T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YENİ KESİKLİ DAĞILIMLAR İÇİN TİP-I SAĞDAN SANSÜRLÜ ÖRNEKLEME DAYALI PARAMETRE TAHMİNİ

Mehtap KOCA YILMAZ

YÜKSEK LİSANS TEZİ

İstatistik Anabilim Dalını

Temmuz-2019

KONYA

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Mehtap KOCA YILMAZ

Tarih: 18.07.2019

TEZ KABUL VE ONAYI

Mehtap KOCA YILMAZ tarafından hazırlanan "Yeni Kesikli Dağılımlar İçin Tip-I Sağdan Sansürlü Örnekleme Dayalı Parametre Tahmini" adlı tez çalışması 18/07/2019 tarihinde aşağıdaki jüri tarafından oy birliği ~~/ oy çokluğu~~ ile Selçuk Üniversitesi Fen Bilimleri Enstitüsü İSTATİSTİK Anabilim Dalı'nda YÜKSEK LİSANS olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Doç.Dr.Murat ERİŞOĞLU	....
Danışman Dr.Öğr.Üyesi Yunus AKDOĞAN	....
Üye Dr.Öğr.Üyesi Demet SEZER	....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mustafa YILMAZ
FBE Müdürü

Bu tez çalışması bilimsel Araştırma Projeleri Koordinatörlüğü tarafından 19201043 nolu proje ile desteklenmiştir.

ÖZET

YÜKSEK LİSANS TEZİ

YENİ KESİKLİ DAĞILIMLAR İÇİN TİP-I SAĞDAN SANSÜRLÜ ÖRNEKLEME DAYALI PARAMETRE TAHMİNİ

Mehtap KOCA YILMAZ

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Dr.Öğr.Üyesi Yunus AKDOĞAN

2019, 50 Sayfa

Jüri

Dr.Öğr.Üyesi Yunus AKDOĞAN
Doç.Dr. Murat ERİŞOĞLU
Dr.Öğr.Üyesi Demet SEZER

Bu tez çalışmasında, düzgün geometrik, geometrik-sıfırdan budanmış Poisson ve Binom-Lindley dağılımın bilinmeyen parametreleri için tip I sağdan sansürlü örnekleme dayalı parametre tahmini en çok olabilirlik, oranlar metodu ve modifiye edilmiş en çok olabilirlik yöntemleri kullanarak elde edilmiştir. Simulasyon çalışmasında tahmin edici yöntemlerinin performansı hata kareler ortalaması ve yan kriterlerine göre kıyaslanmıştır. Son olarak çalışmanın amacını daha iyi anlamak açısından her dağılım için gerçek örnek çalışmaya dahil edilmiştir.

Anahtar Kelimeler: En çok olabilirlik, Kesikli dağılımlar, modifiye edilmiş en çok olabilirlik, oranlar metodu, Tip-I sansürlü örneklem.

ABSTRACT

MS THESIS

PARAMETER ESTIMATION FOR SOME DISCRETE DISTRIBUTIONS BASED ON TYPE-I RIGHT CENSORED SAMPLE

Mehtap KOCA YILMAZ

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCEOF SELÇUK UNIVERSITY

THE DEGREE OF MASTER OF SCIENCE IN PHILOSOPHY STATISTICS

Advisor: Assit.Prof. Yunus AKDOĞAN

2019, 50 Pages

Jury

Assist. Prof. Yunus AKDOĞAN
Assoc. Prof. Murat ERİŞOĞLU
Assist. Prof. Demet SEZER

In this thesis, parameter estimation based on type I right censored sampling for unknown parameters Uniform Geometric, zero truncated Poisson and Binomial Lindley distribution is obtained using maximum likelihood, proportion method and modified maximum likelihood methods. In the simulation study, the performance of the estimator methods are compared according to the mean squares errors and bias criteria. Finally, in order to better understand the purpose of the study, the real data for each distribution is included in this thesis.

Keywords: Discrete distributions, maximum likelihood estimation, modified maximum likelihood estimation, proportions method, type-I censored sample.

ÖNSÖZ

Bu tez çalışması sürecinde değerli bilgi ve görüşlerini benimle paylaşan ve çalışmam süresince yardımlarını hiçbir zaman esirgemeyen değerli hocam sayın Yunus AKDOĞAN' a, maddi ve manevi her konuda destek olan ablam Alev Yüksel Aydar ve abim Ali Osman Aydar'a, bugünlere gelmemi sağlayan haklarını hiç bir zaman ödeyemeyeceğim annem ve babama, tanıdığım ilk günden bugüne kadar iyi-kötü her anımda yanımda olan canım eşim Adem Yılmaz'a ve varlığıyla hayatıma mutluluk katan biricik oğlum Yiğit Selim Yılmaz'a teşekkür ederim.

Mehtap KOCA YILMAZ
KONYA-2019

İÇİNDEKİLER

1. GİRİŞ VE KAYNAK ARAŞTIRMASI

Yaşam zamanı, güvenirlik analizinde ilgilenilen rasgele değişkendir. Gerçek hayatta bu değişken genellikle sürekli rasgele değişken olarak ele alınır. Literatüre baktığımız zaman güvenirlik ve yaşam zamanı deneylerinde çok sayıda sürekli model kullanılmıştır. Yaşam zamanlarının genelde sürekli olduğu gerçeğinin yanında kesikli olduğu durumlarla karşılaşılacağı muhakkaktır. Örneğin; kanser hastasının yaşayacağı Hafta sayısı, ya da çok ağır iş üstlenen bir devir dayım makinasının döngü sayısı, bir fotokopi makinasının tonerinin basım adet sayısı ve makinenin bozulmasında etkili şok sayısı gibi değişkenler kesikli rasgele değişkenlerdir. Bir elektrik lambasının anahtarı için açma-kapama sayısı, beyin tümörü olan hastanın yaşayacağı hafta sayısı yine kesikli bir rastgele değişkendir. Yaşam zamanlarının sürekli rasgele değişkenlerle ölçmenin imkansız ya da sakıncalı olabileceği durumlarda kesikli dağılımlar kullanılmıştır. Örneğin bir cep telefonunun pil ömrünün kaç kez şarj edildikten sonra biteceği veya bir meyve sıkacağının kaç kez açıp kapattıktan sonra bozulacağı gibi yaşam zamanlarını sürekli rasgele değişkenlerle ölçmek mümkün değildir. Bu gibi durumlar günlük yaşamda çoğu kez karşımıza çıkmaktadır. Bahsi geçen olayların varlığı ve bu olayların modellemesi ve modelleme üzerinden bazı sonuçların çıkarılması, kesikli yaşam zamanları ile ilgili istatistiksel analizler, önceleri çok bilinen Poisson, Geometrik ve Negatif Binom dağılımları gibi kesikli dağılımlar kullanılarak yapılmaktaydı. Bu dağılımların gerçek verileri modellemede yetersizliği yeni kesikli dağılımların önerilme ihtiyacını doğurmuştur. Bu doğrultuda birçok kesikli dağılım önerilmiş ve önerilen dağılımların beklenen değer, varyans, medyan gibi konum ve ölçek parametreleri ile dağılım, moment çıkaran ve bozulma(hazard) fonksiyonu gibi dağılımsal özellikleri incelenmiştir. Literatüre bakıldığında sürekli dağılımlardan çeşitli yöntemler kullanılarak yeni kesikli dağılımlar önerilmiştir. İlk olarak Nakagawa ve Osaki tarafından 1975 yılında Weibull dağılımını kesikli Weibull dağılımı olarak kesikleştiren çalışmanın ardından birçok sürekli dağılım aynı yada benzer yöntemlerle kesikli dağılım olarak sunulmuştur. (Nakagawa ve Osaki, 1975). Yine Stein ve Dattero 1984 yılında bir başka tipte Weibull dağılımını kesikleştirerek literatüre kazandırmışlardır. (Stein ve Dattero, 1984) 2003 yılına gelindiğinde Roy tarafından kesikli normal dağılım önerilmiştir. (Roy, 2003) Bir yıl aradan sonra yine Roy tarafından kesikli Rayleigh dağılımı elde edilmiştir. (Roy, 2004) 2009 yılına gelindiğinde Krishna ve Pundir tarafından aynı çalışmada hem kesikli Burr hem de kesikli Pareto dağılımlarını önermişlerdir. (Krishna ve Singh Pundir, 2009). (Jazi ve ark., 2010), çalışmalarında sürekli ters Weibull'un kesikli bir versiyonu olan kesikli ters Weibull dağılımı önermişlerdir. (Chakraborty ve Chakravarty, 2012), sürekli olasılık dağılımlarının kesikleştirilmesi için genel yaklaşım kullanılarak sürekli gamma dağılımına karşılık gelen kesikli gamma dağılımını önermişlerdir. Yine bu çalışmaları takiben, (Nekoukhou ve ark., 2013) yılında Kesikli Chen dağılımını, (Déniz, 2013), Binom-Poisson dağılımını, (Barbiero, 2014) yılında kesikli Skew-Laplace dağılımı, (Bakouch ve ark., 2014a) yılında kesikli Lindley dağılımını, (Akdoğan ve ark., 2016) yılında Düzgün-Geometrik dağılımını, (Akdoğan ve ark., 2019) yılında Binom-Lindley dağılımını, son olarak (Kuş ve ark., 2018) yılında geometrik-sıfırdan budanmış Poisson dağılımını elde etmişlerdir. Elde edilen çok sayıda yeni dağılıma ait çalışmalarda parametre tahmini ve gerçek veri uygulaması yapılmış ancak sansürlü veri ve güvenirlik analizi verilmemiştir.

Sansürlenmiş veriler genelde yaşam zamanı modellemelerinde kullanılmaktadır. Sansürleme, sağdan sansürleme ve soldan sansürleme olmak üzere iki ana gruba ayrılmaktadır. Gözlemlenecek yaşam süreleri henüz sona ermediği durumlarda sansürlenecek gözlemler, bulunulan zaman noktasından daha sonraki bir zamanda yani zaman ölçeğinde daha sağda bir yerde gerçekleşeceğinden sağdan sansürleme ismini almaktadır. Tip I sansürlemede her gözlemin bir sansürleme zamanının olduğu düşünülür. Belirli bir zamanda başlayıp daha önceden belirlenen bir bitirme zamanında bitirilmesi sonucu ortaya çıkan sansürleme türüdür. Önceden belirlenen zamana kadar ilgilenilen olayın henüz gerçekleşmediği gözlem noktaları, gözlemlerin durma zamanından sonra ne kadar daha devam edeceği gözlemci tarafından bilinmediğinden bu gözlemler Tip I sansürleme ile sansürlenmiş örneklem adını alırlar.

İstatistiksel sonuç çıkarımı yapılırken tüm birimlerin bozulma zamanlarını gözlemlemek her zaman mümkün olmayabilir. Örneğin bir tedavinin faydalı olup olmadığı araştırılırken tüm hastaların ölüm sürelerini tespit etmek çok zaman alabilir yada bunu tespit etmek hem maliyetli hemde uzun zaman gerektirebilir. Yaşam analizi sadece hastalık ve yaşamı etkileyen faktörlerin analizinde faydalanılan bir yöntem değildir. Ölüm veya hayatta kalma süreleri ile sınırlandırılmamalıdır. Örneğin; evli çiftlerin evli kalma süreleri, şirketlerin aldıkları iş makinelerinin bozulma süreleri veya işlevini göremez duruma gelme süreleri, makinelerin ardışık iki kez bozulma süreleri arasında geçen süre, elektronik parçaların yaşam sürelerinin analiz edilmesi gibi birçok alanda kullanılır. Tüm birimlerin başarısız olma zamanları tespit edilemediği durumlarda sansürlenmiş verilere ihtiyaç duyulur. Bir sistemin veya bir deneyin gerçekleşmesi aşamasında gözlenemeyen birimlerle ilgili verilerin yok sayılması işlemine sansürleme denilir. Son zamanlarda birçok alanda sansürlenmiş verilerle karşılaşılmaktadır. Daha önceki çalışmalarda, kesikli dağılımlarda yapılan çalışmalarda istatistiksel sonuç çıkarımı için Tip-I ve Tip-II sansürleme tipi kullanılmaktadır. (Kulasekera, 1994), Kesikli Weibull dağılımının parametrelerinin tahminini Tip-I sansürlü örnekleme dayalı olarak incelemiştir. (Bakouch ve ark., 2014a), kesikli Lindley dağılımı için Tip-I ve Tip-II sansürlü örnekleme dayalı parametre tahmini çalışılmıştır. (Akdoğan ve ark., 2014) çalışmalarında Tip I sansürlü örnekleme dayalı kesikli Burr dağılımı parametrelerinin en çok olabilirlik, oranlar ve modifiye edilmiş en çok olabilirlik tahmin edicilerini elde etmişlerdir. Bu tahmincilerin yan ve hata kareler ortalamaları açısından karşılaştırmak için Monte Carlo simülasyon çalışması yapmışlardır. Elde edilen teorik sonuçlar için sayısal bir örnek de sunmuşlardır.

