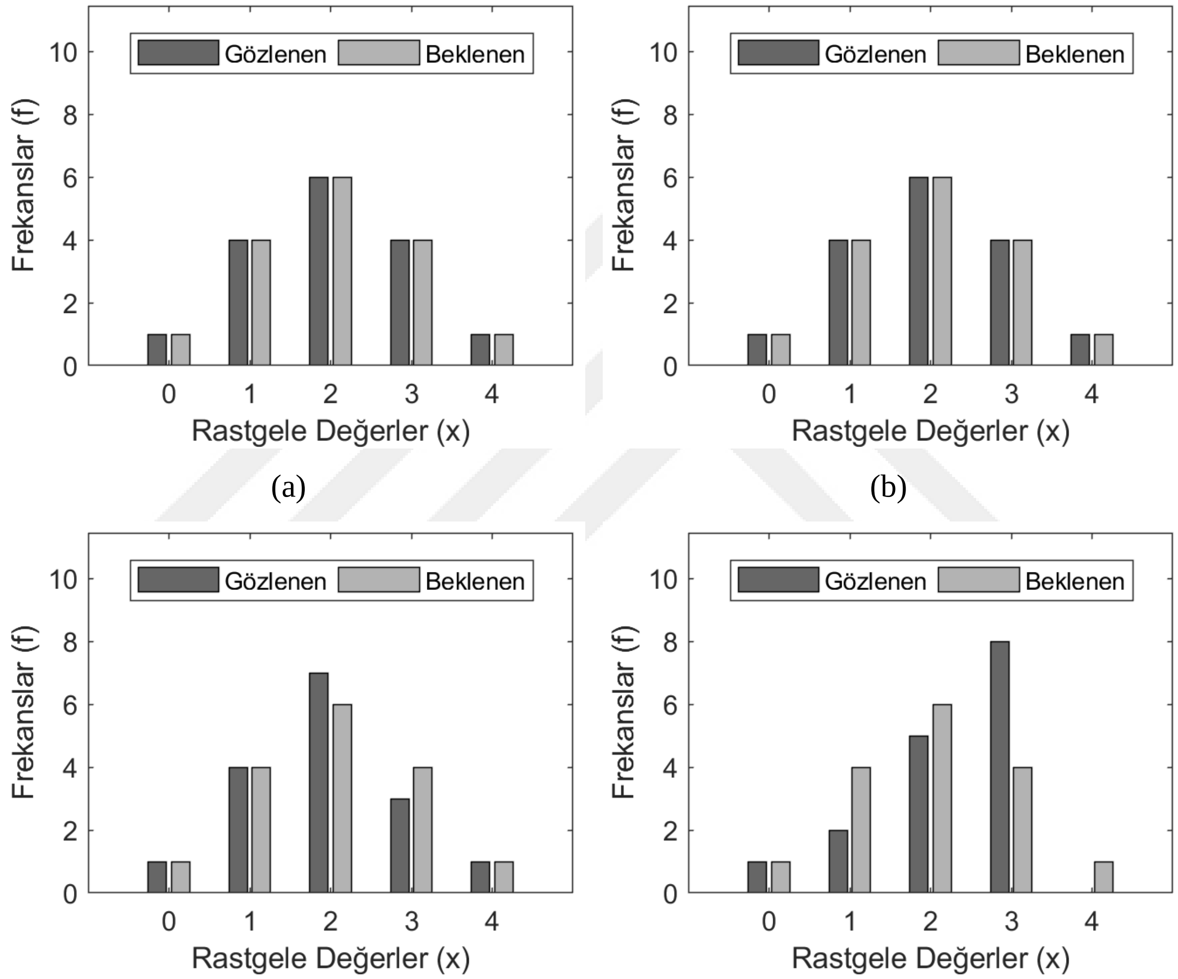


Şekil 9. Düzgün dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 200); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

3.1.3. Binom Dağılımından Rastgele Örnekleme

Olasılık yoğunluk fonksiyonu $f(x) = \binom{n}{x} p^x q^{n-x}$; $x = 0,1,2,3,n$; $p = 0.5$, $n = 4$ olan bir kitledeki değerler [0,4] aralığında tanımlansın. Bu binom dağılımın oluşturduğu kitleden seçilen 16 birim için beklenen ve gözlenen frekansların histogramı

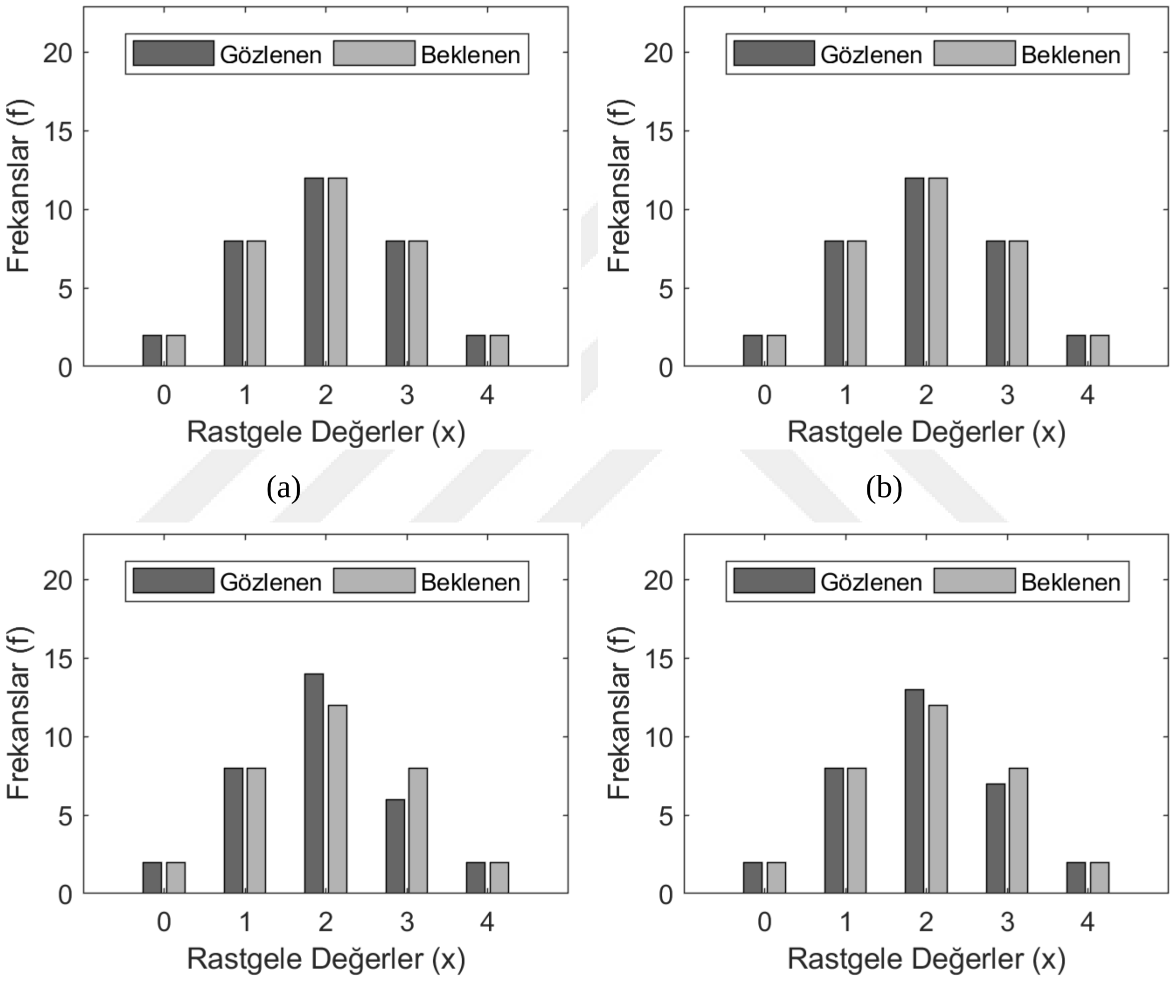
Şekil 10'da verilmektedir. Şekil 10 (a) ve (b)'de bulanıklaştırma değeri $m = 0, 1/3$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 10 (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



(a) (b) (c) (d)

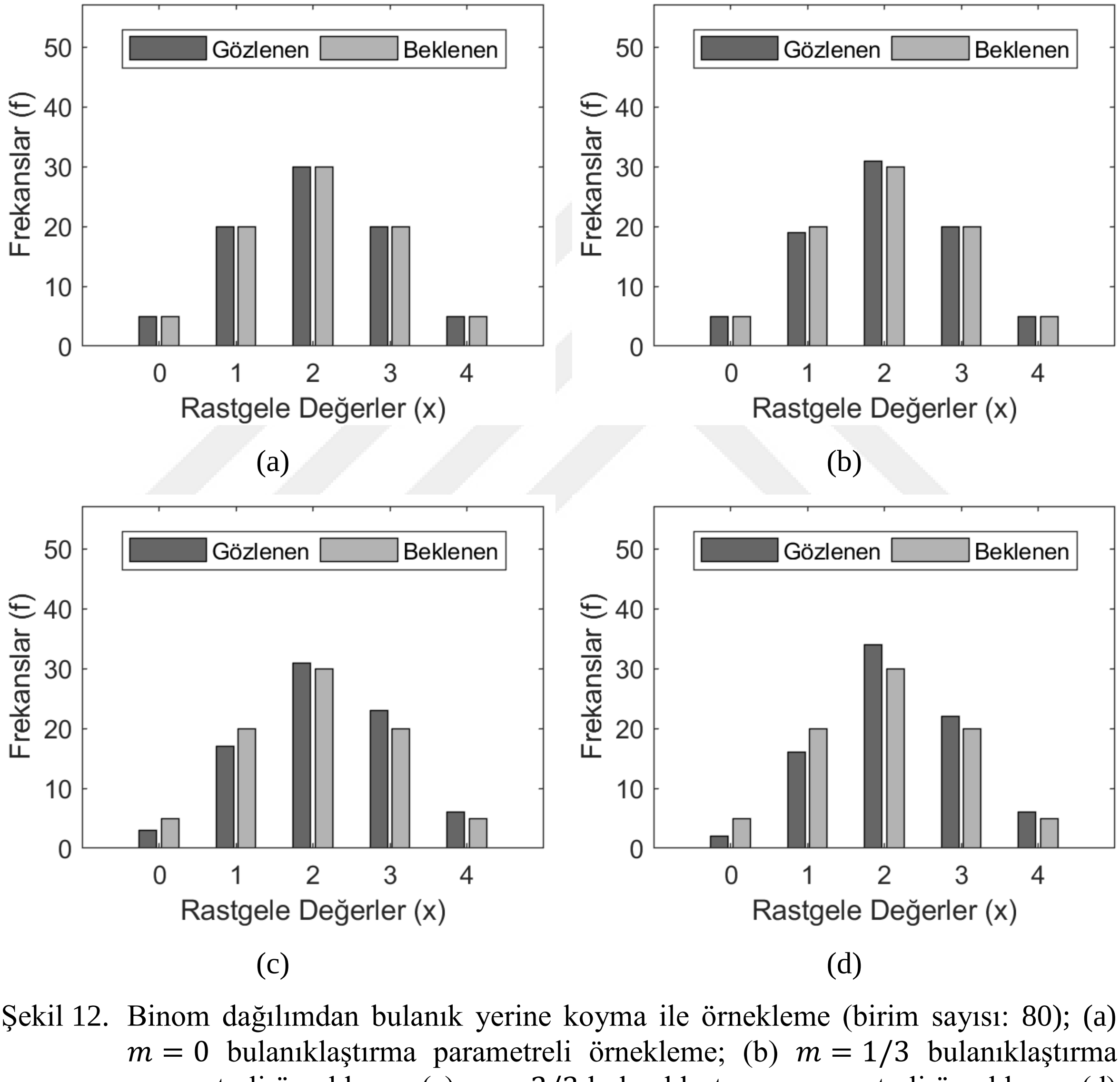
Şekil 10. Binom dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 16); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \binom{n}{x} p^x q^{n-x}\ ;\ \ x = 0,1,2,3,n;$ $p = 0.5,\ n = 4$ olan bir kitledeki değerler [0,4] aralığında tanımlansın. Bu binom dağılımın oluşturduğu kitleden seçilen 32 birim için beklenen ve gözlenen frekansların histogramı Şekil 11'de verilmektedir. Şekil 11 (a) ve (b)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 11(c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



