



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

BAZI KESİKLİ DAĞILIMLAR İÇİN İSTATİSTİKSEL SONUÇ ÇIKARIMI

Yunus AKDOĞAN

DOKTORA TEZİ

İstatistik Anabilim Dalı

Eylül-2016
KONYA
Her Hakkı Saklıdır

ÖZET

DOKTORA TEZİ

BAZI KESİKLİ DAĞILIMLAR İÇİN İSTATİSTİKSEL SONUÇ ÇIKARIMI

Yunus AKDOĞAN

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof. Dr. Coşkun KUŞ

2016, 82 Sayfa

Jüri
Prof. Dr. Coşkun KUŞ
Prof. Dr. Halil AYDOĞDU
Doç. Dr. Buğra SARAÇOĞLU
Doç. Dr. İsmail KINACI
Yrd. Doç. Dr. Aydın KARAKOCA

Bu tez çalışmasında, Düzgün-geometrik ve Geometrik-sıfırdan budanmış poisson isimli iki yeni kesikli dağılım önerilmiştir. Dağılımların özellikleri ve istatistiksel sonuç çıkarımı tartışılmıştır. Literatürde olan kesikli dağılımların (Burr XII, Chen ve Pareto vb.) Tip-I sansürlü örnekleme dayalı parametre tahmini üzerine çalışılmıştır. Tahmin edicilerin, yan ve hata kareler ortalaması bakımından performanslarını karşılaştırmak için simülasyon çalışması yapılmıştır. Elde edilen dağılımların modelleme kabiliyetini anlamak için gerçek veri uygulamaları verilmiştir.

Anahtar Kelimeler: Kesikli Dağılımlar, Monte Carlo Simülasyon, Tahmin, Tip-I Sansürlü Örneklem.

ABSTRACT

Ph.D THESIS

STATISTICAL INFERENCE FOR SOME DISCRETE DISTRIBUTIONS

Yunus AKDOĞAN

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCEOF SELÇUK UNIVERSITY
THE DOCTOR OF PHILOSOPHY
INSTATISTICS

Advisor: Prof. Dr. Coşkun KUŞ

2016, 82 Pages

Jury
Prof. Dr. Coşkun KUŞ
Prof. Dr. Halil AYDOĞDU
Assoc. Prof. Dr. Buğra SARAÇOĞLU
Assoc. Prof. Dr. İsmail KINACI
Assist. Prof. Dr. Aydın KARAKOCA

In this thesis, two new discrete distributions are proposed called Uniform-Geometric distribution and Geometric-zero-truncated poisson distribution. Several distributional properties and statistical inferences are discussed. Estimation on parameters based on Type I censored sample are studied for several existing discrete distribution's in literatüre (Burr XII, Chen and Pareto etc.). Simulation study is performed to compare the performance of the different estimates in terms of bias and mean square errors. Real data applications are also presented to understand the modelling abaility of new introduced discrete distributions.

Keywords: Discrete Distributions, Estimation, Monte Carlo Simulation, Type-I Censored Sample.

ÖNSÖZ

Çalışmamın planlanmasında, araştırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteğini esirgemeyen, engin bilgi ve tecrübelerinden yararlandığım, yönlendirme ve bilgilendirmeleriyle çalışmamı bilimsel temeller ışığında şekillendiren danışman hocam Prof. Dr. Coşkun KUŞ'a, tez konularımda ilerleme konusunda yardımcı olan tez izleme komitesi Prof. Dr. Halil AYDOĞDU hocama, bilgi ve yorumları ile katkı sağlayan değerli Doç. Dr. İsmail KINACI hocama, simülasyon çalışmalarında yardımlarını esirgemeyen değerli arkadaşlarım Arş. Gör. Kadir KARAKAYA ve Fatih ŞAHİN'e, proje kapsamında tezin kaynak ve kırtasiye desteğini üstlenen Fen Bilimleri Enstitüsü Müdürlüğü ve Bilimsel Araştırma Projesi Koordinatörlüğü'ne teşekkür ederim. Son olarak bana güvenen ve desteklerini hiç bir zaman esirgemeyen canım aileme, değerli eşim Esra'ya ve bana uğur getiren canım kızım Nilüfer Rana' ya teşekkür ediyorum.

Yunus AKDOĞAN
KONYA-2016

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

KISALTMALAR ..... **ix**

1. GİRİŞ ve KAYNAK ARAŞTIRMASI **1**

2. TEMEL KAVRAMLAR **4**

2.1. Yaşam ve Güvenirlik Fonksiyonu 4
2.2. Bozulma Oranı Fonksiyonu 4
2.3. Çarpıklık Katsayısı 5
2.4. Basıklık Katsayısı 5
2.5. Asimptotik Normallik 5
2.6. Nokta Tahmini 6
2.6.1. En Çok Olabilirlik Tahmini (MLE) 6
2.6.2. Momentler Tahmini 7
2.6.3. Oranlar Methodu Tahmini 7
2.6.4. En küçük kareler tahmini (LSE) 8
2.6.5. Uyarlanmış En Çok Olabilirlik Tahmini 8
2.7. Aralık Tahmini 9
2.8. Fisher Bilgi Matrisi 10
2.9. Hata Kareler Ortalaması ve Yan 10
2.10. LerchPhi Fonksiyonu 11
2.11. Bazı Kesikli Dağılımlar 11
2.11.1. Kesikli Weibull Dağılımı 11
2.11.2. Kesikli Rayleigh Dağılımı 11
2.11.3. Kesikli Burr XII Dağılımı 12
2.11.4. Kesikli Pareto Dağılımı 12
2.11.5. Kesikli Chen Dağılımı 12
2.12. Tip-I Sansürlü Örneklem 13
2.13. Kaplan Meier Tahmin Edicisi 13
2.14. Akaike ve Bayes Bilgi Kriteri 14

3. YENİ DÜZGÜN-GEOMETRİK DAĞILIM **15**

3.1. Dağılıma ait Bazı Özellikler 17
3.1.1. Yaşam Fonksiyonu ve Dağılım Fonksiyonu 17
3.1.2. Moment Üreten ve Olasılık Üreten Fonksiyonu 17
3.1.3. Beklenen Değer, Varyans, Çarpıklık ve Basıklık 18
3.1.4. Bozulma (Hazard) Oranı Fonksiyonu 21

3.2. Parametre Tahmini 22
3.2.1. En Çok Olabilirlik Tahmin Edicisi 22
3.2.2. Oranlar Metodu Tahmini 23
3.2.3. Momentler Yöntemi Tahmini 24
3.2.4. Simülasyon Çalışması 25
3.2.5. Gerçek Veri Uygulaması 26

4. YENİ GEOMETRİK-SIFIRDAN BUDANMIŞ POİSSON DAĞILIMI 30

4.1. Dağılıma ait Bazı Özellikler 31
4.1.1. Beklenen Değer, Varyans, Çarpıklık ve Basıklık 32
4.1.2. Kuantil Fonksiyonu 39
4.1.3. Yaşam, Klasik ve Alternatif Bozulma Oranı Fonksiyonu 40
4.2. Parametre Tahmini 42
4.2.1. En Çok Olabilirlik Tahmin Edicisi 42
4.2.2. En Küçük Kareler Tahmin Edicisi 43
4.2.3. Momentler Yöntemi Tahmini 44
4.2.4. Simülasyon Çalışması 44
4.2.5. Gerçek Veri Uygulaması 46

5. BAZI KESİKLİ DAĞILIMLAR İÇİN SANSÜRLÜ ÖRNEKLEM DURUMUNDA PARAMETRE TAHMİNİ 49

5.1. Kesikli Chen Dağılımı 49
5.1.1. En Çok Olabilirlik Tahmini 49
5.1.2. Oranlar Tahmini 50
5.1.3. En Küçük Kareler Tahmini 51
5.1.4. Simülasyon Çalışması 51
5.1.5. Gerçek Veri Uygulaması 54
5.2. Kesikli Burr Dağılımı 55
5.2.1. En Çok Olabilirlik Tahmini 55
5.2.2. Oran Yöntemi Tahmini 56
5.2.3. Uyarlanmış En Çok Olabilirlik Tahmini 57
5.2.4. Simülasyon Çalışması 58
5.2.5. Gerçek Veri Uygulaması 60
5.3. Kesikli Pareto Dağılımı 61
5.3.1. En Çok Olabilirlik Tahmini 61
5.3.2. Oran Yöntemi Tahmini 62
5.3.3. Uyarlanmış En Çok Olabilirlik Tahmini 63
5.3.4. Simülasyon Çalışması 63
5.3.5. Gerçek Veri Uygulaması 65

6. SONUÇLAR VE ÖNERİLER 66

KISALTMALAR

HKO: Hata Kareler Ortalaması
MLE: En Çok Olabilirlik Tahmini
MPE: Oranlar Metodu Tahmini
MME: Momentler Metodu Tahmini
UMMLE: Güncellenmiş Uyarlanmış En Çok Olabilirlik Tahmini
LSE: En Küçük Kareler Tahmini
MMLE: Uyarlanmış En Çok Olabilirlik Tahmini
UG: Düzgün-Geometrik Dağılım
$GZTP$: Geometrik-Sıfırdan Budanmış Poisson
AIC: Akeike Bilgi Kriteri
BIC: Bayes Bilgi Kriteri
DFR: Azalan Bozulma Oranı Fonksiyonu
IFR: Artan Bozulma Oranı Fonksiyonu

SİMGELER

$\mathbb{R}$: Reel Sayılar Kümesi
$\mathbb{R}^+$: Pozitif Reel Sayılar Kümesi
$L(\boldsymbol{\gamma};\mathbf{X})$: Olabilirlik Fonksiyonu
$\ell(\boldsymbol{\gamma})$: Log-Olabilirlik Fonksiyonu
$\nu(.)$: Gösterge Fonksiyonu

1. GİRİŞ ve KAYNAK ARAŞTIRMASI

Yaşam ve güvenilirlik analizinde ilgilenilen rasgele değişken yaşam zamanıdır ve genellikle sürekli rasgele değişken olarak ele alınır. Elektronik parçalarda dayanma süresi, canlılarda yaşam süresi süreklidir. En çok bilinen yaşam zamanı dağılımları üstel, gamma, log-normal ve Weibull olarak verilebilir. Bu dağılımlar sürekli olarak ele alınan yaşam zamanlarının modellemesinde kullanılmaktadır. Son zamanlarda birçok sürekli yaşam zamanı dağılımları önerilmiştir. Bunlardan bazıları (Kuş, 2007), (Barreto-Souza ve ark., 2011), (Lu ve Daimin, 2012), (Torabi ve Narges, 2012), (Bidram ve ark., 2013), (Mahmoudi ve Mitra, 2012), (Percontini ve ark., 2013), (Rezaei ve ark., 2013), (Mahmoudi ve Afsaneh, 2013), (Ristić ve Saralees, 2014), (Wang ve Ibrahim, 2015), (Delgarm ve Mohammad, 2015) ve (Bagheri ve ark., 2016) olarak sıralanabilir. Bu çalışmalarda önerilen dağılımların dağılımsal özellikleri incelenmiş, tam ve(veya) sansürlü örneklem durumunda parametre tahmini konusu ele alınmıştır.

Her ne kadar yaşam zamanı sürekli bir rasgele değişken olarak düşünülse de kesikli olduğu durumlar da vardır. Hafta sayısı, döngü sayısı, adım sayısı ve şok sayısı gibi değişkenler kesikli rasgele değişkenlerdir. Bir elektrik lambasının anahtarı için açma-kapama sayısı, beyin tümörü olan hastanın yaşayacağı hafta sayısı kesikli bir rastgele değişkendir.

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. Tarih sırasıyla bulunan dağılımlar, Kesikli Weibull(Nakagawa ve Osaki, 1975), Kesikli Tip II Weibull(Stein ve Dattero, 1984), Kesikli Normal(Roy, 2003), Kesikli Rayleigh(Roy, 2004), Keslikli Laplace(Inusah ve Kozubowski, 2006), Kesikli Burr ve Kesikli Pareto(Krishna ve Pundir, 2009), Kesikli Ters Weibull(Jazi ve ark., 2010), Kesikli modifiye edilmiş Weibull(Noughabi ve ark., 2011), Kesikli Genelleştirilmiş Üstel (Nekoukhou ve ark., 2012), Kesikli Gamma(Chakraborty ve Chakravarty, 2012), Kesikli ikinci tip Genelleştirilmiş Üstel(Nekoukhou ve ark., 2013b), Kesikli Beta-Üstel(Nekoukhou ve ark., 2013a), Kesikli Bathtube-shape dağılımı(Noughabi ve ark., 2013), Kesikli Lindley

dağılımı(Bakouch ve ark., 2014), Kesikli ters Rayleigh(Hussain ve Ahmad, 2014), Kesikli Gumbel(Chakraborty ve Chakravarty, 2014), Üstelleştirilmiş kesikli Weibull(Nekoukhou ve Bidram, 2015) şeklindedir. Bu dağılımlar bilinen sürekli rasgele değişkenlerin tam değerleri alınarak elde edilmiştir.

Ayrıca bilinen iki dağılımın birleştirilmesi (compounding) ile de yeni dağılımlar önerilmiştir. Hu ve ark. Binom dağılımı ile Poisson dağılımını birleştirerek yeni bir dağılım önermiştir(Hu ve ark., 2007). Daha sonra Déniz, aynı yöntemle düzgün ve poisson dağılımını birleştirerek literatüre yeni bir dağılım kazandırmıştır(Déniz, 2013).

İstatistiksel sonuç çıkarımı yapmak için tam örneklem almak her zaman mümkün olmayabilir. Bir klinikte tedavi gören hastalara ilişkin veriler, eksiksiz gözlenemeyebilir veya pahalı bir elektronik parçanın yaşam zamanı hakkında bilgi edinmek için yapılan yaşam testinde, parçaların hepsinin bozulmalarının gözlenmesi, maliyeti ve test zamanını artıracağından istenmeyebilir. Bu tip durumlarda sansürlenmiş örneklem elde edilir. Tıp, biyoloji, sigortacılık, mühendislik, kalite kontrol ve birçok alanda sansürlenmiş verilerle karşılaşılmaktadır.

Güvenilirlik teorisi ve yaşam analizinde yaşam testleri çeşitli sansür şemaları altında yapılmakta ve sansürlü örneklemler elde edilmektedir. Kesikli dağılımlar için 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., 2014), Kesikli Lindley Dağılımı için Tip-I ve Tip-II sansürlü örnekleme dayalı parametre tahmini yapmıştır.

Bu tezde, dağılım birleştirme yöntemleri kullanılarak kesikli olan Düzgün-Geometrik ve Geometrik-Sıfırdan budanmış Poisson dağılımları önerilmiştir. Önerilen dağılımların dağılımsal özellikleri incelenmiş ve bu dağılımların parametreleri için istatistiksel sonuç çıkarımı yapılmıştır. Parametre tahmini için, En Çok Olabilirlik (ML), Uyarlanmış En Çok Olabilirlik (MML), En küçük Kareler (LS), Oranlar (MP) ve Momentler yöntemleri (MM) kullanılmıştır. Bu tahmin edicilerin performansları Monte Carlo Simülasyonu ile karşılaştırılmıştır. ML tahmin edicilerinin asimptotik normallik özelliği kullanılarak dağılım parametrelerinin asimptotik güven aralıkları inşa edilmiştir.

Bu tez çalışmasının ikinci bölümünde, tezde kullanılan temel kavramlar üzerinde durulmuş ve literatürde bulunan bazı kesikli dağılımlar tanıtılmıştır. Üçüncü ve dördüncü bölümde ise dağılım birleştirme yöntemleri kullanılarak yeni kesikli dağılımlar önerilmiş ve parametrelerinin istatistiksel sonuç çıkarımları verilmiştir.

Beşinci bölümde, daha önce literatürde olan Kesikli Burr, Kesikli Chen ve Kesikli Pareto dağılımlarının parametrelerinin Tip-I sansürlü örnekleme dayalı ML, MML, MP, LS, MM tahmin edicileri elde edilmiş, hata kareler ortalamaları ve yanları simülasyon yardımıyla karşılaştırılmıştır. Sonuçlar ve öneriler Altıncı Bölümde verilerek tez çalışması sonlandırılmıştır.

2. TEMEL KAVRAMLAR

Bu bölümde tezde kullanılan temel kavramlar verilmiştir.

2.1. Yaşam ve Güvenirlik Fonksiyonu

Bir X rasgele değişkeninin yaşam fonksiyonu,

$$\bar{F}(x) = P(X \geq x),\ x \geq 0$$

şeklinde tanımlansın. Yaşam fonksiyonu monoton artmayan bir fonksiyondur ve

$$\bar{F}(0) = \lim_{x \to 0} \bar{F}(x) = 1$$

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

özelliklerini sağlar (Leemis, 1986).

2.2. Bozulma Oranı Fonksiyonu

Sürekli bir X rasgele değişkenin bozulma oranı fonksiyonu

$$h(x) = \lim_{m \to 0} \frac{P(x \leq X < x+m \mid X > x)}{m}$$

$$= \lim_{m \to 0} \frac{1}{m} \frac{P(x \leq X < x+m,\ X > x)}{P(X > x)}$$

$$= \frac{1}{\bar{F}(x)} \lim_{m \to 0} \frac{P(x \leq X < x+m)}{m}$$

$$= \frac{f(x)}{\bar{F}(x)}$$

biçiminde tanımlanır, burada f, olasılık (yoğunluk) fonksiyonudur(Leemis, 1986). Kesikli halde de bozulma oranı fonksiyonu benzer biçimde

$$h(x) = \frac{f(x)}{\bar{F}(x)}$$

şeklinde tanımlanır. Bozulma oranı fonksiyonu genellikle sabit bozulma oranı, artan bozulma oranı (IFR) ya da azalan bozulma oranı (DFR) şeklinde olmaktadır.

2.3. Çarpıklık Katsayısı

Beklenen değeri μ, varyansı σ^2 olan bir X rasgele değişkeninin üçüncü momenti sonlu olsun. Bu durumda X rasgele değişkeninin *Çarpıklık Katsayısı*

$$\gamma_1 = E\left(\frac{X-\mu}{\sigma}\right)^3$$

şeklinde tanımlanır.

Çarpıklık katsayısı, olasılık (yoğunluk) fonksiyonunun ortalama etrafındaki asimetrikliğin ölçüsü olarak yorumlanır(Casella ve Berger, 1990). $\gamma_1 > 0$ ise dağılım sağa çarpıktır, $\gamma_1 < 0$ ise dağılım sola çarpıktır denir(Roussas, 1973). X rasgele değişkeninin olasılık yoğunluk fonksiyonu $f(x)$, $a \in R$ noktası etrafında simetrik ise $\gamma_1 = 0$ dır.

2.4. Basıklık Katsayısı

Beklenen değeri μ, varyansı σ^2 olan bir X rasgele değişkeninin dördüncü momenti sonlu olsun. Bu durumda X rasgele değişkeninin *Basıklık Katsayısı*

$$\gamma_2 = E\left(\frac{X-\mu}{\sigma}\right)^4$$

şeklinde tanımlanır(Roussas, 1973).

Yukarıdaki tanım ele alındığında, normal dağılım için basıklık katsayısı 3 olarak elde edilir.

Basıklık Katsayısı olasılık (yoğunluk) fonksiyonunun basıklık derecesinin ölçüsü olarak yorumlanır. $\gamma_2 > 3$ normal dağılıma göre sivri, $\gamma_2 < 3$ normal dağılıma göre daha basıktır denir.

2.5. Asimptotik Normallik

(X_n) rasgele değişkenlerin bir dizisi , Z, standart normal dağılıma sahip bir rasgele değişken ve $\xrightarrow{d}$, dağılımda yakınsamayı göstermek üzere,

$$\frac{X_n - a_n}{b_n} \xrightarrow{d} Z$$

olacak şekilde reel sayıların (a_n) ve pozitif reel sayıların (b_n) dizileri varsa, (X_n) dizisine asimptotik normal veya daha açık olarak "a_n ortalaması" ve "b_n^2 varyansı" ile asimptotik normal dizisi denir ve $X_n \sim AN(a_n, b_n^2)$ biçiminde gösterilir. Buradaki a_n sayısı X_n'in beklenen değeri ve b_n^2 sayısı X_n nin varyansı olmayabilir. Bu değerler sırasıyla X_n'in asimptotik ortalama ve asimptotik varyans değerleridir. (Akdi, 2005).

2.6. Nokta Tahmini

Dağılımı bilinen fakat parametreleri bilinmeyen bir kitlenin parametrelerinin tahmin edilmesi istatistik biliminin en önemli problemlerindendir. Kitle parametreleri, kitleden alınan bir örneklem yardımıyla oluşturulan istatistiklerle tahmin edilir. Bu şekilde elde edilen tahminlere nokta tahmini denir. Tezde kullanılan nokta tahmin yöntemleri aşağıdaki bölümlerde verilmiştir.

2.6.1. En Çok Olabilirlik Tahmini (MLE)

Parametresini tahmin etmek istenilen kitle $f(x,\gamma),\ \gamma \in \Gamma$ dağılımına sahip olsun. Burada γ kitle parametresini, Γ, parametre uzayını temsil etmektedir. Bu kitleden alınan ve her biri aynı $f(x,\gamma),\ \gamma \in \Gamma$ dağılımına sahip $X_1, X_2, \ldots, X_n$ rasgele değişkenlerine n-birimlik bir *örneklem* denir. Örneklemin olasılık (yoğunluk) fonksiyonu,

$$L(\mathbf{x},\gamma) = f(x_1, x_2, \ldots, x_n, \gamma)$$

biçimindedir. $L(\mathbf{x},\gamma)$, γ'nın bir fonksiyonu olarak düşünüldüğünde olabilirlik fonksiyonu adını alır ve $\ell(\gamma;\mathrm{x})$ biçiminde gösterilir(Roussas, 1973).