Bu tez çalışmasında yeni elde edilen bu dağılımlar için Tip-I sansürlü örneklem durumu verilmeye çalışılmıştır. Yine tez bu çalışmasında düzgün geometrik, geometrik-sıfırdan budanmış Poisson ve Binom-Lindley dağılımı ele alınarak, bu dağılımların Tip-I sansürlü örneklem durumunda parametre tahmini önerilmiştir. Parametre tahmini için en çok olabilirlik, oranlar ve modifiye edilmiş en çok olabilirlik yöntemleri kullanılmıştır. Yine sansürleme yöntemlerinin etkinliği için gerçek örnekler üzerinde çalışmalar yapılmıştır. Tez çalışmasının ikinci bölümünde kullanılan bazı temel kavramlar verilmiş, üçüncü bölümde literatüre yeni girmiş düzgün geometrik, geometrik-sıfırdan budanmış Poisson ve Binom-Lindley dağılımı için parametre tahmini, simülasyon çalışması ve gerçek veri uygulaması verilmiştir. Son olarak tez çalışması sonuç ve öneriler kısmı ile sonlandırılmıştır.

2. TEMEL KAVRAMLAR

Bu bölümde tez çalışmasında dağılımlarda ve parametre tahmini yapılırken kullanılan bazı temel kavramlar verilmiştir.

2.1. Yaşam ve Güvenilirlik Fonksiyonu

Belirli bir hastalığa uygulanan tedavi çeşitleriyle hastanın ne kadar süre yaşayacağını tahmin etmek ya da çeşitli tedavi yöntemlerinin hastanın yaşam süresine olan etkisini araştırmak için geliştirilmiş yöntemler ‘yaşam analizi’ olarak adlandırılır. Yaşam analizi bazı uygulamalarda güvenilirlik analizi olarak da adlandırılmaktadır.

Bir bireyin ya da bir nesnenin belirli bir başlangıç zamanı ile ölümü arasında geçen zamana yaşam süresi denir. Herhangi bir birimin ya da bireyin yaşam süresinin bir değerden büyük olma olasılığına ise yaşam fonksiyonu adı verilir. Bir olayın meydana gelene kadar geçen süre X olsun, yaşam fonksiyonu;

$$S(x) = F(X > x) = 1 - F(x) \ , \ x \geq 0$$

olarak gösterilir.

2.2. Bozulma Oranı Fonksiyonu

x, zamanına kadar ölmediği bilinen bir bireyin ani ölüm olasılığı bozulma oranı fonksiyonu olarak adlandırılır ve $h(x)$ ile gösterilir.

$$\begin{aligned} h(x) &= \lim_{x \to 0} \frac{P(x \leq X < x+m \mid X > x)}{m} \\ &= \lim_{m \to 0} \frac{1}{m} \frac{P(x \leq X < x+m,\ X > x)}{P(X > x)} \\ &= \frac{1}{\bar{F}(x)} \lim_{m \to 0} \frac{P(x \leq X < x+m)}{m} \\ &= \frac{f(x)}{\bar{F}(x)} \end{aligned}$$

şeklindedir.

2.3. Dağılım Fonksiyonu

X rasgele değişkeninin dağılım fonksiyonu,

$$F(x) = P(X \leq x) = \int_0^x f(x)\,dx$$

şeklinde tanımlanır. Dağılım fonksiyonu azalmayan ve sağdan sürekli bir fonksiyondur. Burada f, olasılık yoğunluk fonksiyonudur. Ayrıca

$$\lim_{x \to 0} F(x) = 0$$

$$\lim_{x \to \infty} F(x) = 1$$

dir.

2.4. Tip I Sansürlü Örneklem

Gözlemlenen olayın durma zamanından sonra ne kadar daha devam edeceği gözlemci tarafından bilinmediği durumlarda, ilgili olayın gerçekleşmediği gözlem noktaları Tip I sansürleme ile sansürlenir.

$L_1, L_2, ..., L_n$ sabitleri sırasıyla $Y_1, Y_2, ..., Y_n$ rasgele değişkenlerine ait sağdan sansür zamanlarını göstersin. Ayrıca $i = 1, 2, ..., n$ için

$$X_i = \min(Y_i, L_i)$$

$$\delta_i = \begin{cases} 1 & , \quad Y_i \leq L_i \\ 0 & , \quad Y_i > L_i \end{cases}$$

şeklinde tanımlanan $X_1, X_2, ..., X_n$ örneklemine Tip I sansürlü örneklem denir. Burada δ_i, Y_i ‘nin sansürlü olup olmadığını göstermektedir. (x_i, δ_i) ‘nin ortak olasılık yoğunluk fonksiyonu

$$P(x_i, \delta_i) = f(x_i)^{\delta_i}\, p(X_i > L_i)^{1-\delta_i}$$

$$= f(x_i)^{\delta_i}\, S(X_i)^{1-\delta_i}$$

şeklinde tanımlanır.

2.5. Nokta tahmini

Örneklem istatistiğine dayanılarak ana kitle parametresinin tek bir değer ile tahmin edilmesine nokta tahmini denir.

Örneklemin bilinmeyen parametre içermeyen fonksiyonuna istatistik denir. İstatistikler aynı zamanda birer rasgele değişkendir. Bir istatistik bir parametreyi veya parametrenin bir fonksiyonunu tahmin etmek amacıyla kullanıldığında tahmin edici adını alır. Tahmin edicinin aldığı değere de tahmin denir.

2.5.1. En Çok Olabilirlik Tahmini

$f(x, y), \gamma \in \Gamma$ dağılımına sahip kitlenin parametresi tahmin edilmek istensin. Burada Γ parametre uzayını , γ kitle parametresini temsil etmektedir. Bu kitleden alınan ve her biri aynı dağılıma sahip $X_1, X_2, ..., X_n$ rasgele değişkenlerin dizisine örneklem adı verilir.

Örneklemin olasılık (yoğunluk) fonksiyonu,

$$L(x, \gamma) = f(x_1, x_2, ..., x_n, \gamma)$$

şeklindedir. $L(x, \gamma)$, γ'nın bir fonksiyonu olarak düşünüldüğünde olabilirlik fonksiyonu adını alır. $X_1, X_2, \ldots, X_n$, $f(x, y), \gamma \in \Gamma \subset R^r$ dağılımından alınmış n birimlik örneklem olmak üzere $L(\hat{\gamma} \mid x) = \sup_{\gamma \in \Gamma} (L(\gamma \mid x))$ olacak şekilde elde edilen $\hat{\gamma} = \hat{\gamma}(X_1, X_2, ..., X_n)$ istatistiğine γ'nın en çok olabilirlik tahmin edicisi denir.

2.5.2. Oranlar Metodu Tahmini

Oranlar metodu kesikli Weibull dağılımının parametrelerinin tahmini için Khan ve arkadaşları tarafından 1989 yılında önerilmiştir. (Khan ve ark., 1989) Oranlar tahmini metodu daha iyi anlaşılabilmesi için teorisi düzgün geometrik dağılım üzerinden anlatılacaktır.

$X_1, X_2, ..., X_n$ düzgün geometrik dağılımdan alınmış rasgele örneklem olsun, $i = 1, 2, ..., n$ için $v(.)$ indikatör fonksiyonu

$$v(X_i)=\begin{cases}1 & , \quad X_i=1 \\ 0 & , \quad X_i>1\end{cases}$$

biçimindedir. $Y=\frac{1}{n}\sum_{i=1}^{n} v(X_i)$, 1 değerinin örneklemdeki değerini gösterir. Y oranı, $P(X=1)$ olasılığının tutarlı ve yansız bir tahminidir. $P(X=1)=f(1)=Y$ eşitliğinin çözümü ile tek parametreli bir dağılım parametre tahmini yapılabilir. Parametre sayısı iki olduğu durumlarda, örneklemdeki 2 değerinin oranı hesaplanıp iki değişkenli iki denklem sisteminin çözümü ile oran tahmin edicileri bulunur. 1 ve 2 değerleri yerine bazen farklı değerler alınabilir. Bu durum dağılımın rasgele değişkeninin aldığı değere göre değişir.

2.5.3. Modifiye Edilmiş En Çok Olabilirlik Tahmini

$X_1, X_2, ..., X_n$, herhangi bir dağılımdan alınmış örneklem olmak üzere en çok olabilirlik tahmin edicisi, olabilirlik denklemlerinin çözümünden elde edildiğinden her zaman analitik sonuçlar elde edilemeyebilir. Modifiye edilmiş en çok olabilirlik yöntemi, en çok olabilirlik yönteminde olabilirlik denklemlerindeki analitik çözüm elde edilmesine engel olan doğrusal olmayan fonksiyonları Taylor seri açılımı ile doğrusal fonksiyonlara yaklaştırıp olabilirlik denklemlerini analitik olarak çözülebilir hale getirir ve elde edilen Modifiye edilmiş olabilirlik denklem sistemini çözer. Tahmin yapılacak dağılımın ilgili parametreleri q ve β, olabilirlik denklemleri $f_1(q,\beta)=0$ ve $f_2(q,\beta)=0$ olsun. $\tilde{q}$ ve $\tilde{\beta}$ parametrelerin oran tahmin edicisini ifade etmek üzere, $f_1(q,\beta)=0$ ve $f_2(q,\beta)=0$ de verilen $f_1(q,\beta)$ ve $f_2(q,\beta)$ fonksiyonlarının $(\tilde{q},\tilde{\beta})$ civarında ikinci merteben Taylor polinomları aşağıdaki gibi elde edilir.

$$\left(f_1(q,\beta)+(\hat{q}_*-q)\frac{\partial f_1(q,\beta)}{\partial q}+\left(\hat{\beta}_*-\beta\right)\frac{\partial f_1(q,\beta)}{\partial \beta} \right)\Bigg|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} =0$$

$$\left(f_2(q,\beta)+(\hat{q}_*-q)\frac{\partial f_2(q,\beta)}{\partial q}+\left(\hat{\beta}_*-\beta\right)\frac{\partial f_2(q,\beta)}{\partial \beta} \right)\Bigg|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} =0$$

Bu durumda parametrelerin modifiye edilmiş en çok olabilirlik tahmin edicileri

$$\hat{\beta}_* = \left. \frac{-f_2(q,\beta)\frac{\partial f_1(q,\beta)}{\partial q} + \frac{\partial f_2(q,\beta)}{\partial q} f_1(q,\beta) - \frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}\beta + \frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q}\beta}{\frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q} - \frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}} \right|_{\substack{q=\tilde{q} \\ \beta=\tilde{\beta}}}$$

ve

$$\hat{q}_* = -\left. \frac{f_1(q,\beta) - \frac{\partial f_1(q,\beta)}{\partial q} q + \frac{\partial f_1(q,\beta)}{\partial \beta}\hat{\beta}_* - \frac{\partial f_1(q,\beta)}{\partial \beta}\beta}{\frac{\partial f_1(q,\beta)}{\partial q}} \right|_{\substack{q=\tilde{q} \\ \beta=\tilde{\beta}}}$$

şeklinde elde edilir.

2.6. Newton-Raphson Yöntemi

Newton-Raphson yöntemi doğrusal olmayan denklemlerin çözümünde kullanılan iteratif yöntemlerden biridir. Newton-Raphson yönteminde öncelikle gerçek köke yakın bir başlangıç noktası seçilir. Başlangıç noktasının fonksiyonda aldığı değerden bir teğet çizilir ve tanjant doğrusu elde edilir. Bu tanjant doğrusunun x eksenini kestiği noktadaki değer bulunur. Bulunan bu değer fonksiyonun köküne önceki bulunan değerden daha yakındır. Önceden belirlenen bir yakınsama kriterine ulaşılana kadar yöntem tekrar edilir.

Newton – Raphson yöntemi geometrik olarak incelenecek olursa $f(x)=0$ fonksiyonunun başlangıç yaklaşık kökü x_0 olmak üzere fonksiyonun $(x_0, f(x_0))$ noktasındaki teğetinin denklemi

$$y - f(x_0) = f'(x_0)(x - x_0)$$

olarak yazılabilir. Bu teğetin x eksenini kestiği nokta ilk kök yaklaşımı olur ve

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

elde edilir. Bu şekilde ardışık yaklaşımlar kullanılarak, gerçek köke ulaşılır

2.7. LerchPhi Fonksiyonu

Lerchphi fonksiyonu

$$LerchPhi(z,a,v)=\sum_{n=0}^{\infty}\frac{z^n}{(v+n)^a}\;\left[|z|<1,\ v\neq 0,-1,\ldots\right]$$

şeklinde tanımlanır.

2.8. Hata Kareler Ortalaması ve Yan

$X_1, X_2, \ldots, X_n$ parametresi $\theta\ (\theta \in \Theta)$ olan kitleden alınan bir örneklem, T de θ nın herhangi bir tahmin edicisi olsun. T nin beklenen değeri mevcut ve $E(T)=\theta$ ise T ye θ için yansız tahmin edici denir. θ parametresinin bir T tahmin edicisi için yan

$$Yan_T(\theta)=E(T)-\theta$$

ve θ parametresinin bir T tahmin edicisi için hata kareler ortalaması (HKO)

$$HKO_T(\theta)=E(T-\theta)^2$$

şeklinde tanımlanır.

2.9. Bazı Kesikli Dağılımlar

Bu alt bölümde tezin ana bölümünde adı geçen literatürdeki kesikli dağılımlar tanıtılacaktır.