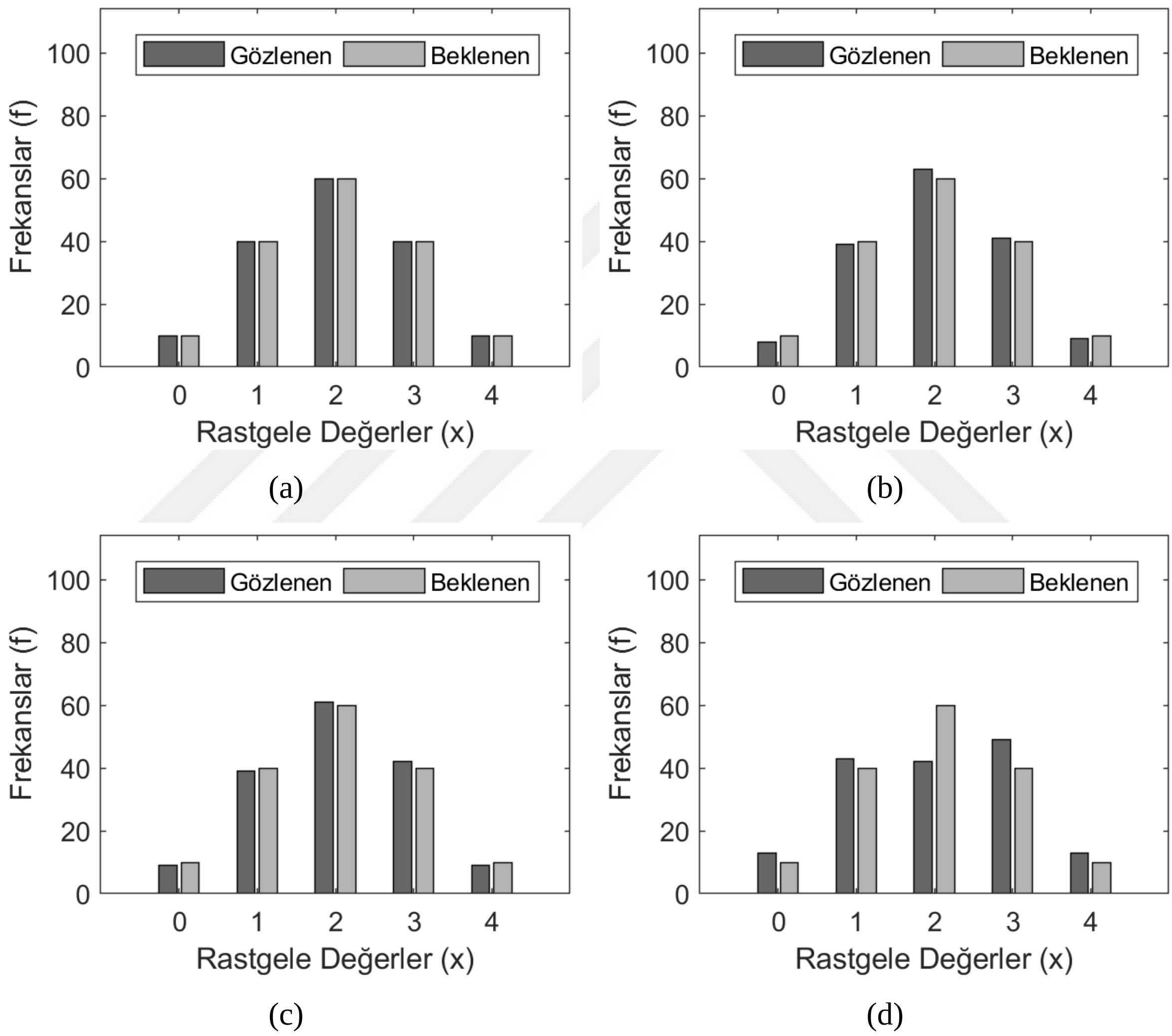
Şekil 11. Binom dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 32); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \binom{n}{x} p^x q^{n-x}$; $x = 0,1,2,3, n$; $p = 0.5$, $n = 4$ olan bir kitledeki değerler [0,4] aralığında tanımlansın. Bu binom dağılımın oluşturduğu kitleden seçilen 80 birim için beklenen ve gözlenen frekansların histogramı Şekil 12'de verilmektedir. Şekil 12 (b)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 12(c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 12. Binom dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 80); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \binom{n}{x} p^x q^{n-x}$; $x = 0,1,2,3,n;$ $p = 0.5,$ $n = 4$ olan bir kitledeki değerler [0,4] aralığında tanımlansın. Bu binom dağılımın oluşturduğu kitleden seçilen 160 birim için beklenen ve gözlenen frekansların histogramı Şekil 13'te verilmektedir. Şekil 13 (a) ve (b)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olduğu görülmektedir. Şekil 13 (c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.

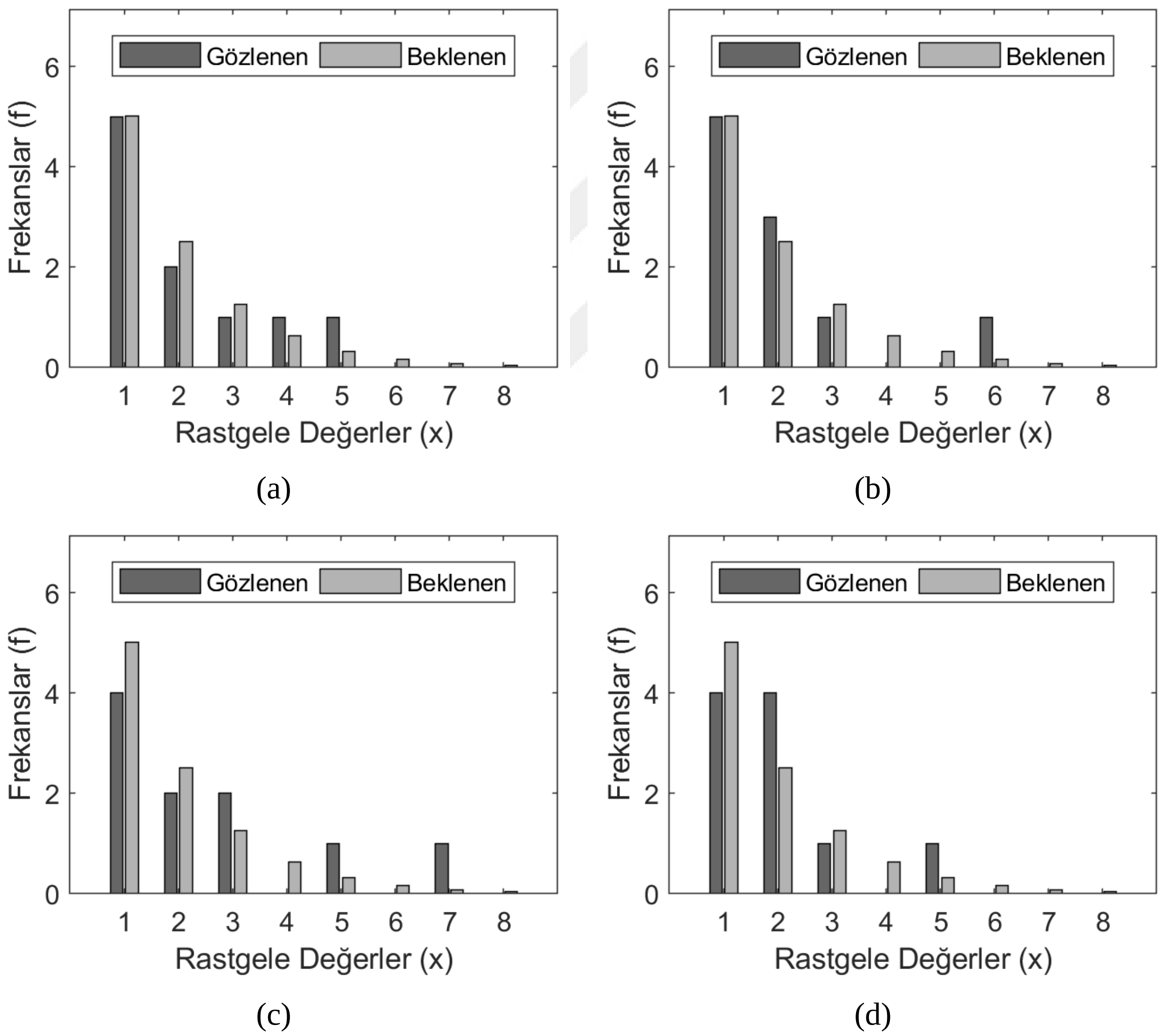


(a) (b) (c) (d)

Şekil 13. Binom dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 160); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