Örneklemin bilinmeyen parametre içermeyen Borel ölçülebilir bir 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(Roussas, 1973).

$X_1, X_2, \ldots, X_n$, $f(x,\gamma)$ $,\gamma \in \Gamma \subset R^r$ dağılımından alınmış tam veya sansürlü örneklem olmak üzere $L(\gamma;\mathrm{x})$ i supremum yapan $\hat{\gamma} = \hat{\gamma}(X_1, X_2 ..., X_n)$ istatistiğine γ 'nın ML tahmin edicisi denir.

2.6.2. Momentler Tahmini

$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k) \in \Theta \subset \mathbb{R}^k$ olmak üzere, $k(k \geq 1)$ bileşenli $\boldsymbol{\theta}$ parametre vektörünü tahmin edilmek istensin. $X_1, X_2, \ldots, X_n$ olasılık (yoğunluk) fonksiyonu $f(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta$ olan dağılımdan bir örneklem olmak üzere, var olması halinde, dağılım kitle momentleri,

$$\mu_r = E_{\boldsymbol{\theta}}\left(X_1^{\,r}\right) \quad , \quad r = 1, 2, 3, \ldots$$

ve örneklem momentleri,

$$\hat{\mu}_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \quad , \quad r = 1, 2, 3, \ldots$$

olsun. Kitle momentleri ile örneklem momentlerinden ilk k tanesinin eşitlenmesiyle elde edilen ve $\theta_1, \theta_2, \ldots, \theta_k$ bilinmeyenlerine göre k tane denklemden oluşan denklem sisteminin çözümü olan, $\hat{\theta}_i(X_1, X_2, \ldots, X_n)$, $i = 1,2,...,k$ istatistiklerine momentler tahmin edicileri denir (Akdi, 2005).

2.6.3. Oranlar Methodu Tahmini

Bu metot (Khan ve ark., 1989) tarafından kesikli Weibull dağılımının parametrelerinin tahmini için önerilmiştir. Bu tahmin yöntemi kesikli dağılımların parametreleri için kullanılmaktadır.

$X_1, X_2, \ldots, X_n$, en küçük aldığı değer "1" olan herhangi bir kesikli dağılımından alınmış örneklem ve $i = 1, 2, \ldots, n$ için $\nu(.)$ gösterge fonksiyonu

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

şeklinde olsun. $Y = \frac{1}{n}\sum_{i=1}^{n} \nu(X_i)$, X rasgele değişkeninin 1 değerini alması olasılığının bir tahmin edicisi olacaktır. Y tahmin edicisi X'in f olasılık fonksiyonunun 1 deki

değerine(parametreye bağlı) eşitlenerek elde edilen

$$f(1) = Y$$

denklemin çözümü oranlar tahmini olarak adlandırılır. Parametre sayısı birden fazla olduğunda momentler tahminindeki gibi adımlar uygulanır.

2.6.4. En küçük kareler tahmini (LSE)

$X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$, F dağılım fonksiyonuna sahip kitleden alınan örnekleme ait sıra istatistikleri olsun. F_n^*, ampirik dağılım fonksiyonu olmak üzere

$$F_n^*\left(X_{(i)}\right) = F\left(X_{(i)}\right) + \varepsilon_i,\ i = 1,2,\ldots,n$$

modeli yazılabilir, burada ε, hata terimidir. LSE tahmin edicisi hata kareler toplamını

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left\{ \frac{i}{n} - F\left(X_{(i)}\right) \right\}^2$$

minimum yapan değeridir.

2.6.5. Uyarlanmış En Çok Olabilirlik Tahmini

$X_1, X_2, \ldots, X_n$, herhangi bir dağılımdan alınmış örneklem olsun. ML tahmin edicisi, olabilirlik denklemlerinin çözümünden elde edildiğinden bazen analitik sonuçlar elde edilemeyebilir(parametreye göre lineer olmayan olabilirlik fonksiyonunun varlığı). Bu durumda, ML tahminlerinin elde edilmesinde iteratif yöntemler kullanılmaktadır. Bu yöntemlerin kullanılmasında da bazı sorunlarla karşılaşılabilmektedir. Başlangıç değerlerinin seçimi yakınsamayı etkileyebilir veya olabilirlik fonksiyonu tek tepeli olmayabilir. MML tahmin edicileri, oran tahmin edicileri kullanarak analitik olarak elde edilir ve iteratif yöntemlerin kullanılması zorunluluğu ortadan kaldırılır. MML tahmin edicileri, ML denklemleri üzerinden parametreye göre Taylor açılımının ilk iki teriminin alınması ve ortak çözülmesi ile elde edilir. Tahmin yapılacak dağılımın ilgili parametreleri q ve β, olabilirlik denklemlerinin

$$\begin{aligned} f_1(q,\beta) &= 0 \\ f_2(q,\beta) &= 0 \end{aligned} \qquad (2.6.1)$$

olduğu varsayılsın.

$\tilde{q}$ ve $\tilde{\beta}$ parametrelerin oran tahmin edicisini ifade etmek üzere, Eşitlik(2.6.1)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.

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

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

Bu durumunda parametrelerin MML tahmin edicileri

$$\hat{\beta}_* = \left.\frac{-f_2(q,\beta)\frac{\partial f_1(q,\beta)}{\partial q}+\frac{\partial f_2(q,\beta)}{\partial q}f_1(q,\beta)-\frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}\beta+\frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q}\beta}{\frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q}-\frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}}\right|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} \qquad (2.6.2)$$

ve

$$\hat{q}_* = \left.-\frac{f_1(q,\beta)-\frac{\partial f_1(q,\beta)}{\partial q}q+\frac{\partial f_1(q,\beta)}{\partial \beta}\hat{\beta}_*-\frac{\partial f_1(q,\beta)}{\partial \beta}\beta}{\frac{\partial f_1(q,\beta)}{\partial q}}\right|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} \qquad (2.6.3)$$

şeklinde elde edilir. Elde edilen MML tahmin edicilerini her defasında tekrar denklem (2.6.2) ve (2.6.3) da yerine yazılması durumunda birkaç adımda MML tahmin edicisi ML tahmini ile hemen hemen aynı olmaktadır(Tiku ve Akkaya, 2004).

2.7. Aralık Tahmini

$X_1, X_2, \ldots, X_n$, $f(x,\gamma)$ $,\gamma \in \Gamma \subset R$ dağılımından alınmış tam veya sansürlü örneklem olsun. Rasgele aralık, en az bir sınır noktası rasgele değişken olan sonlu veya sonsuz aralıktır.

$L:\Re^n \to \Re$ ve $U:\Re^n \to \Re$, $\forall x \in \Re^n$ için $L(x) \le U(x)$ koşulunu sağlayan Borel ölçülebilir iki fonksiyon olmak üzere, L ve U fonksiyonları yardımıyla oluşturulan $\left[L(X_1, X_2, \ldots, X_n), U(X_1, X_2, \ldots, X_n)\right]$ aralığı aşağıdaki (2.7.1) eşitsizliğini sağlarsa, γ parametresi için $1-\alpha$ $(0<\alpha<1)$ anlam seviyeli güven aralığı adını alır.

$$P_\theta\left[L\left(X_1,X_2,\ldots,X_n\right)\le\gamma\le U\left(X_1,X_2,\ldots,X_n\right)\right]\ge 1-\alpha,\quad \forall\gamma\in\Gamma \tag{2.7.1}$$

2.8. Fisher Bilgi Matrisi

$X_1,X_2,...,X_n$ örneklemi, olasılık (yoğunluk) fonksiyonu $f(x;\boldsymbol{\gamma}),\boldsymbol{\gamma}\in\Re^p$ olan kitleden alınan n birimlik bir örneklem olsun. Örneklemin olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$L\left(\boldsymbol{\gamma};\mathbf{x}\right)=f\left(\mathbf{x};\boldsymbol{\gamma}\right),\ \boldsymbol{\gamma}\in\Gamma\subset\Re^p \tag{2.8.1}$$

$$\ell\left(\boldsymbol{\gamma}\right)=\log\left(L\left(\boldsymbol{\gamma};\mathbf{x}\right)\right),\quad \boldsymbol{\gamma}\in\Gamma\subset\Re^p \tag{2.8.2}$$

şeklindedir. Bu örneklem için Fisher bilgi matrisi

$$\begin{aligned} I(\boldsymbol{\gamma}) &= -E\left\{\left(\frac{\partial}{\partial\boldsymbol{\gamma}}\log\left(L\left(\boldsymbol{\gamma};\mathbf{X}\right)\right)\right)\right\} \\ &= -E\left\{\left(\frac{\partial}{\partial\boldsymbol{\gamma}}\ell\left(\boldsymbol{\gamma}\right)\right)\right\} \\ &= -\begin{pmatrix} \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_1{}^2} & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_1\partial\gamma_2} & \cdots & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_1\partial\gamma_p} \\ \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_2\partial\gamma_1} & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_2{}^2} & \cdots & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_2\partial\gamma_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_p\partial\gamma_1} & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_p\partial\gamma_2} & \cdots & \frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_p{}^2} \end{pmatrix} \end{aligned} \tag{2.8.3}$$

şeklinde tanımlanır(Wu ve Kuş, 2009).

2.9. Hata Kareler Ortalaması ve Yan

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

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

şeklinde tanımlanır. θ parametresinin bir T tahmin edicisi için hata kareler ortalaması(HKO) da

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

biçiminde tanımlanır. Hata kareler ortalaması

$$HKO_T(\theta) = Var(T) + \left(Yan(T,\theta)\right)^2$$

şeklinde de ifade edilebilir.

2.10. LerchPhi Fonksiyonu

LerchPhi fonksiyonu (Gradshteyn ve Ryzhik, 2007)

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

biçiminde tanımlanır.

2.11. Bazı Kesikli Dağılımlar

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

2.11.1. Kesikli Weibull Dağılımı

X, kesikli Weibull dağılımına sahip bir rasgele değişken olduğunda X in olasılık ve dağılım fonksiyonu sırasıyla,

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

ve

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

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

2.11.2. Kesikli Rayleigh Dağılımı

X, kesikli Rayleigh dağılımına sahip bir rasgele değişken olduğunda X in olasılık ve dağılım fonksiyonu sırasıyla,

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

ve

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

şeklinde tanımlanır (Roy, 2004). Burada $0<q<1$ dağılımın parametresidir.

2.11.3. Kesikli Burr XII Dağılımı

X, kesikli Burr XII dağılımına sahip bir rasgele değişken olduğunda X in 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

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

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

2.11.4. Kesikli Pareto Dağılımı

X, kesikli Pareto dağılımına sahip bir rasgele değişken olduğunda X in olasılık ve dağılım fonksiyonu sırasıyla,

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

ve

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

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

2.11.5. Kesikli Chen Dağılımı

X, kesikli Chen dağılımına sahip bir rasgele değişken olduğunda X in olasılık ve dağılım fonksiyonu sırasıyla,

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

ve dağılım fonksiyonu

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

şeklinde tanımlanır (Noughabi ve ark., 2013). Burada $0 < q < 1$ ve $\beta > 0$ dağılımın parametreleridir.

2.12. Tip-I Sansürlü Örneklem

$Y_1, Y_2, \ldots, Y_n$, bağımsız ve olasılık fonksiyonu $f_Y(y;\theta), \theta \in \Re^p$ olan kitleden alınmış n birimlik örneklem olsun ve $L_1, L_2, \ldots, L_n$ sabitleri sırasıyla her bir $Y_1, Y_2, \ldots, Y_n$ rasgele değişkenlerine ilişkin sağdan sansür zamanlarını göstersin. Ayrıca $i = 1, \ldots, n$ için

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

ve

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

tanımlansın. Burada δ_i, yaşam zamanı Y_i 'nin sansürlü olup olmadığını göstermektedir. Bu şekilde tanımlanan $X_1, X_2, \ldots, X_n$ örneklemine Tip-I sansürlü örneklem denir. Bu durumda $\left(X_1, \delta_1\right), \left(X_2, \delta_2\right), \ldots, \left(X_n, \delta_n\right)$ 'in ortak olasılık fonksiyonu

$$f\left(x_1, x_2, \ldots, x_n; \theta\right) = \prod_{i=1}^{n} f_Y\left(x_i; \theta\right)^{\delta_i} \left(1 - F_Y\left(x_i; \theta\right)\right)^{1-\delta_i}$$

(5.2.1)

şeklinde elde edilir (Lawless, 1982).

2.13. Kaplan Meier Tahmin Edicisi

$X_1, X_2, ..., X_n$, olasılık (yoğunluk) fonksiyonu $f(x;\theta), \theta \in \Re^p$ olan kitleden alınan n birimlik bir örneklem olsun. Yaşam testine tabi tutulan n tane özdeş parçanın ölüm zamanları sıra istatistikleri $t_{(1)} \le t_{(2)} \le \cdots \le t_{(n)}$ şeklinde verilsin. $t_{(i)}$ zamanında ölen bireylerin sayısını n_i ve risk altında bulunan birey sayısını ise r_i olsun. Bu durumunda r_i : $t_{(i)}$ zamanında yaşayan birey sayısıdır. n_i / r_i ifadesi ise $t_{(i)}$ zamanından önce sağ olan

bir bireyin $t_{(i)}$ zamanında ölmesi olasılığıdır. Bu durumunda yaşam fonksiyonunun Kaplan-Meier tahmin edicisi

$$\hat{S}(t)=\begin{cases}\prod_{t_{(i)}\le t}\left[1-\dfrac{n_i}{r_i}\right] & , \quad t_{(i)}\le t\\ 1 & , \quad t_{(i)}>t\end{cases}$$

şeklinde ifade edilir (Lawless, 1982).

2.14. Akaike ve Bayes Bilgi Kriteri

$X_1, X_2, ..., X_n$, olasılık (yoğunluk) fonksiyonu $f(x;\boldsymbol{\gamma}), \boldsymbol{\gamma}\in\Re^p$ olan kitleden alınan n birimlik bir örneklem olsun. Örnekleme ait olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$L(\boldsymbol{\gamma};\mathbf{x})=f(\mathbf{x};\boldsymbol{\gamma}),\ \boldsymbol{\gamma}\in\Gamma\subset\Re^p \tag{2.8.1}$$

$$\ell(\boldsymbol{\gamma})=\log\left(L(\boldsymbol{\gamma};\mathbf{x})\right),\quad \boldsymbol{\gamma}\in\Gamma\subset\Re^p \tag{2.8.2}$$

şeklinde olsun. $\hat{\boldsymbol{\gamma}}$, $\boldsymbol{\gamma}$ nın ML tahmin edicisi olmak üzere Akaike Bilgi Kriteri(AIC)

$$AIC=-2\ell(\hat{\boldsymbol{\gamma}})+2p$$

ve Bayes Bilgi Kriteri(BIC)

$BIC=-2\ell(\hat{\boldsymbol{\gamma}})+p\log(n)$ şeklindedir(Fabozzi ve ark., 2014).

3. YENİ DÜZGÜN-GEOMETRİK DAĞILIM

Bu bölümde, kesikli düzgün dağılım ile geometrik dağılım birleştirilmiş ve yeni bir dağılım olan Düzgün-Geometrik Dağılımı elde edilmiştir. N, p parametreli Geometrik dağılımına sahip ve $X \mid N = n$, n parametreli $\left(U(n), n > 0\right)$ kesikli düzgün dağılıma sahip iki rasgele değişken olmak üzere olasılık fonksiyonları sırasıyla

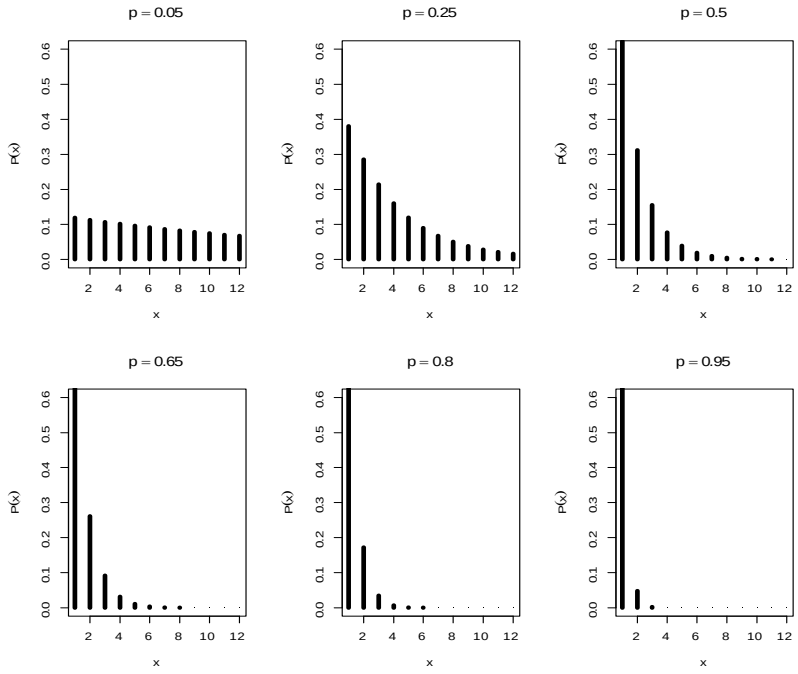
$$P(X = x \mid N = n) = \frac{1}{n}, n = 1, ..., n$$

$$P(N = n) = p(1-p)^{n-1}$$

verilmiştir. X rasgele değişkeninin marjinal dağılımı ise

$$\begin{aligned} P(X = x) &= \sum_{n=x}^{\infty} P(X = x \mid N = n) P(N = n) \\ &= \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} \\ &= \sum_{j=0}^{\infty} \frac{1}{x+j} p(1-p)^{x+j-1} \\ &= p(1-p)^{x-1} \sum_{j=0}^{\infty} \frac{(1-p)^{j}}{x+j} \\ &= p(1-p)^{x-1} \, LerchPhi\left[(1-p), 1, x\right], x = 1, 2, \ldots, \quad 0 < p < 1, \end{aligned} \tag{3.1.1}$$

olarak bulunur. Burada $LerchPhi(z, a, v)$, LerchPhi fonksiyonudur. Düzgün-Geometrik dağılımı olarak isimlendirilen yeni dağılımının olasılık fonksiyonu (3.1.1) denkleminde verilmiştir ve $X \sim UG(p)$ şeklinde gösterilir. Aşağıda farklı p değerleri için $UG(p)$ dağılımın olasılık fonksiyonunun grafikleri çizilmiştir.



Şekil 3.1.1. Farklı p parametreleri için Düzgün-Geometrik Dağılımı Grafikleri

Şekil 3.1.1 deki grafiklere göre, p nin farklı durumları için x değeri arttıkça olasılık fonksiyonunun azalan olduğu görülmektedir. Bu durum düzgün-geometrik dağılımın modunun 1 ve tek olduğunu ifade eder. İfadelerin ispatı için $P(X=x+1)<P(X=x)$ ifadesinin gösterilmesi gerekir. (Gradshteyn and Ryzhik, 2007) den alınan

$$LerchPhi(z,a,\nu)=z^m LerchPhi(z,a,\nu+m)+\sum_{n=0}^{m-1}(\nu+n)^{-a}z^n,$$

eşitliği yardımıyla $P\left(X=x+1\right)$ olasılığı

$$\begin{aligned}P(X=x+1)&=p(1-p)^x LerchPhi(1-p,1,x+1)\\&=p(1-p)^x\left[\frac{LerchPhi(1-p,1,x)}{1-p}-\frac{1}{x(1-p)}\right]\\&=p(1-p)^{x-1}LerchPhi(1-p,1,x)-\frac{p(1-p)^{x-1}}{x}\\&=P(X=x)-\frac{p(1-p)^{x-1}}{x}.\end{aligned}$$

şeklinde elde edilir. Burada $x=1,2,\ldots$ için $p(1-p)^{x-1}/x>0$ olduğundan $P(X=x+1)<P(X=x)$ ifadesi doğrudur. Buna göre $UG\left(p\right)$ dağılımının modu 1 ve tek tepelidir.

3.1. Dağılıma ait Bazı Özellikler

Tezin bu bölümünde $UG(p)$ dağılımına ait dağılımsal özellikler verilmiştir.

3.1.1. Yaşam Fonksiyonu ve Dağılım Fonksiyonu

Aşağıda verilen (3.1.2) eşitliği göz önüne alınsın.

$$\begin{aligned}\sum_{n=0}^{\infty} n\frac{(1-p)^n}{n+k} &= \sum_{n=k}^{\infty}(n-k)\frac{(1-p)^{n-k}}{n}\\ &= (1-p)^{-k}\sum_{n=k}^{\infty} n\frac{(1-p)^n}{n} - k(1-p)^{-k}\sum_{n=k}^{\infty}\frac{(1-p)^n}{n}\\ &= \frac{1}{p} - kLerchPhi(1-p,1,k).\end{aligned} \tag{3.1.2}$$

Eşitlik (3.1.2) yardımıyla $UG(p)$ dağılımının yaşam fonksiyonu

$$\begin{aligned}S(x) = P(X \ge x) &= \sum_{k=x}^{\infty}\sum_{n=k}^{\infty}\frac{p(1-p)^{n-1}}{n}\\ &= \sum_{n=x}^{\infty}\sum_{k=x}^{n}\frac{p(1-p)^{n-1}}{n}\\ &= \sum_{n=x}^{\infty}(n-x+1)\frac{p(1-p)^{n-1}}{n}\\ &= p(1-p)^{x-1}\left[\frac{1}{p} - (x-1)LerchPhi(1-p,1,x)\right],\ x=1,2,...\end{aligned}$$

şeklinde elde edilir. Ayrıca yaşam dağılım fonksiyonu

$$\begin{aligned}F(x) &= P(X \le x)\\ &= 1 - P(X \ge x+1)\\ &= 1 - p(1-p)^x\left[\frac{1}{p} - x\ LerchPhi(1-p,1,x+1)\right],\ x=1,2,...\end{aligned}$$

olarak bulunur.