2.9.1. Kesikli Weibull Dağılımı

X, kesikli Weibull dağılımına sahip bir rasgele değişken olmak üzere *X*’in olasılık ve dağılım fonksiyonu sırasıyla,

$$f(x)=P(X=x)=q^{(x-1)^\beta}-q^{x^\beta},\ x=1,2,\ldots$$

ve

$$F(x)=1-q^{x^\beta},\ x=1,2,\ldots$$

biçiminde tanımlanır. Burada $0<q<1$ ve $\beta>0$ dağılımın parametreleridir. (Nakagawa ve Osaki, 1975)

2.9.2. Kesikli Burr XII Dağılımı

X , kesikli Burr XII dağılımına sahip bir rasgele değişken olmak üzere X rasgele değişkeninin olasılık ve dağılım fonksiyonu sırasıyla,

$$P(X=x)=q^{\log(1+x^{\beta})}-q^{\log(1+(1+x^{\beta}))}, x=0,1,2,\ldots$$

ve

$$F(x)=1-q^{\log(1+x^{\beta})}, x=0,1,2,\ldots$$

biçiminde tanımlanır. Burada $0<q<1$ ve $\beta>0$ dağılımın parametreleridir. (Krishna ve Singh Pundir, 2009).

2.9.3. Kesikli Pareto Dağılımı

X , kesikli Pareto dağılımına sahip bir rasgele değişkeninin olasılık ve dağılım fonksiyonu sırasıyla,

$$P(X=x)=q^{\log(1+x)}-q^{\log(1+(1+x))},\ x=0,1,2,\ldots$$

ve

$$F(x)=1-q^{\log(1+x)}, x=0,1,2,\ldots$$

biçiminde tanımlanır. Burada $0<q<1$ dağılımın parametresidir. (Krishna ve Singh Pundir, 2009)

2.9.4. Düzgün Geometrik Dağılım

$X \mid N=n$, n parametreli $(U(n), n>0)$ kesikli düzgün dağılıma sahip ve N , p parametreli Geometrik dağılımına sahip olsun. Bu durumda X rasgele değişkeninin dağılımı

$$P(X=x)=\sum_{n=x}^{\infty}P(X=x \mid N=n)P(N=n)$$

$$=\sum_{n=x}^{\infty}\frac{1}{n}p(1-p)^{n-1}$$

$$=\sum_{j=0}^{\infty}\frac{1}{x+j}p(1-p)^{x+j-1}$$

$$= p(1-p)^{x-1}\sum_{j=0}^{\infty}\frac{(1-p)^j}{x+j}$$

$$= p(1-p)^{x-1}\, LerchPhi\left[(1-p),1,x\right], x=1,2,\ldots,\quad 0<p<1,$$

şeklinde tanımlanır. Burada $LerchPhi(z,a,v)$, LerchPhi fonksiyonudur. Düzgün geometrik dağılımı $UG(p)$ şeklinde gösterilir. (Akdoğan ve ark., 2016).

2.9.5. Geometrik-Sıfırdan Budanmış Poisson Dağılımı (GZTP)

$N \sim Sıfırdan\ Budanmış\ Poisson(\lambda)$ ve $Y_1,Y_2,\ldots,Y_N \sim$ iid $Geo(p)$ birbirinden bağımsız rasgele değişkenler olsun, $X = \text{maks}(Y_1,Y_2,\ldots,Y_N)$ dönüşümünün dağılımı

$$\begin{aligned} F(x) = P(X \le x) &= P\left(Maks(Y_1,Y_2,\ldots,Y_N) \le x\right) \\ &= P(Y_1 \le x, Y_2 \le x,\ldots,Y_N \le x) \\ &= P(Y_1 \le x)P(Y_2 \le x)\ldots P(Y_N \le x) \\ &= (1-q^x)(1-q^x)\ldots(1-q^x) \\ &= (1-q^x)^n \end{aligned}$$

şeklindedir. Aynı işlemler $F(x-1)$ için uygulandığında, $P(X=x \mid N=n)$ rasgele değişkeninin olasılık fonksiyonu

$$\begin{aligned} P(X=x \mid N=n) &= F(x)-F(x-1) \\ &= (1-q^x)^n-(1-q^{x-1})^n \end{aligned}$$

şeklinde elde edilir ve $N \sim Sıfırdan\ Budanmış\ Poisson(\lambda)$ olmak üzere olasılık fonksiyonu

$$P(N=n)=\frac{e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})},\ n=1,2,\ldots$$

şeklindedir. Bu durumda

$$f(x)=P(X=x)=\sum_{n=1}^{\infty}P(X=x \mid N=n)P(N=n)$$

$$=\sum_{n=1}^{\infty}(1-q^x)^n-(1-q^{x-1})^n\,\frac{e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})}$$

$$= \frac{e^{-\lambda}}{\left(1-e^{-\lambda}\right)}\left(\sum_{n=1}^{\infty}\frac{\left(\lambda\left(1-q^{x}\right)\right)^{n}}{n!}-\sum_{n=1}^{\infty}\frac{\left(\lambda\left(1-q^{x-1}\right)\right)^{n}}{n!}\right)$$

$$= \frac{1}{\left(1-e^{-\lambda}\right)}\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right)$$

şeklinde elde edilir. Burada $\theta = e^{-\lambda}$ alındığında olasılık fonksiyonu

$$P\left(X = x\right) = \frac{1}{\left(1-e^{-\lambda}\right)}\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right), \qquad x = 1, 2, \ldots$$

ve dağılım fonksiyonu

$$F\left(x\right) = \begin{cases} 0 & , \quad x \leq 0 \\ \frac{\theta^{\left(q^{x}-1\right)}-1}{\theta-1} & , \quad x = 1, 2, \ldots \end{cases}$$

biçimindedir. X rasgele değişkeninin dağılımı, Geometrik-sıfırdan budanmış Poisson (GZTP) dağılımı olarak adlandırılır ve $X \sim GZTP\left(q, \theta\right)$ biçiminde gösterilir. Burada $0 < q < 1$ ve $0 < \theta < 1$ dağılımın parametreleridir. (Akdoğan ve ark., 2019)

3. BAZI KESİKLİ DAĞILIMLAR İÇİN SANSÜRLÜ ÖRNEKLEM DURUMUNDA PARAMETRE TAHMİNİ

3.1. Düzgün Geometrik Dağılımı

Düzgün geometrik dağılımı Akdoğan ve arkadaşları tarafından önerilmiştir. *X*, *UG* dağılımına sahip rasgele değişken olsun, *X* rasgele değişkeninin olasılık fonksiyonu ve dağılım fonksiyonu sırasıyla,

$$f(x) = p(1-p)^{x-1} LerchPhi\left[(1-p),1,x\right],\ x=1,2,\dots,\ 0<p<1$$

$$F(x) = 1-p(1-p)^{x}\left[\frac{1}{p} - xLerchPhi(1-p,1,x+1)\right],\quad x=1,2,\dots$$

şeklindedir. Burada dağılımın parametreleri $p \in (0,1)$ dir. Dağılım kısaca $UGD(p)$ şeklinde gösterilir. (Akdoğan ve ark., 2016). Bu bölümde düzgün geometrik dağılımı için Tip I sansürlü örnekleme dayalı parametre konusu incelenmiştir.

3.1.1. En Çok Olabilirlik Tahmini

$Y_1, Y_2, \dots, Y_n$ n birimlik bağımsız ve aynı dağılımlı örneklem, $L_1, L_2, \dots, L_n$ her bir rasgele değişken için sağdan sansür zamanını göstersin. Bu durumda

$$X_i = \min(Y_i, L_i)\ i=1,2,\dots,n$$

şeklinde tanımlanmak üzere $X_1, X_2, \dots, X_n$, örneklemine Tip I sansürlü örneklem denir.

$$\delta_i = \begin{cases} 1 & , \quad X_i = Y_i \\ 0 & , \quad \text{Diğer Yerlerde} \end{cases}$$

Gösterge fonksiyonu kullanılarak olabilirlik ve log-olabilirlik fonksiyonu sırasıyla,

$$\ell(p) = \log\left\{\prod_{i=1}^{n} f(x_i)^{\delta_i}\left(1-F(x_i)\right)^{1-\delta_i}\right\}$$

$$= \sum_{i=1}^{n}\delta_i\left(\log\left(f(x_i)\right)\right) + \sum_{i=1}^{n}(1-\delta_i)\left(\log\left(1-F(x_i)\right)\right)$$

$$= \sum_{i=1}^{n} (1-\delta_i) \log\left(p(1-p)^x \left[\frac{1}{p} - LerchPhi(1-p,1,x+1) \right] \right)$$

elde edilmiştir. Olabilirlik denklemi ise

$$\frac{\partial \log(\ell(p))}{\partial p} = \sum_{i=1}^{n} \delta_i \left[\frac{x(1-p)+2p-1}{p(1-p)} - \frac{1}{p(1-p)LerchPhi\left[(1-p),1,x\right]} \right]$$
$$+ \sum_{i=1}^{n} (1-\delta_i) \left[\frac{xD(p)}{p} - \frac{p+p^2\left[x-1-(x+1)xLerchPhi(1-p,1,x+1)\right]}{p^3(1-p)\left[\frac{1}{p} - xLerchPhi(1-p,1,x+1)\right]} \right] = 0 \qquad 3.1.1$$

şeklindedir. Elde edilen (3.1.1) denkleminin çözümü analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır. (Lagarias ve ark., 1998)

3.1.2. Oranlar tahmini

Khan ve arkadaşları kesikli Weibull dağılımının parametrelerinin tahmin edilmesi için oranlar yöntemini önermiştir. Tezin bu kısmında bu yöntem düzgün geometrik dağılım için kullanılacaktır. $X_1, X_2, ..., X_n$ $UG(p)$'den rastgele örneklem olsun.

$$v(X_i) = \begin{cases} 1 & , \quad X_i = 1 \\ 0 & , \quad X_i > 1 \end{cases}$$

burada $Y = \frac{1}{n}\sum_{i=1}^{n} v(X_i)$ örneklemdeki 1'in oranını belirtmektedir. $f(1) = pLerchPhi\left[(1-p,1,1)\right]$ olasılığı *Y*'nin yansız ve tutarlı bir oranlar tahmincisidir. Bu durumda, p değeri şu şekilde tanımlanabilir $Y = pLerchPhi\left[(1-p,1,1)\right]$ ve $\frac{p\log(p)}{p-1} - Y = 0$ denklemin çözümü p oranının tahminini göstermektedir. (Khan ve ark., 1989)

3.1.3. Modifiye Edilmiş En Çok Olabilirlik Tahmini

Bu bölümde dağılımın parametrelerinin modifiye edilmiş en çok olabilirlik tahmin edicileri Tip-I sansürlü örnekleme dayalı olarak elde edilecektir. $X_1, X_2, \ldots, X_n$, $UG(p)$ dağılımından alınmış Tip-I sansürlü örneklem olsun. Log-olabilirlik denklemlerinin p parametresine göre türevi $f(p)$, parametreye göre Taylor serisinin ilk iki terimi

$$\left(f(p) + (\hat{p}_* - p)\frac{\partial f(p)}{\partial p} \right)\Bigg|_{p=\tilde{p}} = 0$$

ele alınır. Burada $\tilde{p}$, parametrenin oran tahmin edicisidir. (Kulasekera, 1994) Yukarıdaki denklem çözüldüğünde p parametresinin modifiye edilmiş en çok olabilirlik tahmin edicisi

$$\hat{p}_* = -\frac{f(p) - \frac{\partial f(p)}{\partial p} p}{\frac{\partial f_1(p)}{\partial p}}\Bigg|_{p=\tilde{p}}$$

olarak elde edilir. Burada Log-olabilirlik denklemlerinin p parametresine göre türevi $f(p)$ olmak üzere,

$$\frac{\partial \log(\ell(p))}{\partial p} = \sum_{i=1}^{n} \delta_i \left[\frac{x(1-p)+2p-1}{p(1-p)} - \frac{1}{p(1-p)\,LerchPhi\left[(1-p),1,x\right]} \right]$$
$$+\sum_{i=1}^{n}(1-\delta_i)\left[\frac{xD(p)}{p} - \frac{p+p^2\left[x-1-(x+1)\,xLerchPhi(1-p,1,x+1)\right]}{p^3(1-p)\left[\frac{1}{p} - xLerchPhi(1-p,1,x+1)\right]} \right] = 0$$

şeklinde elde edilmiştir. Elde edilen denklemin çözümü türevlerin karmaşık olmasından dolayı analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır. (Lagarias ve ark., 1998).

3.1.4. Simulasyon Çalışması

Simülasyon çalışmasında düzgün geometrik dağılımına sahip rasgele örnekleminden en çok olabilirlik, oranlar ve modifiye edilmiş en çok olabilirlik tahmin edici yöntemleri kullanılarak parametre tahmini ve parametre tahminlerine ilişkin yan ve hata kareler ortalaması değerleri elde edilmiştir. Simülasyon çalışmasında $p = 0.3$ ve $p = 0.8$ değerleri için 10000 tekrar sonucunda elde edilen değerler Tablo 1-4' te sunulmuştur.