3.1.4. Geometrik Dağılımdan Rastgele Örnekleme

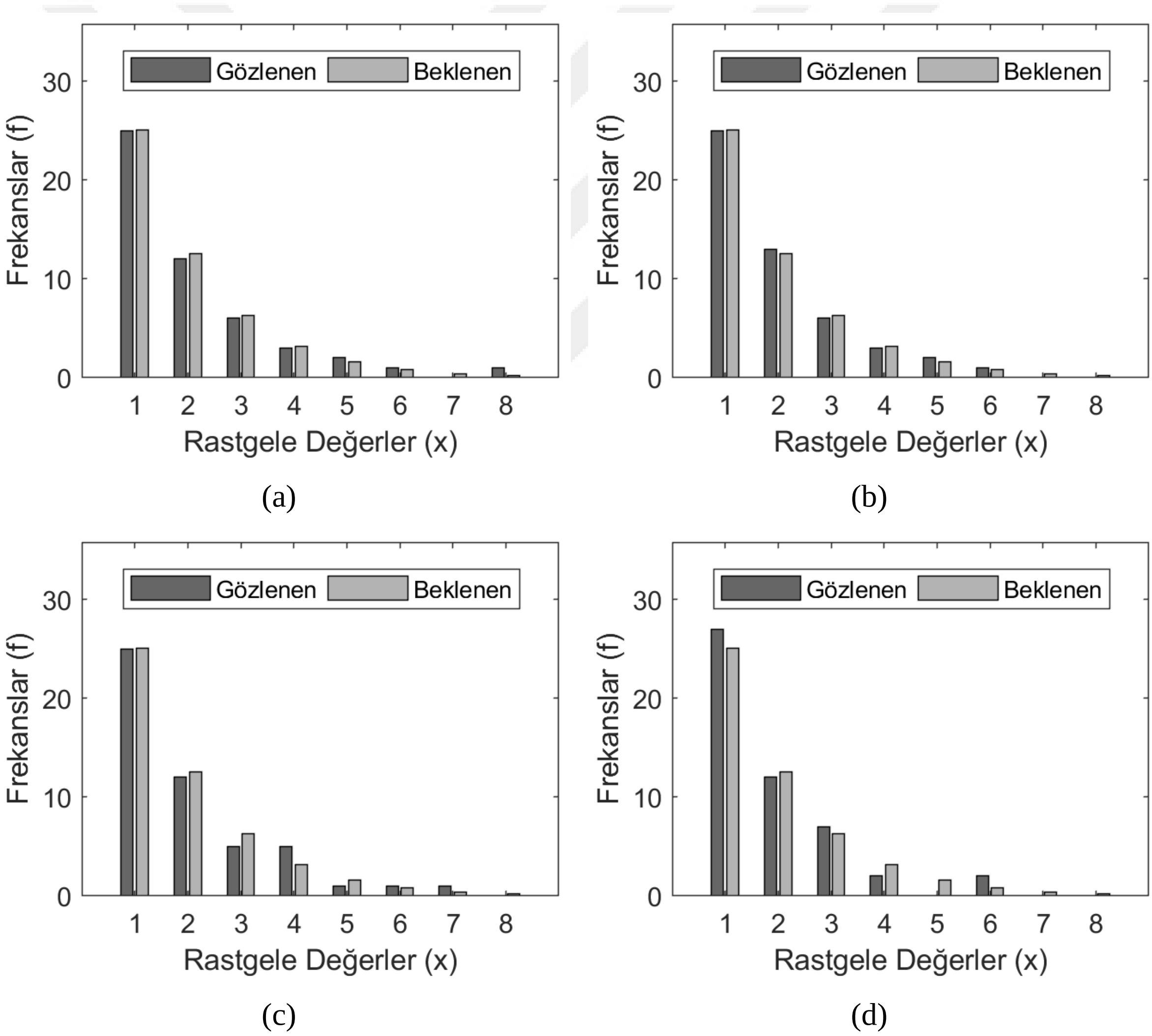
Olasılık yoğunluk fonksiyonu $f(x) = p(1-p)^{x-1}$; $x = 1,2 \ldots ,8$, $p = 0.5$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu geometrik dağılımın oluşturduğu kitleden seçilen 10 birim için beklenen ve gözlenen frekansların histogramı Şekil 14'te verilmektedir. Şekil 14 (a) ve (b)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olması beklenmektedir. Oluşan hata ise olasılık değerlerinin 10 ile çarpımında tam değerler olmadığından kaynaklanmaktadır.



Şekil 14. Geometrik dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 10); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

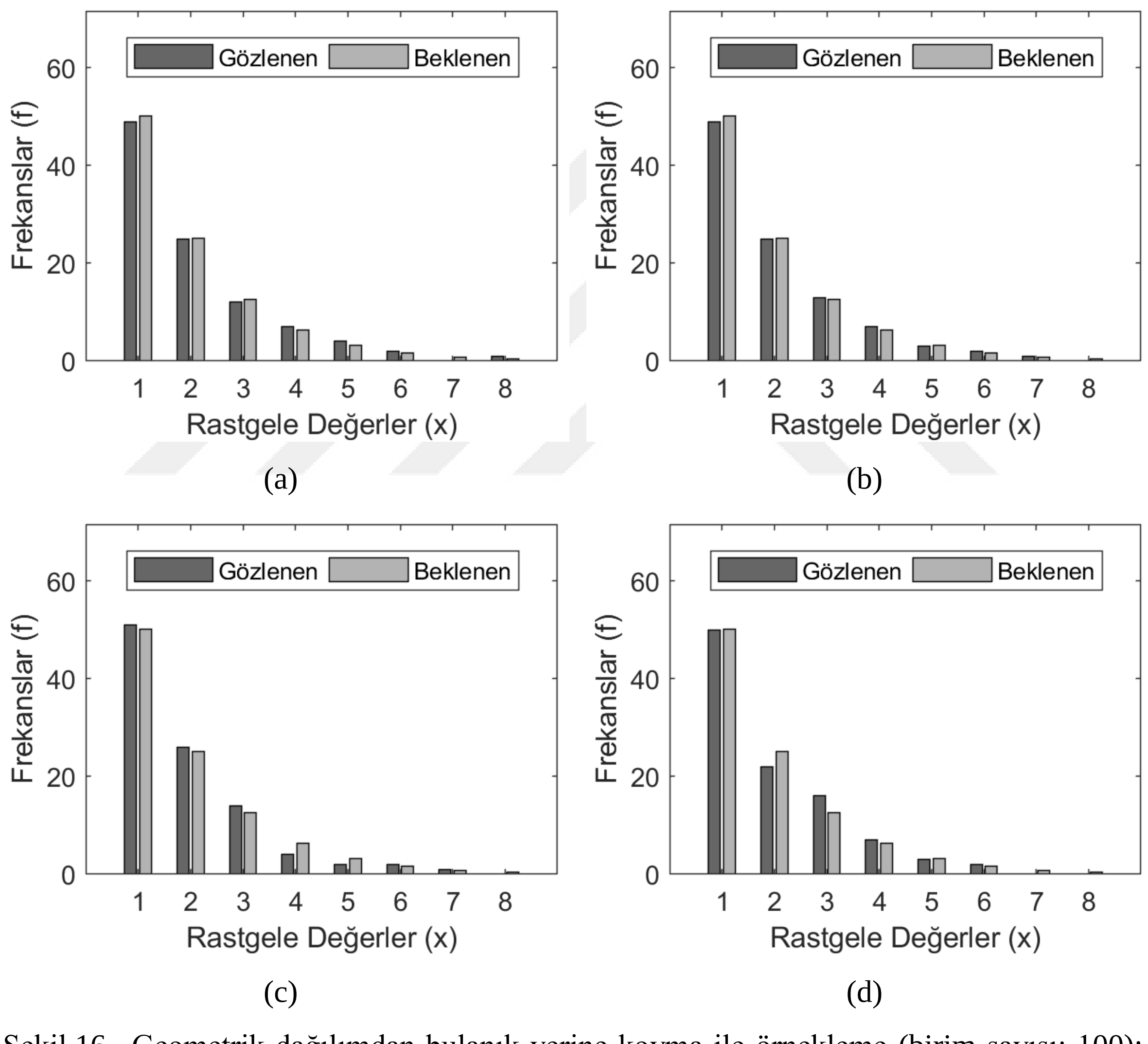
Şekil 14(c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.

Olasılık yoğunluk fonksiyonu $f(x) = p(1-p)^{x-1}$; $x = 1,2 \dots ,8,\ p = 0.5$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu geometrik dağılımın oluşturduğu kitleden seçilen 50 birim için beklenen ve gözlenen frekansların histogramı Şekil 15'te verilmektedir. Şekil 15(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



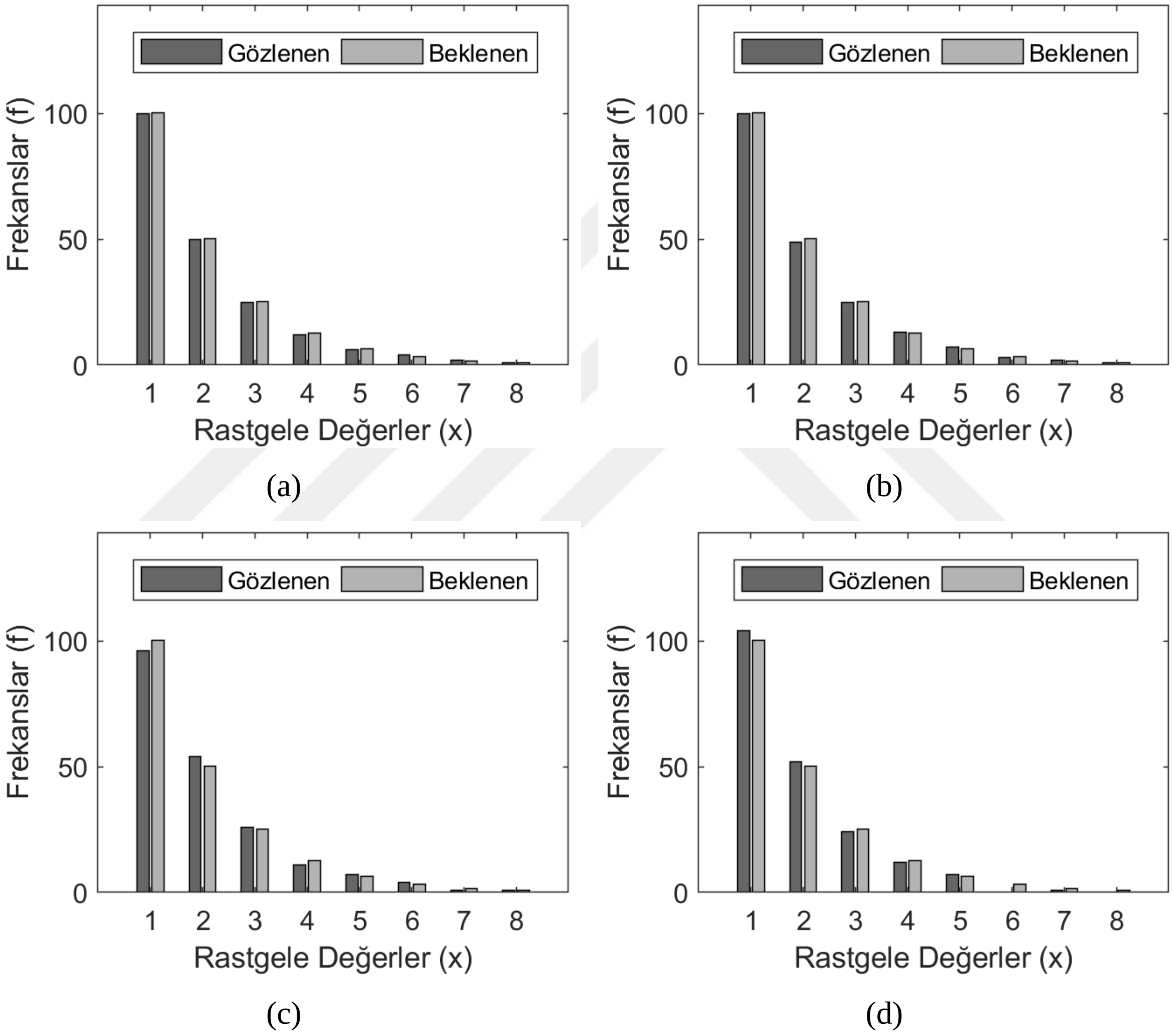
Şekil 15. Geometrik dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 50); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = p(1-p)^{x-1}$; $x = 1,2 \ldots ,8,\ p = 0.5$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu geometrik dağılımın oluşturduğu kitleden seçilen 100 birim için beklenen ve gözlenen frekansların histogramı Şekil 16'da verilmektedir. Şekil 16(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 16. Geometrik dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 100); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