3.1.2. Moment Üreten ve Olasılık Üreten Fonksiyonu

$X \sim UG(p)$ olmak üzere, X rasgele değişkeninin moment çıkaran fonksiyonunu

$$
\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \sum_{x=1}^{\infty} e^{tX} \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(\sum_{x=1}^{n} \frac{e^{tx}}{n} \right) p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(\frac{e^{t} - e^{t(n+1)}}{n(1-e^{t})} \right) p(1-p)^{n-1} \\
&= \frac{pe^{t}}{(1-p)(1-e^{t})} \left[\sum_{n=1}^{\infty} \frac{(1-p)^{n}}{n} - \sum_{n=1}^{\infty} \frac{((1-p)e^{t})^{n}}{n} \right] \\
&= \frac{pe^{t}}{(1-p)(1-e^{t})} \left[\ln(1-(1-p)e^{t}) - \ln p \right]. \qquad t < -\ln(1-p),
\end{aligned}
$$

şeklinde elde edilir. Burada

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x), \ |x| < 1$$

eşitliği kullanılmıştır. Moment çıkaran fonksiyonu yardımıyla dağılımın olasılık üreten fonksiyonu

$$
\begin{aligned}
\psi_X(t) &= E(t^X) \\
&= M_X(\ln(t)) \\
&= \frac{pt}{(1-p)(1-t)} \left[\ln(1-(1-p)t) - \ln p \right], \quad t < \frac{1}{1-p}.
\end{aligned}
$$

olarak elde edilir.

3.1.3. Beklenen Değer, Varyans, Çarpıklık ve Basıklık

$X \sim UG(p)$ olmak üzere, X rasgele değişkeninin beklenen değeri aşağıdaki gibi elde edilir:

$$
\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(\sum_{x=1}^{n} \frac{x}{n} \right) p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(\frac{n+1}{2} \right) p(1-p)^{n-1}
\end{aligned}
$$

$$= \frac{p}{2(1-p)} \sum_{n=1}^{\infty} (n+1)(1-p)^n$$
$$= \frac{p}{2(1-p)} \times \frac{1-p^2}{p^2}$$
$$= \frac{1+p}{2p}.$$

Düzgün-Geometrik dağılımının ikinci, üçüncü ve dördüncü momentleri sırasıyla

$$E(X^2) = \sum_{x=1}^{\infty} x^2 \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} = \sum_{n=1}^{\infty} \left(\sum_{x=1}^{n} \frac{x^2}{n} \right) p(1-p)^{n-1}$$
$$= \sum_{n=1}^{\infty} \frac{(2n+1)(n+1)}{6} p(1-p)^{n-1}$$
$$= \frac{p}{6(1-p)} \sum_{n=1}^{\infty} (2n+1)(n+1)(1-p)^n$$
$$= \frac{p^2+p+4}{6p^2},$$

$$E(X^3) = \sum_{x=1}^{\infty} x^3 \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} = \sum_{n=1}^{\infty} \left(\sum_{x=1}^{n} \frac{x^3}{n} \right) p(1-p)^{n-1}$$
$$= \sum_{n=1}^{\infty} \frac{(n+1)^2 n}{4} p(1-p)^{n-1}$$
$$= \frac{p}{4(1-p)} \sum_{n=1}^{\infty} n(n+1)^2 (1-p)^n$$
$$= \frac{3-p}{2p^3},$$

ve

$$E(X^4) = \sum_{x=1}^{\infty} x^4 \sum_{n=x}^{\infty} \frac{1}{n} p(1-p)^{n-1} = \sum_{n=1}^{\infty} \left(\sum_{x=1}^{n} \frac{x^4}{n} \right) p(1-p)^{n-1}$$
$$= \sum_{n=1}^{\infty} \frac{(2n+1)(n+1)(3n^2+3n-1)}{30} p(1-p)^{n-1}$$
$$= \frac{p}{30(1-p)} \sum_{n=1}^{\infty} (2n+1)(n+1)(3n^2+3n-1)(1-p)^n$$
$$= \frac{-126p+14p^2-p^3+144-p^4}{30p^4}$$

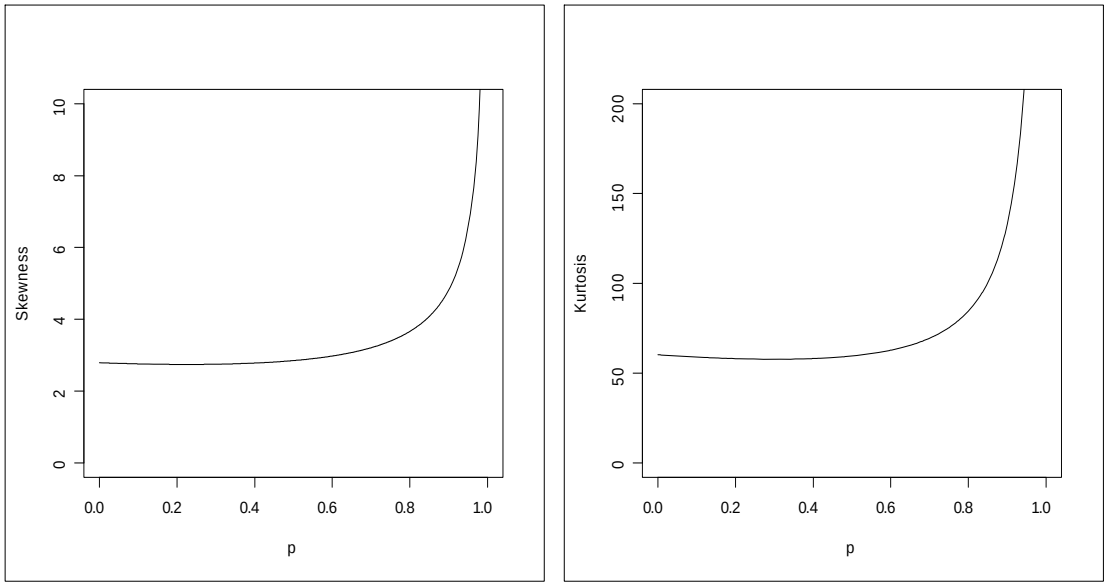
olarak elde edilir. Elde edilen momentler yardımıyla varyans, çarpıklık ve basıklık katsayısı sırasıyla

$$
\begin{aligned}
Var(X) &= E\left(X^2\right) - E(X)^2 \\
&= \frac{p^2+p+4}{6p^2} - \left(\frac{1+p}{2p}\right)^2 \\
&= \frac{5-4p-p^2}{12p^2},
\end{aligned}
$$

$$
\begin{aligned}
\gamma_1(X) &= \frac{E\left(X-E(X)\right)^3}{\left[E\left(X-E(X)\right)^2\right]^{3/2}} \\
&= \frac{E\left(X^3\right)-3E(X)E\left(X^2\right)+3\left(E(X)\right)^2E(X)-\left(E(X)\right)^3}{\left[Var(X)\right]^{3/2}} \\
&= \frac{54(p-1)(p-3)}{\left(-3(p+5)(p-1)\right)^{3/2}} > 0,
\end{aligned}
$$

$$
\begin{aligned}
\gamma_2(X) &= \frac{E\left(X-E(X)\right)^4}{\left[E\left(X-E(X)\right)^2\right]^2} - 3 \\
&= \frac{E\left(X^4\right)-4E(X)E\left(X^3\right)+6\left(E(X)\right)^2E\left(X^2\right)}{\left[Var(X)\right]^2} \\
&\quad - \frac{4\left(E(X)\right)^3E(X)+\left(E(X)\right)^4}{\left[Var(X)\right]^2} \\
&= \frac{6\left(p^3-23p^2+213p-251\right)}{(p+5)^2(p-1)} > 0
\end{aligned}
$$

şeklinde elde edilir. Dağılımın basıklık ve çarpıklık katsayılarının p değerine karşılık grafikleri Şekil 3.1.2 de verilmiştir. Şekil 3.1.2 den dağılımın basıklığı ve çarpıklığı özellikle p parametresinin 0.8 den büyük olduğu durumlarda hızlı bir şekilde artmaktadır. Ayrıca Düzgün-Geometrik Dağılım sağa çarpık bir dağılımdır.



Şekil 3.1.2. *UG*(*p*) dağılımının basıklık ve çarpıklık grafikleri

3.1.4. Bozulma (Hazard) Oranı Fonksiyonu

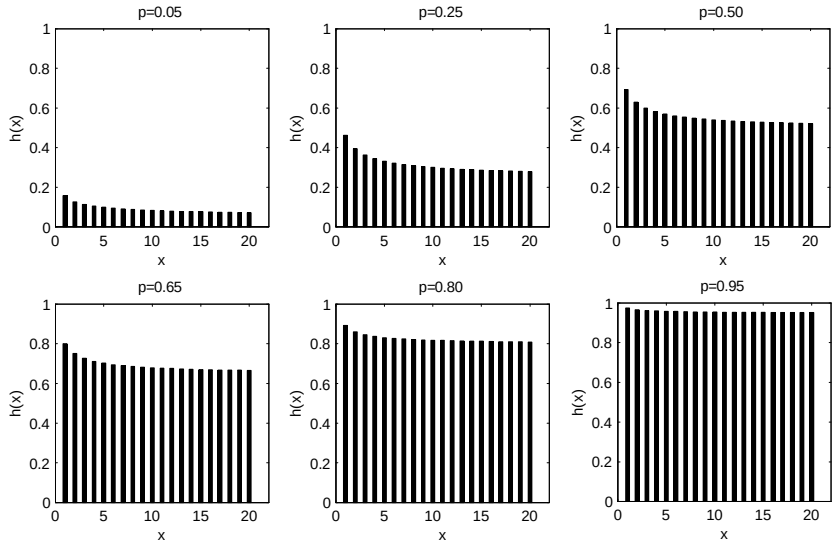
X, Düzgün-Geometrik dağılımına sahip bir rasgele değişken ise X in bozulma (Hazard) oranı fonksiyonu

$$h(x)=P(X=x\mid X\geq x)=\frac{P(X=x)}{P(X\geq x)}$$

$$=\frac{p(1-p)^{x-1}\,LerchPhi(1-p,1,x)}{p(1-p)^{x-1}\left[\frac{1}{p}-(x-1)\,LerchPhi(1-p,1,x)\right]}$$

$$=\frac{pLerchPhi(1-p,1,x)}{1-p(x-1)\,LerchPhi(1-p,1,x)}$$

şeklinde yazılabilir. Bozulma oranı fonksiyonun limitleri $\lim_{x\to 1}h(x)=-\frac{p\log(p)}{1-p}$ ve $\lim_{x\to\infty}h(x)=p$ şeklindedir. Şekil 3.1.3 de bazı p değerleri için bozulma oranı fonksiyonu grafikleri verilmiştir. Şekil 3.1.3 den Bozulma oranı fonksiyonun azalan olduğu anlaşılmaktadır ancak burada ispatlanamamıştır.



Şekil 3.1.3. $UG(p)$ dağılımının bozulma oranı fonksiyonu grafikleri

3.2. Parametre Tahmini

Bu bölümde, Düzgün-Geometrik dağılımının p parametresinin ML, MP ve MM tahmin edicileri elde edilmiştir.

3.2.1. En Çok Olabilirlik Tahmin Edicisi

$X_1, X_2, \ldots, X_n$, $UG(p)$ dağılımından bir örneklem olmak üzere, olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$L(p) = p^n (1-p)^{\sum_{i=1}^{n}(x_i)-n} \prod_{i=1}^{n} LerchPhi\left[(1-p),1,x_i\right]$$

ve

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

şeklinde yazılabilir. Log-olabilirlik fonksiyonunun p parametresine göre türevi sıfıra eşitlenerek

$$\frac{\partial \log\left(L(p)\right)}{\partial p} = \frac{n}{p} - \frac{\sum(x_i - 1)}{1-p} + \sum_{i=1}^{n} \frac{px\left(LerchPhi\left((1-p),1,x_i\right)\right) - 1}{p(1-p)\left(LerchPhi\left((1-p),1,x_i\right)\right)}$$

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

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

$$= \frac{1}{p(1-p)} \left[n - \sum_{i=1}^{n} \frac{1}{LerchPhi(1-p,1,x_i)} \right] = 0. \qquad (3.1.3)$$

olabilirlik denklemi (3.1.3) elde edilir. (3.1.3) denkleminin analitik çözümü olmadığından, denklemin çözümü için Newton-Raphson metodu kullanılabilir. Denklemin çözümü, p parametresinin ML tahmini olan $\hat{p}$ yı verir. Fisher bilgisi

$$I(p) = -\frac{\partial^2 \log\left(L(p)\right)}{\partial p^2} = \frac{n}{p^2} + \frac{\sum(x_i - 1)}{(1-p)^2}$$
$$- \sum_{i=1}^{n} \left[\frac{p^2 x_i (1+x_i) LerchPhi(1-p,1,x_i) - x_i p - 2p + 1}{p^2 (1-p)^2 LerchPhi(1-p,1,x_i)} \right.$$
$$\left. + \left(\frac{p x_i LerchPhi(1-p,1,x_i) - 1}{p(1-p) LerchPhi(1-p,1,x_i)} \right)^2 \right].$$

olmak üzere tahmin edicinin asimptotik dağılımı, ortalaması sıfır, varyansı $1/I(p)$ olan Normal dağılımdır. Bu bilgiler ışığında, p parametresinin yaklaşık $100(1-\alpha)\%$ güven aralığı

$$P\left(\hat{p} - \frac{z_{\alpha/2}}{\sqrt{I(\hat{p})}} < p < \hat{p} + \frac{z_{\alpha/2}}{\sqrt{I(\hat{p})}} \right) = 1 - \alpha, \qquad (3.1.4)$$

olarak elde edilmiştir.

3.2.2. Oranlar Metodu Tahmini

$X_1, X_2, \ldots, X_n$, $UG(p)$ dağılımından alınmış bir örneklem olmak üzere, $i = 1, 2, \ldots, n$ için

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

gösterge fonksiyonu tanımlansın. O halde $Y = \frac{1}{n} \sum_{i=1}^{n} \nu(X_i)$, X rasgele değişkeninin 1'e eşit olması olasılığının bir tahmin edicisi olacaktır. Ayrıca Y oranı, $f(1) = -\frac{p \log(p)}{(1-p)}$ olasılığın yansız ve tutarlı bir tahmin edicisidir. p parametresinin MP tahmini

$$y = -\frac{p\log(p)}{(1-p)} \tag{3.1.5}$$

(3.1.5) eşitliğinin çözümünden elde edilir. Denklemin çözümü için nümerik yöntemler kullanılabilir.

3.2.3. Momentler Yöntemi Tahmini

$X_1, X_2, \ldots, X_n$, $UG(p)$ dağılımından alınmış örneklem olmak üzere, p parametresinin MM tahmin edicisi örneklem ortalamasının dağılımın beklenen değerine eşitlenmesiyle

$$E(X) = \frac{p+1}{2p} = \overline{X} \tag{3.1.6}$$

şeklinde elde edilir. Yukarıda verilen eşitlik (3.1.5) çözüldüğünde p parametresinin MM tahmin edicisi

$$\tilde{p} = \frac{1}{2\overline{X} - 1}.$$

olarak edilir. Merkezi limit teoreminden

$$\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{d} N\left(0, \sigma^2\right),$$

yazılabilir, burada $\mu = \frac{1+p}{2p}$ ve $\sigma^2 = \frac{5-4p-p^2}{12p^2}$ dır. Delta kuralı kullanılarak

$$\sqrt{n}\left(\tilde{p} - p\right) \xrightarrow{d} N\left(0, \frac{p^2\left(5-4p-p^2\right)}{3}\right).$$

elde edilir. Böylece $\tilde{p}$ nin p için asimptotik yansız ve tutarlı olduğu açıktır. Ayrıca p parametresinin $0 < \alpha < 1$ değeri için $100(1-\alpha)\%$ lık güven aralığı

$$P\left(\tilde{p} - z_{1-\alpha/2}\frac{S}{\sqrt{n}} < p < \tilde{p} + z_{1-\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha, \tag{3.1.7}$$

şeklinde oluşturulur. Burada kitle varyansının tahmini eşitlik (3.1.8) deki gibidir.

$$S = \sqrt{\frac{\tilde{p}^2\left(5-4\tilde{p}-\tilde{p}^2\right)}{3}} \tag{3.1.8}$$

3.2.4. Simülasyon Çalışması

Simülasyon çalışmasında, $UG(p)$ dağılımının p parametresinin ML, MP ve MM tahmin edicilerinin karşılaştırılması amaçlanmıştır. Simülasyon çalışması için $UG(p)$ dağılımından farklı hacimlerde 10000 rasgele örneklem üretilmiş, elde edilen tahminlere ilişkin tahmin edicilerin yanları ve HKO'ları tablolaştırılmıştır. Simülasyon çalışmasında, farklı parametre değerleri ve anlam seviyeleri için güven aralıkları uzunluğu ve kapsama olasılıkları Tablo 3.1.1-3.1.2 de verilmiştir.

Tablo 3.1.1. Tahmin edicilerin HKO ve yanları

		MLE		MPE		MME	
p	n	HKO	Yan	HKO	Yan	HKO	Yan
0.2	50	0.0012	0.0045	0.0034	-0.0023	0.0015	-0.0045
	100	0.0006	0.0019	0.0015	0.0015	0.0008	-0.0072
	200	0.0003	0.0010	0.0007	0.0008	0.0004	-0.0081
	300	0.0002	0.0006	0.0005	0.0006	0.0003	-0.0085
	500	0.0001	0.0005	0.0003	0.0005	0.0002	-0.0086
0.5	50	0.0047	0.0087	0.0071	0.0044	0.0048	0.0088
	100	0.0023	0.0042	0.0035	0.0023	0.0024	0.0042
	200	0.0012	0.0018	0.0018	0.0008	0.0012	0.0018
	300	0.0008	0.0010	0.0012	0.0004	0.0008	0.0009
	500	0.0005	0.0010	0.0007	0.0006	0.0005	0.0010
0.8	50	0.0037	-0.0421	0.0057	0.0029	0.0049	0.0064
	100	0.0017	-0.0233	0.0029	0.0008	0.0025	0.0027
	200	0.0008	-0.0112	0.0015	0.0007	0.0013	0.0014
	300	0.0006	-0.0076	0.0010	0.0005	0.0009	0.0010
	500	0.0004	-0.0038	0.0006	0.0003	0.0006	0.0006

Yukarıdaki Tablo 3.1.1 de n örneklem hacminin artması durumunda tahmin edicilerin hem yanları hem de HKO düşmektedir. Yine ML ve MM kullanılarak elde edilen tahminlerin yanları ve HKO su, MP tahminini kullanarak elde edilen tahminlerin yan ve HKO dan daha küçüktür. Büyük örneklem durumunda ise ML, MM ve MP tahmin edicilerin yan ve HKO su birbirine yaklaşmaktadır.

Tablo 3.1.2. $UG(p)$ dağılımının Kapsama Olasılıkları ve Aralık uzunlukları

		$\alpha = 0.05$				$\alpha = 0.10$			
		MLE		MME		MLE		MME	
p	n	Aralık Uzunluğu	Kapsama Olasılığı	Aralık Uzunluğu	Kapsama Olasılığı	Aralık Uzunluğu	Kapsama Olasılığı	Aralık Uzunluğu	Kapsama Olasılığı
0.2	25	0.1913	0.9533	0.1920	0.9520	0.1603	0.8990	0.1608	0.8937
	50	0.1325	0.9455	0.1329	0.9445	0.1109	0.8978	0.1112	0.8940
	100	0.0927	0.9478	0.0930	0.9468	0.0776	0.8975	0.0778	0.9005
	500	0.0413	0.9495	0.0414	0.9499	0.0345	0.8965	0.0346	0.8959
0.5	25	0.3726	0.9288	0.3726	0.9322	0.3113	0.8703	0.3114	0.8666
	50	0.2647	0.9441	0.2648	0.9442	0.2215	0.8848	0.2216	0.8908
	100	0.1874	0.9441	0.1875	0.9430	0.1568	0.8966	0.1568	0.8950
	500	0.0839	0.9506	0.0839	0.9507	0.0702	0.8925	0.0702	0.8919
0.8	25	0.3708	0.9883	0.3706	0.9883	0.3102	0.8128	0.3100	0.8128
	50	0.2649	0.9298	0.2649	0.9298	0.2217	0.9015	0.2216	0.9006
	100	0.1912	0.9345	0.1912	0.9345	0.1600	0.8809	0.1600	0.8816
	500	0.0869	0.9464	0.0869	0.9464	0.0727	0.8963	0.0727	0.8920

Simülasyon çalışması sonucunda, Eşitlik (3.1.4) ve (3.1.7) de verilen asimptotik güven aralıklarının farklı parametre değerleri için kapsama olasılıkları ve aralık uzunlukları verilmiştir. n örneklem hacminin artması ile her iki güven aralığının uzunluğu azalmakla birlikte kapsama olasılıkları da istenilen nominal değere ulaşmaktadır. İki güven aralığının kapsama olasılıkları ve aralık uzunlukları $n \geq 25$ için hemen hemen aynıdır.