Tablo 1. UG dağılımın parametresi için yan değerleri

			MLE	Oranlar	MMLE
		L			
$p = 0.3$	50	3	0,0675	0,0041	0,0811
	100	3	0,0669	0,0011	0,0808
	200	3	0,0273	0,0008	0,0174
	300	3	0,0228	0,0005	0,0031
	500	3	0,0112	0,0004	0,0013
$p = 0.3$	50	5	-0,0262	0,0041	-0,0285
	100	5	-0,0256	0,0011	-0,0284
	200	5	-0,0250	0,0008	-0,0282
	300	5	-0,0236	0,0005	-0,0281
	500	5	-0,0182	0,0001	-0,0251
$p = 0.3$	50	∞	0,0057	0,0025	0,0006
	100	∞	0,0030	0,0011	0,0005
	200	∞	0,0017	0,0008	0,0003
	300	∞	0,0012	0,0005	0,0003
	500	∞	0,0006	0,0001	0,0002

Tablo 2. UG dağılımın parametresi için HKO değerleri

			MLE	Oranlar	MMLE
		L			
$p = 0.3$	50	3	0,0052	0,0047	0,0076
	100	3	0,0051	0,0023	0,0073
	200	3	0,0025	0,0012	0,0026
	300	3	0,0021	0,0008	0,0022
	500	3	0,0015	0,0004	0,0016
$p = 0.3$	50	5	0,0035	0,0047	0,0042
	100	5	0,0021	0,0023	0,0025
	200	5	0,0014	0,0012	0,0016
	300	5	0,0012	0,0008	0,0014
	500	5	0,0010	0,0005	0,0011
$p = 0.3$	50	∞	0,0024	0,0047	0,0025
	100	∞	0,0011	0,0023	0,0012
	200	∞	0,0006	0,0012	0,0006
	300	∞	0,0004	0,0008	0,0004
	500	∞	0,0002	0,0003	0,0002

Tablo 3. UG dağılımın parametresi için yan değerleri

			MLE	Oranlar	MMLE
		L			
$p = 0.8$	50	3	0,0259	0,0018	0,0127
	100	3	0,0106	0,0018	0,0103
	200	3	0,0026	0,0005	0,0020
	300	3	0,0012	0,0004	0,0019
	500	3	0,0006	0,0002	0,0014
$p = 0.8$	50	5	0,0159	0,0006	0,0056
	100	5	0,0068	0,0003	0,0031
	200	5	0,0012	0,0005	0,0019
	300	5	0,0009	0,0005	0,0013
	500	5	0,0004	0,0004	0,0007
$p = 0.8$	50	∞	0,0094	0,0033	0,0085
	100	∞	0,0038	0,0020	0,0051
	200	∞	0,0016	0,0009	0,0023
	300	∞	0,0009	0,0005	0,0014
	500	∞	0,0003	0,0001	0,0005

Tablo 4. UG dağılımın parametresi için HKO değerleri

			MLE	Oranlar	MMLE
		L			
$p = 0.8$	50	3	0,0106	0,0055	0,0049
	100	3	0,0030	0,0028	0,0026
	200	3	0,0012	0,0014	0,0012
	300	3	0,0009	0,0009	0,0009
	500	3	0,0005	0,0006	0,0005
$p = 0.8$	50	5	0,0080	0,0056	0,0060
	100	5	0,0025	0,0029	0,0025
	200	5	0,0012	0,0014	0,0012
	300	5	0,0008	0,0009	0,0008
	500	5	0,0005	0,0006	0,0005
$p = 0.8$	50	∞	0,0089	0,0056	0,0046
	100	∞	0,0024	0,0028	0,0024
	200	∞	0,0012	0,0014	0,0012
	300	∞	0,0008	0,0010	0,0008
	500	∞	0,0005	0,0005	0,0005

Tablo 1-4'den itibaren tüm tahminlerin yanlı olduğunu ancak bu tahminlerin asimptotik olarak yansız olduğunu görülmüştür. Oranlar tahmini hata karaler ortalaması her zaman küçük olduğundan en çok olabilirlik tahmini(MLE) ve modifiye edilmiş en çok olabilirlik tahmininden daha iyi performans göstermiştir. Ayrıca, örneklem büyüklüğü arttığında parametrelerinin yansızlığı ve hata kareler ortalaması değerleri beklenildiği gibi düşer.

3.1.6. Gerçek Veri Uygulaması

Bu bölümde iki gerçek veri seti Tip-I sansürlü örnekleme dönüştürülmüş ve sansürlü örneklem durumunda parametre tahmini verilmeye çalışılmıştır. İlk veri seti Xie ve Goh tarafından literatüre kazandırılmıştır (Xie ve Goh, 1993). Endüstriyel bir süreç içindeki kusurlara ilişkin denetimlerin sayıları üzerinden yapılacaktır. Bu verinin düzgün-geometrik dağılama uygunluğu 2016 yılında Akdoğan ve ark. Tarafından çalışılmıştır. Akdoğan, (2016). Bu veri seti aşağıda verilmiştir:

1 1 1 1 1 1 2 2 2 23 3 3 4 4 4 5 5 7 9 11 13 14 14 17 18 26 29.

Veriye ait en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 5'te verilmiştir.

Tablo 5. Veriye ait $UG(p)$ dağılımının parameter tahminleri

	Tam Örneklem	L=3	L=5
	p	p	p
MLE	0.0742	0.0469	0.0683
Oranlar	0.0772	0.0772	0.0772
MMLE	0.0742	0.0476	0,0763

İkinci veri seti, Weinberg ve Gladen tarafından elde edilen gebelik için aybaşı döngü sayısı verisidir. Bu veri 586 bayan üzerinden 1844 döngü kaydedilerek oluşmuştur. (Weinberg ve Gladen, 1986) Elde edilen veri aşağıdaki Tablo 6 de verilmiştir.

Tablo 6. Gebelik için aybaşı döngü sayısı

Döngü Sayısı	1	2	3	4	5	6	7	8	9	10	11	12
Kadın Sayısı	227	123	72	42	21	31	11	14	6	4	7	28

Gerçek veri için en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 7'de verilmiştir.

Tablo 7. Veriye ait $UG(p)$ dağılımının parameter tahminleri

	Tam Örneklem	L=3	L=5
	p	p	p
MLE	0.1889	0.1178	0.1485
Oranlar Yöntemi	0.1883	0.1883	0.1883
MMLE	0.1889	0.1178	0.1485

Birinci ve ikinci gerçek veri analizinde, en çok olabilirlik, oranlar yöntemi ve modifiye edilmiş en çok olabilirlik tahmin edicileri tam örneklem için hemen hemen aynı sonucu vermiştir. Oranlar yöntemi, sansürlü örneklem durumunda en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicileriden daha iyi performans göstermektedir.

3.2. Geometrik-Sıfırdan Budanmış Poisson Dağılımı (GZTP)

Geometrik-sıfırdan budanmış Poisson dağılımı (*GZTP*) dağılımı Akdoğan ve arkadaşları tarafından önerilmiştir. $X_1, X_2, ..., X_n$ Geometrik-Sıfırdan Budanmış Poisson dağılımına sahip rasgele değişken olsun, X rasgele değişkeninin olasılık fonksiyonu ve dağılım fonksiyonu sırasıyla;

$$f(x) = \frac{1}{\left(1-e^{-\lambda}\right)}\left(e^{-\lambda q^{x}} - e^{-\lambda q^{x-1}}\right), \quad x = 1, 2, ...$$

$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{e^{-\lambda\left(q^{x-1}\right)-1}}{e^{\lambda}-1}, & x = 1, 2, ... \end{cases}$$

şeklindedir. Burada $0 < q < 1$ ve $\lambda > 0$ dağılımın parametreleridir. (Akdoğan ve ark., 2019).

3.2.1. En Çok Olabilirlik Tahmini

$Y_1, Y_2, ..., Y_n$ bağımsız ve aynı $GZTP(q, \theta)$ dağılımından alınmış örneklem ve $L_1, L_2, ..., L_n$ her bir rasgele değişken için sağdan sansür zamanı olmak üzere (Bu çalışmada bu zamanlar eşit kabul edilmiştir.)

$$X_i = \min\left(Y_i, L_i\right) \; i = 1, 2, ..., n$$

şeklinde tanımlanan yeni rasgele değişken üzerinden

$$\delta_i = \begin{cases} 1 \quad , & X_i = Y_i \\ 0 \quad , & \text{Diğer Yerlerde} \end{cases}$$

indikatör fonksiyonu elde edilir. Bu verilenlerle kesikli dağılımlarda Tip I sağdan sansürlü örneklem durumunda olabilirlik fonksiyonu

$$L(q, \theta) = \prod_{i=1}^{n} f(x_i)^{\delta_i} (1 - F(x_i))^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[\frac{1}{1-\theta}\left(\theta^{q^{x_i}} - \theta^{q^{x_i-1}}\right)\right]^{\delta_i} \left[1 - \frac{\theta^{q^{x_i}} - \theta}{1-\theta}\right]^{1-\delta_i}$$

$$= \left(\frac{1}{1-\theta}\right)^{n\delta_i} \left(\frac{1}{1-\theta}\right)^{n(1-\delta_i)} \prod_{i=1}^{n} \left(\theta^{q^{x_i}} - \theta^{q^{x_i-1}}\right)^{\delta_i} \left(1 - \theta^{q^{x_i}}\right)^{1-\delta_i}$$

$$=\left(\frac{1}{1-\theta}\right)^{n}\prod_{i=1}^{n}\left(\theta^{q^{x_i}}-\theta^{q^{x_i-1}}\right)^{\delta_i}\left(1-\theta^{q^{x_i}}\right)^{1-\delta_i}$$

ve log-olabilirlik fonksiyonu

$$\ell(q,\theta)=-n\log(1-\theta)+\sum\delta_i\log(\theta^{q^{x_i}}-\theta^{q^{x_i-1}})+\sum\left(1-\delta_i\right)\log(1-\theta^{q^{x_i}})$$

şeklinde elde edilir. En çok olabilirlik tahmini için log-olabilirlik fonksiyonunun elde edilen olabilirlik denklemleri aşağıda verilmiştir.

$$\begin{aligned}\frac{\partial\ell(q,\theta)}{\partial q}&=\sum_{i=1}^{n}\frac{\delta_i\log(\theta)\left\{-q^{x_i-1}\left(x_i-1\right)\theta^{q^{x_i-1}}+\theta^{q^{x_i}}q^{x_i}x_i\right\}}{q\left(\theta^{q^{x_i}}-\theta^{q^{x_i-1}}\right)}\\&+\sum_{i=1}^{n}\left(1-\delta_i\right)\frac{\theta^{q^{x_i}}q^{x_i}x_i\log(\theta)}{q(\theta^{q^{x_i}}-1)}=0\end{aligned}\tag{3.2.1}$$

$$\begin{aligned}\frac{\partial\ell(q,\theta)}{\partial\theta}&=\left(\frac{n}{1-\theta}\right)+\sum_{i=1}^{n}\frac{\delta_i\left\{\theta^{q^{x_i}}q^{x_i}-\theta^{q^{x_i-1}}q^{x_i-1}\right\}}{\theta\left(\theta^{q^{x_i}}-\theta^{q^{x_i-1}}\right)}\\&+\sum_{i=1}^{n}\left(1-\delta_i\right)\frac{\theta^{q^{x_i}}q^{x_i}x_i\log(\theta)}{q\left(\theta^{q^{x_i}}-1\right)}=0\end{aligned}\tag{3.2.2}$$

elde edilen (3.2.1) ve (3.2.2) denklemlerinin çözümü analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır. (Lagarias ve ark., 1998)

3.2.2. Oranlar tahmini

Kesikli dağılımlar için kullanılabilen bu yöntemi ilk defa Khan ve arkadaşları kesikli Weibull dağılımı için kullanmışlardır.(Khan ve ark., 1989) Bu tez çalışmasının bu bölümünde aynı yöntem Geometrik-Sıfırdan Budanmış Poisson Dağılımı için kullanılacaktır. $X_1, X_2, ..., X_n$ bağımsız ve aynı *GZTP(q,θ)* dağılımından alınmış örneklem olmak üzere

$$\Phi_1 = \begin{cases} 1 & , \quad X_i = 1 \\ 0 & , \quad X_i > 1 \end{cases}$$

İndikatör fonksiyonu yardımıyla örneklemdeki "1" değerlerinin sayısını ve $Y = \frac{1}{n}\sum_{i=1}^{n}\Phi_1\left(X_i\right)$ eşitlikle örneklemdeki "1" lerin oranı kolayca elde edilir. Hesaplanan bu oran $f(1) = \frac{1}{1-\theta}\left(\theta^q - \theta\right)$ olasılığının yansız ve tutarlı tahmin edicisidir. GZTP dağılımında 2 parametre bulunduğundan aynı işlemler örneklemdeki "2" değerleri içinde uygulanır. Buna göre

$$\Phi_1 = \begin{cases} 1 & , \quad X_i = 2 \\ 0 & , \quad X_i \neq 2 \end{cases}$$

İndikatör (gösterge) fonksiyonu tanımlanarak sansürlü örneklemdeki "2"lerin sayısını ve $Z = \frac{1}{n}\sum_{i=1}^{n}\Phi_2\left(X_i\right)$ eşitlikle örneklemdeki "2" değerlerinin oranı elde edilir. 2'lerin oranı aynı zamanda $f(2) = \frac{1}{1-\theta}\left(\theta^{q2} - \theta^q\right)$ olasılığının yansız ve tutarlı tahmin edicisidir. Böylece elde edilen iki denklem aşağıda verilmiştir.

$$\frac{1}{1-\theta}\left(\theta^q - \theta\right) = Y$$

$$\frac{1}{1-\theta}\left(\theta^{q2} - \theta^q\right) = Z$$

İki bilinmeyenli iki denklemin çözümü oranlar tahmini olarak elde edilir. Bu yöntem basit olmasının yanında bazı sıkıntılar doğurabilmektedir. Örneğin sansürleme zamanının "1" olduğu durumlarda örneklemde hiç "2" değeri gelmeyeceği için denklem sistemi durmayacaktır. Bu sebeple oranlar tahmini hesaplanamaz.