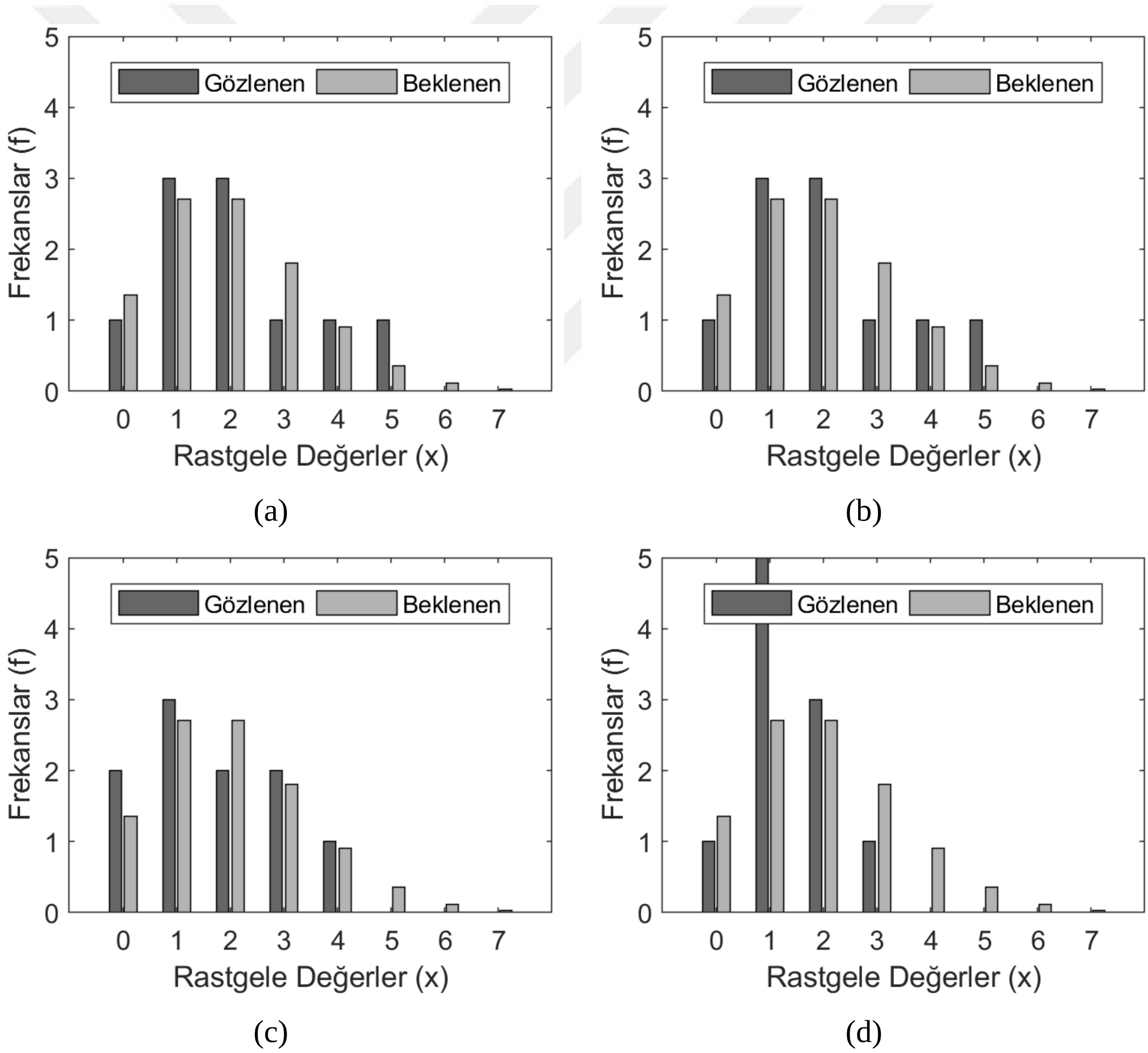
Olasılık yoğunluk fonksiyonu $f(x) = p(1-p)^{x-1}$; $x = 1,2 \dots ,8,\ p = 0.5$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu geometrik dağılımın oluşturduğu kitleden seçilen 200 birim için beklenen ve gözlenen frekansların histogramı Şekil 17'de verilmektedir. Şekil 17(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 17. Geometrik dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 200); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

3.1.5. Poisson Dağılımından Rastgele Örnekleme

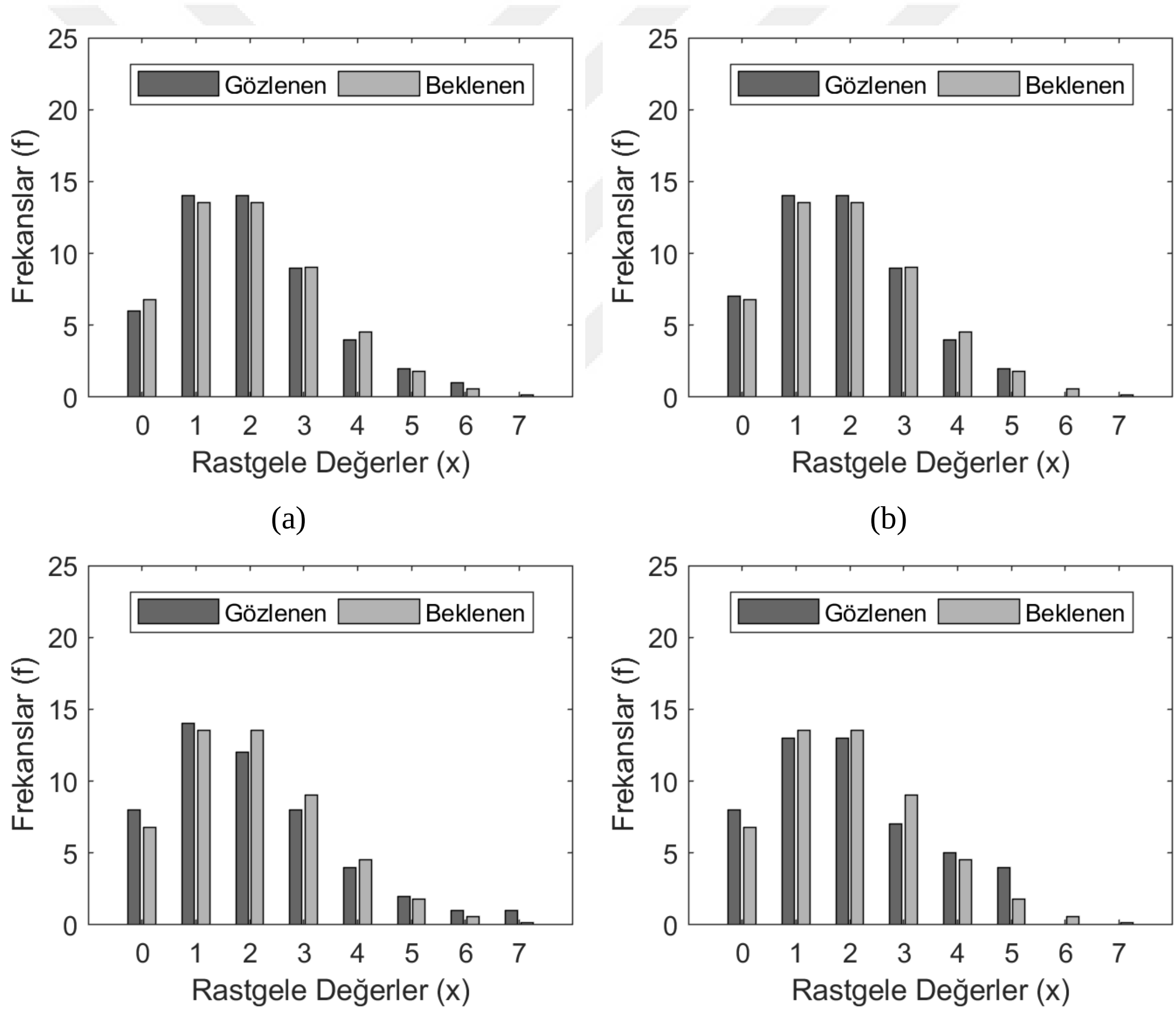
Olasılık yoğunluk fonksiyonu $f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$; $x = 0,1,2 \dots,7,\ \lambda = 2$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu poisson dağılımın oluşturduğu kitleden seçilen 10 birim için beklenen ve gözlenen frekansların histogramı Şekil 18'de verilmektedir. Şekil 18 (a) ve (b)'de bulanıklaştırma değeri $m = 0$ olan benzetimden elde edilen gözlenen ve beklenen değerlerin aynı değer ve hatanın (sapmanın) 0 olması beklenmektedir. Oluşan hata ise olasılık değerlerinin 10 ile çarpımında tam değerler olmadığından kaynaklanmaktadır.



Şekil 18. Poisson dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 10); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Şekil 18(c) ve (d)'de ise bulanıklaştırma parametreleri $m = 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.