3.2.5. Gerçek Veri Uygulaması

Bu bölümde iki gerçek veri seti üzerinde yeni elde edilen Düzgün-Geometrik dağılımının, modelleme performansı Geometrik dağılımla kıyaslanmıştır. İlk veri seti (Xie M. ve T.N., 1993) tarafından literatüre kazandırılmıştır. Endüstriyel bir süreç içindeki kusurlara ilişkin denetimlerin sayıları üzerinden yapılacaktır. Bu verinin geometrik dağılama uygunluğu 2002 yılında Bracquemound ve ark. tarafından çalışılmıştır. Bu veri seti aşağıda verilmiştir:

1 1 1 1 1 1 2 2 2 2 3 3 3 4 4 4 5 5 7 9 1113 14 14 17 18 26 29.

Bu veri seti üzerinde Düzgün-Geometrik dağılımın uygunluğunu test etmek için χ^2 uyum iyiliği testi(Goodness of fit test) kullanılacaktır. $H_0 : X \sim UG(p)$ hipotezinin testi için bilinmeyen p parametresinin ML tahmini kullanarak $\hat{p} = 0.07420$ olarak elde

edilir. Düzgün-Geometrik ve geometrik dağılımın olasılık fonksiyonları sırasıyla

$$\hat{p}_i = P(X=i) = \hat{p}(1-\hat{p})^{i-1} LerchPhi(1-\hat{p},1,i), \quad i=1,2,...$$

ve

$$f_{Ge}(x;p) = (1-p)p^{x-1}, \quad x=1,2,..., \ 0<p<1$$

şeklindedir. Bu fonksiyonlardan yararlanarak beklenen olasılıklar ve bu olasılıklar yardımıyla beklenen frekanslar $\hat{e}_i = n\hat{p}_i$ hesaplanır. Bu değerler Tablo 3.1.3 de verilmiştir. Tablo 3.1.3 den χ^2 istatistiği

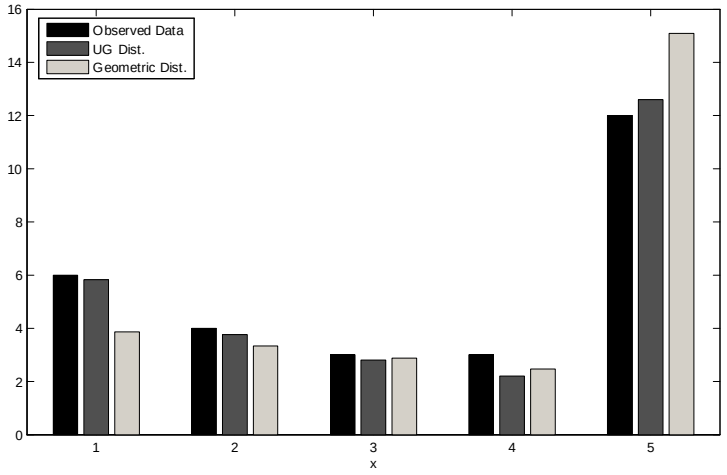
$$\chi^2 = \sum_{i=1}^{5} \frac{(o_i - \hat{e}_i)^2}{\hat{e}_i} = 0.46918 < 7.8147 = \chi^2_{0.95}(3),$$

olarak hesaplanır. Burada o_i ve $\hat{e}_i$ sırasıyla gözlenen ve beklenen frekansları göstermektedir ve $\chi^2_a(b), b$ serbestlik dereceli Ki-Kare dağılımının a sol kuyruk olasılıklı kuantil değeridir. Test sonucuna göre H_0 hipotezini $\alpha = 0.05$ anlam seviyesinde red edilemez. Aynı işlemler Geometrik dağılımı için uygulandığında, geometrik dağılımının p parametresinin ML tahmini $\hat{p} = 0.13793$ olup beklenen ve gözlenen frekanslar Tablo 3.1.3 de verilmiştir. Buradan Ki-kare test istatistiği değeri $\chi^2 = 1.4591 < 7.8147 = \chi^2_{0.95}(3)$ olup H_0 hipotezini $\alpha = 0.05$ anlam seviyesinde red edilemez kararına varılmıştır.

Tablo 3.1.3. $UG(p)$ ve Geometrik dağılımı için ilk veriye ait χ^2 değerleri

x	o_i	$UG(p)$			Geometrik Dağılım		
		$\hat{p}_i$	$\hat{e}_i$	$\frac{(o_i-\hat{e}_i)^2}{\hat{e}_i}$	$\hat{p}_i$	$\hat{e}_i$	$\frac{(o_i-\hat{e}_i)^2}{\hat{e}_i}$
1	6	0.2084	5.8363	0.0046	0.1379	3.8620	1.1835
2	4	0.1342	3.7590	0.0154	0.1139	3.3293	0.1350
3	3	0.0999	2.7974	0.0146	0.1025	2.8701	0.0058
4	3	0.0787	2.2040	0.2875	0.0884	2.4743	0.1117
>4	12	0.4787	13.4033	0.1470	0.4477	12.5358	0.0229
Toplam	n=28	1	28	0.4692	1	28	1.4591

Şekil 3.1.4 de gözlenen frekanslar, fit edilmiş geometrik ve düzgün-geometrik olasılıkları verilmiştir. Buradan düzgün-geometrik dağılımın geometrik dağılıma kıyasla verinin geldiği dağılımı daha iyi modellediği gözlenmiştir.



Şekil 3.1.4. İlk veri setinde gözlenmiş ve beklenen frekans grafikleri

İlk verinin geldiği dağılımı her iki dağılımda modelleyebilmektedir. Sonuçlara göre hangi dağılım daha iyi modellediğini göstermek için bir takım kriterlere bakmak gerekir. Burada ilk bakılacak χ^2 istatistik değeridir ve bu değerin küçük olması daha iyi modellediğini göstermektedir. $UG(p)$ dağılımı geometrik dağılımından daha küçük χ^2 istatistiğine sahip olduğundan daha iyi modellediği görülür. Bir başka kriter ise log-olabilirlik değerine bakılmasıdır ve bu değerinin büyük olması istenir. $UG(p)$ dağılıma ait log-olabilirlik değeri -80,30386 bulunurken, Geometrik dağılımına ait log-olabilirlik değeri ise -81,44154 bulunmuştur.

İkinci veri seti, Weinbergand ve Gladen(Weinberg ve Gladen, 1986) tarafından elde edilen gebelik için aybaşı döngü sayısı verisidir. Bu veri 586 bayan üzerinden 1844 döngü kaydedilerek oluşmuştur. Elde edilen veri aşağıdaki Tablo 3.1.4 te verilmiştir.

Tablo 3.1.4. 586 kadın için hamilelikteki aybaşı döngü sayısı

Döngü	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

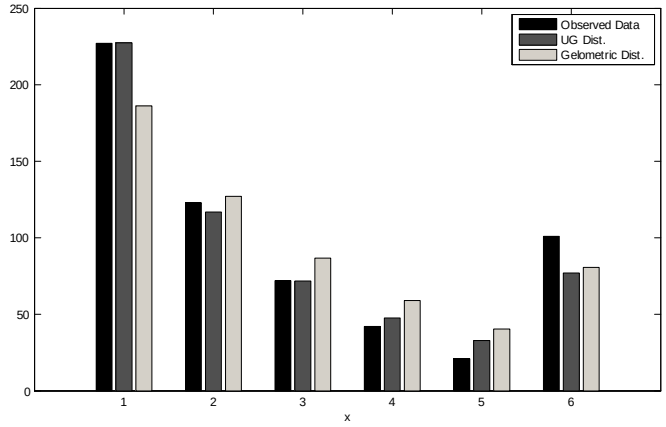
Düzgün-Geometrik dağılımı için p parametresinin ML tahmini $\hat{p} = 0.1889$ olarak elde edilmiş ve log-olabilirlik değeri -1140,675, Geometrik dağılımı p parametresinin ML tahmini $\hat{p} = 0.3178$ olarak elde edilmiş ve log-olabilirlik değeri -1152,8511bulunmuştur.

Tablo 3.1.5 te χ^2 istatistiği için beklenen ve gözlenen frekanslar verilmiştir.

Tablo 3.1.5. $UG(p)$ ve geometrik dağılımının ikinci veriye ait χ^2 değerleri

x	o_i	$UG(p)$			Geometrik Dağılım		
		$\hat{p}_i$	$\hat{e}_i$	$\frac{(o_i-\hat{e}_i)^2}{\hat{e}_i}$	$\hat{p}_i$	$\hat{e}_i$	$\frac{(o_i-\hat{e}_i)^2}{\hat{e}_i}$
1	227	0.3881	227.4037	0.0007	0.3178	186.2234	8.9290
2	123	0.1992	116.7376	0.3359	0.2168	127.0420	0.1286
3	72	0.1226	71.8542	0.0003	0.1480	86.6710	2.4834
4	42	0.0812	49.5827	0.6550	0.1010	59.1280	4.9616
5	21	0.0560	38.8169	4.2551	0.0688	40.3379	9.2705
>5	101	0.1529	81.6049	0.1469	0.1478	86.5959	2.3959
Total	$n = 586$	1	586	5.2471	1	586	28.1690

Düzgün-Geometrik dağılım ikinci veri setinde, hem χ^2 istatistiği hem de log-olabilirlik değerine baktığımızda Geometrik dağılıma göre verinin geldiği dağılımı daha iyi modelleyebilmektedir.



Şekil 3.1.5. İkinci veri setine ait gözlenmiş ve beklenen frekans grafikleri

Hem birinci hem de ikinci gerçek veri seti dağılımlarını $UG(p)$ dağılımı geometrik dağılımdan daha iyi modellediği görülmüştür.

4. YENİ GEOMETRİK-SIFIRDAN BUDANMIŞ POİSSON DAĞILIMI

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

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

elde edilir. Aynı işlemler $F(x-1)$ için elde edildiğinde, $P(X = x \mid N = n)$ 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

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

şeklinde verilir. Bu durumda

$$\begin{aligned} f(x) = P(X = x) &= \sum_{n=1}^{\infty} P(X = x \mid N = n) P(N = n) \\ &= \sum_{n=1}^{\infty} (1-q^x)^n - (1-q^{x-1})^n \frac{e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})} \\ &= \frac{e^{-\lambda}}{(1-e^{-\lambda})} \left(\sum_{n=1}^{\infty} \frac{(\lambda(1-q^x))^n}{n!} - \sum_{n=1}^{\infty} \frac{(\lambda(1-q^{x-1}))^n}{n!} \right) \\ &= \frac{1}{(1-e^{-\lambda})} \left(e^{-\lambda q^x} - e^{-\lambda q^{x-1}} \right), \qquad x = 1, 2, \ldots \end{aligned}$$

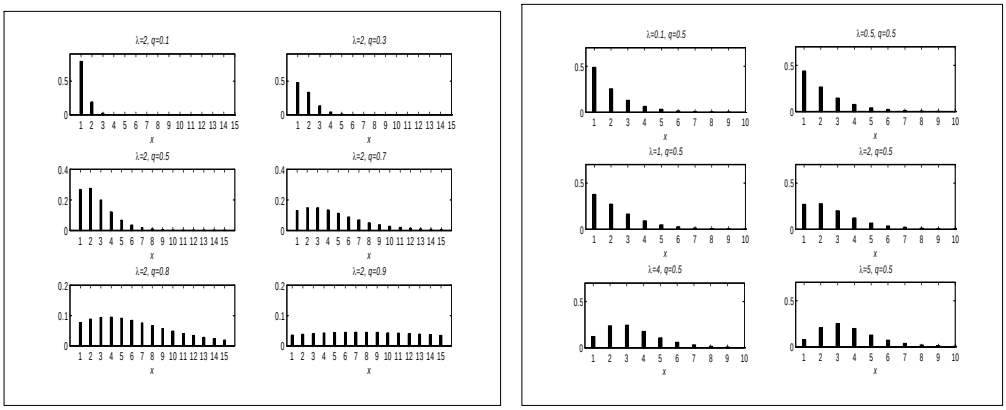
$$P(X = x) = \frac{1}{(1-e^{-\lambda})} \left(e^{-\lambda q^x} - e^{-\lambda q^{x-1}} \right), \qquad x = 1, 2, \ldots$$

ve dağılım fonksiyonu

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

şeklinde elde edilir. X rasgele değişkeninin dağılımı, Geometrik-sıfırdan budanmış Poisson dağılımı olarak adlandırılır ve $X\sim GZTP(q,\lambda)$ olarak gösterilir. Burada $0<q<1$ ve $\lambda>0$ dağılımın parametreleridir.

Şekil 4.1 de $GZTP(q,\lambda)$ dağılımının bazı parametreler için olasılık fonksiyonunun grafikleri verilmiştir.



Şekil 4.1. $GZTP(q,\lambda)$ dağılımının farklı λ , q parametreleri için olasılık fonksiyonu grafikleri

4.1. Dağılıma ait Bazı Özellikler

Bu bölümde yeni elde edilen $GZTP(q,\lambda)$ dağılımının dağılımsal bazı özellikleri üzerinde durulacaktır.

Teorem 4.1. $GZTP(q,\lambda)$ dağılımı herhangi q ve λ parametreleri için log-konkavdır.

İspat: Bir olasılık fonksiyonunu $P_x^2\ge P_{x-1}P_{x+1}$ eşitsizliğini sağlarsa log-konkavdır (Keilson ve Gerber, 1971). $GZTP(q,\lambda)$ dağılımının olasılık fonksiyonu bu eşitsizlikte yerine yazıldığında

$$\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right)^{2}\ge\left(e^{-\lambda q^{x-1}}-e^{-\lambda q^{x-2}}\right)\left(e^{-\lambda q^{x+1}}-e^{-\lambda q^{x}}\right).$$

olup

$$g(x;q,\lambda)=\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right)^{2}-\left(e^{-\lambda q^{x-1}}-e^{-\lambda q^{x-2}}\right)\left(e^{-\lambda q^{x+1}}-e^{-\lambda q^{x}}\right)$$

fonksiyonunun pozitif değerler aldığı gösterildiğinde(nan-negative) ispat tamamlanacaktır. Bazı cebirsel işlemler sonunda $g(x;q,\lambda)$ fonksiyonu $x\in\Re^{+}$ değerleri için azalandır. Ayrıca $\lim_{x\to+\infty} g(x;q,\lambda)=0$ olup $g(x;q,\lambda)$ fonksiyonu negatif olmayan bir fonksiyondur ve bu da ispatı tamamlar.

Sonuç: $GZTP(q,\lambda)$ dağılımının olasılık fonksiyonun log-konkav olmasından dolayı tek tepelidir ve dağılımın modu

$$\text{mod}(X)=\begin{cases}\lfloor m\rfloor & ; \quad f(\lfloor m\rfloor)>f(\lfloor m\rfloor+1)\\ \lfloor m\rfloor+1 & ; \quad f(\lfloor m\rfloor)<f(\lfloor m\rfloor+1)\end{cases}$$

şeklindedir. Burada $m=\log\left(\frac{\log(q)}{(1-q)\log\left(e^{-\lambda}\right)}\right)/\log(q)$.

İspat: $GZTP(q,\lambda)$ dağılımının olasılık fonksiyonunun log-konkav olduğu Teorem 4.1 de gösterilmiştir. Olasılık fonksiyonun log-konkav olmasından dolayı dağılımın olasılık fonksiyonu tek tepelidir(Keilson ve Gerber, 1971). Olasılık fonksiyonu x değişkenine göre sürekli bir fonksiyon olarak düşünüldüğünde maksimum yapan nokta $m=\log\left(\frac{\log(q)}{(1-q)\log\left(e^{-\lambda}\right)}\right)/\log(q)$ şeklinde elde edilir. Buda ispatı tamamlar.

4.1.1. Beklenen Değer, Varyans, Çarpıklık ve Basıklık

$X\sim GZTP(q,\lambda)$ olmak üzere, X rasgele değişkeninin beklenen değeri

$$E(X)=\frac{1}{1-e^{-\lambda}}\sum_{x=1}^{\infty}x\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right) \qquad (4.2.1)$$

şeklindedir. Beklenen değerin hesaplanması için (4.2.1) deki seri toplamının elde edilmesi gerekmektedir ve bu serinin toplamı burada elde edilememiştir. Bu nedenle beklenen değerin hesaplanması için oran testi metodu kullanılmıştır (Ratio Test Method). $S=\sum_{n=1}^{\infty}a_n$ ve $S_n=\sum_{k=1}^{n}a_k$, n. kısmi toplamlar olmak üzere oran testi metodu aşağıdaki gibi ifade edilir (Braden, 1992).

$\{a_n\}$ pozitif azalan bir dizi ve $\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=L<1$ olsun, $\frac{a_{n+1}}{a_n}$ azalan bir dizi ise

$$S_n + a_n\left(\frac{L}{1-L}\right) < S < S_n + \frac{a_{n+1}}{1-\frac{a_{n+1}}{a_n}}$$

yazılabilir. Yukarıdaki yaklaşım $GZTP(q,\lambda)$ dağılımın beklenen değeri için uygulanırsa

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

$$= \frac{\sum_{x=1}^{c} a_x + \sum_{x=c+1}^{\infty} a_x}{1-e^{-\lambda}}$$

$$= \frac{\sum_{x=1}^{c} a_x + \sum_{y=1}^{\infty} a_{y+c}}{1-e^{-\lambda}},$$

elde edilir. Burada $a_x = x\left(e^{-\lambda q^x} - e^{-\lambda q^{x-1}}\right)$ ve $c = \lfloor \arg\max(a_x) \rfloor$ şeklindedir. Teorem 4.1 den $GZTP(q,\lambda)$ dağılımının olasılık fonksiyonu tek tepelidir. Kolayca gösterilebilir ki $a_x = xf(x)$ de tek tepelidir ve $x \geq c$ için a_x azalandır.

$\lim_{n\to\infty} \frac{a_{n+c+1}}{a_{n+c}} = q < 1$ olup $\frac{a_{n+c+1}}{a_{n+c}}$ ifadesi q'ya göre azalandır. Buradan beklenen değer için alt ve üst sınır

$$\frac{\frac{q}{1-q}a_{n+c} + \sum_{x=1}^{n+c} a_x}{1-e^{-\lambda}} \leq E(X) \leq \frac{\frac{a_{n+c+1}}{1-\frac{a_{n+c+1}}{a_{n+c}}} + \sum_{x=1}^{n+c} a_x}{1-e^{-\lambda}}$$

şeklinde elde edilir. Bu yöntem diğer momentler için kullanılmış ve Tablo 4.1-4.7 de verilmiştir. Elde edilen momentler yardımıyla dağılımın beklenen değeri, varyansı, çarpıklık ve basıklık katsayıları için yaklaşık alt sınır ve üst sınır hesaplanmıştır.

Tablo 4.1. Beklenen değerin gerçek, yaklaşık, alt ve üst sınır değerleri

(q,λ)	$E(X)$	n	$E(X)$ Yaklaşık değeri	$E(X)$ Alt sınır	$E(X)$ Üst Sınır
(0.3,2)	1.7997151003	5	1.7993766884	1.7989899838	1.7997633932
		10	1.7997142548	1.7997133442	1.7997151653
		15	1.7997150982	1.7997150960	1.7997151004
		20	1.7997151003	1.7997151003	1.7997151003
(0.5,2)	2.7194339620	5	2.7116771654	2.7008500136	2.7225043172
		10	2.7191750259	2.7188684061	2.7194816456
		15	2.7194256573	2.7194163137	2.7194350008
		20	2.7194336988	2.7194334105	2.7194339870
		50	2.7194339620	2.7194339620	2.7194339620
(0.8,2)	7.3433586356	5	7.2709297316	6.6218080726	7.9200513908
		10	7.2760098324	7.1206461299	7.4313735348
		15	7.3169854757	7.2730163538	7.3609545977
		20	7.3340874125	7.3207103315	7.3474644935
		50	7.3433586356	7.3433586356	7.3433586356
(0.3,5)	2.3314642857	5	2.3312413457	2.3309910793	2.3314916121
		10	2.3314637320	2.3314631392	2.3314643248
		15	2.3314642844	2.3314642829	2.3314642858
		20	2.3314642857	2.3314642857	2.3314642857
(0.5,5)	3.6795124171	5	3.6708662414	3.6590567668	3.6826757160
		10	3.6792286275	3.6788966754	3.6795605796
		15	3.6795033488	3.6794932126	3.6795134850
		20	3.6795121301	3.6795118170	3.6795124432
		50	3.6795124171	3.6795124171	3.6795124171
(0.8,5)	10.3715943115	5	10.2727957683	9.5103299111	11.0352616212
		10	10.2921570783	10.1152364277	10.4690777314
		15	10.3412252852	10.2921480801	10.3903024904
		20	10.3610415078	10.3462252766	10.3758577390
		50	10.3715943115	10.3715943115	10.3715943115

Tablo 4.2**.** İkinci moment değerinin gerçek, yaklaşık, alt ve üst sınır değerleri

(q,λ)	$E(X^2)$	n	$E(X^2)$ Yaklaşık değeri	$E(X^2)$ Alt sınır	$E(X^2)$ Üst Sınır
(0.3,2)	4.2058576581	5	4.2040654811	4.2019843387	4.2061466235
		10	4.2058506173	4.2058429824	4.2058582521
		15	4.2058576345	4.2058576096	4.2058576594
		20	4.2058576581	4.2058576581	4.2058576581
(0.5,2)	10.3270443806	5	10.2498545069	10.1361104836	10.3635985302
		10	10.3230966012	10.3182884029	10.3279047994
		15	10.3268755137	10.3266826103	10.3270684172
		20	10.3270377017	10.3270303181	10.3270450854
		50	10.3270443806	10.3270443806	10.3270443806
(0.8,2)	82.2033763050	5	80.8082834323	70.7197520803	90.8968147842
		10	80.8203299830	77.6221455576	84.0185144083
		15	81.55151513755	80.4316343401	82.6713959350
		20	81.93452922212	81.5329064413	82.3361520029
		50	82.2033763050	82.2033763050	82.2033763050
(0.3,5)	6.6264871121	5	6.6225857689	6.6180511919	6.6271203458
		10	6.6264717888	6.6264551729	6.6264884048
		15	6.6264870608	6.6264870066	6.6264871150
		20	6.6264871121	6.6264871121	6.6264871121
(0.5,5)	16.9780363808	5	16.8830743334	16.7498609051	17.0162877617
		10	16.9734205753	16.9678924846	16.9789486661
		15	16.9778428269	16.9776235105	16.9780621433
		20	16.9780288093	16.9780204784	16.9780371403
		50	16.9780363808	16.9780363808	16.9780363808
(0.8,5)	140.135363172	5	138.7540798974	118.6356849681	158.8724748267
		10	137.6757476437	131.7412565072	143.6102387802
		15	138.9571438559	136.9329252110	140.9813625019
		20	139.6509419832	138.9330031956	140.3688807708
		50	140.1353631722	140.1353631722	140.1353631722