3.2.3. Modifiye Edilmiş En Çok Olabilirlik Tahmini

Bu bölümde dağılımın parametrelerinin en çok olabilirlik tahmin edicileri Tip-I sansürlü örnekleme dayalı olarak elde edilecektir. $X_1, X_2, \ldots, X_n$, $GZTP(q,\theta)$ dağılımından alınmış Tip-I sansürlü örneklem olsun. Log-olabilirlik denklemlerinin q parametresine göre türevi $f_1(q,\theta)$ ve θ parametresine göre türevi $f_2(q,\theta)$ olduğunda, parametreye göre Taylor serisinin ilk iki terimi

$$\left(f_1(q,\theta) + (\hat{q}_* - q)\frac{\partial f_1(q,\theta)}{\partial q} + \left(\hat{\theta}_* - \theta\right)\frac{\partial f_1(q,\theta)}{\partial \theta} \right)\Bigg|_{\substack{q=\tilde{q}\\ \theta=\tilde{\theta}}} = 0$$

$$\left(f_2(q,\theta) + (\hat{q}_* - q)\frac{\partial f_2(q,\theta)}{\partial q} + \left(\hat{\theta}_* - \theta\right)\frac{\partial f_2(q,\theta)}{\partial \theta} \right)\Bigg|_{\substack{q=\tilde{q}\\ \theta=\tilde{\theta}}} = 0,$$

ele alınır. Burada $\tilde{q}$ ve $\tilde{\theta}$, parametrelerin oranlar tahmin edicileridir. Yukarıdaki denklemler çözüldüğünde en çok olabilirlik tahmin edicileri

$$\hat{\theta}_* = \frac{-f_2(q,\theta)\frac{\partial f_1(q,\theta)}{\partial q} + \frac{\partial f_2(q,\theta)}{\partial q} f_1(q,\theta) - \frac{\partial f_2(q,\theta)}{\partial q}\frac{\partial f_1(q,\theta)}{\partial \theta}\theta + \frac{\partial f_2(q,\theta)}{\partial \theta}\frac{\partial f_1(q,\theta)}{\partial q}\theta}{\frac{\partial f_2(q,\theta)}{\partial \theta}\frac{\partial f_1(q,\theta)}{\partial q} - \frac{\partial f_2(q,\theta)}{\partial q}\frac{\partial f_1(q,\theta)}{\partial \theta}}\Bigg|_{\substack{q=\tilde{q}\\ \theta=\tilde{\theta}}}$$

$$\hat{q}_* = -\frac{f_1(q,\theta) - \frac{\partial f_1(q,\theta)}{\partial q} q + \frac{\partial f_1(q,\theta)}{\partial \theta}\hat{\theta}_* - \frac{\partial f_1(q,\theta)}{\partial \theta}\theta}{\frac{\partial f_1(q,\theta)}{\partial q}}\Bigg|_{\substack{q=\tilde{q}\\ \theta=\tilde{\theta}}}$$

olarak elde edilir. Burada $f_1(q,\theta)$ ve $f_2(q,\theta)$ denklemleri

$$\begin{aligned} f_1(q,\theta) = \frac{\partial \ell(q,\theta)}{\partial q} &= \sum_{i=1}^{n} \frac{\delta_i \log(\theta)\left\{-q^{x_i-1}(x_i-1)\theta^{q^{x_i-1}} + \theta^{q^{x_i}} q^{x_i} x_i\right\}}{q\left(\theta^{q^{x_i}} - \theta^{q^{x_i-1}}\right)} \\ &+ \sum_{i=1}^{n}(1-\delta_i)\frac{\theta^{q^{x_i}} q^{x_i} x_i \log(\theta)}{q(\theta^{q^{x_i}}-1)} = 0 \end{aligned} \qquad (3.2.3)$$

$$\begin{aligned} f_2(q,\theta) = \frac{\partial \ell(q,\theta)}{\partial \theta} &= \left(\frac{n}{1-\theta}\right) + \sum_{i=1}^{n} \frac{\delta_i\left\{\theta^{q^{x_i}} q^{x_i} - \theta^{q^{x_i-1}} q^{x_i-1}\right\}}{\theta\left(\theta^{q^{x_i}} - \theta^{q^{x_i-1}}\right)} \\ &+ \sum_{i=1}^{n}(1-\delta_i)\frac{\theta^{q^{x_i}} q^{x_i} x_i \log(\theta)}{q\left(\theta^{q^{x_i}}-1\right)} = 0 \end{aligned} \qquad (3.2.4)$$

Elde edilen (3.2.3) ve (3.2.4) denklemlerinin çözümü analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır. (Lagorias ve ark. 1998)

3.2.4. Simulasyon Çalışması

Simülasyon çalışmasında geometrik sıfırdan budanmış poisson dağılımına sahip rasgele örnekleminden en çok olabilirlik, oranlar ve modifiye edilmiş en çok olabilirlik tahmin edici yöntemleri kullanılarak parametre tahmini ve parametre tahminlerine ilişkin yan ve hata kareler ortalaması değerleri elde edilmiştir. Simülasyon çalışmasında $q=0.3, \theta=0.3$, $q=0.5, \theta=0.5$, $q=0.3, \theta=0.8$ ve $q=0.8, \theta=0.3$ değerleri için 10000 tekrar sonucunda elde edilen değerler Tablo 8-15' te sunulmuştur.

Tablo 8. GZTP dağılımın parametreleri için yan değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.3	3	100	0.0137	0.3488	0.0167	0.5042	0.0145	0.3499
		3	200	0.0069	0.1746	0.0090	0.2523	0.0077	0.1750
		3	300	0.0037	0.0875	0.0047	0.1260	0.0037	0.0877
		3	400	0.0020	0.0438	0.0025	0.0632	0.0018	0.0441
		3	500	0.0012	0.0228	0.0014	0.0340	0.0007	0.0223
0.3	0.3	∞	100	0.0078	0.0420	0.0098	0.4809	0.0076	0.0440
		∞	200	0.0040	0.0212	0.0050	0.2411	0.0039	0.0221
		∞	300	0.0021	0.0108	0.0026	0.1209	0.0020	0.0112
		∞	400	0.0006	0.0061	0.0015	0.0606	0.0011	0.0057
		∞	500	0.0004	0.0032	0.0008	0.0308	0.0006	0.0030
0.5	0.5	3	200	0.0059	0.4025	0.0160	0.7131	0.0062	0.4118
		3	300	0.0031	0.2013	0.0082	0.3566	0.0032	0.2110
		3	400	0.0017	0.1007	0.0043	0.1734	0.0017	0.1056
		3	500	0.0009	0.0005	0.0022	0.0875	0.0009	0.0530
0.5	0.5	∞	200	-0.0021	0.0525	0.0191	0.8120	0.0024	0.0550
		∞	300	-0.0012	0.0264	0.0096	0.4062	0.0013	0.0271
		∞	400	-0.0007	0.0133	0.0050	0.2032	0.0007	0.0136
		∞	500	-0.0004	0.0067	0.0026	0.1017	0.0004	0.0069

Tablo 9. GZTP dağılımın parametreleri için HKO değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.3	3	200	0.0031	0.0766	0.0055	0.0920	0.0032	0.0922
		3	300	0.0016	0.0385	0.0029	0.0461	0.0017	0.0462
		3	400	0.0009	0.0194	0.0016	0.0232	0.0009	0.0232
		3	500	0.0005	0.0089	0.0009	0.0117	0.0005	0.0113
0.3	0.3	∞	200	0.0021	0.1137	0.0050	0.0904	0.0022	0.1150
		∞	300	0.0011	0.0558	0.0026	0.0451	0.0012	0.0560
		∞	400	0.0006	0.0280	0.0014	0.0226	0.0007	0.0300
		∞	500	0.0004	0.0141	0.0008	0.0114	0.0004	0.0016
0.5	0.5	3	200	0.0043	17.5250	0.0103	5.5399	0.0045	13.4556
		3	300	0.0023	0.9760	0.0053	0.0912	0.0023	0.0944
		3	400	0.0012	0.4882	0.0027	0.0456	0.0013	0.0473
		3	500	0.0007	0.2441	0.0014	0.0229	0.0007	0.0237
0.5	0.5	∞	200	0.0019	0.1013	0.0110	0.4470	0.0021	0.1015
		∞	300	0.0011	0.0507	0.0056	0.2236	0.0012	0.0508
		∞	400	0.0006	0.0254	0.0029	0.1120	0.0007	0.0254
		∞	500	0.0004	0.0129	0.0016	0.0056	0.0004	0.0127

Tablo 10. GZTP dağılımın parametreleri için yan değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.3	5	100	0.0123	0.0944	0.0136	0.5081	0.0124	0.0945
		5	200	0.0067	0.0473	0.0069	0.2510	0.0062	0.0473
		5	300	0.0034	0.0263	0.0035	0.1256	0.0032	0.0237
		5	400	0.0018	0.0132	0.0018	0.630	0.0017	0.0116
		5	500	0.0010	0.0067	0.0009	0.316	0.0009	0.0059
0.3	0.3	∞	100	0.0078	0.0420	0.0098	0.4809	0.0076	0.0440
		∞	200	0.0040	0.0212	0.0050	0.2411	0.0039	0.0221
		∞	300	0.0021	0.0108	0.0026	0.1209	0.0020	0.0112
		∞	400	0.0006	0.0061	0.0015	0.0606	0.0011	0.0057
		∞	500	0.0004	0.0032	0.0008	0.0308	0.0006	0.0030
0.5	0.5	5	100	0.0032	1.0584	0.0234	1.3690	0.0033	1.0610
		5	200	0.0017	0.0960	0.0119	0.0980	0.0017	0.0950
		5	300	0.0009	0.0481	0.0060	0.0499	0.0009	0.0476
		5	400	0.0005	0.0242	0.0031	0.0251	0.0005	0.0239
		5	500	0.0003	0.0122	0.0016	0.0126	0.0003	0.0121
0.5	0.5	∞	200	-0.0021	0.0525	0.0191	0.8120	0.0024	0.0550
		∞	300	-0.0012	0.0264	0.0096	0.4062	0.0013	0.0271
		∞	400	-0.0007	0.0133	0.0050	0.2032	0.0007	0.0136
		∞	500	-0.0004	0.0067	0.0026	0.1017	0.0004	0.0069

Tablo 11. GZTP dağılımın parametreleri için HKO değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.3	5	200	0.0023	0.0942	0.0053	0.0976	0.0024	0.0948
		5	300	0.0012	0.0472	0.0027	0.0484	0.0013	0.0475
		5	400	0.0007	0.0238	0.0014	0.0243	0.0007	0.0240
		5	500	0.0004	0.0120	0.0008	0.0124	0.0005	0.0121
0.3	0.3	∞	200	0.0021	0.1137	0.0050	0.0904	0.0022	0.1150
		∞	300	0.0011	0.0558	0.0026	0.0451	0.0012	0.0560
		∞	400	0.0006	0.0280	0.0014	0.0226	0.0007	0.0300
		∞	500	0.0004	0.0141	0.0008	0.0114	0.0004	0.0016
0.5	0.5	5	200	0.0026	0.0940	0.0089	0.0910	0.0027	0.0904
		5	300	0.0014	0.0471	0.0046	0.0456	0.0014	0.0453
		5	400	0.0008	0.0236	0.0024	0.0229	0.0008	0.0226
		5	500	0.0005	0.0119	0.0013	0.0116	0.0006	0.0114
0.5	0.5	∞	200	0.0019	0.1013	0.0110	0.4470	0.0021	0.1015
		∞	300	0.0011	0.0507	0.0056	0.2236	0.0012	0.0508
		∞	400	0.0006	0.0254	0.0029	0.1120	0.0007	0.0254
		∞	500	0.0004	0.0129	0.0016	0.0056	0.0004	0.0127

Tablo 12. GZTP dağılımın parametreleri için yan değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.8	3	200	0.0080	1.4450	0.0137	1.3803	0.0081	1.4500
		3	300	0.0041	0.0988	0.0069	0.0999	0.0043	0.0920
		3	400	0.0022	0.0495	0.0036	0.0498	0.0021	0.0461
		3	500	0.0013	0.0297	0.0019	0.0252	0.0012	0.0232
0.3	0.8	∞	200	-0.0007	0.6350	0.0118	1.3114	0.0009	0.6410
		∞	300	-0.0004	0.3177	0.0060	0.0988	0.0005	0.3211
		∞	400	-0.0002	0.1588	0.0031	0.0495	0.0003	0.1620
		∞	500	-0.0001	0.0794	0.0016	0.0248	0.0002	0.0821
0.3	0.8	5	200	0.0017	0.0962	0.0081	0.0921	0.0017	0.0980
		5	300	0.0009	0.0482	0.0042	0.0462	0.0009	0.0492
		5	400	0.0005	0.0243	0.0022	0.0232	0.0005	0.0247
		5	500	0.0003	0.0122	0.0012	0.0161	0.0003	0.0124
0.3	0.8	∞	200	-0.0007	0.6350	0.0118	1.3114	0.0009	0.6410
		∞	300	-0.0004	0.3177	0.0060	0.0988	0.0005	0.3211
		∞	400	-0.0002	0.1588	0.0031	0.0495	0.0003	0.1620
		∞	500	-0.0001	0.0794	0.0016	0.0248	0.0002	0.0821