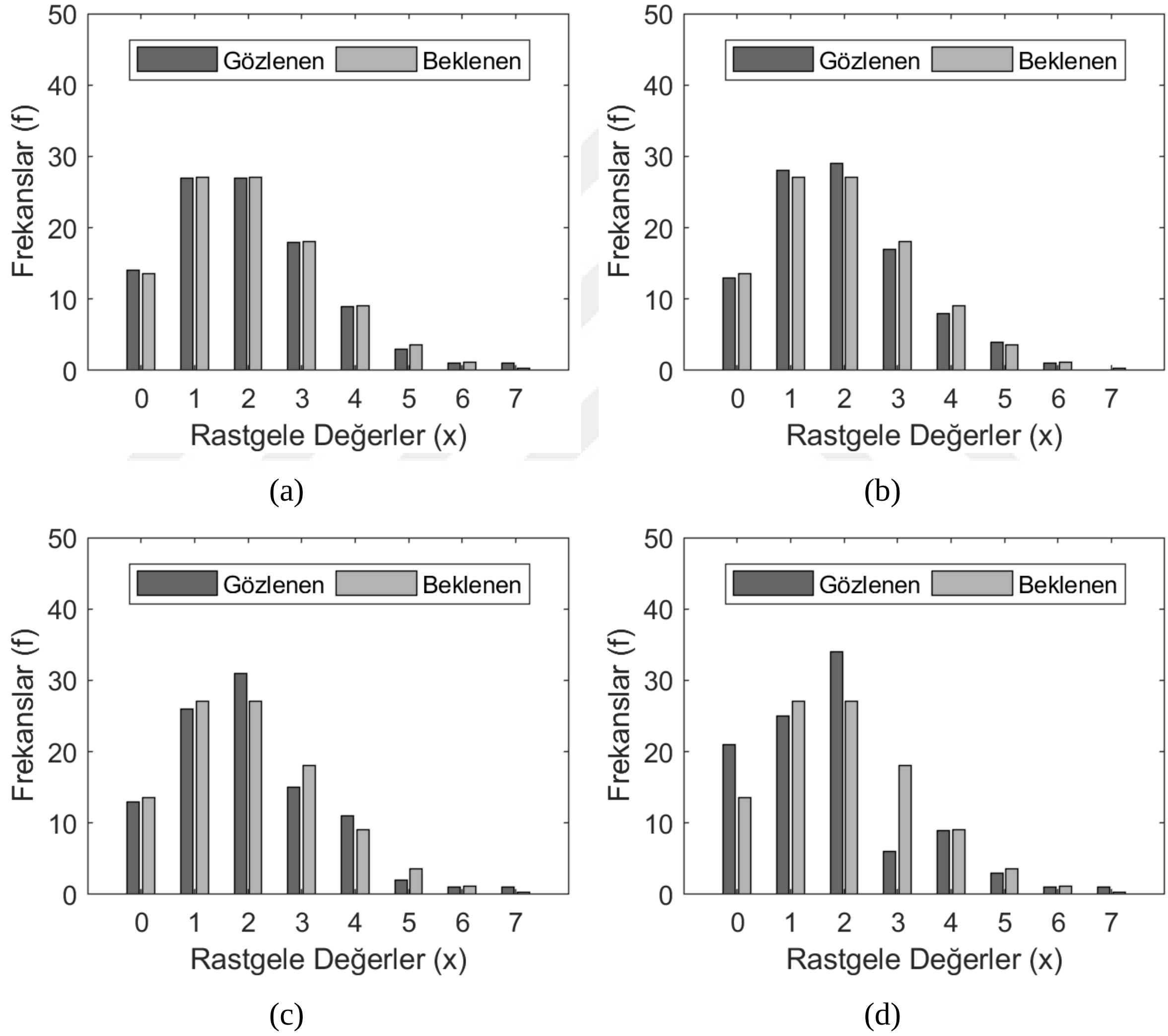
Olasılık yoğunluk fonksiyonu $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x = 0,1,2 \ldots ,7,\ \lambda = 2$ olan bir kitledeki değerler [0,9] aralığında tanımlansın. Bu poisson dağılımın oluşturduğu kitleden seçilen 50 birim için beklenen ve gözlenen frekansların histogramı Şekil 19'da verilmektedir. Şekil 19(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



(a) (b) (c) (d)

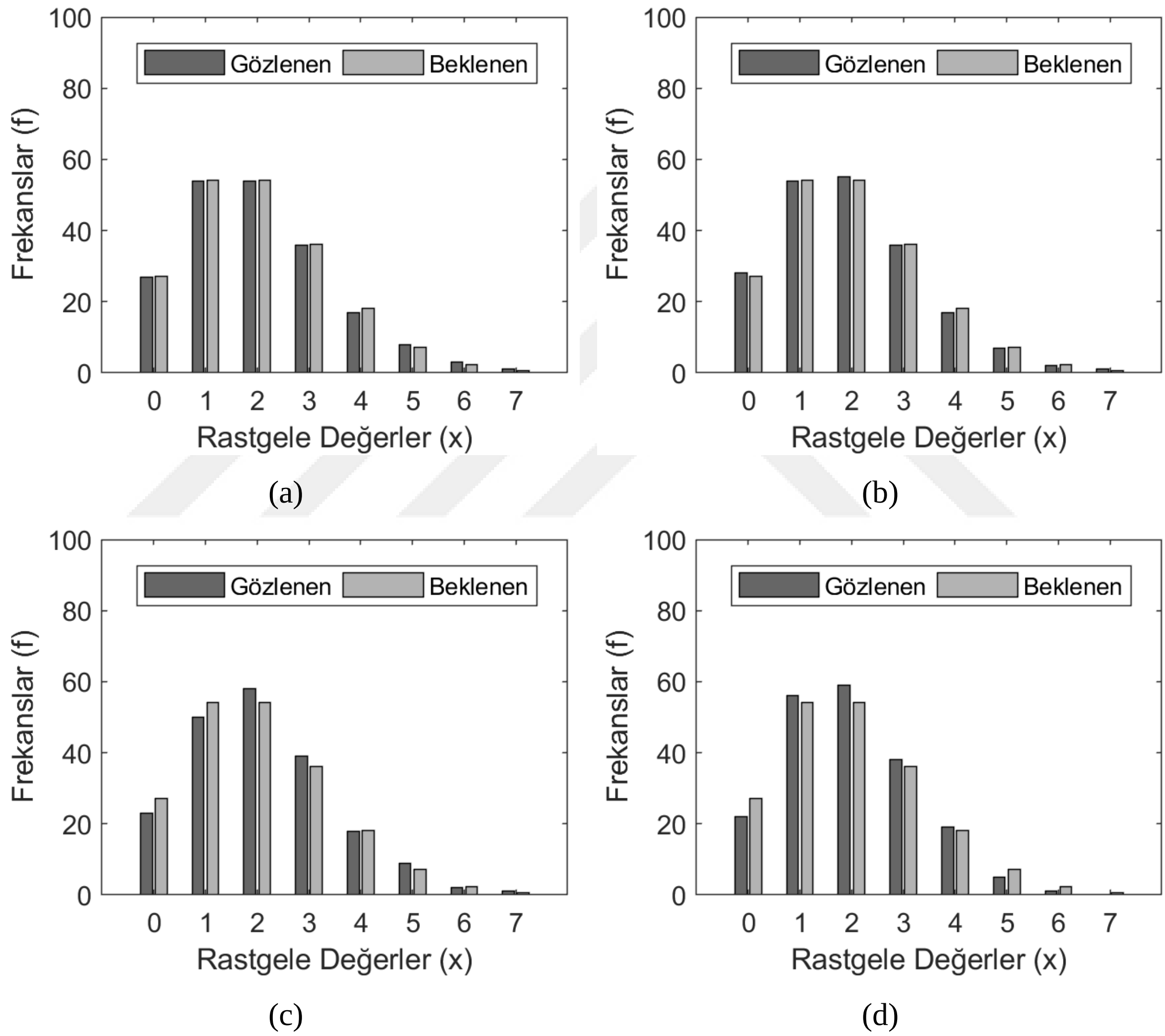
Şekil 19. Poisson dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 50); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x = 0,1,2 \dots,7,\ \lambda = 2$ olan bir kitledeki değerler $[0,7]$ aralığında tanımlansın. Bu poisson dağılımın oluşturduğu kitleden seçilen 100 birim için beklenen ve gözlenen frekansların histogramı Şekil 20'de verilmektedir. Şekil 20(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 20. Poisson dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 100); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

Olasılık yoğunluk fonksiyonu $f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$; $x = 0,1,2 \ldots ,7,\ \lambda = 2$ olan bir kitledeki değerler $[0,7]$ aralığında tanımlansın. Bu poisson dağılımın oluşturduğu kitleden seçilen 200 birim için beklenen ve gözlenen frekansların histogramı Şekil 21'de verilmektedir. Şekil 21(a)-(d)'de ise bulanıklaştırma parametreleri $m = 0, 1/3, 2/3, 1$ olan benzetimden elde edilen gözlenen ve beklenen değerler giderek değişkenlik göstermeye başlamıştır.



Şekil 21. Poisson dağılımdan bulanık yerine koyma ile örnekleme (birim sayısı: 200); (a) $m = 0$ bulanıklaştırma parametreli örnekleme; (b) $m = 1/3$ bulanıklaştırma parametreli örnekleme; (c) $m = 2/3$ bulanıklaştırma parametreli örnekleme; (d) $m = 1$ bulanıklaştırma parametreli örnekleme.

3.2. Bulanık Yerine Koymalı Örneklemin Başarım Ölçütü

Bulanık yerine koyma tekniği ile rastgele örneklemede beklenen ve gözlenen frekanslar arasında hatanın normal dağıldığı bilinmektedir (Thompson, 2012). Ancak toplam hata bulunmak istendiğinde bu hataların varyansları beklenen frekansla ilişki olduğundan göreceli bir ölçeğe ihtiyaç duyulmaktadır. Hataların göreceli olarak hesaplanabilmesi için ki-kare istatistiği tercih edilmiştir. Ki-kare istatistiği,

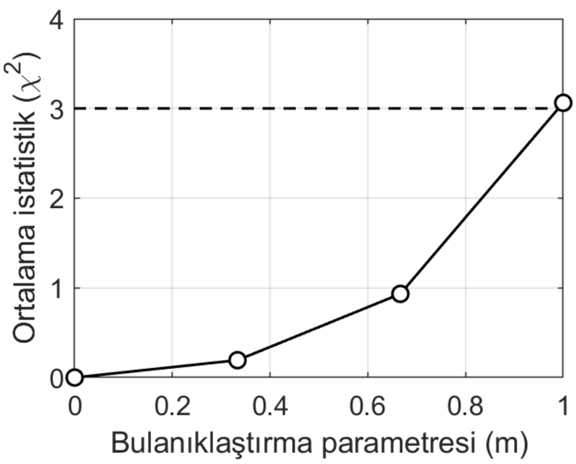
$$\chi_N^2 = \sum_{j=1}^{N} \frac{\left(O_j - E_j\right)^2}{E_i} \tag{22}$$

biçiminde verilmektedir. Burada O_j gözlenen frekansları gösterirken, E_i beklenen frekansları göstermektedir.

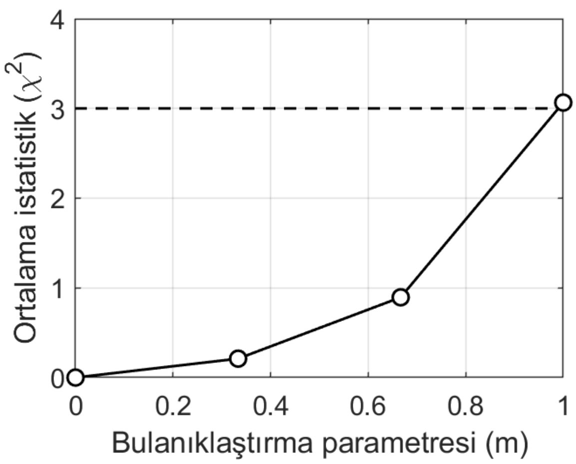
Bölüm 3.1.'de 5 ayrı dağılım fonksiyonu ile bulanık yerine koyma ile örneklemler yapıldı.

3.2.1. Keyfi Dağılımdan Bulanık Örneklemenin Başarımı

Bu örneklemeler keyfi dağılım için 1000 kez tekrarlanarak ki-kare istatistikleri elde edildi. Bu istatistiklerin ortalaması alınarak grafikleri aşağıda verilmiştir.