Tablo 4.3. Üçüncü moment değerinin gerçek, yaklaşık, alt ve üst sınır değerleri

(q,λ)	$E(X^3)$	n	$E(X^3)$ Yaklaşık değeri	$E(X^3)$ Alt sınır	$E(X^3)$ Üst Sınır
(0.3,2)	4.2058576581	5	12.4939463170	12.4651874874	12.5227051466
		10	12.5173569130	12.5171956171	12.5175182089
		15	12.5175026653	12.5175019543	12.5175033763
		20	12.5175033330	12.5175033330	12.5175033330
(0.5,2)	10.32704438065		50.7385559932	49.6143845106	51.8627274759
		10	51.3868949458	51.3221591789	51.4516307128
		15	51.4355163156	51.4322141707	51.4388184606
		20	51.4382221534	51.4380684745	51.4383758323
		50	51.4383597889	51.4383597889	51.4383597889
(0.8,2)	82.203376305	5	1234.3291513899	988.1171170565	1480.5411857287
		10	1207.9207956454	1120.8654736891	1294.9761176911
		15	1220.2442316775	1185.9287110829	1254.5597522721
		20	1229.4822052060	1215.7443688879	1243.2200415241
		50	1238.0022259723	1238.0022259723	1238.0022259723
(0.3,5)	6.6264871121	5	22.4005692799	22.3777449063	22.4233936536
		10	22.4197998532	22.4196794701	22.4199202362
		15	22.4199096698	22.4199091564	22.4199101833
		20	22.4199101539	22.4199101539	22.4199101539
(0.5,5)	16.978036380	5	94.1855161955	92.7842105337	95.5868218571
		10	95.0618558767	94.9834400421	95.1402717113
		15	95.1221566441	95.1182408428	95.1260724454
		20	95.1254002612	95.1252206850	95.1255798374
		50	95.1255619862	95.1255619862	95.1255619862
(0.8,5)	140.1353631725		2341.7423202778	1874.0130670372	2809.4715735183
		10	2290.3131450151	2129.4031997199	2451.2230903103
		15	2314.6551147951	2251.9216422846	2377.3885873056
		20	2332.2265513194	2307.2631869384	2357.1899157004
		50	2348.0089599203	2348.0089599203	2348.0089599203

Tablo 4.4. Dördüncü moment değerinin gerçek, yaklaşık, alt ve üst sınır değerleri

(q,λ)	$E(X^4)$	n	$E(X^4)$ Yaklaşık değeri	$E(X^4)$ Alt sınır	$E(X^4)$ Üst Sınır
(0.3,2)	4.2058576581	5	45.7071913705	45.5629353351	45.85144740589
		10	45.8222523917	45.8211354742	45.8233693093
		15	45.8232532560	45.8232469405	45.8232595716
		20	45.8232591363	45.8232591065	45.8232591660
(0.5,2)	10.3270443806	5	315.3307616449	308.2237682178	322.4377550720
		10	319.3497615114	318.8402364352	319.8592865875
		15	319.7245847128	319.6933720546	319.7557973711
		20	319.7497853119	319.7480849866	319.7514856372
		50	319.7513018869	319.7513018869	319.7513018869
(0.8,2)	82.2033763050	5	23810.6143308767	18061.5690816387	29559.6595801147
		10	22906.6886555535	20692.1893184298	25121.1879926773
		15	23096.4997056349	22145.8802567577	24047.1191545122
		20	23302.7777572645	22889.4464241834	23716.1090903457
		50	23543.7971502919	23543.7971502919	23543.7971502919
(0.3,5)	6.6264871121	5	88.6398519815	88.3258385838	88.9538653793
		10	88.8902494939	88.8878187157	88.8926802721
		15	88.8924277008	88.8924139562	88.8924414455
		20	88.8924405597	88.8924405597	88.8924405597
(0.5,5)	16.9780363808	5	624.6818416154	609.1818766184	640.1818066123
		10	633.4092014310	632.3002216227	634.5181812393
		15	634.2249116427	634.1569825256	634.2928407598
		20	634.2797563248	634.2760558623	634.2834567873
		50	634.2830568858	634.2830568858	634.2830568858
(0.8,5)	140.1353631722	5	47493.6794176137	36919.9576946154	58067.4011406399
		10	46164.3723006129	42062.7170481901	50266.0275530357
		15	46607.9363184560	44850.0816650891	48365.7909718229
		20	47017.9461841499	46256.1762641423	47779.7161041575
		50	47474.2645743976	47474.2645743976	47474.2645743976

Tablo 4.5. Varyans değerin gerçek, yaklaşık, alt ve üst sınır değerleri

(q,λ)	$Var(X)$	n	Yaklaşık $Var(X)$	(q,λ)	$Var(X)$	n	Yaklaşık $Var(X)$
(0.3,2)	0.9668832158	5	0.9663090140	(0.3,5)	1.1907613964	5	1.1878995568
		10	0.9668792184			10	1.1907486552
		20	0.9668832158			20	1.1907613964
(0.5,2)	2.9317233071	5	2.8966614573	(0.5,5)	3.4392247529	5	3.4078153711
		10	2.9317180597			10	3.4392192935
		50	2.9317233071			50	3.4392247529
(0.8,2)	28.2784602533	5	27.9418642698	(0.8,5)	32.5653946091	5	33.2237469999
		20	28.1456910482			20	32.2997608568
		50	28.2779548653			50	32.5644642142
		150	28.2784602533			150	32.5653946091

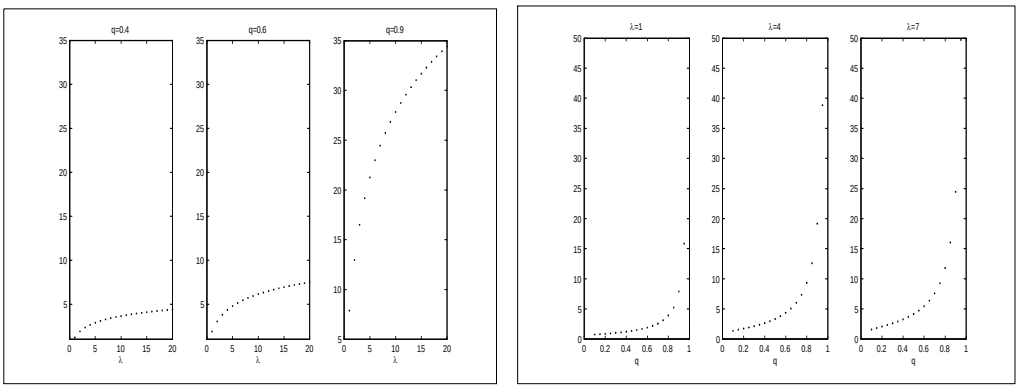
Tablo 4.6**.** Çarpıklık değerin gerçek, yaklaşık, Alt ve Üst sınır değerleri

(q,λ)	$Skewness(X)$	n	Yak. $Skewness(X)$	(q,λ)	$Skewness(X)$	n	Yak. $Skewness(X)$
(0.3,2)	1.5439885033	5	1.5293884914	(0.3,5)	1.0913310183	5	1.0974467318
		10	1.5438744811			10	1.0913286139
		20	1.5439885033			20	1.0913310183
(0.5,2)	1.4760392637	5	1.4620813735	(0.5,5)	1.1516492654	5	1.1686130051
		10	1.4760226997			10	1.1516370129
		50	1.4760392637			50	1.1516492654
(0.8,2)	1.4565690343	5	1.6365220116	(0.8,5)	1.1788315851	5	1.3763783197
		20	1.4392972052			20	1.1750124067
		50	1.4564145782			50	1.1786652455
		150	1.4565690343			150	1.1788315851

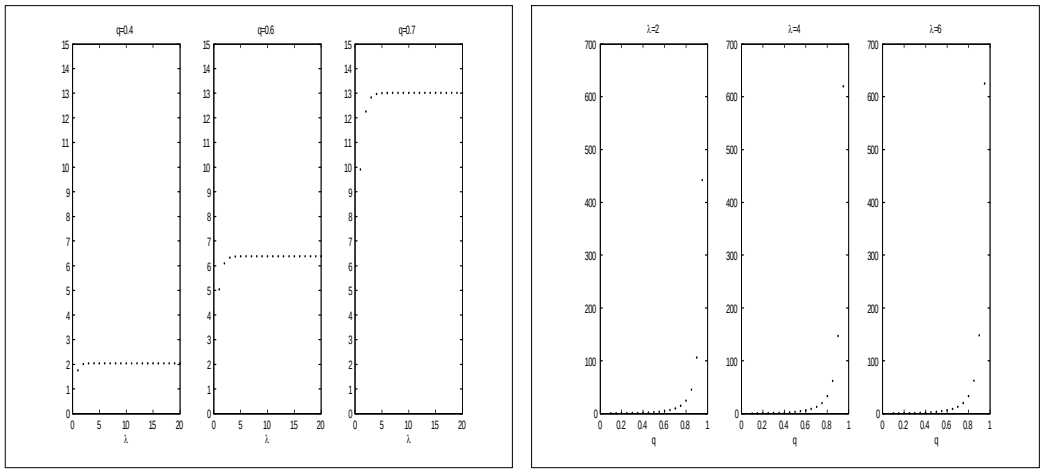
Tablo 4.7**.** Basıklık değerin gerçek, yaklaşık, Alt ve Üst sınır değerleri

(q,λ)	$Kurtosis(X)$	n	Yak. $Kurtosis(X)$	(q,λ)	$Kurtosis(X)$	n	Yak. $Kurtosis(X)$
(0.3,2)	6.3909103067	5	6.4108993095	(0.3,5)	5.1377052475	5	4.9970349940
		10	6.3908145204			10	5.1365329801
		20	6.3909103066			20	5.1377052475
(0.5,2)	6.3265775593	5	6.2994312225	(0.5,5)	5.3682140159	5	5.0740361546
		10	6.3265408209			10	5.3680842138
		50	6.3265775593			50	5.3682140159
(0.8,2)	6.3183347918	5	6.2224533983	(0.8,5)	5.4654383959	5	4.8882486478
		20	6.2211167752			20	5.3577364298
		50	6.3173067981			50	5.4644336039
		150	6.3183347918			150	5.4654383959

Tablo 4.6 ve 4.7 den $GZTP(q,\lambda)$ dağılımının sağa çarpık ve normal dağılıma göre basık olduğu görülmektedir. Dağılıma ait farklı parametreler için beklenen değer ve varyans grafikleri aşağıdaki Şekil 4.2 ve Şekil 4.3 te verilmiştir.



Şekil 4.2. λ ve q parametrelerinin farklı değerleri için beklenen değer grafikleri



Şekil 4.3. λ ve q parametrelerinin farklı değerleri için varyans grafikleri

4.1.2. Kuantil Fonksiyonu

Bu bölümde, dağılımın kuantil fonksiyonu elde edilerek dağılıma ait medyan, birinci ve üçüncü çeyreklik gibi özellikler verilecektir. $Q(u)$, $GZTP(q,\lambda)$ dağılımının kuantil fonskiyonu olmak üzere, $F(Q(u))=u$ denkleminin çözümü dağılımın kuantil fonksiyonunu verir. $GZTP(q,\lambda)$ dağılımı için

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

eşitliğinin çözümünden elde edilen kuantil fonksiyonu

$$Q(u)=\frac{\log(1-\lambda^{-1}\log(1+ue^{\lambda}-u))}{\log(q)}.$$

şeklinde elde edilir. Dağılım kuantilinin tek olması için $Q(u)$ fonksiyonun monoton artan yada azalan olması gereklidir. Açıktır ki, $u\in(0,1)$, $\lambda>0$ ve $q\in(0,1)$ için $Q(u)$ fonksiyonu u ya göre artandır. x_z, $X\sim GZTP(q,\lambda)$ dağılımının z. kuantili olmak üzere,

$$x_z = \begin{cases} \left\lfloor \frac{\log\left(1-\lambda^{-1}\log\left(1+ze^{\lambda}-z\right)\right)}{\log(q)} \right\rfloor + 1 & , \quad \lfloor Q(z) \rfloor \neq Q(z) \\ \left[\frac{\log\left(1-\lambda^{-1}\log\left(1+ze^{\lambda}-z\right)\right)}{\log(q)}, \frac{\log\left(1-\lambda^{-1}\log\left(1+ze^{\lambda}-z\right)\right)}{\log(q)} + 1 \right) & , \quad \lfloor Q(z) \rfloor = Q(z), \end{cases}$$

şeklinde elde edilir ve $F\left(x_z^-\right) \leq p \leq F\left(x_z\right)$ eşitsizliğini sağlar. Burada $\lfloor x \rfloor$, x değişkenin tam kısmını göstermektedir. $GZTP(q,\lambda)$ dağılımının medyanı

$$x_{0.5} = \begin{cases} \left\lfloor \frac{\log\left(1-\lambda^{-1}\log\left(1+0.5e^{\lambda}-0.5\right)\right)}{\log(q)} \right\rfloor + 1 & , \quad \lfloor Q(0.5) \rfloor \neq Q(0.5) \\ \left[\frac{\log\left(1-\lambda^{-1}\log\left(1+0.5e^{\lambda}-0.5\right)\right)}{\log(q)}, \frac{\log\left(1-\lambda^{-1}\log\left(1+0.5e^{\lambda}-0.5\right)\right)}{\log(q)} + 1 \right) & , \quad \lfloor Q(0.5) \rfloor = Q(0.5). \end{cases}$$

şeklinde yazılabilir.

4.1.3. Yaşam, Klasik ve Alternatif Bozulma Oranı Fonksiyonu

X, $GZTP(q,\lambda)$ dağılımına sahip rasgele değişken olmak üzere dağılımın yaşam fonksiyonu ve bozulma oranı fonksiyonu sırasıyla,

$$\begin{aligned} S(x) &= P(X \geq x) = 1 - P(X < x) \\ &= 1 - P(X \leq x-1) \\ &= \frac{e^{\lambda} - e^{-\lambda\left(q^{x-1}-1\right)}}{e^{\lambda}-1}. \end{aligned} \tag{4.2.2}$$

ve

$$\begin{aligned} h_0(x) &= P(X = x \mid X \geq x) \\ &= \frac{P(X=x)}{P(X \geq x)} \\ &= \frac{e^{-\lambda q^x} - e^{-\lambda q^{x-1}}}{1 - e^{-\lambda q^{x-1}}}, \quad x = 1,2,\ldots \end{aligned}$$

şeklindedir.

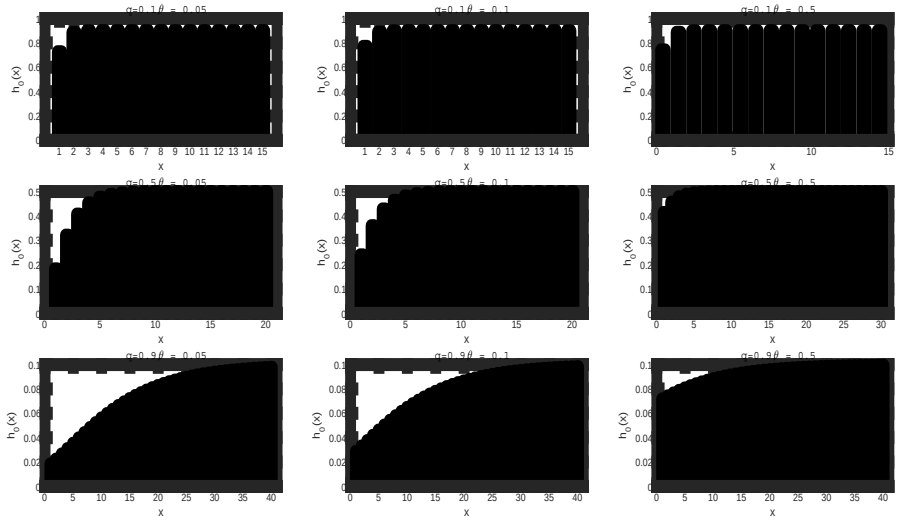
Teorem 4.2. $GZTP(q,\lambda)$ dağılımının bozulma oranı fonksiyonu artandır.

İspat: Teorem 4.1 de $GZTP(q,\lambda)$ dağılımın log-konkav olduğu gösterilmişti. Log-konkav olasılık fonksiyonu aynı zamanda artan bozulma oranına sahiptir(Keilson ve

Gerber, 1971). Bir diğer yol ise, her $x \in \Re^+$ değeri için bozulma oranı fonksiyonun birinci mertebeden türevinin sıfırdan büyük olduğunu gösterilmesidir. $h_0(x)$ fonksiyonun x e göre türevi

$$\frac{d}{dx}h_0(x) = \frac{(e^{-\lambda q^x}q^x \log e^{-\lambda} \log q - e^{-\lambda q^{x-1}}q^{x-1} \log e^{-\lambda} \log q)(1-e^{-\lambda q^{x-1}})}{(1-e^{-\lambda q^{x-1}})^2}$$
$$+\frac{e^{-\lambda q^{x-1}}q^{x-1} \log e^{-\lambda} \log q(e^{-\lambda q^x} - e^{-\lambda q^{x-1}})}{(1-e^{-\lambda q^{x-1}})^2}.$$

şeklinde olup ifadedeki iki terim de pozitifdir ve $\frac{d}{dx}h_0(x) > 0$ dır. Bu durumunda bozulma oranı fonksiyonu $x \in \Re^+$ değeri için artandır.



Şekil 4.4. λ ve q parametrelerinin farklı değerleri için klasik bozulma oranı fonksiyonu grafikleri

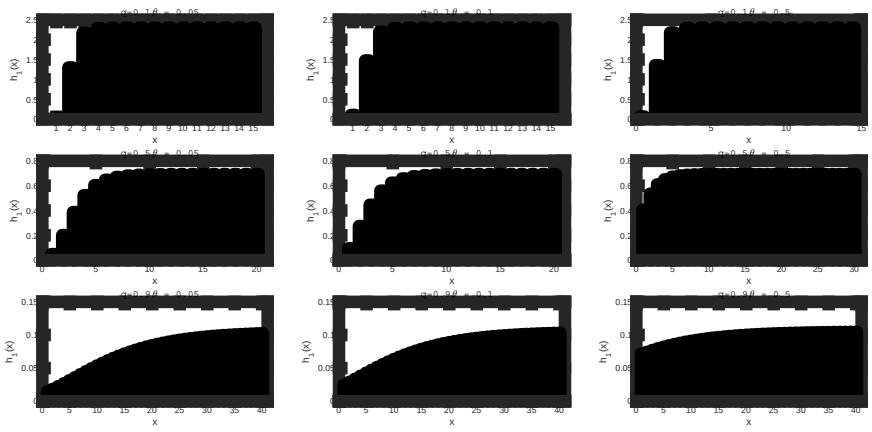
Literatürde klasik bozulma oranı fonksiyonu yanında önerilen farklı bozulma oranı fonksiyonları mevcuttur. Bu kısımda alternatif bozulma oranı fonksiyonu da olarak verilmiştir. (Roy ve Gupta, 1992) tarafından tanımlanan alternatif bozulma oranı fonksiyonu $h_1(x)$,

$$h_1(x) = \log\left\{\frac{S(x-1)}{S(x)}\right\} \tag{4.2.3}$$

şeklindedir. $GZTP(q,\lambda)$ dağılımının alternatif bozulma oranı fonksiyonu, Eşitlik (4.2.2) ün Eşitlik (4.2.3) de yazılmasıyla

$$h_1(x) = \log\left(\frac{1-e^{-\lambda q^{x-2}}}{1-e^{-\lambda q^{x-1}}}\right), \quad x = 1,2,\ldots$$

olarak elde edilir. $GZTP(q,\lambda)$ dağılımı için alternatif bozulma oranı fonksiyonu, klasik bozulma oranı fonksiyonu gibi artandır. Alternatif bozulma oranı fonksiyonu grafiği farklı parametre değerleri için Şekil 4.5 te verilmiştir.



Şekil 4.5. λ ve q parametrelerinin farklı değerleri için Alternatif Bozulma Oranı fonksiyonu grafikleri

4.2. Parametre Tahmini

Bu bölümde, $GZTP(q,\lambda)$ dağılımın parametrelerinin ML, LS ve MM tahmin edicileri incelenmiştir.