Tablo 13. GZTP dağılımın parametreleri için HKO değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.3	0.8	5	200	0.0054	124,7534	0.0092	12.5745	0.0055	122.5436
		5	300	0.0028	0.0904	0.0047	0.0963	0.0028	0.0984
		5	400	0.0015	0.0453	0.0024	0.0482	0.0015	0.0493
		5	500	0.0008	0.0227	0.0013	0.0243	0.0008	0.0249
0.3	0.8	∞	200	0.0035	1.1146	0.0089	11.2475	0.0036	1.0234
		∞	300	0.0018	0.0990	0.0045	0.0970	0.0019	0.0980
		∞	400	0.0009	0.0496	0.0023	0.0485	0.0011	0.0510
		∞	500	0.0005	0.0249	0.0012	0.0243	0.0006	0.0261
0.3	0.8	5	200	0.0042	17.9101	0.0099	14.1799	0.0043	12.2094
		5	300	0.0022	0.0980	0.0051	0.0940	0.0021	0.0984
		5	400	0.0012	0.0499	0.0026	0.0473	0.0012	0.0493
		5	500	0.0007	0.0251	0.0014	0.0241	0.0007	0.0247
0.3	0.8	∞	200	0.0035	1.1146	0.0089	11.2475	0.0036	1.0234
		∞	300	0.0018	0.0990	0.0045	0.0970	0.0019	0.0980
		∞	400	0.0009	0.0496	0.0023	0.0485	0.0011	0.0510
		∞	500	0.0005	0.0249	0.0012	0.0243	0.0006	0.0261

Tablo 14. GZTP dağılımın parametreleri için yan değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.8	0.3	3	200	0.0118	0.7616	0.0183	1.5668	0.0120	0.7710
		3	300	0.0595	0.3080	0.0092	0.0982	0.0061	0.3895
		3	400	0.0301	0.1550	0.0047	0.0492	0.0033	0.1948
		3	500	0.0150	0.0751	0.0024	0.0250	0.0017	0.974
0.8	0.3	∞	200	-0.0011	-0.4790	0.0080	1.3857	0.0012	0.4800
		∞	300	-0.0006	-0.2396	0.0042	0.0092	0.0007	0.2401
		∞	400	-0.0004	-0.1197	0.0022	0.0047	0.0004	0.1222
		∞	500	-0.0002	-0.0599	0.0013	0.0024	0.0002	0.0603
0.8	0.3	5	200	0.0063	0.0657	0.0102	1.4314	0.0064	0.0658
		5	300	0.0034	0.0329	0.0057	0.0980	0.0033	0.0330
		5	400	0.0018	0.0164	0.0029	0.0491	0.0017	0.0116
		5	500	0.0009	0.0083	0.0016	0.0250	0.0009	0.0059
0.8	0.3	∞	200	-0.0011	-0.4790	0.0080	1.3857	0.0012	0.4800
		∞	300	-0.0006	-0.2396	0.0042	0.0092	0.0007	0.2401
		∞	400	-0.0004	-0.1197	0.0022	0.0047	0.0004	0.1222
		∞	500	-0.0002	-0.0599	0.0013	0.0024	0.0002	0.0603

Tablo 15. GZTP dağılımın parametreleri için HKO değerleri

				MLE		Oranlar Tahmini		MMLE	
q	θ	L	n	(q)	(θ)	(q)	(θ)	(q)	(θ)
0.8	0.3	3	200	0.0058	10.2300	0.0116	10.4327	0.0059	10.21
		3	300	0.0030	0.0980	0.0059	0.0990	0.0031	0.0960
		3	400	0.0016	0.0491	0.0031	0.0496	0.0016	0.0490
		3	500	0.0009	0.0246	0.0017	0.0249	0.0009	0.0250
0.8	0.3	∞	200	0.0003	0.2516	0.0117	10.5467	0.0004	0.2517
		∞	300	0.0002	0.1259	0.0059	0.0990	0.0003	0.1259
		∞	400	0.0001	0.0631	0.0030	0.0495	0.0002	0.0615
		∞	500	0.0001	0.0316	0.0016	0.0248	0.0001	0.0306
0.8	0.3	5	200	0.0021	9.8981	0.0115	10.7824	0.0022	9.9596
		5	300	0.0011	0.0980	0.0058	0.0943	0.0012	0.0910
		5	400	0.0007	0.0491	0.0030	0.0471	0.0007	0.0456
		5	500	0.0004	0.0248	0.0016	0.0234	0.0003	0.0229
0.8	0.3	∞	200	0.0003	0.2516	0.0117	10.5467	0.0004	0.2517
		∞	300	0.0002	0.1259	0.0059	0.0990	0.0003	0.1259
		∞	400	0.0001	0.0631	0.0030	0.0495	0.0002	0.0615
		∞	500	0.0001	0.0316	0.0016	0.0248	0.0001	0.0306

Tablo 8-15'den itibaren tüm tahminlerin yanlı olduğunu ancak bu tahminlerin asimptotik olarak yansız olduğunu görülmüştür. Her üç tahmin edicinin küçük örneklem durumunda kötü performans sergilediği görülmüştür. En çok olabilirlik tahmini(MLE) ve modifiye edilmiş en çok olabilirlik hata karaler ortalaması her zaman küçük olduğundan Oranlar tahmininden daha iyi performans göstermiştir. Ayrıca, örneklem büyüklüğü arttığında parametrelerinin yansızlığı ve hata kareler ortalaması değerleri beklenildiği gibi düşmektedir.

3.2.5. Gerçek Veri Uygulaması

Bu bölümde üç gerçek veri seti Tip-I sansürlü örnekleme dönüştürülmüş ve sansürlü örneklem durumunda parametre tahmini verilmeye çalışılmıştır. İlk veri seti Xie ve Lai tarafından verilmiş ve analiz edilmiştir. (Xie ve Lai, 1995) Eşit uzunluktaki zaman aralıklarında bozulma sayılarını gösteren veri seti Tablo 16 da verilmiştir. Bu verinin geometrik-sıfırdan budanmış Poisson dağılımına uygunluğu 2019 yılında Akdoğan ve arkadaşları tarafından çalışılmıştır. (Akdoğan ve ark., 2019) Bu veri seti aşağıda verilmiştir:

Tablo 16. Eşit uzunluklu zaman aralıklarında gerçekleşen bozulma sayıları

Zaman	Bozulma sayısı	Zaman	Bozulma sayısı	Zaman	Bozulma sayısı
1	53	7	22	13	13
2	29	8	16	14	5
3	29	9	18	15	5
4	36	10	8	16	4
5	13	11	22	17	1
6	25	12	11	18	1
	$n = 311$				

Veriye ait en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 17'de verilmiştir.

Tablo 17. Veriye ait $GZTP(q,\theta)$ dağılımının parameter tahminleri

	Tam örneklem	L=3	L=5
MLE	$\hat{q} = 0.7904$, $\hat{\theta} = 0.3100$	$\hat{q} = 0.8475$, $\hat{\theta} = 0.8778$	$\hat{q} = 0.7821$, $\hat{\theta} = 0.3487$
Oranlar	$\hat{q} = 0.7876$, $\hat{\theta} = 0.3164$	$\hat{q} = 0.7876$, $\hat{\theta} = 0.3164$	$\hat{q} = 0.7876$, $\hat{\theta} = 0.3164$
MMLE	$\hat{q} = 0.7904$, $\hat{\theta} = 0.3100$	$\hat{q} = 0.8375$, $\hat{\theta} = 0.9278$	$\hat{q} = 0.7894$, $\hat{\theta} = 0.3562$

İkinci veri seti Makcutek tarafından verilmiş ve analiz edilmiştir. (Makcutek, 2008). Bu veri soleven dil ailesindeki yazı birim frekanslarını göstermektedir. Bu verinin geometrik-sıfırdan budanmış Poisson dağılımına uygunluğu 2019 yılında Akdoğan ve arkadaşları tarafından çalışılmıştır. (Akdoğan ve ark., 2019) Bu veri seti aşağıda verilmiştir:

Tablo 18. Soleven dil ailesindeki yazı birim frekansları

i	$f(i)$	i	$f(i)$	i	$f(i)$	i	$f(i)$	i	$f(i)$
1	32036	6	16088	11	13034	16	7446	21	2606
2	31891	7	16084	12	10517	17	6413	22	2554
3	31122	8	15221	13	10514	18	5361	23	2463
4	27150	9	14668	14	10216	19	5055	24	1675
5	22905	10	14043	15	9568	20	4608	25	497
									$N = 313735$

Veriye ait en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 19'da verilmiştir.

Tablo 19. Veriye ait $GZTP(q,\theta)$ dağılımının parameter tahminleri

	Tam örneklem	L=3	L=5
MLE	$\hat{q} = 0.8378$, $\hat{\lambda} = 0.3228$	$\hat{q} = 0.5964$, $\hat{\lambda} = 0.6374$	$\hat{q} = 0.8153$, $\hat{\lambda} = 0.3319$
Oranlar	$\hat{q} = 0.8586$, $\hat{\lambda} = 0.3596$	$\hat{q} = 0.8586$, $\hat{\lambda} = 0.3596$	$\hat{q} = 0.8586$, $\hat{\lambda} = 0.3596$
MMLE	$\hat{q} = 0.8384$, $\hat{\lambda} = 0.3223$	$\hat{q} = 0.6072$, $\hat{\lambda} = 0.6274$	$\hat{q} = 0.8168$, $\hat{\lambda} = 0.3329$

Üçüncü veri seti Bakouch ve arkadaşları (Bakouch ve ark., 2014b) tarafından verilmiş ve analiz edilmiştir. Bu veri lösemi hastalığı tedevisi sırasında rasgele seçilen yirmi hastanın haftalık remisyon sürelerini göstermektedir. Bu verinin geometrik-sıfırdan budanmış Poisson dağılımına uygunluğu 2019 yılında Akdoğan ve arkadaşları tarafından çalışılmıştır. (Akdoğan ve ark., 2019). Yirmi hastaya ait veriler şöyledir:

1 3 3 6 7 7 10 12 14 15 18 19 22 26 28 29 34 40 48 49.

Veriye ait en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 20'de verilmiştir.

Tablo 20. Veriye ait $GZTP(q,\theta)$ dağılımının parameter tahminleri

	Tam örneklem	L=3	L=5
MLE	$\hat{q} = 0.9301$, $\hat{\theta} = 0.2252$	$\hat{q} = 0.7794$, $\hat{\theta} = 0.1654$	$\hat{q} = 0.8832$, $\hat{\theta} = 0.2045$
Oranlar	$\hat{q} = 0.9457$, $\hat{\theta} = 0.2109$	$\hat{q} = 0.9457$, $\hat{\theta} = 0.2109$	$\hat{q} = 0.9457$, $\hat{\theta} = 0.2109$
MMLE	$\hat{q} = 0.9298$, $\hat{\theta} = 0.2264$	$\hat{q} = 0.7659$, $\hat{\theta} = 0.1763$	$\hat{q} = 0.8997$, $\hat{\theta} = 0.2037$

Verilen üç gerçek veri analizinde, en çok olabilirlik, oranlar yöntemi ve modifiye edilmiş en çok olabilirlik tahmin edicileri tam örneklem için hemen hemen aynı sonucu vermiştir. Oranlar yöntemi, sansürlü örneklem durumunda en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicileriden daha iyi performans göstermektedir.