(a)



(b)

Şekil 22. Keyfi dağılıma göre yapılan benzetimdeki hatalar; (a) birim sayısı 10 olduğundaki sonuçlar; (b) birim sayısı 50 olduğundaki sonuçlar.

Elde edilen ki-kare istatistiklerine göre 1000 kez yapılan ki-kare uyum iyiliği testlerinin başarımları Tablo 7'de verilmiştir.

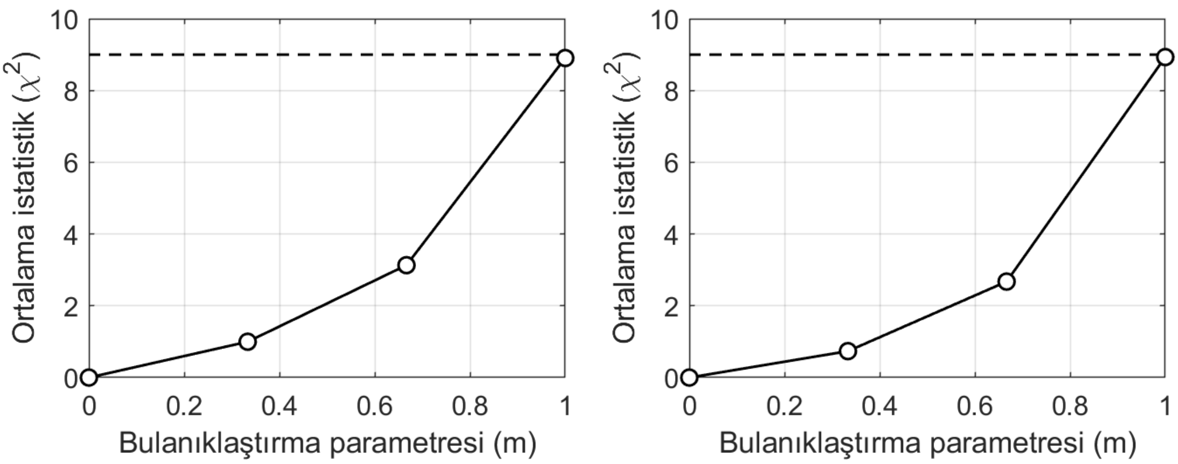
Tablo 7. Keyfi dağılımdan bulanık örneklemenin ki-kare uyum iyiliği sonuçları

		m			
		0	1/3	2/3	1
Kabul Sayısı	n=10	1000	1000	1000	901
	n=50	1000	1000	1000	892
Ortalama p	n=10	1.0000	0.9660	0.8015	0.4882
	n=50	1.0000	0.9706	0.8282	0.4980

Yukarıdaki tabloda $m = 1$ için yapılan ki-kare testinde $\alpha = 0.1$ alındığından 1000 denemede yaklaşık 900 kabul edilmesi yöntemin başarımını ortaya koymaktadır. Ayrıca m değeri küçüldükçe hipotezlerin kabul değerleri 1000 olmaktadır. p değerleri ise m değerleri küçüldükçe 0.5'ten 1'e doğru yakınsamaktadır.

3.2.2. Düzgün Dağılımdan Bulanık Örneklemenin Başarımı

Bu örneklemeler düzgün dağılım için 1000 kez tekrarlanarak ki-kare istatistikleri elde edildi. Bu istatistiklerin ortalaması alınarak grafikleri aşağıda verilmiştir.



(a) (b)

Şekil 23. Düzgün dağılıma göre yapılan benzetimdeki hatalar; (a) birim sayısı 10 olduğundaki sonuçlar; (b) birim sayısı 50 olduğundaki sonuçlar.

Elde edilen ki-kare istatistiklerine göre 1000 kez yapılan ki-kare uyum iyiliği testlerinin başarımları Tablo 8'de verilmiştir.

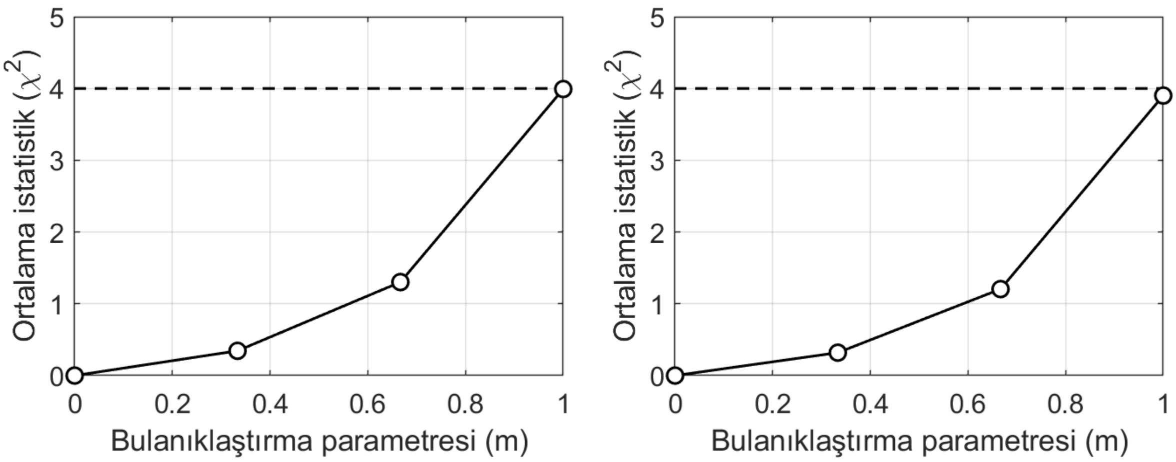
Tablo 8. Düzgün dağılımdan bulanık örneklemenin ki-kare uyum iyiliği sonuçları

		m			
		0	1/3	2/3	1
Kabul Sayısı	n=10	1000	1000	1000	916
	n=50	1000	1000	1000	892
Ortalama p	n=10	1.0000	0.9959	0.9290	0.4996
	n=50	1.0000	0.9995	0.9574	0.4954

Yukarıdaki tabloda $m = 1$ için yapılan ki-kare testinde $\alpha = 0.1$ alındığından 1000 denemede yaklaşık 900 kabul edilmesi yöntemin başarımını ortaya koymaktadır. Ayrıca m değeri küçüldükçe hipotezlerin kabul değerleri 1000 olmaktadır. p değerleri ise m değerleri küçüldükçe 0.5'ten 1'e doğru yakınsamaktadır.

3.2.3. Binom Dağılımdan Bulanık Örneklemenin Başarımı

Bu örneklemeler binom dağılımı için 1000 kez tekrarlanarak ki-kare istatistikleri elde edildi. Bu istatistiklerin ortalaması alınarak grafikleri aşağıda verilmiştir.



(a) (b)

Şekil 24. Binom dağılıma göre yapılan benzetimdeki hatalar; (a) birim sayısı 16 olduğundaki sonuçlar; (b) birim sayısı 32 olduğundaki sonuçlar.

Elde edilen ki-kare istatistiklerine göre 1000 kez yapılan ki-kare uyum iyiliği testlerinin başarımları Tablo 9'da verilmiştir.

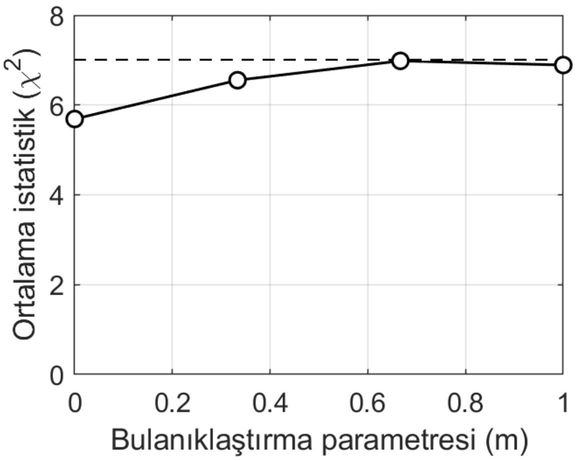
Tablo 9. Binom dağılımdan bulanık örneklemenin ki-kare uyum iyiliği sonuçları

		m			
		0	1/3	2/3	1
Kabul Sayısı	n=16	1000	1000	1000	910
	n=32	1000	1000	1000	912
Ortalama p	n=16	1.0000	0.9714	0.8420	0.5017
	n=32	1.0000	0.9805	0.8644	0.5082

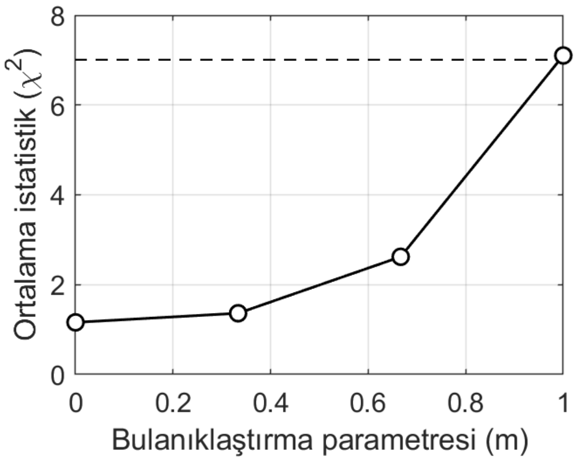
Yukarıdaki tabloda $m = 1$ için yapılan ki-kare testinde $\alpha = 0.1$ alındığından 1000 denemede yaklaşık 900 kabul edilmesi yöntemin başarımını ortaya koymaktadır. Ayrıca m değeri küçüldükçe hipotezlerin kabul değerleri 1000 olmaktadır. p değerleri ise m değerleri küçüldükçe 0.5'ten 1'e doğru yakınsamaktadır.

3.2.4. Geometrik Dağılımdan Bulanık Örneklemenin Başarımı

Bu örneklemeler geometrik dağılım için 1000 kez tekrarlanarak ki-kare istatistikleri elde edildi. Bu istatistiklerin ortalaması alınarak grafikleri aşağıda verilmiştir.



(a)



(b)

Şekil 25. Geometrik dağılıma göre yapılan benzetimdeki hatalar; (a) birim sayısı 10 olduğundaki sonuçlar; (b) birim sayısı 100 olduğundaki sonuçlar.

Elde edilen ki-kare istatistiklerine göre 1000 kez yapılan ki-kare uyum iyiliği testlerinin başarımları Tablo 10'da verilmiştir.