4.2.1. En Çok Olabilirlik Tahmin Edicisi

$X_1, X_2, \ldots, X_n$, $GZTP(q,\lambda)$ dağılımından elde edilen örneklem olmak üzere, q ve λ parametrelerinin tahmini için olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$L(q,\lambda)=\prod_{i=1}^{n}\left(\frac{1}{\left(1-e^{-\lambda}\right)}\left(e^{-\lambda q^{x_i}}-e^{-\lambda q^{x_i-1}}\right)\right)$$

$$\ell(q,\lambda)=n\log\left(\frac{1}{1-e^{-\lambda}}\right)+\sum_{i=1}^{n}\log\left(e^{-\lambda q^{x_i}}-e^{-\lambda q^{x_i-1}}\right),$$

şeklindedir. Log-olabilirlik fonksiyonunun parametrelere göre türevinin sıfıra eşitlenmesi ile elde edilen olabilirlik denklemleri

$$\frac{\partial \log\left(l\left(q,\lambda\right)\right)}{\partial \lambda} = -n\left(\frac{e^{-\lambda}}{1-e^{-\lambda}}\right) + \sum_{i=1}^{n}\left(\frac{-q^{x_i}e^{-\lambda q^{x_i}} + q^{x_i-1}e^{-\lambda q^{x_i-1}}}{e^{-\lambda q^{x_i}} - e^{-\lambda q^{x_i-1}}}\right) = 0 \qquad (4.2.3)$$

$$\frac{\partial \log\left(l\left(q,\lambda\right)\right)}{\partial q} = -\sum_{i=1}^{n}\frac{\lambda\left(-q^{(x_i-1)}\left(x_i-1\right)e^{\left(-\lambda q^{(x_i-1)}\right)} + q^{x_i}x_i e^{-\lambda q^{x_i}}\right)}{q\left(e^{-\lambda q^{x_i}} - e^{-\lambda q^{x_i-1}}\right)} = 0 \qquad (4.2.4)$$

biçiminde yazılabilir. (4.2.3) ve (4.2.4) denklemlerin analitik olarak çözümü elde edilemediğinden, denklem çözümleri için iteratif yöntemler kullanılabilir. Bu denklem sisteminin çözümünden elde edilen $(\hat{q},\hat{\lambda})$ ML tahmin edicilerinin asimptotik dağılımı, ortalama vektörü $(q,\lambda)'$, varyans kovaryans matirisi $\mathbf{I}_0^{-1}(\hat{q},\hat{\lambda})$ olan iki değişkenli normal dağılımdır. Burada $\mathbf{I_0^{-1}}(\hat{q},\hat{\lambda})$ ise gözlenen Fisher bilgi matrisinin tersidir. Buradan q ve λ parametrelerinin $100(1-\alpha)\%$ asimptotik güven aralığı sırasıyla

$$P\left(\hat{q} - z_{\frac{\alpha}{2}}\sqrt{V_{11}} < p < \hat{q} + z_{\frac{\alpha}{2}}\sqrt{V_{11}}\right) = 1-\alpha$$

ve

$$P\left(\hat{\lambda} - z_{\frac{\alpha}{2}}\sqrt{V_{22}} < \lambda < \hat{\lambda} + z_{\frac{\alpha}{2}}\sqrt{V_{22}}\right) = 1-\alpha$$

şeklindedir, burada V_{11} ve V_{22}, $\mathbf{I}_0^{-1}(\hat{q},\hat{\lambda})$ matrisinin diagonal elemanlarıdır.

4.2.2. En Küçük Kareler Tahmin Edicisi

$x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ q ve λ parametreli $GZTP(q,\lambda)$ dağılımına sahip kitleden alınmış sıra istatistiklerinin gözlenen değeri olsun. Bu durumda

$$F\left(x_{(i)}\right) = \frac{e^{-\lambda\left(q^{x(i)}-1\right)}-1}{e^{\lambda}-1},\ i = 1,2,\ldots,n. \qquad (4.2.5)$$

şeklinde yazılabilir. (4.2.5) deki $F\left(x_{(i)}\right)$ yerine ampirik dağılım fonksiyonu yazıldığında

$$\frac{i}{n} = \frac{e^{-\lambda\left(q^{x(i)}-1\right)}-1}{e^{\lambda}-1} + \varepsilon_i,\ i = 1,2,\ldots,n, \qquad (4.2.6)$$

modeli yazılabilir. Burada, ε_i, hata terimini, F^* ise ampirik dağılım fonksiyonunu

göstermektedir. (4.2.6) modelinde ε_i hata terimlerinin kareleri toplamını minimum yapan q ve λ değerleri LS tahmini olacaktır.

4.2.3. Momentler Yöntemi Tahmini

$X_1, X_2, \ldots, X_n$, $GZTP(q,\lambda)$ dağılımından alınmış rasgele örneklem olmak üzere, birinci ve ikinci örneklem momenti, dağılımın birinci ve ikinci kitle momentine eşitlenerek

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

$$\sum_{x=1}^{\infty} x^2\left(\frac{1}{\left(1-e^{-\lambda}\right)}\left(e^{-\lambda q^{x}}-e^{-\lambda q^{x-1}}\right)\right)=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

elden edilen denklem sisteminin çözümü momentler tahminini verecektir. İki bilinmeyenli iki denklemin çözümünde iteratif yöntemler kullanılarak çözümler elde edilmiştir.

4.2.4. Simülasyon Çalışması

Bu bölümde, Simülasyon yardımıyla ML, MM ve LS tahmin edicilerinin tam örneklem durumunda HKO ve yan kriterlerine göre performansları kıyaslanmıştır. Elde edilen sonuçlar farklı n ve (q,λ) parametreleri için Tablo 4.8 ve 4.9 da verilmiştir. Simülasyon çalışmasında 10000 tekrar kullanılmış ve örneklem hacmi n $(n=50,100,300,500)$ şeklinde alınmıştır.

Tablo 4.8. Parametre Tahmin edicilerin Yan tahminleri

		MLE		MME		LSE	
(q,λ)	n	$\hat{p}$	$\hat{\lambda}$	$\tilde{p}$	$\tilde{\lambda}$	$\bar{p}$	$\bar{\lambda}$
(0.2,2)	50	0.0049	0.1418	0.0094	0.3012	0.0062	0.1792
	100	0.0031	0.0744	0.0055	0.1583	0.0034	0.0998
	300	0.0009	0.0285	0.0016	0.0514	0.0010	0.0328
	500	0.0005	0.0121	0.0009	0.0279	0.0004	0.0132
(0.3,5)	50	0.0056	0.3738	0.0096	0.6610	0.0054	0.4207
	100	0.0027	0.1073	0.0047	0.3166	0.0027	0.1967
	300	0.0008	0.0504	0.0015	0.0987	0.0008	0.0604
	500	0.0005	0.0293	0.0008	0.0541	0.0003	0.0286
(0.5,2)	50	0.0037	0.0942	0.0115	0.2540	-0.0003	0.0583
	100	0.0023	0.0693	0.0071	0.1622	0.0008	0.0577
	300	0.0019	0.0370	0.0034	0.0673	0.0016	0.0362
	500	0.0003	0.0151	0.0013	0.0342	0.0002	0.0146
(0.7,3)	50	0.0041	0.3270	0.0117	0.5562	-0.0016	0.2687
	100	0.0034	0.1880	0.0072	0.3040	0.0012	0.1668
	300	0.0005	0.0465	0.0018	0.0859	0.0004	0.0427
	500	0.0005	0.0306	0.0014	0.0535	0.0001	0.0294

Tablo 4.9. Parametre Tahmin edicilerin HKO tahminleri

		MLE		MME		LSE	
(q,λ)	n	$\hat{p}$	$\hat{\lambda}$	$\hat{p}$	$\hat{\lambda}$	$\hat{p}$	$\hat{\lambda}$
(0.2,2)	50	0.0010	0.8286	0.0013	1.2071	0.0013	0.9998
	100	0.0005	0.4021	0.0007	0.5967	0.0006	0.4606
	300	0.0002	0.1195	0.0002	0.1865	0.0002	0.1321
	500	0.0001	0.0736	0.0001	0.0119	0.0001	0.0815
(0.3,5)	50	0.0011	2.1478	0.0016	3.8685	0.0013	2.9504
	100	0.0005	0.8743	0.0008	1.6046	0.0006	1.1230
	300	0.0002	0.2567	0.0003	0.4671	0.0002	0.3235
	500	0.0001	0.1475	0.0002	0.2638	0.0001	0.1867
(0.5,2)	50	0.0042	1.1331	0.0051	1.5048	0.0051	1.2963
	100	0.0021	0.4962	0.0026	0.6778	0.0023	0.5323
	300	0.0007	0.1529	0.0009	0.2188	0.0007	0.1606
	500	0.0004	0.0914	0.0006	0.1325	0.0004	0.0959
(0.7,3)	50	0.0044	2.2897	0.0052	3.0230	0.0052	2.5288
	100	0.0021	0.9256	0.0027	1.2521	0.0024	0.9992
	300	0.0007	0.2763	0.0009	0.3710	0.0008	0.2942
	500	0.0004	0.1598	0.0005	0.2153	0.0005	0.1708

Tablo 4.8 de verilen sonuçlara göre tahmin ediciler, genel olarak yanlı ancak asimptotik olarak yazsızdır. Tablo 4.9 a bakıldığında tahmin edicilerin HKO su, ML ve LS yönteminde hemen hemen aynı iken MM yönteminden daha iyi olduğu görülmüştür. Tablo 4.8 ve Tablo 4.9 da örneklem hacmi arttığında tahmin edicilerin yan ve HKO nda düşüş gözlenmiştir.

4.2.5. Gerçek Veri Uygulaması

Bu bölümde, yeni elde edilen $GZTP(q,\lambda)$ dağılımının gerçek verilerin geldiği dağılımı modelleme kabiliyeti bakımından aşağıda verilen dağılımlarla kıyaslanacaktır.

1. İlk dağılım Nakagawa and Osaki(Nakagawa ve Osaki, 1975) tarafından önerilen kesikli Weibull (DW) dağılımıdır. Bu dağılımın olasılık fonksiyonu

$$f_{DW}(x;p,\alpha,\gamma)=p^{x^{\alpha}}-p^{(x+1)^{\alpha}};\ \ x=0,1,2,...,\ \alpha>0,\ 0<p<1$$

şeklindedir.

2. İkinci dağılım Nekoukhou ve Bidram(Nekoukhou ve Bidram, 2015) tarafından önerilen üstelleştirilmiş kesikli Weibull (EDW) dağılımıdır. Bu dağılımın olasılık fonksiyonu

$$f_{EDW}(x;p,\alpha,\gamma)=(1-p^{(x+1)^{\alpha}})^{\gamma}-(1-p^{x^{\alpha}})^{\gamma};\ \ x=0,1,2,...,\ \alpha>0,\ \gamma>0,\ 0<p<1$$

biçimindedir.

3. Üçüncü dağılım Nekoukhou ve arkadaşları (Nekoukhou ve ark., 2012) tarafından önerilen Kesikli genelleştirilmiş üstel (DGE) dağılımıdır. Bu dağılımın olasılık fonksiyonu

$$f_{DGE}(x;\alpha,p)=kp^{x-1}(1-p^{x})^{\alpha-1},\ \ x=1,2,...,\ \alpha>0,\ 0<p<1$$

olarak verilir.

4. Dördüncü dağılım Geometrik dağılımdır ve olasılık fonksiyonu

$$f_{Ge}(x;p)=(1-p)p^{x-1},\ \ x=1,2,...,\ 0<p<1$$

şeklinde yazılabilir.

5. Beşinci dağılım sıfırdan budanmış Poisson (ZTP) dağılımıdır. Olasılık fonksiyonu

$$P(X=x)=\frac{e^{-l}l^{x}}{x!\left(1-e^{-l}\right)};\quad x=1,2,...\quad (l>0)$$

şeklindedir.

Dağılımların modelleme performans kriterleri olarak log-olabilirlik değeri, Akaike bilgi kriteri(AIC) ve Kolmogorov- Smirnov(K-S) değerleri ele alınmıştır. Bu değerlerden log-olabilirlik değerinin yüksek, AIC ve K-S değerinin küçük olması kıyaslamada kullanılacaktır.

Birinci Veri Seti: Bu veri seti Xie ve Lai(Xie ve Lai, 1995) tarafından verilmiş ve analiz edilmiştir. Eşit uzunluktaki zaman aralıklarında bozulma sayılarını gösteren veri seti Tablo 4.10 da, veri analiz edildiğinde elde edilen sonuçlar ise Tablo 4.11 de verilmiştir.

Tablo 4.10.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
	$n = 311$				

Tablo 4.11.Birinci veri seti için dağılımlara ait MLE, Log-olabilirlik, AIC, ve K-S değerleri.

Model	MLE	$l(\hat{\theta})$	AIC	K-S
GZTP	$\hat{q} = 0.7904$, $\hat{\theta} = 0.3100$	-835.2550	1674.5	0.1255
DW	$\hat{p} = 0.9579$, $\hat{\alpha} = 1.5861$	-851.0519	1706.1	0.4720
EDW	$\hat{p} = 0.9581$, $\hat{\alpha} = 1.5881$ $\hat{\gamma} = 0.9979$	-851.0520	1708.1	0.1213
DGE	$\hat{\alpha} = 1.3715$, $\hat{p} = 0.8106$	−838.4162	1680.8	0.3831
Ge	$\hat{p} = 0.8335$	−841.1116	1684.2	0.8335
ZTP	$\hat{\lambda} = 5.9914$	−996.0346	1994.1	0.4720

İkinci veri seti: Bu veri seti Macutek(Makcutek, 2008)tarafından elde edilmiştir. Veri soleven dil ailesindeki yazı birim frekanslarını göstermektedir. Tablo 4.12 de verilen veri seti Nekoukhou ve ark.(Nekoukhou ve ark., 2012) tarafından da analiz edilmiştir.

Tablo 4.12. 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
									$N = 313735$

Veri analiz edildiğinde elde edilen sonuçlar Tablo 4.13 te verilmiştir.

Tablo 4.13. İkinci veri seti için dağılımlara ait MLE, Log-olabilirlik, AIC, ve K-S değerleri.

Model	MLE	$l(\hat{\theta})$	AIC	K-S
GZTP	$\hat{q}=0.8378$, $\hat{\lambda}=0.3228$	-9.3114E05	1.8623E06	0.0960
DW	$\hat{p}=0.9626$, $\hat{\alpha}=1.4739$	-9.4503E05	1.8901E06	0.1005
EDW	$\hat{p}=0.8517$, $\hat{\alpha}=1.0407$ $\hat{\gamma}=1.9099$	-9.4394E05	1.8879E06	0.0887
DGE	$\hat{\alpha}=1.3950$, $\hat{p}=0.8808$	-9.4396E05	1.8879E06	0.2858
Ge	$\hat{p}=0.8715$	-9.3638E05	1.8728E06	0.8814
ZTP	$\hat{\lambda}=7.7787$	-1.2318E06	2.4635E06	0.5062

Üçüncü veri seti: Bu veri seti (Bakouch ve ark., 2014) tarafından verilmiştir. Veri lösemi hastalığı tedevisi sırasında rasgele seçilen yirmi hastanın haftalık remisyon sürelerini göstermektedir. 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. Veri analiz edildiğinde elde edilen sonuçlar Tablo 4.14 te verilmiştir.

Tablo 4.14. Üçüncü veri seti için dağılımlara ait MLE, Log-olabilirlik, AIC, ve K-S değerleri.

Model	MLE	$l(\hat{\theta})$	AIC	K-S
GZTP	$\hat{q}=0.9301$, $\hat{\theta}=0.2252$	-78.3397	160.6793	0.0853
DW	$\hat{p}=0.9857$, $\hat{\alpha}=1.3740$	-78.6119	161.2237	0.0866
EDW	$\hat{p}=0.9998$, $\hat{\alpha}=2.4424$ $\hat{\gamma}=0.4261$	-78.4807	162.9614	0.0720
DGE	$\hat{\alpha}=1.3481$, $\hat{p}=0.9408$	-78.6044	161.2089	0.6006
Ge	$\hat{p}=0.9488$	-78.9391	159.8781	0.9499
ZTP	$\hat{\lambda}=19.5500$	-152.7180	307.4360	0.5617

Yukarıda elde edilen sonuçlara göre her üç veri seti için, $GZTP(q,\lambda)$ dağılımı diğer dağılımlar içerisinde en küçük AIC değerine sahiptir. En yüksek log-olabilirlik değerine yine $GZTP(q,\lambda)$ dağılımı sahiptir. K-S değerine göre elde edilen yeni dağılım, EDW dağılımdan sonra ikinci küçük değere sahip olduğu görülmektedir.

5. BAZI KESİKLİ DAĞILIMLAR İÇİN SANSÜRLÜ ÖRNEKLEM DURUMUNDA PARAMETRE TAHMİNİ

5.1. Kesikli Chen Dağılımı

Kesikli Chen dağılımı, (Noughabi ve ark., 2013) tarafından elde edilmiştir. Bu bölümünde kesikli Chen dağılımı için tam ve sansürlü örneklem durumunda parametre tahmini problemi incelenmiştir. X kesikli Chen dağılımına sahip rasgele değişken olduğunda X rasgele değişkeninin olasılık fonksiyonu ve dağılım fonksiyonu sırasıyla,

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

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

şeklindedir. Burada dağılımın parametreleri $q \in (0,1)$, $\beta > 0$ dır ve dağılım kısaca $DCD(q,\beta)$ biçiminde gösterilir. Bu bölümde Tip-I sansürlü örnekleme dayalı parametre konusu tartışılmıştır.

5.1.1. En Çok Olabilirlik Tahmini

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

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

şeklinde tanımlanmak üzere $X_1, X_2, \ldots, X_n$ 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

$$L(q,\beta) = \prod_{i=1}^{n} \left(p(x_i)^{\delta_i} S(x_i)^{1-\delta_i} \right) \tag{5.1.1}$$

ve

$$\begin{aligned} \ell(q,\beta) &= \log\left\{ \prod_{i=1}^{n} \left(p(x_i)^{\delta_i} S(x_i)^{1-\delta i} \right) \right\} \\ &= \sum_{i=1}^{n} \delta_i \log\left(p(x_i) \right) + \sum_{i=1}^{n} (1-\delta_i) \log\left(S(x_i) \right) \\ &= \sum_{i=1}^{n} \delta_i \log\left(q^{e^{x_i^\beta}-1} - q^{e^{(x_i+1)^\beta}-1} \right) + \sum_{i=1}^{n} (1-\delta_i) \log\left(q^{e^{(x_i+1)^\beta}-1} \right) \end{aligned} \tag{5.1.2}$$

şeklinde yazılabilir. Olabilirlik denklemleri ise aşağıdaki gibi elde edilir:

$$\begin{aligned}\frac{\partial \log\left(l\left(q,\beta\right)\right)}{\partial \beta} = \sum_{i=1}^{n}\delta_i \Bigg\{ & \frac{x_i^{\beta}\log\left(x_i\right)\exp\left(x_i^{\beta}\right)\log\left(q\right)q^{\exp\left(x_i^{\beta}\right)-1}}{q^{e^{x_i^{\beta}}-1}-q^{e^{(x_i+1)^{\beta}}-1}} \\ & - \frac{\left(x_i+1\right)^{\beta}\log\left(x_i+1\right)\exp\left(\left(x_i+1\right)^{\beta}\right)\log\left(q\right)q^{\exp\left((x_i+1)^{\beta}\right)-1}}{q^{e^{x_i^{\beta}}-1}-q^{e^{(x_i+1)^{\beta}}-1}} \\ & + \left(x_i+1\right)^{\beta}\log\left(x_i+1\right)\exp\left(\left(x_i+1\right)^{\beta}\right)\log\left(q\right)\Bigg\} = 0,\end{aligned} \tag{5.1.3}$$

$$\begin{aligned}\frac{\partial \log\left(l\left(q,\beta\right)\right)}{\partial q} = \frac{1}{q}\sum_{i=1}^{n}\left(1-\delta_i\right)\Bigg\{ & \frac{q^{\exp\left(x_i^{\beta}\right)}\left(\exp\left(x_i^{\beta}\right)-1\right)}{q^{e^{x_i^{\beta}}}-q^{e^{(x_i+1)^{\beta}}}} \\ & + \frac{q^{\exp\left((x_i+1)^{\beta}\right)}\left(1-\exp\left(\left(x_i+1\right)^{\beta}\right)\right)}{q^{e^{x_i^{\beta}}}-q^{e^{(x_i+1)^{\beta}}}} + \left(\exp\left(\left(x_i+1\right)^{\beta}\right)-1\right)\Bigg\} = 0.\end{aligned} \tag{5.1.4}$$

(5.1.3) ve (5.1.4) denklemleri analitik olarak çözülemediğinden ML tahminlerini elde etmek için nümerik yöntemler kullanılır.

5.1.2. Oranlar Tahmini

Bu yöntem (Khan ve ark., 1989) tarafından Kesikli Weibull dağılımının parametrelerinin tahmininde kullanılmıştır. Bu yöntem $DCD\left(q,\beta\right)$ için de kullanılacaktır. $X_1, X_2, \ldots, X_n$ bağımsız ve $DCD\left(q,\beta\right)$ den alınmış Tip-I sansürlü örneklem olmak üzere, aşağıdaki gösterge fonksiyonu tanımlansın:

$$\nu_0\left(X_i\right) = \begin{cases} 1, & X_i = 0 \\ 0, & X_i > 0. \end{cases}$$

Bu durumda $Y = \frac{1}{n}\sum_{i=1}^{n}\nu_0\left(X_i\right)$, $f\left(0\right) = 1-q^{e-1}$ olasılığının bir tahmin edicisidir. Y, $f\left(0\right)$ a eşitlenip q yalnız bırakıldığında, q parametresinin MP tahmin edicisi

$$\tilde{q} = \left(1-Y\right)^{\left(e-1\right)^{-1}}$$

olarak elde edilir. Aşağıdaki gösterge fonksiyonu göz önüne alınsın:

$$\nu_1(X_i)=\begin{cases}1, & X_i=1\\ 0, & X_i\neq 1\end{cases}$$

Bu durumda $Z=\frac{1}{n}\sum_{i=1}^{n}\nu_1(X_i)$, $f(1)=q^{\exp(1)-1}-q^{\exp(2^{\beta})-1}$ olasılığının yansız ve tutarlı bir tahmin edicisidir. Z, $f(1)$ a eşitlenip β yalnız bırakıldığında, β parametresinin oranlar tahmin edicisi

$$\tilde{\beta}=\frac{\log\left\{\log\left[\log\left(\tilde{q}^{\exp(1)}-\tilde{q}Z\right)-\log\left(\log\left(\tilde{q}\right)\right)\right]\right\}}{\log(2)},$$

olarak elde edilir. Bu yöntem bazı durumlarda çalışmayabilir. Örneğin; sansürlemenin $L_i=1,\ i=1,\ldots,n$ olduğu durumlarda birlerin sayısını hesaplanamadığından için β parametresinin oran tahmini elde edilemez.