3.3. Binom Kesikli Lindley Dağılımı

Bu bölümde, Kuş ve arkadaşlarının önerdiği Binom-Kesikli Lindley dağılımının Tip 1 sağdan sansürlü örneklem durumunda parametre tahmini yapılacaktır. Binom-Kesikli Lindley dağılımına sahip rasgele değişken olmak üzere olasılık fonksiyonu ve dağılım fonksiyonu sırasıyla

$$f(x)=\frac{p^{2x}\left[\left(p^{3}-(1-p)(1-p+x)\right)\log p+(1-p)\left(1-p(1-p)\right)\right]}{(1-\log p)\left(1-p(1-p)\right)^{x+2}}$$

ve

$$F(x)=1-\frac{\left\{\left[1+(-2-x)\log p\right]p^{2x+2}+\left(p^{2x+3}(1-p)\right)(\log p-1)\right\}\left(p^{2}-p+1\right)^{(-2-x)}}{1-\log p}$$

şeklindedir. Burada $p\in(0,1)$ ‘dir. Dağılım kısaca $BDL(p)$ olarak gösterilmektedir. (Kuş ve ark., 2018)

3.3.1. En Çok Olabilirlik Tahmini

$Y_1,Y_2,...,Y_n$ bağımsız ve aynı $BDL(p)$ dağılımına sahip örneklem olsun ve $L_1,L_2,...,L_n$ her bir rasgele değişken için sağdan sansür zamanı olmak üzere (Bu çalışmada zamanlar eşit alınmıştır.)

$$X_i=\min(Y_i,L_i)\quad i=1,2,...,n$$

Şeklinde tanımlanan yeni rasgele değişken üzerinden

$$\delta_i=\begin{cases}1 & , \quad X_i=Y_i\\ 0 & , \quad X_i\neq Y_i\end{cases}$$

İndikatör fonksiyonu tanımlansın. Böylece $BDL(p)$ dağılımının Tip 1 sağdan sansürlü örneklem için olabilirlik fonksiyonu

$$L(p)=\prod_{i=1}^{n} f(x_i)^{\delta_i}\left(1-F(x_i)\right)^{1-\delta_i}$$

$$=\prod_{i=1}^{n}\left\{\frac{p^{2x_i}\left[\left(p^3-(1-p)(1-p+x)\right)\log p+(1-p)\left(1-p(1-p)\right)\right]}{(1-\log p)\left(1-p(1-p)\right)^{x+2}}\right\}^{\delta_i}$$

$$\left\{\frac{\left[\left(1+(-2-x_i)\log p\right)p^{2x_i+2}-\left(p^{2x_i+3}(1-p)(\log p-1)\right)\right]\left(p^2-p+1\right)^{-2-x_i}}{1-\log p}\right\}^{(1-\delta_i)}$$

ve log-olabilirlik fonksiyonu

$$\ell(p)=\sum_{i=1}^{n}\delta_i\log\left\{\frac{p^{2x_i}\left[\left(p^3-(1-p)(1-p+x)\right)\log p+(1-p)\left(1-p(1-p)\right)\right]}{(1-\log p)\left(1-p(1-p)\right)^{x+2}}\right\}$$
$$+\sum_{i=1}^{n}(1-\delta_i)\log\left\{\frac{\left[\left(1+(-2-x_i)\log p\right)p^{2x_i+2}-\left(p^{2x_i+3}(1-p)(\log p-1)\right)\right]\left(p^2-p+1\right)^{-2-x_i}}{1-\log p}\right\}$$

şeklinde elde edilir. Burada olabilirlik denklemi

$$\frac{\partial\ell(p)}{\partial p}=\sum_{i=1}^{n}\delta_i\left\{\frac{p\left[\left(3p-4+p^3-p^2\right)\log(p)^2+\left(-6p+6-2p^3+4p^2\right)\log p+2p-1+p^3-2p^2\right]}{\left[\left(p^3-p^2+2p-1\right)\log(p)-p^3+2p^2-2p+1\right]\left(1-\log(p)\right)\left(p^2-p+1\right)}\right\}$$
$$+(1-\delta_i)\left\{\frac{\left[\left(-3p+4-p^3+p^2\right)\log(p)^2+\left(6p-6+2p^3-4p^2\right)\log p-p^3+2p^2-2p+1\right]}{\left[\left(2-p+p^2\right)\log(p)-p^2-1+p\right]\left(\log(p)-1\right)p\left(p^2-p+1\right)}\right\}$$

şeklinde elde edilir. Elde edilen denklemin çözümü analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır. (Lagarias ve ark., 1998)

3.3.2. Oranlar Tahmini

Khan ve ark.(1989) tarafından kesikli dağılımlar için kullanılabilen bu yöntem bu tez çalışmasının bu bölümünde geometrik-sıfırdan Budanmış Poisson Dağılımı için kullanılacaktır. $X_1, X_2, \ldots, X_n$ bağımsız ve aynı *BDL(p)* dağılımından alınmış örneklem olmak üzere

$$\Phi = \begin{cases} 1 & , \quad X_i = 0 \\ 0 & , \quad X_i > 0 \end{cases}$$

indikatör fonksiyonu yardımıyla örneklemdeki "0" değerlerinin sayısını ve $Y = \frac{1}{n}\sum_{i=1}^{n}\Phi\left(X\right)_i$ eşitlikle örneklemdeki "0" lerın oranı kolayca elde edilir. Hesaplanan bu oran $f(0) = \frac{\left[\left(p^3 - (1-p)(1-p)\right)\log p + (1-p)\left(1-p(1-p)\right)\right]}{(1-\log p)\left(1-p(1-p)\right)^2}$ olasılığının yansız ve tutarlı tahmin edicisidir. Böylece elde edilen denklem aşağıda verilmiştir.

$$\frac{\left[\left(p^3 - (1-p)(1-p)\right)\log p + (1-p)\left(1-p(1-p)\right)\right]}{(1-\log p)\left(1-p(1-p)\right)^2} = Y$$

elde edilen denklemin çözümü oranlar tahmini olarak elde edilir. (Khan ve ark., 1989)

3.3.3. Modifiye Edilmiş En Çok Olabilirlik Tahmini

Bu bölümde dağılımın parametrelerinin modifiye edilmiş en çok olabilirlik tahmin edicileri Tip-I sansürlü örnekleme dayalı olarak elde edilecektir. $X_1, X_2, \ldots, X_n$, $DBL\left(p\right)$ dağılımından alınmış Tip-I sansürlü örneklem olsun. Log-olabilirlik denklemlerinin p parametresine göre türevi $f\left(p\right)$, parametreye göre Taylor serisinin ilk iki terimi

$$\left.\left(f\left(p\right) + \left(\hat{p}_* - p\right)\frac{\partial f\left(p\right)}{\partial p}\right)\right|_{p=\tilde{p}} = 0$$

ele alınır. Burada $\tilde{p}$, parametrenin oran tahmin edicisidir. Yukarıdaki denklem çözüldüğünde p parametresinin modifiye edilmiş en çok olabilirlik tahmin edicisi

$$\hat{p}_* = -\left.\frac{f\left(p\right) - \frac{\partial f\left(p\right)}{\partial p}p}{\frac{\partial f_1\left(p\right)}{\partial p}}\right|_{p=\tilde{p}}$$

olarak elde edilir. Burada Log-olabilirlik denklemlerinin p parametresine göre türevi $f\left(p\right)$,

$$\frac{\partial \ell(p)}{\partial p}=\sum_{i=1}^{n}\delta_i\left\{\frac{p\left[\left(3p-4+p^3-p^2\right)\log(p)^2+\left(-6p+6-2p^3+4p^2\right)\log p+2p-1+p^3-2p^2\right]}{\left[\left(p^3-p^2+2p-1\right)\log(p)-p^3+2p^2-2p+1\right]\left(1-\log(p)\right)\left(p^2-p+1\right)}\right\}$$

$$+\left(1-\delta_i\right)\left\{\frac{\left[\left(-3p+4-p^3+p^2\right)\log(p)^2+\left(6p-6+2p^3-4p^2\right)\log p-p^3+2p^2-2p+1\right]}{\left[\left(2-p+p^2\right)\log(p)-p^2-1+p\right]\left(\log(p)-1\right)p\left(p^2-p+1\right)}\right\}$$

şeklinde elde edilmiştir. Elde edilen denklemin çözümü türevlerin karmaşık olmasından dolayı analitik olarak zordur. Bu nedenle iteratif yöntemlerle denklemler çözülecektir. Bu denklemlerin çözümünde Nelder-Mead algoritması kullanılmıştır. Bu algoritma MATLAB yazılımında fminsearch komutu ile çalışmaktadır (Lagarias ve ark., 1998).

3.3.4. Simulasyon Çalışması

Simülasyon çalışmasında Binom kesikli Lindley dağılımına sahip rasgele örnekleminden en çok olabilirlik, oranlar ve modifiye edilmiş en çok olabilirlik tahmin edici yöntemleri kullanılarak parametre tahmini ve parametre tahminlerine ilişkin yan ve hata kareler ortalaması değerleri elde edilmiştir. Simülasyon çalışmasında $p=0.3$, $p=0.5$ ve $p=0.8$ değerleri için 10000 tekrar sonucunda elde edilen değerler Tablo 16-21' de sunulmuştur.

Tablo 16. BDL dağılımının $p=0.3$ parametresi için yan değerleri

	n	L	MLE	Oran Tahmini	MMLE
$p=0.3$	100	3	-0.0025	-0.0017	0.0026
	200	3	-0.0009	-0.0007	0.0010
	300	3	-0.0008	-0.0006	0.0009
	400	3	-0.0006	-0.0005	0.0007
	500	3	-0.0005	-0.0004	0.0004
	100	5	-0.0024	-0.0018	0.0025
	200	5	-0.0009	-0.0006	0.0010
	300	5	-0.0009	-0.0006	0.0010
	400	5	-0.0007	-0.0004	0.0008
	500	5	-0.0004	-0.0003	0.0005
	100	∞	-0.0026	-0.0018	0.0027
	200	∞	-0.0022	-0.0016	0.0023
	300	∞	-0.0018	-0.0014	0.0019
	400	∞	-0.0016	-0.0012	0.0017
	500	∞	-0.0014	-0.0010	0.0015

Tablo 17. BDL dağılımının $p = 0.3$ parametresi için HKOdeğerleri

	n	L	MLE	Oran Tahmini	MMLE
$p = 0.3$	100	3	0.0008	0.0009	0.0009
	200	3	0.0004	0.0004	0.0005
	300	3	0.0002	0.0003	0.0003
	400	3	0.0001	0.0002	0.0002
	500	3	0.0001	0.0002	0.0001
	100	5	0.0007	0.0009	0.0008
	200	5	0.0004	0.0004	0.0004
	300	5	0.0002	0.0003	0.0003
	400	5	0.0001	0.0002	0.0002
	500	5	0.0001	0.0002	0.0001
	100	∞	0.0007	0.0009	0.0006
	200	∞	0.0004	0.0005	0.0004
	300	∞	0.0003	0.0003	0.0003
	400	∞	0.0002	0.0002	0.0002
	500	∞	0.0001	0.0001	0.0001

Tablo 18. BDL dağılımının $p = 0.5$ parametresi için yan değerleri

	n	L	MLE	Oran Tahmini	MMLE
$p = 0.5$	100	3	-0.0011	-0.0003	0.0012
	200	3	-0.0004	0.0003	0.0007
	300	3	-0.0004	0.0003	0.0004
	400	3	0.0003	0.0002	0.0003
	500	3	-0.0002	0.0001	0.0002
	100	5	-0.0022	-0.0016	0.0021
	200	5	-0.0012	-0.0009	0.0011
	300	5	-0.0007	-0.0005	0.0007
	400	5	-0.0004	-0.0004	0.0004
	500	5	-0.0003	-0.0003	0.0003
	100	∞	-0.0028	-0.0020	0.0027
	200	∞	-0.0024	-0.0018	0.0023
	300	∞	-0.0020	-0.0016	0.0020
	400	∞	-0.0018	-0.0014	0.0018
	500	∞	-0.0016	-0.0012	0.0016

Tablo 19. BDL dağılımının $p=0.5$ parametresi için HKOdeğerleri

	n	L	MLE	Oran Tahmini	MMLE
$p=0.5$	100	3	0.0005	0.0008	0.0006
	200	3	0.0002	0.0004	0.0004
	300	3	0.0001	0.0003	0.0003
	400	3	0.0001	0.0002	0.0002
	500	3	0.0000	0.0001	0.0001
	100	5	0.0009	0.0008	0.0010
	200	5	0.0007	0.0006	0.0006
	300	5	0.0005	0.0004	0.0004
	400	5	0.0003	0.0002	0.0003
	500	5	0.0001	0.0001	0.0002
	100	∞	0.0005	0.0008	0.0006
	200	∞	0.0004	0.0006	0.0004
	300	∞	0.0003	0.0004	0.0002
	400	∞	0.0002	0.0003	0.0001
	500	∞	0.0001	0.0002	0.0001

Tablo 20. BDL dağılımının $p=0.8$ parametresi için yan değerleri

	n	L	MLE	Oran Tahmini	MMLE
$p=0.8$	100	3	0.0002	0.0040	0.0003
	200	3	0.0001	0.0015	0.0002
	300	3	0.0001	0.0008	0.0001
	400	3	0.0001	0.0005	0.0001
	500	3	0.0001	0.0004	0.0001
	100	5	-0.0020	-0.0012	0.0023
	200	5	-0.0016	-0.0007	0.0012
	300	5	-0.0012	-0.0004	0.0007
	400	5	-0.0008	-0.0002	0.0004
	500	5	-0.0004	-0.0001	0.0003
	100	∞	-0.0030	-0.0025	0.0029
	200	∞	-0.0027	-0.0022	0.0016
	300	∞	-0.0022	-0.0019	0.0012
	400	∞	-0.0018	-0.0017	0.0007
	500	∞	-0.0012	-0.0013	0.0003

Tablo 21. BDL dağılımının $p = 0.8$ parametresi için HKO değerleri

	n	L	MLE	Oran Tahmini	MMLE
$p = 0.8$	100	3	0.0030	0.0025	0.0031
	200	3	0.0027	0.0022	0.0019
	300	3	0.0022	0.0019	0.0013
	400	3	0.0018	0.0017	0.0011
	500	3	0.0012	0.0013	0.0007
	100	5	0.0019	0.0018	0.0017
	200	5	0.0017	0.0015	0.0009
	300	5	0.0015	0.0012	0.0005
	400	5	0.0013	0.0010	0.0003
	500	5	0.0010	0.0007	0.0001
	100	∞	0.0003	0.0009	0.0004
	200	∞	0.0001	0.0006	0.0003
	300	∞	0.0001	0.0003	0.0002
	400	∞	0.0001	0.0002	0.0001
	500	∞	0.0001	0.0001	0.0001

Tablo 16-21'den itibaren tüm tahminlerin yanlı olduğunu ancak bu tahminlerin asimptotik olarak yansız olduğunu görülmüştür. Her üç tahmin edicinin küçük örneklem durumunda kötü performans sergilediği görülmüştür. En çok olabilirlik tahmini(MLE) ve modifiye edilmiş en çok olabilirlik hata karaler ortalaması her zaman küçük olduğundan Oranlar tahmininden daha iyi performans göstermiştir. Ayrıca, örneklem büyüklüğü arttığında parametrelerinin yansızlığı ve hata kareler ortalaması değerleri beklenildiği gibi düşer.