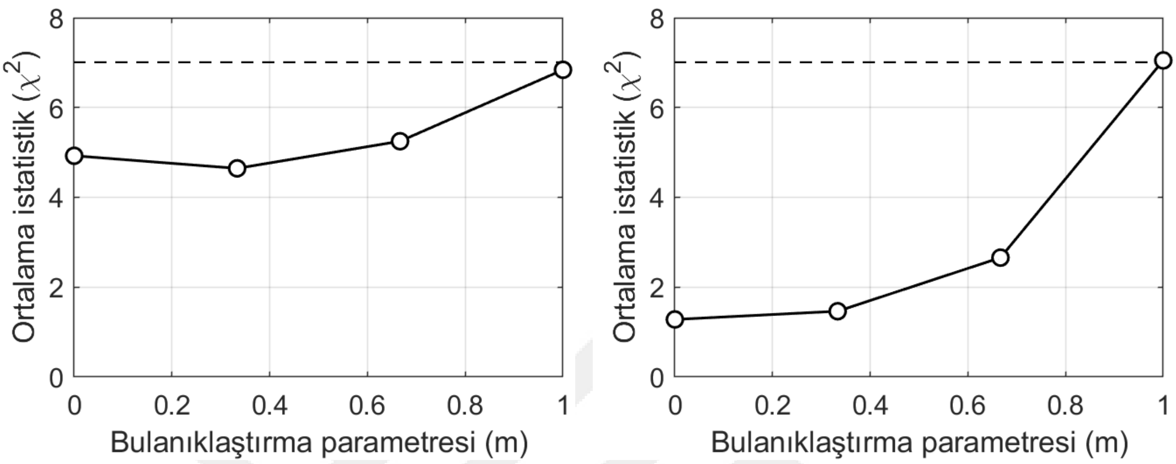
Tablo 10. Geometrik dağılımdan bulanık örneklemenin ki-kare uyum iyiliği sonuçları

		m			
		0	1/3	2/3	1
Kabul Sayısı	n=10	839	827	817	846
	n=100	1000	1000	999	893
Ortalama p	n=10	0.6855	0.6507	0.6415	0.6174
	n=100	0.9881	0.9821	0.8913	0.4979

Yukarıdaki tabloda $m = 1$ için yapılan ki-kare testinde $\alpha = 0.1$ alındığından 1000 denemede yaklaşık 900 kabul edilmesi beklenmektedir. Oysa beklenen frekansların tam olmaması durumunda küçük örneklemde sapmanın yüksek olmasına neden olmaktadır. Ayrıca m değeri küçüldükçe hipotezlerin kabul değerleri 1000 olmaktadır (yüksek örneklem büyüklüğünde). p değerleri ise m değerleri küçüldükçe yüksek örneklem büyüklüğünde 0.5'ten 1'e doğru yakınsamaktadır. Oysa küçük örneklem boyutlarında aynı şekilde sapmalar gözlenmiştir.

3.2.5. Poisson Dağılımdan Bulanık Örneklemenin Başarımı

Bu örneklemeler poisson dağılımı için 1000 kez tekrarlanarak ki-kare istatistikleri elde edildi. Bu istatistiklerin ortalaması alınarak grafikleri aşağıda verilmiştir.



(a) (b)

Şekil 26. Poisson dağılıma göre yapılan benzetimdeki hatalar; (a) birim sayısı 10 olduğundaki sonuçlar; (b) birim sayısı 100 olduğundaki sonuçlar.

Elde edilen ki-kare istatistiklerine göre 1000 kez yapılan ki-kare uyum iyiliği testlerinin başarımları Tablo 11'de verilmiştir.

Tablo 11. Poisson dağılımdan bulanık örneklemenin ki-kare uyum iyiliği sonuçları

		m			
		0	1/3	2/3	1
Kabul Sayısı	n=10	942	948	934	900
	n=50	1000	1000	1000	901
Ortalama p	n=10	0.7723	0.7850	0.7440	0.5867
	n=50	0.9843	0.9752	0.8910	0.5196

Yukarıdaki tabloda $m = 1$ için yapılan ki-kare testinde $\alpha = 0.1$ alındığından 1000 denemede yaklaşık 900 kabul edilmesi beklenmektedir. Ayrıca m değeri küçüldükçe hipotezlerin kabul değerleri 1000 olmaktadır (yüksek örneklem büyüklüğünde).Oysa küçük örneklem büyüklüğünde 1000 değerinden sapmalar gözlenmiştir. p değerleri ise m

değerleri küçüldükçe yüksek örneklem büyüklüğünde 0.5’ten 1’e doğru yakınsamaktadır. Oysa küçük örneklem boyutlarında aynı şekilde sapmalar gözlenmiştir.

3.3. Bulanık Örneklemede Ortalama Tahmini

Bölüm 3.1.’de 5 ayrı dağılım fonksiyonu ile bulanık yerine koyma ile örneklemler yapıldı. Öncelikli olarak bulanık yerine koyma tekniği ile rastgele örneklemede elde örnek değerlerinin ortalaması bulunmaktadır. Bu ortalamaları karşılaştırmak için dağılımın beklenen değeri

$$E(X) = \sum_{i=1}^{N} x_i p_i \tag{23}$$

biçiminde hesaplanırken varyansı,

$$\sigma^2 = Var(X) = E(X^2) - E(X)^2 \tag{24}$$

biçiminde bulunmaktadır. Yerine koyarak için ortalamanın varyansı ise,

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \tag{25}$$

biçiminde verilmektedir.

3.3.1. Keyfi Dağılımdan Bulanık Örneklemenin Başarımı

Keyfi dağılımın beklenen değeri,

$$\begin{aligned} E(X) &= 1 \times 0.3 + 2 \times 0.2 + 3 \times 0.4 + 4 \times 0.1 \\ &= 2.3 \end{aligned} \tag{26}$$

olarak bulunur. Varyansı,

$$Var(X) = 1^2 \times 0.3 + 2^2 \times 0.2 + 3^2 \times 0.4 + 4^2 \times 0.1 - 2.3^2 \\ = 1.01 \tag{27}$$

olarak bulunur. $n = 10$ ve $n = 50$ için yerine koyarak örnekleme için ortalamanın varyansı,

$$\sigma_{\bar{x}}^2 = \frac{1.01}{10} = 0.101 \tag{28}$$

ve

$$\sigma_{\bar{x}}^2 = \frac{1.01}{50} = 0.0202 \tag{29}$$

biçiminde bulunur. Bu örneklemeler keyfi dağılım için 1000 kez tekrarlanarak örneklem ortalamasının ortalamasını ve varyansı bulunmuştur. Elde edilen istatistikleri Tablo 12'de verilmiştir.

Tablo 12. Keyfi dağılımdan bulanık örneklemenin bazı istatistikleri

		m			
		0	1/3	2/3	1
Ortalama Değer	n=10	2.3000	2.2955	2.3132	2.3001
	n=50	2.3000	2.3016	2.3039	2.3056
Ortalama Değerin Varyansı	n=10	0.0000	0.0058	0.0297	0.0967
	n=50	0.0000	0.0014	0.0055	0.0197

Yukarıdaki tabloya göre ortalamanın beklenen değere yakınsadığı ve ortalama değerin varyansının $m = 1$ için beklenen varyansa yakınsadığı görülmektedir. m değeri küçüldükçe ortalama varyansın sıfıra yaklaştığı görülmektedir.

3.3.2. Düzgün Dağılımdan Bulanık Örneklemenin Başarımı

Düzgün dağılımın beklenen değeri,

$$
\begin{aligned}
E(X) &= \sum_{i=1}^{10} i.\frac{1}{10} \\
&= 5.5
\end{aligned}
\tag{30}
$$

olarak bulunur. Varyansı,

$$
\begin{aligned}
Var(X) &= \sum_{i=1}^{10} i^2.\frac{1}{10} - 5.5^2 \\
&= 8.25
\end{aligned}
\tag{31}
$$

olarak bulunur. $n = 10$ ve $n = 50$ için yerine koyarak örnekleme için ortalamanın varyansı,

$$\sigma_{\bar{x}}^2 = \frac{8.25}{10} = 0.825 \tag{32}$$

ve

$$\sigma_{\bar{x}}^2 = \frac{8.25}{50} = 0.1650 \tag{33}$$

biçiminde bulunur. Bu örneklemeler düzgün dağılım için 1000 kez tekrarlanarak örneklem ortalamasının ortalamasını ve varyansı bulunmuştur. Elde edilen istatistikleri Tablo 13'te verilmiştir.

Tablo 13. Düzgün dağılımdan bulanık örneklemenin bazı istatistikleri

	m			
	0	1/3	2/3	1

Ortalama Değer	n=10	5.5000	5.5144	5.4744	5.4982
	n=50	5.5000	5.4952	5.5160	5.4859
Ortalama Değerin Varyansı	n=10	0.0000	0.0804	0.2924	0.8924
	n=50	0.0000	0.0135	0.0514	0.1678

Yukarıdaki tabloya göre ortalamanın beklenen değere yakınsadığı ve ortalama değerin varyansının $m = 1$ için beklenen varyansa yakınsadığı görülmektedir. m değeri küçüldükçe ortalama varyansın sıfıra yaklaştığı görülmektedir.

3.3.3. Binom Dağılımdan Bulanık Örneklemenin Başarımı

Binom dağılımın beklenen değeri,

$$\begin{aligned} E(X) &= np = 4(0.5) \\ &= 2 \end{aligned} \tag{34}$$

olarak bulunur. Varyansı,

$$\begin{aligned} Var(X) &= np(1-p) = 4(0.5)(0.5) \\ &= 1 \end{aligned} \tag{35}$$

olarak bulunur. $n = 16$ ve $n = 32$ için yerine koyarak örnekleme için ortalamanın varyansı,

$$\sigma_{\bar{x}}^2 = \frac{1}{16} = 0.0625 \tag{36}$$

ve

$$\sigma_{\bar{x}}^2 = \frac{1}{32} = 0.0313 \tag{37}$$

biçiminde bulunur. Bu örneklemeler binom dağılım için 1000 kez tekrarlanarak örneklem ortalamasının ortalamasını ve varyansı bulunmuştur. Elde edilen istatistikleri Tablo 14'te verilmiştir.

Tablo 14. Binom dağılımdan bulanık örneklemenin bazı istatistikleri

		m			
		0	1/3	2/3	1
Ortalama Değer	n=16	2.0000	1.9967	1.9961	1.9973
	n=32	2.0000	1.9983	2.0057	2.0024
Ortalama Değerin Varyansı	n=16	0.0000	0.0049	0.0201	0.0614
	n=32	0.0000	0.0025	0.0092	0.0309

Yukarıdaki tabloya göre ortalamanın beklenen değere yakınsadığı ve ortalama değerin varyansının $m = 1$ için beklenen varyansa yakınsadığı görülmektedir. m değeri küçüldükçe ortalama varyansın sıfıra yaklaştığı görülmektedir.

3.3.4. Geometrik Dağılımdan Bulanık Örneklemenin Başarımı

Geometrik dağılımın beklenen değeri,

$$\begin{aligned} E(X) &= 1/p = 1/(0.5) \\ &= 2 \end{aligned} \tag{38}$$

olarak bulunur. Varyansı,

$$\begin{aligned} Var(X) &= \frac{1-p}{p^2} = \frac{0.5}{(0.5)^2} \\ &= 2 \end{aligned} \tag{39}$$

olarak bulunur. $n = 10$ ve $n = 100$ için yerine koyarak örnekleme için ortalamanın varyansı,

$$\sigma_{\bar{x}}^2 = \frac{2}{10} = 0.2 \tag{40}$$

ve

$$\sigma_{\bar{x}}^2 = \frac{2}{100} = 0.02 \tag{41}$$

biçiminde bulunur. Bu örneklemeler geometrik dağılım için 1000 kez tekrarlanarak örneklem ortalamasının ortalamasını ve varyansı bulunmuştur. Elde edilen istatistikleri Tablo 15Tablo 14'te verilmiştir.