5.1.3. En Küçük Kareler Tahmini

$x_{(1)}<x_{(2)}<\cdots<x_{(n)}$, $DCD(q,\beta)$ dağılımına sahip kitleden alınmış sıra istatistiklerinin gözlenen değeri olsun. Bu durumda

$$S\left(x_{(i)}\right)=q^{e^{\left(x_{(i)}+1\right)^{\beta}}-1},\ i=1,2,\ldots,n. \tag{5.1.5}$$

şeklinde yazılabilir, burada $S\left(x_{(i)}\right)$ $DCD(q,\beta)$ dağılımının yaşam fonksiyonudur. (5.1.5) deki $S\left(x_{(i)}\right)$ yerine Kaplan-Maier tahmin edicisi yazıldığında

$$S^{*}\left(x_{(i)}\right)=q^{e^{\left(x_{(i)}+1\right)^{\beta}}-1}+\varepsilon_i,\ i=1,2,\ldots,n. \tag{5.1.6}$$

modeli yazılabilir. Burada, ε_i, hata terimini, S^{*} ise Kaplan-Maier tahmin edicisidir. (5.1.6) modelinde, ε_i hata terimlerinin kareleri toplamını minimum yapan q ve β değerleri LS tahmini olacaktır.

5.1.4. Simülasyon Çalışması

Bu bölümde, Monte Carlo Simülasyonu yardımıyla ML, MP ve LS tahmin edicilerinin sansürlü ve tam örneklem durumunda HKO ve yan kriterlerine göre performansları kıyaslanmıştır. Simülasyon çalışmasında 10000 tekrar kullanılmış ve $L_i=L=2,\ i=1,\ldots,n$, örneklem hacmi $(n=50,100,200,300,500)$ şeklinde

belirlenmiştir. Elde edilen sonuçlar Tablo 5.1.1 ve Tablo 5.1.2 de verilmiştir.

Tablo 5.1.1. Tahmin edicilerin yanları

				MLE		LSE		MPE	
q	β	L	n	$Bias(q)$	$Bias(\beta)$	$Bias(q)$	$Bias(\beta)$	$Bias(q)$	$Bias(\beta)$
0.9	1.5	2	50	-0.00049	0.02259	0.00116	0.00539	-0.00083	0.02052
		2	100	-0.00021	0.01036	0.00071	0.00243	-0.00031	0.00971
		2	200	-0.00021	0.00495	0.00028	0.00118	-0.00024	0.00479
		2	300	-0.00013	0.00352	0.00021	0.00106	-0.00014	0.00347
		2	500	-0.00004	0.00215	0.00016	0.00069	-0.00005	0.00213
0.9	1.5	∞	50	0.00100	0.02811	0.00260	0.01069	0.00067	0.02583
		∞	100	-0.00027	0.01123	0.00065	0.00329	-0.00037	0.01061
		∞	200	-0.00017	0.00585	0.00033	0.00208	-0.00019	0.00571
		∞	300	-0.00008	0.00368	0.00025	0.00122	-0.00009	0.00362
		∞	500	-0.00012	0.00188	0.00008	0.00042	-0.00012	0.00186
0.5	0.5	2	50	-0.00189	0.02138	0.01051	-0.01744	-0.00249	0.02198
		2	100	-0.00133	0.00800	0.00482	-0.01058	-0.00171	0.00521
		2	200	-0.00022	0.00461	0.00308	-0.00441	-0.00047	0.00364
		2	300	-0.00015	0.00211	0.00182	-0.00393	-0.00036	0.00126
		2	500	-0.00010	0.00131	0.00114	-0.00234	-0.00018	0.00066
0.5	0.5	∞	50	0.00015	0.02651	0.00972	-0.02527	-0.00268	0.01157
		∞	100	-0.00015	0.01407	0.00412	-0.01556	-0.00195	0.00426
		∞	200	-0.00012	0.00763	0.00214	-0.00792	-0.00184	0.00238
		∞	300	-0.00012	0.00443	0.00079	-0.00609	-0.00098	0.00103
		∞	500	-0.00011	0.00279	0.00071	-0.00341	-0.00047	0.00086
0.9	0.4	2	50	-0.00050	0.00752	0.00160	0.00973	0.00013	0.01490
		2	100	-0.00025	0.00468	0.00073	0.00421	-0.00011	0.00550
		2	200	-0.00010	0.00160	0.00040	0.00138	0.00010	0.00372
		2	300	-0.00006	0.00070	0.00025	0.00058	0.00005	0.00252
		2	500	-0.00000	0.00090	0.00020	0.00076	0.00002	0.00085
0.9	0.4	∞	50	0.00138	0.00632	0.00055	-0.00232	0.00051	0.01468
		∞	100	0.00034	0.00311	-0.00023	-0.00111	-0.00019	0.00803
		∞	200	0.00029	0.00157	-0.00016	-0.00098	-0.00012	0.00144
		∞	300	0.00019	0.00130	-0.00018	-0.00061	0.00009	0.00239
		∞	500	0.00010	0.00068	-0.00013	-0.00060	0.00002	0.00206

Tablo 5.1.2. Tahmin edici yöntemlerine göre parametrelere ait HKO

				MLE		LSE		MPE	
q	β	L	n	$HKO(q)$	$HKO(\beta)$	$HKO(q)$	$HKO(\beta)$	$HKO(q)$	$HKO(\beta)$
0.9	1.5	2	50	0.00108	0.03119	0.00107	0.03279	0.00111	0.03294
		2	100	0.00053	0.01492	0.00053	0.01515	0.00054	0.01522
		2	200	0.00027	0.00736	0.00027	0.00739	0.00027	0.00741
		2	300	0.00018	0.00499	0.00018	0.00499	0.00018	0.00500
		2	500	0.00011	0.00291	0.00011	0.00291	0.00011	0.00291
0.9	1.5	∞	50	0.00100	0.02943	0.00100	0.03041	0.00103	0.03068
		∞	100	0.00055	0.01514	0.00054	0.01534	0.00055	0.01542
		∞	200	0.00027	0.00740	0.00027	0.00742	0.00027	0.00745
		∞	300	0.00018	0.00484	0.00018	0.00485	0.00018	0.0049
		∞	500	0.00011	0.00293	0.0011	0.00293	0.00011	0.00294
0.5	0.5	2	50	0.00368	0.01965	0.00348	0.01130	0.00371	0.02065
		2	100	0.00190	0.00545	0.00185	0.00553	0.00193	0.00917
		2	200	0.00095	0.00268	0.00094	0.00280	0.00096	0.00454
		2	300	0.00064	0.00170	0.0064	0.00182	0.00065	0.00296
		2	500	0.00038	0.00101	0.00038	0.00109	0.00039	0.00178
0.5	0.5	∞	50	0.00368	0.00988	0.00350	0.01032	0.00378	0.01822
		∞	100	0.00190	0.00420	0.00185	0.00495	0.00196	0.00918
		∞	200	0.00095	0.00195	0.00095	0.00242	0.00099	0.00454
		∞	300	0.00064	0.00123	0.00063	0.00157	0.00066	0.00297
		∞	500	0.00038	0.00073	0.00038	0.00094	0.00039	0.00178
0.9	0.4	2	50	0.00113	0.01580	0.00100	0.01590	0.00109	0.03736
		2	100	0.00053	0.00782	0.00052	0.00782	0.00054	0.01740
		2	200	0.00026	0.00370	0.00026	0.00370	0.00027	0.00813
		2	300	0.00018	0.00251	0.00018	0.00251	0.00018	0.00538
		2	500	0.00011	0.00150	0.00011	0.00151	0.00011	0.00328
0.9	0.4	∞	50	0.00061	0.00092	0.00069	0.00127	0.00111	0.04018
		∞	100	0.00030	0.00042	0.00033	0.00055	0.00054	0.01705
		∞	200	0.00015	0.00021	0.00017	0.00027	0.00027	0.00836
		∞	300	0.00010	0.00013	0.00011	0.00017	0.00018	0.00548
		∞	500	0.00006	0.00008	0.00006	0.00010	0.00011	0.00321

Simülasyon çalışması sonucunda beklendiği gibi örneklem hacmi arttığında q ve β parametrelerinin LS, ML ve MP tahmin edicilerinin hem sansürlü, hem de tam örneklem durumunda HKO ve yanları düşmektedir. Küçük hacimli sansürlü örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO su, tam örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO na göre daha büyüktür. Ancak büyük örneklem durumunda, sansürlü ve tam örneklemden elde edilen parametre tahminlerinin yan ve HKO su hemen hemen aynıdır.

5.1.5. Gerçek Veri Uygulaması

(Phyo, 1973) ten alınan 10-11 yaş 100 çocuğa ait çürük diş verisi kullanılmıştır. Parametre tahminleri Tablo 5.1.4 te verilmiştir.

Tablo 5.1.3. Phyo (1973) diş verisi

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

Tablo 5.1.4. Veriye ait Kesikli Chen Dağılımının Parametre Tahmini

	Parametre Tahminleri	
Yöntem	$\hat{q}$	$\hat{\beta}$
MLE	0.5646	0.4908
LSE	0.5627	0.4444
PE	0.5518	0.4149

Kesikli Chen dağılımının Kesikli Weibull, Kesikli Burr, Geometrik, Poisson ve Kesikli Pareto dağılımları ile olabilirlik fonksiyon değeri, AIC ve BIC kriterlerine göre kıyaslanması Tablo 5.1.5 te verilmiştir.

Tablo 5.1.5. Veriye ait uyum iyiliği tablosu

Diş Sayısı(x)	$L_{\max}$	AIC	BIC
Chen(0.5646,0.4908)	-110.9033	225.8067	231.0170
Weibull (0.3742,0.8951)	-112.1001	228.2001	233.4105
Burr(0.2312,1.4408)	-114.7397	233.4794	238.6897
Geometric (0.4012)	-112.4735	226.9471	229.5523
Pareto(0.1843)	-116.8314	235.6628	238.2679
Poisson(0.6700)	-121.0482	244.0964	246.7015

Tablo 5.1.5' e göre (Phyo, 1973) diş verisinin geldiği dağılımı, Kesikli Chen dağılımının diğer kesikli Weibull, Kesikli Burr, Geometrik, Kesikli Pareto ve Poisson dağılımına göre daha iyi modellediği görülmüştür.

5.2. Kesikli Burr Dağılımı

Kesikli Burr dağılımı, (Krishna ve Pundir, 2009) tarafından bilinen Burr dağılımının kesikleştirilmesi sonucunda elde edilmiştir. Çalışmada elde edilen dağılımın özellikleri ve tam örneklem durumunda parametre tahmini ML tahmin edici yöntemiyle elde edilmiştir. Tezin bu bölümünde Kesikli Burr dağılımı için sansürlü örneklem durumunda parametre tahmini ML, oran ve MML yöntemleri ile elde edilecektir.

Kesikli Burr dağılımının olasılık ve dağılım fonksiyonu sırasıyla

$$f(x) = q^{\log\left(1+x^{\beta}\right)} - q^{\log\left(1+(1+x)^{\beta}\right)}, \quad x = 0,1,2,\ldots$$

ve

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

elde edilmiştir. Burada $0 < q < 1$ ve $\beta > 0$ dağılım parametreleridir. Dağılım kısaca $DBD(q,\beta)$ şeklinde gösterilecektir. Bu bölümde Tip-I sansürlü örnekleme dayalı parametre konusu tartışılmıştır.

5.2.1. En Çok Olabilirlik Tahmini

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

$$X_i = \min\left(Y_i, L_i\right),\ i = 1,\ldots,n$$

şeklinde tanımlanmak üzere $X_1, X_2, \ldots, X_n$ örneklemine Tip-I sansürlü örneklem olsun.

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

gösterge fonksiyonu kullanılarak olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$\begin{aligned} L(q,\beta) &= \prod_{i=1}^{n} f(x_i)^{\delta_i} \left(1 - F(x_i)\right)^{1-\delta_i} \\ &= \prod_{i=1}^{n} \left(q^{\log\left(1+x^{\beta}\right)} - q^{\log\left(1+(1+x)^{\beta}\right)} \right)^{\delta_i} \left(q^{\log\left(1+x^{\beta}\right)} \right)^{1-\delta_i} \end{aligned} \tag{5.2.1}$$

ve

$$\begin{aligned}\ell(q,\beta) &= \log\left\{\prod_{i=1}^{n} f(x_i)^{\delta_i}\left(1-F(x_i)\right)^{1-\delta_i}\right\} \\ &= \sum_{i=1}^{n}\delta_i \log\left(f(x_i)\right)+\sum_{i=1}^{n}(1-\delta_i)\log\left(1-F(x_i)\right) \\ &= \sum_{i=1}^{n}\delta_i \log\left(q^{\log\left(1+x_i^{\beta}\right)}-q^{\log\left(1+(1+x_i)^{\beta}\right)}\right)+\sum_{i=1}^{n}(1-\delta_i)\log\left(q^{\log\left(1+x_i^{\beta}\right)}\right)\end{aligned} \tag{5.2.2}$$

şeklinde yazılabilir. Olabilirlik denklemleri ise aşağıdaki gibi elde edilir.

$$\begin{aligned}\frac{\partial \ell(q,\beta)}{\partial q} &= \sum_{i=0}^{n}\frac{\delta_i}{q}\left(\frac{q^{\log\left(1+x_i^{\beta}\right)}\log\left(1+x_i^{\beta}\right)-q^{\log\left(1+x_i^{\beta}\right)}\log\left(1+\left(1+x_i\right)^{\beta}\right)}{q^{\log\left(1+x_i^{\beta}\right)}-q^{\log\left(1+(1+x_i)^{\beta}\right)}}\right) \\ &+\sum_{i=0}^{n}\frac{(1-\delta_i)\log\left(1+x_i^{\beta}\right)}{q}=0,\end{aligned} \tag{5.2.3}$$

$$\begin{aligned}\frac{\partial \ell(q,\beta)}{\partial \beta} &= \sum_{i=0}^{n}\delta_i\frac{q^{\log\left(1+x_i^{\beta}\right)}x_i^{\beta}\log(x_i)\log(q)}{1+x_i^{\beta}}\left(\left(q^{\log\left(1+x_i^{\beta}\right)}-q^{\log\left(1+(1+x_i)^{\beta}\right)}\right)\right)^{-1} \\ &-\sum_{i=0}^{n}\delta_i\frac{q^{\log\left(1+(1+x_i)^{\beta}\right)}(1+x)^{\beta}\log(1+x_i)\log(q)}{\left(1+(1+x_i)^{\beta}\right)\left(\left(q^{\log\left(1+x_i^{\beta}\right)}-q^{\log\left(1+(1+x_i)^{\beta}\right)}\right)\right)} \\ &+\sum_{i\neq 0}^{n}\frac{(1-\delta_i)x_i^{\beta}\log(x_i)\log(q)}{1+x_i^{\beta}}=0.\end{aligned} \tag{5.2.4}$$

(5.1.3) ve (5.1.4) denklemleri analitik olarak çözülemediğinden ML tahminlerini elde etmek için nümerik yöntemler kullanılır.

5.2.2. Oran Yöntemi Tahmini

Bu yöntem (Khan ve ark., 1989) tarafından Kesikli Weibull dağılımının parametrelerinin tahmininde kullanılmıştır. Bu yöntem $DBD(q,\beta)$ için de kullanılacaktır. $X_1, X_2, \ldots, X_n$ bağımsız ve $DBD(q,\beta)$ den alınmış Tip-I sansürlü örneklem olmak üzere, aşağıdaki gösterge fonksiyonu tanımlansın:

$$\nu_0(X_i)=\begin{cases}1, & X_i=0\\ 0, & X_i>0.\end{cases}$$

Bu durumda $Y=\frac{1}{n}\sum_{i=1}^{n}\nu_0(X_i)$, $f(0)=1-(q)^{\log(2)}$ olasılığının bir tahmin edicisidir. Y, $f(0)$ a eşitlenip q yalnız bırakıldığında, q parametresinin oranlar tahmin edicisi

$$\tilde{q}=(1-Y)^{1/\log(2)}.$$

olarak elde edilir. Aşağıdaki gösterge fonksiyonu göz önüne alınsın:

$$\nu_1(X_i)=\begin{cases}1, & X_i=1\\ 0, & X_i\neq 1\end{cases}$$

Bu durumda $Z=\frac{1}{n}\sum_{i=1}^{n}\nu_1(X_i)$, $f(1)=q^{\log(2)}-q^{\log\left(1+2^{\beta}\right)}$ olasılığının yansız ve tutarlı bir tahmin edicisidir. Z, $f(1)$ a eşitlenip β yalnız bırakıldığında, β parametresinin oranlar tahmin edicisi

$$\tilde{\beta}=\log\left(\exp\left\{\frac{\log\left(\tilde{q}^{\log(2)}-Z\right)}{\log(\tilde{q})}\right\}-1\right)(\log(2))^{-1},$$

olarak elde edilir. Bu yöntem bazı durumlarda çalışmayabilir. Örneğin; sansürlemenin $L_i=1$ olduğu durumlarda birlerin sayısını hesaplayamadığımız için β parametresinin oran tahmini elde edilemez.

5.2.3. Uyarlanmış En Çok Olabilirlik Tahmini

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

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

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

ele alınır. Burada $\tilde{q}$ ve $\tilde{\beta}$, parametrelerin MP edicileridir. Yukarıdaki denklemler çözüldüğünde MML tahmin edicileri

$$\hat{\beta}_{*}=\left.\frac{-f_2(q,\beta)\frac{\partial f_1(q,\beta)}{\partial q}+\frac{\partial f_2(q,\beta)}{\partial q}f_1(q,\beta)-\frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}\beta+\frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q}\beta}{\frac{\partial f_2(q,\beta)}{\partial \beta}\frac{\partial f_1(q,\beta)}{\partial q}-\frac{\partial f_2(q,\beta)}{\partial q}\frac{\partial f_1(q,\beta)}{\partial \beta}}\right|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} \tag{5.2.5}$$

$$\hat{q}_{*}=\left.-\frac{f_1(q,\beta)-\frac{\partial f_1(q,\beta)}{\partial q}q+\frac{\partial f_1(q,\beta)}{\partial \beta}\hat{\beta}_{*}-\frac{\partial f_1(q,\beta)}{\partial \beta}\beta}{\frac{\partial f_1(q,\beta)}{\partial q}}\right|_{\substack{q=\tilde{q}\\ \beta=\tilde{\beta}}} \tag{5.2.6}$$

olarak elde edilir.

5.2.4. Simülasyon Çalışması

Tezin bu bölümünde ML, MP, MML ve UMML tahmin edicilerinin Tip-I sansürlü örneklem durumunda, HKO ve yan kriterlerine göre karşılaştırılması verilmiştir. 10000 tekrarla elde edilen sonuçlar Tablo 5.2.1-5.2.4 da verilmiştir. Simülasyon çalışmasında örneklem hacmi $(n=100,300,500)$ ve $L_i=L=3,5,7,\quad i=1,...,n$ şeklinde belirlenmiştir.