3.3.5. Gerçek Veri Uygulaması

Bu bölümde verilen gerçek veri seti Tip-I sansürlü örnekleme dönüştürülmüş ve sansürlü örneklem durumunda parametre tahmini verilmeye çalışılmıştır. Bu veri seti (Phyo, 1973) ten alınan 10-11 yaş 100 çocuğa ait çürük diş verisidir. Bu veri seti Tablo 22 de verilmiştir:

Tablo 22. Phyo (1973) diş verisi

Diş Sayısı(x)	0	1	2	3	4
Frekans	64	17	10	6	3

Veriye ait en çok olabilirlik tahmini, oranlar tahmini ve modifiye edilmiş en çok olabilirlik tahmini Tablo 23'de verilmiştir.

Tablo 23. Veriye ait $BDL(p)$ dağılımının parametre tahminleri

	Tam örneklem	$L=2$	$L=3$
Yöntem			
MLE	0.1842	0.2117	0.1931
Oranlar Tahmini	0.1824	0.1824	0.1824
MMLE	0.1798	0.2009	0.1875

Gerçek veri analizinde, en çok olabilirlik, oranlar yöntemi ve modifiye edilmiş en çok olabilirlik tahmin edicileri tam örneklem için hemen hemen aynı sonucu vermiştir. Oranlar yöntemi, sansürlü örneklem durumunda en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicileriden daha iyi performans göstermektedir.

5. SONUÇLAR VE ÖNERİLER

5.1. Sonuçlar

Bu tez çalışmasında literatüre yeni giren Akdoğan ve arkadaşlarının 2016 yılında önerdiği düzgün-geometrik ve yine Akdoğan ve arkadaşlarının 2019 yılında önerdiği geometrik-sıfırdan budanmış poisson ve 2019 yılında Kuş ve arkadaşlarının önerdiği Binom-kesikli Lindley dağılımlarının Tip-I sağdan sansürlü örneklem durumda parametre tahmini elde edilmiştir. Parametre tahmini için en çok olabilirlik yöntemi, oranlar yöntemi, modifiye edilmiş en çok olabilirlik yöntemi metotlarından yararlanılmıştır. Elde edilen tahmin edicilerin performansları kıyaslamak için yan ve hata kareler ortalaması kriterini kullanılarak simülasyon çalışması yapılmıştır. Ayrıca gerçek veri uygulaması her teorik sonuç için verişmiştir. Düzgün-geometrik dağılım için yapılan simülasyon çalışmasında farklı parametreler durumunda bias ve hata kareler ortalaması kriterleri incelenmiş ve sonuçlar incelendiğinde küçük örneklem durumunda oranlar tahmini, hem en çok olabilirlik tahmini hem de modifiye edilmiş en çok olabilirlik tahmininden daha iyi performans göstermiştir. Ayrıca sansürleme zamanı arttığında beklenildiği gibi sonuçlar tam örnekleme yaklaşmıştır. Yine örneklem hacmi büyüdüğünde oranlar, en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicilerinin hem yan hem de hata karaler ortalaması birbirine çok yakındır. Yani örneklem hacmi büyükken her üç tahmin edicilerde kullanılabilir. Bu durum gerçek veri analizinde kendini göstermiştir. Bir diğer incelenen dağılım geometrik-sıfırdan budanmış Poisson dağılımı için yapılan simülasyon çalışmasında farklı parametreler durumunda bias ve hata kareler ortalaması kriterleri incelenmiş ve sonuçlar incelendiğinde küçük örneklem durumunda oranlar tahmini, hem en çok olabilirlik tahmini hem de modifiye edilmiş en çok olabilirlik tahmininden daha iyi performans göstermiştir. Ayrıca sansürleme zamanı arttığında beklenildiği gibi sonuçlar tam örnekleme yaklaşmıştır. Yine örneklem hacmi büyüdüğünde oranlar, en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicilerinin hem yan hem de hata karaler ortalaması birbirine çok yakındır. Yani örneklem hacmi büyükken her üç tahmin edicilerde geometrik-sıfırdan budanmış Poisson dağılımı için kullanılabilir. Bu durum gerçek veri analizinde kendini göstermiştir. Gerçek veri analizinde ayrıca sansürleme zamanı arttığında sonuçların tam örnekleme daha çok yaklaştığı görülmüştür. Son olarak incelenen dağılım ise Binom-kesikli Lindley dağılımı olmuş ve yapılan simülasyon çalışmasında farklı parametreler durumunda yan ve hata kareler ortalaması kriterleri incelenmiş ve sonuçlar incelendiğinde küçük örneklem durumunda oranlar tahmini, hem en çok olabilirlik tahmini hem de modifiye edilmiş en çok olabilirlik tahmininden daha iyi performans göstermiştir. Binom-kesikli Lindley dağılımı küçük örneklem durumunda hem düzgün-geometrik hem de geometrik-sıfırdan budanmış Poisson dağılımdan daha iyi performans göstermiştir. Gerçek veri uygulamasında, verilerin her üç dağılıma uyması durumunda Binom-kesikli Lindley dağılımı yan ve hata kareler ortalaması kriterine göre tercih edilebilir. Ayrıca sansürleme zamanı arttığında beklenildiği gibi sonuçlar tam örnekleme yaklaşmıştır. Yine örneklem hacmi büyüdüğünde oranlar, en çok olabilirlik ve modifiye edilmiş en çok olabilirlik tahmin edicilerinin hem yan hem de hata karaler ortalaması birbirine çok yakındır. Yani örneklem hacmi büyükken her üç tahmin edicilerde Binom-kesikli Lindley dağılımı için kullanılabilir. Bu durum gerçek veri analizinde kendini göstermiştir. Gerçek veri analizinde ayrıca sansürleme zamanı arttığında sonuçların tam örnekleme daha çok yaklaştığı görülmüştür. Elde edilen sonuçlar göstermiştir ki, tam örneklem ile sansürlü örnekleme dayalı tahminler bir birine yakındır. Bu sonuç yardımıyla yaşam zamanı birimleri kesikli olan deneylerde sansürlü örneklem kullanmak, zaman ve maliyetin çok önemli olduğu günümüzde oldukça yararlıdır.

5.2. Öneriler

Kesikli dağılımlar üzerinden Tip-I sansürlü örnekleme dayalı parametre tahmini ilk kez 1994 yılında (Kulasekera, 1994) başlanmış ancak geçen sürede elde edilen dağılımlar için çalışılmamıştır. Uzun aradan sonra Bakouch ve ark., (2014) yılında kesikli Lindley dağılımı için Tip-I ve karma sansürlü örnekleme dayalı parametre tahmini çalışmıştır. Yine aynı 2014 yılında Akdoğan ve arkadaşları kesikli Burr dağılımı için tip-I sansürlü örneklem durumunu çalışmıştır. Bu tez çalışmasında literatüre yeni giren düzgün-geometrik, geometrik-sıfırdan budanmış Poisson ve Binom-kesikli Lindley dağılımı sansürlü örneklem durumunda bu alandaki eksiklik giderilmeye çalışılmıştır. Yine bu çalışmada sansürlü örneklem türlerinden sadece tip-I sağdan sansürlü örneklem verilmiştir. Ancak tip-II ve karma sansürlü örneklem türlerinde çalışmalar yoktur. Bakouch ve arkadaşlarının çalışmalarında kesikli Lindley dağılımının karma sansürlü örneklem durumunda parametre tahmini çalıştığı görülmüştür. Bu çalışma tektir. Bakouch ve ark., (2014) çalışmasından esinlenerek, literatüre yeni girmiş kesikli dağılımlarda karma sansürlü örneklem için parametre tahmini çalışılabilir. Gelecek çalışmalarda yine elde edilen kesikli dağılımlar için stres-dayanılılık güvenirliği çalışılabilir.

KAYNAKLAR

Akdoğan, Y., Kuş, C. ve Kınacı, İ., 2014, Poınt Estımatıon Of Parameters In Dıscrete Burr Dıstrıbutıon Based On Type I Censored Sample, *Journal of the Turkish Statistical Association*, 7 (3), 80-86.

Akdoğan, Y., Kuş, C., Asgharzadeh, A., Kınacı, İ. ve Sharafi, F., 2016, Uniform-Geometric distribution *Journal of Statistical Computation and Simulation*, 86 (9), 1754-1770.

Akdoğan, Y., Kuş, C., Bidram, H. ve Kınacı, İ., 2019, Geometric-Zero Truncated Poisson Distribution: Properties And Applications, *Gazi University Journal of Science*, 32, in Press.

Bakouch, H. S., Jazi, M. A. ve Nadarajah, S., 2014a, A new discrete distribution. , *Statistics,*, 48 ((1),), 200–240.

Bakouch, H. S., Jazi, M. A. ve Nadarajah, S., 2014b, A new discrete distribution, *Statistics*, 48, 200-240.

Barbiero, A., 2014, An alternative discrete skew Laplace distribution, *Statistical Methodology*, 16, 47-67.

Chakraborty, S. ve Chakravarty, D., 2012, Discrete gamma distributions: Properties and parameter estimations. Communications in Statistics *Theory and Methods*, 41 (18), 3301–3324.

Déniz, E. G., 2013, A New Discrete Distribution; Proporties and Applications in madical care, *Journal of Applied Statisitcs*, 40 (12), 2760-2770.

Jazi, M. A., Lai, C. D. ve Alamatsaz, M. H., 2010, A discrete inverse Weibull distribution and estimation of its parameters, *Statistical Methodology*, 7 (2), 121–132.

Khan, M. S. A., Khalique, A. ve Abouammoh, A. M., 1989, On Estimating Parameters in a Discrete Weibull Distribution, *IEEE Transactions on Reliability*, 38 (3), 348–350.

Krishna, H. ve Singh Pundir, P., 2009, Discrete Burr and discrete Pareto distributions, *Statistical Methodology*, 6 (2), 177–188.

Kulasekera, K. B., 1994, Approximate MLE's of the Parameters of a Disc.Weibull Dist.with Type I Cen. Data *Microelection Reliability*, 34 (7).

Kuş, C., Akdoğan, Y., Asgharzadeh, A., Kınacı, İ. ve K., K., 2018, Binomial-discrete Lindley distribution, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 68 (1), 401-411.

Lagarias, J. C., Reeds, J. A., Wright, M. H. ve P.E., W., 1998, Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions, *SIAM Journal of Optimization*, 9 (1), 112-147.

Makcutek, J., 2008, A generalization of the geometric distribution and its applicationin quantitative linguistics, *Romanian Rep. Phys.*, 60, 501-509.

Nakagawa, T. ve Osaki, S., 1975, The Discrete Weibull Distribution, *IEEE Transactions on Reliability*, R-24 (5), 300–301.

Nekoukhou, V., Alamatsaz, M. H. ve Bidram, H., 2013, Discrete generalized exponential distribution of a second type, *Statistics*, 47 (4), 876–887.

Phyo, I., 1973, Use of a chain binomial in epidemiology of caries, *Journal of Dental Research*, 52, 750-752.

Roy, D., 2003, The discrete normal distribution, *Communications in Statistics Theory and Methods*, 32 (10), 1871-1883.

Roy, D., 2004, Discrete Rayleigh distribution, *IEEE Transactions on Reliability*, 53 (2).

Stein, W. E. ve Dattero, R., 1984, A New Discrete Weibull Distribution, *IEEE Transactions on Reliability*, R-33 (2), 196–197.

Weinberg, C. R. ve Gladen, B. C., 1986, The beta-geometric distribution applied to comparative fecundability studies, *Biometrics*, 547-560.

Xie, M. ve Goh, T. N., 1993, İmprovement detection by control charts for high yield processes, *İnternational Journal of Quality and Reliability Management*, 10, 24-31.

Xie, M. ve Lai, C. D., 1995, Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function, *Reliability Engineering and System Safety*, 52, 87-93.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	Mehtap Koca Yılmaz
Uyruğu	**:**	T.C
Doğum Yeri ve Tarihi	**:**	Kartal / 22.08.1988
Telefon	**:**	05558506555
e-mail	**:**	mehtapkoca88@gmail.com.tr

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Lise	:	**Fatin Rüştü Zorlu Lisesi, Kartal, İstanbul**	**2006**
Üniversite	:	**Selçuk Üniversitesi, Selçuklu, Konya**	**2011**
Yüksek Lisans	:	**Selçuk Üniversitesi, Selçuklu, Konya**	**Devam ediyor**

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2013	**Pendik Belediyesi**	**İstatistikçi**

YABANCI DİLLER

İngilizce

YAYINLAR*

Koca Yılmaz Mehtap, Akdoğan Yunus, Karakaya Kadir (2018). Parameter Estimation for Uniform-Geometric Distribution Based on Censored Sample. İnternational Statistics Days Conference 2018 (ISDC2018) [Yüksek Lisans Tezinden Elde Edilen Bildiri]