Tablo 15. Geometrik dağılımdan bulanık örneklemenin bazı istatistikleri

		m			
		0	1/3	2/3	1
Ortalama Değer	n=10	2.1089	2.1741	2.2108	1.9772
	n=100	2.0031	1.9939	2.0080	1.9708
Ortalama Değerin Varyansı	n=10	0.0738	0.0963	0.1361	0.1780
	n=100	0.0007	0.0021	0.0056	0.0181

Yukarıdaki tabloya göre ortalamanın beklenen değere yakınsadığı ve ortalama değerin varyansının $m = 1$ için beklenen varyansa yakınsadığı görülmektedir. m değeri küçüldükçe ortalama varyansın sıfıra yaklaştığı görülmektedir. Ancak küçük örneklem boyutunda bu sapmalar olmaktadır.

3.3.5. Poisson Dağılımdan Bulanık Örneklemenin Başarımı

Poisson dağılımın beklenen değeri,

$$E(X) = \lambda = 2 \tag{42}$$

olarak bulunur. Varyansı,

$$Var(X) = \lambda = 2 \tag{43}$$

olarak bulunur. $n = 10$ ve $n = 100$ için yerine koyarak örnekleme için ortalamanın varyansı,

$$\sigma_{\bar{x}}^2 = \frac{2}{10} = 0.2 \tag{44}$$

ve

$$\sigma_{\bar{x}}^2 = \frac{2}{100} = 0.02 \tag{45}$$

biçiminde bulunur. Bu örneklemeler poisson dağılım için 1000 kez tekrarlanarak örneklem ortalamasının ortalamasını ve varyansı bulunmuştur. Elde edilen istatistikleri Tablo 16Tablo 14'da verilmiştir.

Tablo 16. Poisson dağılımdan bulanık örneklemenin bazı istatistikleri

		m			
		0	1/3	2/3	1
Ortalama Değer	n=10	3.1257	3.0952	3.1367	3.0102
	n=100	3.0127	3.0087	3.0160	2.9962
Ortalama Değerin Varyansı	n=10	0.0828	0.0864	0.1028	0.1809
	n=100	0.0011	0.0022	0.0060	0.0203

Yukarıdaki tabloya göre ortalamanın beklenen değere yakınsadığı ve ortalama değerin varyansının $m = 1$ için beklenen varyansa yakınsadığı görülmektedir. m değeri küçüldükçe ortalama varyansın sıfıra yaklaştığı görülmektedir. Ancak küçük örneklem boyutunda bu sapmalar olmaktadır.

4. SONUÇLAR

Günümüzde artan veriler ve artan bu verilerin büyük bir çoğunluğunun depolandığı bir çağda yaşıyoruz. Buna rağmen verilerin birçoğuna gelecekte olabileceği veya ulaşmaktaki zorluklardan dolayı çıkarım yapmak için örneklemeye ihtiyaç duyulmaktadır. Örneklem yapılırken genelde rastgele bir yaklaşım izlenmektedir. Bu ise seçilen örneklemin bazen kitleyi yansıtmadığı durumların ortaya çıkmasına neden olmaktadır. Kitle ile örneklem arasındaki bu farklılık rastgele gerçekleştiği için çoğu zaman bunlara müdahil olunamıyor. Bu çalışmada örneklem ile kitle arasında uyumun veya uyumsuzluğun derecesini bir bulanıklaştırma parametresi yardımıyla ayarlanabilmektedir.

Bulgular bölümünde 5 ayrı dağılım fonksiyonu ile bulanık yerine koyma ile örneklem yapıldı. Bu örneklemlerde bulanıklaştırma parametresi 0'dan 1'e gittikçe sapmaların arttığı gözlenmiştir. Bulanıklaştırma parametresinin 0 olması durumunda hatanın sıfır olması beklenirken bazı uygulamalarda sıfır olmadığı görüldü. Bu ise beklenen frekansların tam sayı olmadığından kaynaklanmaktadır. Eğer örneklem büyüklüğü yeterince büyük seçilirse veya Binom dağılım örneğinde olduğu gibi iyi seçilirse olasılık değerleri tam değer olacağından hata sıfırlanacaktır. Bu ise yerine koymadan örneklem istediği koşulları ortaya koyacaktır. Buna rağmen örneklem büyüklüğü farklı seçilmesinde bile hatalarda değişimlerin bulanıklaştırma parametresine bağlı olduğu görülmektedir.

Bulgular bölümünde önerilen yöntemin başarım ölçütü için yapılan benzetimlerde bulanıklaştırma parametresi $m = 0$ olduğunda hatanın sıfır veya sıfıra yakın çıktığı görülmektedir. Bulanıklaştırma parametresi $m = 1$ olduğunda ise hatanın beklenen değeri $N - 1$ serbestlik derecesine sahip ki-kare dağılımın beklenen değeri olan $N - 1$ değerine yakınsadığı görülmüştür.

MATLAB yazılım paketinde rastgele örnekleme,

```
y = randsample(population, n, replacement)
```

biçiminde veya

```
y = randsample(population, n, true, w)
```

biçiminde gösterilmektedir. Burada `population` kitleyi gösterirken `n` bu kitleden çekilecek birim sayısı vermektedir. `replacement` parametresi ise `true` veya `false` gibi iki farklı değer alabiliyor. True olduğu zaman yerine koyma, `false` olduğu zaman yerine koymama koşuluna göre örneklem yapılmaktadır. Eğer `replacement` parametresi `true` olduğunda w ağırlık (olasılık) değerleri verilebilmektedir. Bu parametre `false` olduğunda ağırlık değerleri verilmemektedir. Ayrıca örneklem sayısı kitle sayısından büyük olamamaktadır. Bu çalışmada. `replacement` parametresi bulanık verilebildiği gibi ağırlık değerleri de verilebilmektedir. Ayrıca örneklem sayısı istendiği kadar verilebilmektedir. Bu ise uygulamacıların bu konudaki sorunlarına çözüm oluşturabileceği amaçlanmaktadır.

MATLAB yazılım paketinde önerilen rastgele örnekleme komutu,

```
% eşit olasılıklı yerine koymadan (varsayılan m=0)
y = frandsample(population, n)

% eşit olasılıklı bulanık yerine koyma
y = frandsample(population, n, m)

% keyfi olasılıklı bulanık yerine koyma
y = frandsample(population, n, m, p)
```

biçiminde kullanılabilir. Burada

$m = 0$	olduğunda yerine koymama,
$m \in [0,1]$	olduğunda bulanık yerine koyma
$m = 1$	olduğunda tam yerine koyma

işlemi gerçekleştirilir.

5. KAYNAKLAR

Bowling, A., 2014. Research methods in health: investigating health and health services, McGraw-Hill Education (UK).

Cochran, W. G., 2007. Sampling techniques, John Wiley and Sons.

Çıngı, H., 1990. Örnekleme kuramı, Hacettepe Üniversitesi.

Devroye, L., 2013. Non-Uniform Random Variate Generation, Springer New York.

Fellegi, I. P., 1963. Sampling with varying probabilities without replacement: rotating and non-rotating samples, Journal of the American Statistical Association, 58, 301, 183-201.

Goodman, L. A., 1961. Snowball sampling, The annals of mathematical statistics, 148-170.

Grundy, P., 1954. A method of sampling with probability exactly proportional to size, Journal of the Royal Statistical Society. Series B (Methodological), 236-238.

Hansen, M. H. ve Hurwitz, W. N., 1943. On the theory of sampling from finite populations, The Annals of Mathematical Statistics, 14, 4, 333-362.

Hartley, H. ve Rao, J., 1962. Sampling with unequal probabilities and without replacement, The Annals of Mathematical Statistics, 350-374.

Hausman, J. A. ve Wise, D. A., 1981. Stratification on endogenous variables and estimation: The Gary income maintenance experiment, Structural analysis of discrete data with econometric applications, 365-391.

Horvitz, D. G. ve Thompson, D. J., 1952. A generalization of sampling without replacement from a finite universe, Journal of the American statistical Association, 47, 260, 663-685.

Imbens, G. W. ve Lancaster, T., 1996. Efficient estimation and stratified sampling, Journal of Econometrics, 74, 2, 289-318.

Kılıç, S., 2013. Örnekleme yöntemleri, Journal of Mood Disorders, 3, 1, 44-6.

Kothari, C. R., 2004. Research methodology: Methods and techniques, New Age International.

Madow, W. G., 1949. On the theory of systematic sampling, II, The Annals of Mathematical Statistics, 333-354.

Madow, W. G. ve Madow, L. H., 1944. On the theory of systematic sampling, I, The Annals of Mathematical Statistics, 15, 1, 1-24.

Moser, C. A., 1952. Quota sampling, Journal of the Royal Statistical Society. Series A (General), 115, 3, 411-423.

Neyman, J., 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection, Journal of the Royal Statistical Society, 97, 4, 558-625.

Olken, F. ve Rotem, D., 1986. Simple random sampling from relational databases.

Sedgwick, P., 2013. Convenience sampling, Bmj, 347, f6304, f6304.

Sen, A. R., 1953. On the estimate of the variance in sampling with varying probabilities, Journal of the Indian Society of Agricultural Statistics, 5, 1194, 127.

Thompson, S. K., 2012. Sampling, John Wiley & Sons, New Jersey.

Tillé, Y., 2006. Simple Random Sampling, Sampling Algorithms, 41-62.

Yamane, T., 2001. Temel Örnekleme Yöntemleri (Çevirenler: Esin, A., Aydın, C. Bakır, MA, Gürbüzsel, E.), Literatür Yayıncılık, İstanbul.

Yates, F. ve Grundy, P., 1953. Selection without replacement from within strata with probability proportional to size, Journal of the Royal Statistical Society. Series B (Methodological), 253-261.

ÖZGEÇMİŞ

Ebru NAZ, 31 Mayıs 1991 tarihinde Trabzon'da doğdu. Liseyi 2005-2009 yıllar arasında Trabzon Tevfik Serdar Anadolu Lisesi'nde okudu. Üniversiteyi 2009-2014 yılları arasında Ankara Hacettepe Üniversitesi İstatistik Bölümünü hazırlık programı ile beraber başarı ile tamamladı. 2013 Mart ayı ile Temmuz ayı arasında Erasmus Programı ile Almanya Bremen Üniversitesi Matematik Bölümünde programı başarı ile tamamladı. İş hayatına 01 Şubat 2015 yılında Garanti Bankası Genel Müdürlüğünde şubesiz bankacılıkta Dijital Pazarlama ve Satış Uzmanı olarak göreve başladı. Görevini 30 Haziran 2015'te sonlandırdı. 2015 yılında Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü İstatistik ve Bilgisayar Bilimleri Anabilim dalında tezli yüksek lisans programına başladı. Uluslararası sempozyumda sunulmuş bir bildirisi bulunmaktadır.