Tablo 5.2.1. DB dağılımı $(q=0.3,\beta=0.5)$ için parametrelerin tahmin edicilerinin HKO

		MLE		MPE		UMMLE		MMLE	
L	n	$\hat{q}$	$\hat{\beta}$	$\widetilde{q}$	$\widetilde{\beta}$	$\hat{q}_{**}$	$\hat{\beta}_{**}$	$\hat{q}_{*}$	$\hat{\beta}_{*}$
3	100	0.0024	0.0161	0.0024	0.0266	0.0025	0.0169	0.0024	0.0204
3	300	0.0008	0.0063	0.0008	0.0097	0.0008	0.0063	0.0008	0.0061
3	500	0.0005	0.0038	0.0005	0.0058	0.0005	0.0038	0.0006	0.0052
5	100	0.0023	0.0118	0.0023	0.0262	0.0024	0.0123	0.0026	0.0152
5	300	0.0008	0.0042	0.0008	0.0095	0.0008	0.0042	0.0008	0.0042
5	500	0.0005	0.0024	0.0005	0.0057	0.0005	0.0024	0.0006	0.0029
7	100	0.0024	0.0116	0.0024	0.0256	0.0024	0.0113	0.0027	0.0135
7	300	0.0008	0.0037	0.0008	0.0093	0.0008	0.0037	0.0007	0.0036
7	500	0.0005	0.0022	0.0005	0.0056	0.0005	0.0022	0.0004	0.0022
∞	100	0.0024	0.0116	0.0024	0.0254	0.0024	0.0112	0.0024	0.0124
∞	300	0.0008	0.0024	0.0008	0.0091	0.0008	0.0024	0.0006	0.0038
∞	500	0.0005	0.0015	0.0005	0.0055	0.0005	0.0015	0.0005	0.0015

Tablo 5.2.2. DB dağılımı $(q=0.5,\beta=1.5)$ için tahmin edicilerinin HKO

		MLE		MPE		UMMLE		MMLE	
L	n	$\hat{q}$	$\hat{\beta}$	$\tilde{q}$	$\tilde{\beta}$	$\hat{q}_{**}$	$\hat{\beta}_{**}$	$\hat{q}_{*}$	$\hat{\beta}_{*}$
3	100	0.0032	0.0878	0.0032	0.1252	0.0031	0.0860	0.0072	0.0686
3	300	0.0011	0.0323	0.0011	0.0402	0.0011	0.0323	0.0028	0.0519
3	500	0.0006	0.0150	0.0006	0.0229	0.0006	0.0120	0.0010	0.0282
5	100	0.0032	0.0831	0.0032	0.1249	0.0032	0.0792	0.0055	0.0395
5	300	0.0011	0.0256	0.0011	0.0396	0.0011	0.0256	0.0010	0.0243
5	500	0.0006	0.0149	0.0006	0.0254	0.0006	0.0149	0.0010	0.0237
7	100	0.0031	0.0748	0.0031	0.1210	0.0035	0.0772	0.0052	0.0287
7	300	0.0011	0.0245	0.0011	0.0400	0.0011	0.0245	0.0012	0.0250
7	500	0.0007	0.0150	0.0007	0.0236	0.0007	0.0150	0.0009	0.0202
∞	100	0.0030	0.0722	0.0031	0.1190	0.0035	0.0779	0.0044	0.0280
∞	300	0.0011	0.0236	0.0011	0.0388	0.0011	0.0236	0.0009	0.0226
∞	500	0.0006	0.0141	0.0006	0.0235	0.0006	0.0141	0.0008	0.0146

Tablo 5.2.3. DB dağılımı $(q=0.3,\beta=0.5)$ için tahmin edicilerinin yanları

		MLE		MPE		UMMLE		MMLE	
L	n	$\hat{q}$	$\hat{\beta}$	$\tilde{q}$	$\tilde{\beta}$	$\hat{q}_{**}$	$\hat{\beta}_{**}$	$\hat{q}_{*}$	$\hat{\beta}_{*}$
3	100	0.0071	0.0466	0.0071	0.0463	0.0065	0.0447	-0.0075	0.0395
3	300	0.0005	0.0033	0.0005	0.0029	0.0005	0.0033	-0.0012	0.0028
3	500	0.0005	0.0015	0.0005	0.0016	0.0005	0.0015	-0.0010	-0.0014
5	100	0.0018	0.0140	0.0018	0.0322	0.0015	0.0132	-0.0013	0.0134
5	300	0.0005	0.0031	0.0005	0.0010	0.0005	0.0031	-0.0010	0.0024
5	500	0.0004	0.0013	0.0004	0.0013	0.0004	0.0013	-0.0009	-0.0013
7	100	0.0020	0.0116	0.0020	0.0135	0.0012	0.0083	-0.0010	0.0111
7	300	0.0002	0.0019	0.0010	0.0040	0.0002	0.0019	-0.0009	0.0021
7	500	0.0004	0.0012	0.0004	0.0016	0.0004	0.0012	-0.0008	0.0010
∞	100	0.0016	0.0113	0.0016	0.0115	0.0011	0.0079	-0.0010	-0.0083
∞	300	0.0010	0.0013	0.0010	0.0040	0.0010	0.0013	-0.0009	-0.0018
∞	500	0.0003	0.0010	0.0004	0.0015	0.0003	0.0010	-0.0007	0.0010

Tablo 5.2.4. DB dağılımı $(q = 0.5, \beta = 1.5)$ için tahmin edicilerinin yanları

		MLE		MPE		UMMLE		MMLE	
L	n	$\hat{q}$	$\hat{\beta}$	$\tilde{q}$	$\tilde{\beta}$	$\hat{q}_{**}$	$\hat{\beta}_{**}$	$\hat{q}_{*}$	$\hat{\beta}_{*}$
3	100	0.0023	0.0366	0.0013	0.0360	-0.0031	0.0320	-0.0027	0.0415
3	300	0.0007	0.0123	0.0007	0.0115	0.0007	0.0123	-0.0014	-0.0221
3	500	-0.0002	0.0091	-0.0002	0.0086	-0.0002	0.0091	0.0009	0.0182
5	100	0.0011	0.0362	0.0011	0.0354	0.0012	0.0319	-0.0015	-0.0302
5	300	0.0005	0.0137	0.0005	0.0117	0.0005	0.0137	-0.0013	0.0140
5	500	0.0005	0.0085	0.0005	0.0070	0.0005	0.0085	-0.0009	-0.0074
7	100	0.0006	0.0307	0.0006	0.0344	0.0010	0.0266	-0.0014	0.0220
7	300	0.0003	0.0118	0.0010	0.0102	0.0003	0.0118	-0.0009	0.0123
7	500	0.0002	0.0086	0.0006	0.0071	0.0002	0.0084	-0.0005	-0.0073
∞	100	0.0005	0.0284	0.0004	0.0310	0.0009	0.0211	-0.0012	0.0116
∞	300	-0.0002	0.0096	-0.0002	0.0063	-0.0002	0.0096	-0.0004	0.0071
∞	500	0.0002	0.0084	0.0002	0.0059	0.0002	0.0084	-0.0004	0.0069

Tablo 5.2.3 ve 5.2.4 de verilen sonuçlara göre, örneklem hacmi arttığında q ve β parametrelerinin ML, MML ve MP tahmin edicilerinin Tip-I sansürlü ve tam örneklem durumunda HKO ve yan değerleri azalmaktadır. Küçük hacimli Tip-I sansürlü örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO su, tam örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO suna göre daha büyüktür. Ancak büyük hacimli örneklem durumunda, Tip-I sansürlü ve tam örneklemden elde edilen parametre tahminlerinin yan ve HKO'su hemen hemen aynıdır. β parametresinin ML ve MML tahmin edicilerinin HKO ve yan değerleri, MP tahmin edicilerin HKO ve yan değerlerinden daha küçük olduğu görülmüştür. q parametresinin bütün tahmin edici yöntemlerinden elde edilen HKO ve yan değerleri hemen hemen aynıdır. Bir diğer sonuç ise L sansürleme zamanı ve örneklem hacmi arttığında HKO ve yan değerleri düşmektedir.

5.2.5. Gerçek Veri Uygulaması

Tablo 5.1.3 deki veriler bu bölümde kullanılacaktır. Bu veri üzerinden elde edilen tahminler Tablo 5.2.5 te verilmiştir.

Tablo 5.2.5. Veriye ait Kesikli Burr Dağılımının Parametre Tahmini

	Tam örneklem		$L = 2$		$L = 3$	
Yöntem	$\hat{q}$	$\hat{\beta}$	$\hat{q}$	$\hat{\beta}$	$\hat{q}$	$\hat{\beta}$
MLE	0.2312	1.4408	0.2322	1.4820	0.2314	1.4124
MPE	0.2290	1.0604	0.2290	1.0604	0.2290	1.0604
MMLE	0.2163	1.3240	0.2197	1.3529	0.2195	1.3125
UMMLE	0.2312	1.4408	0.2322	1.4820	0.2314	1.4124

5.3. Kesikli Pareto Dağılımı

Kesikli Pareto dağılımı, Pareto dağılımının kesikleştirilmesi sonucunda elde edilmiştir (Krishna ve Pundir, 2009) çalışmada dağılımın özellikleri ve tam örneklem durumunda parametre tahmini ML tahmin edici yöntemiyle elde edilmiştir. Tezin bu bölümünde Kesikli Pareto dağılımı için Tip-I sansürlü örneklem durumunda parametre tahminleri ML, MP ve MML yöntemleri ile elde edilecektir.

Kesikli Pareto dağılımının olasılık ve dağılım fonksiyonu sırasıyla

$$f\left(x\right)=q^{\log(1+x)}-q^{\log(x+2)},\quad x=0,1,2,\ldots$$

ve

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

elde edilmiştir. Burada $0<q<1$ dağılım parametresidir. Dağılım kısaca $DPD\left(q\right)$ şeklinde gösterilecektir. Bu bölümde Tip-I sansürlü örnekleme dayalı parametre konusu verilmiştir.

5.3.1. En Çok Olabilirlik Tahmini

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

$$X_i=\min\left(Y_i,L_i\right),\ i=1,\ldots,n$$

şeklinde tanımlanmak üzere $X_1,X_2,\ldots,X_n$ örneklemine Tip-I sansürlü örneklem olsun

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

gösterge fonksiyonu kullanılarak olabilirlik ve log-olabilirlik fonksiyonu sırasıyla

$$\begin{aligned}L\left(q\right)&=\prod_{i=1}^{n}f\left(x_i\right)^{\delta_i}\left(1-F\left(x_i\right)\right)^{1-\delta_i}\\&=\prod_{i=1}^{n}\left(q^{\log(1+x)}-q^{\log(x+2)}\right)^{\delta_i}\left(q^{\log(1+x)}\right)^{1-\delta_i}\end{aligned}\qquad(5.3.1)$$

ve

$$\begin{aligned}\ell(q) &= \log\left\{\prod_{i=1}^{n} f(x_i)^{\delta_i}\left(1-F(x_i)\right)^{1-\delta_i}\right\} \\ &= \sum_{i=1}^{n}\delta_i \log\left(f(x_i)\right)+\sum_{i=1}^{n}(1-\delta_i)\log\left(1-F(x_i)\right) \\ &= \sum_{i=1}^{n}\delta_i \log\left(q^{\log(1+x_i)}-q^{\log(x_i+2)}\right)+\sum_{i=1}^{n}(1-\delta_i)\log\left(q^{\log(1+x_i)}\right)\end{aligned} \tag{5.3.2}$$

şeklinde yazılabilir. Olabilirlik denklemi ise aşağıdaki gibi elde edilir.

$$\begin{aligned}\frac{\partial \ell(q,\beta)}{\partial q} &= \sum_{i=0}^{n}\frac{\delta_i}{q}\left(\frac{q^{\log(1+x_i)}\log(1+x_i)-q^{\log(1+x_i)}\log(x_i+2)}{q^{\log(1+x_i)}-q^{\log(x_i+2)}}\right) \\ &+\sum_{i=0}^{n}\frac{(1-\delta_i)\log(1+x_i)}{q}=0.\end{aligned} \tag{5.3.3}$$

(5.3.3) denklemi analitik olarak çözülemediğinden ML tahminlerini elde etmek için nümerik yöntemler kullanılır.

5.3.2. Oran Yöntemi Tahmini

Bu yöntem (Khan ve ark., 1989) tarafından Kesikli Weibull dağılımının parametrelerinin tahmininde kullanılmıştır. Bu yöntem $DPD(q)$ için de kullanılacaktır. $X_1, X_2, \ldots, X_n$ bağımsız ve $DPD(q)$ den alınmış Tip-I sansürlü örneklem olmak üzere, aşağıdaki gösterge fonksiyonu tanımlansın:

$$\nu_0(X_i)=\begin{cases}1, & X_i=0\\ 0, & X_i>0.\end{cases}$$

Bu durumda $Y=\frac{1}{n}\sum_{i=1}^{n}\nu_0(X_i)$, $f(0)=1-(q)^{\log(2)}$ olasılığının bir tahmin edicisidir. Y, $f(0)$ a eşitlenip q yalnız bırakıldığında, q parametresinin oranlar tahmin edicisi

$$\tilde{q}=(1-Y)^{1/\log(2)}$$

olarak elde edilir.

5.3.3. Uyarlanmış En Çok Olabilirlik Tahmini

Bu bölümde dağılımın parametrelerinin MML tahmin edicileri Tip-I sansürlü örnekleme dayalı olarak elde edilecektir. $X_1, X_2, \ldots, X_n$, $DPD(q)$ dağılımından alınmış Tip-I sansürlü örneklem olsun. Log-olabilirlik denklemlerinin q parametresine göre türevi $f_1(q,\beta)$, parametreye göre Taylor serisinin ilk iki terimi

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

ele alınır. Burada $\tilde{q}$, parametrenin MP edicisidir. Yukarıdaki denklem çözüldüğünde q parametresinin MML tahmin edicisi

$$\hat{q}_* = -\left. \frac{f(q) - \frac{\partial f(q)}{\partial q} q}{\frac{\partial f_1(q)}{\partial q}} \right|_{q=\tilde{q}} \tag{5.2.5}$$

olarak elde edilir.

5.3.4. Simülasyon Çalışması

Tezin bu bölümünde ML, MP, MML ve UMML tahmin edicilerinin Tip-I sansürlü örneklem durumunda, HKO ve yan kriterlerine göre karşılaştırılması verilmiş ve Simülasyon çalışmasında 10000 deneme yapılmıştır. Farklı parametreler ve örnek hacimleri için Simülasyon çalışması sonuçları Tablo 5.3.1-5.3.2 de verilmiştir.

Tablo 5.3.1. Tahmin Edicilerin HKO Tahmini

		MLE	MPE	UMMLE	MMLE	MLE	MPE	UMMLE	MMLE
L	n	$\hat{q}$	$\tilde{q}$	$\hat{q}_{**}$	$\hat{q}_*$	$\hat{q}$	$\tilde{q}$	$\hat{q}_{**}$	$\hat{q}_*$
		$q = 0.3$				$q = 0.5$			
3	50	0.0038	0.0065	0.0032	0.0515	0.0032	0.0065	0.0034	0.0615
3	100	0.0024	0.0024	0.0025	0.0315	0.0025	0.0033	0.0025	0.0574
3	200	0.0013	0.0015	0.0012	0.0310	0.0016	0.0017	0.0014	0.0328
3	300	0.0008	0.0008	0.0008	0.0008	0.0011	0.0011	0.0011	0.0028
3	500	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0010
5	50	0.0032	0.0056	0,0033	0,0105	0,0053	0,0062	0,0038	0,0095
5	100	0.0023	0.0023	0.0024	0.0026	0.0032	0.0032	0.0032	0.0055
5	200	0.0010	0.0011	0,0011	0,0012	0.0023	0.0022	0.0022	0.0027
5	300	0.0008	0.0008	0.0008	0.0008	0.0011	0.0011	0.0011	0.0010
5	500	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0010
7	50	0,0035	0,0061	0,0030	0,0050	0.0065	0.0063	0.0065	0.0098
7	100	0.0024	0.0024	0.0024	0.0027	0.0031	0.0031	0.0035	0.0052
7	200	0.0011	0.0014	0.0012	0.0014	0.0015	0.0015	0.0014	0.0025
7	300	0.0008	0.0008	0.0008	0.0007	0.0011	0.0011	0.0011	0.0012

7	500	0.0005	0.0005	0.0005	0.0004	0.0007	0.0007	0.0007	0.0009
∞	50	0.0033	0.0058	0,0029	0,0047	0.0062	0.0057	0.0061	0.0088
∞	100	0.0024	0.0024	0.0024	0.0024	0.0030	0.0031	0.0035	0.0044
∞	200	0.0010	0.0012	0.0010	0.0011	0.0013	0.0014	0.0013	0.0021
∞	300	0.0008	0.0008	0.0008	0.0006	0.0011	0.0011	0.0011	0.0009
∞	500	0.0005	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0008

Tablo 5.3.2. Tahmin Edicilerin yan tahmini

		MLE	MPE	UMMLE	MMLE	MLE	MPE	UMMLE	MMLE
L	n	$\hat{q}$	$\widetilde{q}$	$\hat{q}_{**}$	$\hat{q}_{*}$	$\hat{q}$	$\widetilde{q}$	$\hat{q}_{**}$	$\hat{q}_{*}$
		$q=0.3$				$q=0.5$			
3	50	-0.0256	0.0038	0.0168	0.0761	-0.0296	0.0038	0.0168	0.1761
3	100	-0.0231	0.0024	-0.0078	-0.0232	0.0287	0.0009	-0.0076	0.1096
3	200	-0.0236	0.0012	-0.0042	0.0162	-0.0279	0.0009	-0.0051	0.1041
3	300	0.0005	0.0005	0.0005	-0.0012	0.0007	0.0007	0.0007	-0.0014
3	500	0.0005	0.0005	0.0005	-0.0010	-0.0002	-0.0002	-0.0002	0.0009
5	50	0.0036	0,0025	0,0030	-0,0035	0,0040	0,0025	0,0030	-0.0084
5	100	0.0018	0.0018	0.0015	-0.0013	0.0011	0.0011	0.0012	-0.0015
5	200	0.0009	0,0012	0,0010	-0,0011	0,0007	0,0008	0,0007	-0.0014
5	300	0.0005	0.0005	0.0005	-0.0010	0.0005	0.0005	0.0005	-0.0013
5	500	0.0004	0.0004	0.0004	-0.0009	0.0005	0.0005	0.0005	-0.0009
7	50	0.0031	0.0028	0.0029	-0,0050	0.0015	0.0012	0.0021	-0.0065
7	100	0.0020	0.0020	0.0012	-0.0010	0.0006	0.0006	0.0010	-0.0014
7	200	0.0010	0.0015	0.0009	-0,0009	0.0005	0.0005	0.0005	-0.0011
7	300	0.0002	0.0010	0.0002	-0.0009	0.0003	0.0004	0.0003	-0.0009
7	500	0.0004	0.0004	0.0004	-0.0008	0.0002	0.0006	0.0002	-0.0005
∞	50	0.0025	0.0030	0.0024	-0.0012	0,0009	0.0012	0.0010	-0.0015
∞	100	0.0016	0.0016	0.0011	-0.0010	0.0005	0.0004	0.0009	-0.0012
∞	200	0.0012	0.0013	0.0011	-0.0010	0,0004	0,0003	0.0004	-0.0008
∞	300	0.0010	0.0010	0.0010	-0.0009	0.0002	0.0002	0.0002	-0.0004
∞	500	0.0003	0.0004	0.0003	-0.0007	0.0002	0.0002	0.0002	-0.0004

Tablo 5.3.1 ve 5.3.2 de verilen sonuçlara göre, örneklem hacmi arttığında q ve β parametrelerinin ML, MML ve MP tahmin edicilerinin Tip-I sansürlü ve tam örneklem durumunda HKO ve yan değerleri azalmaktadır. Küçük hacimli Tip-I sansürlü örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO su, tam örneklem durumunda elde edilen parametre tahminlerinin yan ve HKO suna göre daha büyüktür. Ancak büyük hacimli örneklem durumunda, Tip-I sansürlü ve tam örneklemden elde edilen parametre tahminlerinin yan ve HKO'su hemen hemen aynıdır. β parametresinin ML ve MML tahmin edicilerinin HKO ve yan değerleri, MP tahmin edicilerin HKO ve yan değerlerinden daha küçük olduğu görülmüştür. q parametresinin bütün tahmin edici yöntemlerinden elde edilen HKO ve yan değerleri hemen hemen

aynıdır. Bir diğer sonuç ise L sansürleme zamanı ve örneklem hacmi arttığında HKO ve yan değerleri düşmektedir.

5.3.5. Gerçek Veri Uygulaması

Tablo 5.1.3 deki veriler bu bölümde kullanılacaktır. Bu veri üzerinden elde edilen tahminler Tablo 5.3.3 te verilmiştir.

Tablo 5.3.3. Veriye ait Kesikli Pareto Dağılımının Parametre Tahmini

	Tam örneklem	$L=2$	$L=3$
Yöntem	$\hat{q}$	$\hat{q}$	$\hat{q}$
MLE	0.1842	0.2117	0.1931
MPE	0.1824	0.2112	0.1912
MMLE	0.1798	0.2009	0.1875
UMMLE	0.1842	0.2117	0.1931

6. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, literatürde kullanılan dağılım birleştirme teknikleri kullanılarak iki yeni kesikli dağılım önerilmiş ve özellikleri incelenmiştir. Elde edilen dağılımların parametre tahmini için ML, MML, MP, LS ve MM tahmin edicileri tartışılmış, parametrelerin güven aralıkları ve bunların kapsama olasılıkları Simülasyon çalışmalarıyla incelenmiştir. Gerçek veri uygulamalarında verilerin geldiği dağılımı modellerken yeni önerilen kesikli dağılımların diğer literatürde olan kesikli dağılımlarla kıyaslanmasında log-olabilirlik, AIC, BIC, Ki-Kare istatistiği gibi kriterlere bakılmış ve önerilen yeni dağılımların, gerçek verileri diğer literatürdeki dağılımlardan daha iyi modellediği görülmüştür. Ayrıca literatüre yeni girmiş bazı kesikli dağılımlarının (Kesikli Chen dağılımı, Kesikli Burr dağılımı ve Kesikli Pareto dağılımı) parametrelerinin Tip-I sansürlü örneklem durumunda parametre tahmini elde edilmiştir. Elde edilen tahmin edicilerin HKO ve yan değerleri Monte Carlo simülasyonu kullanılarak kıyaslanmıştır. Ayrıca gerçek veri uygulaması, verilen her dağılım için yapılmıştır.

Tez çalışmasının, 3. ve 4. Bölümünde tartışılan birleştirme teknikleri kullanılarak yeni kesikli dağılımlar önerilebilir. Önerilen iki yeni kesikli dağılım için Tip-I sansürlü örneklem durumunda parametre tahmini incelenebilir. Önerilen yeni dağılımların parametrelerinin Bayes tahmin edicileri Thirney-Kadene ve Lindley yaklaşımları kullanılarak elde edilebilir.

Kaynaklar

Akdi, Y., 2005, Matematiksel İstatistiğe Giriş, *Ankara*, Bıçaklar Kitapevi, p.

Bagheri, S. F. E., Bahrami, S. ve Ganjali, M., 2016, The generalized modified Weibull power series distribution: Theory and applications, *Computational Statistics & Data Analysis*, 94, 136-160.

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

Barreto-Souza, W., Morais, A. L. ve Cordeiro, G. M., 2011, The Weibull-geometric distribution, *Journal of Statistical Computation and Simulation*, 81 (5), 645-657.

Bidram, H., Javad, B. ve Mina, T., 2013, A new generalized exponential geometric distribution, *Communications in Statistics-Theory and Methods*, 42 (3), 528-542.

Braden, B., 1992, Calculating sums of infinite series, *The American Mathematical Monthly*, 7, 649-655.

Casella, G. ve Berger, R. L., 1990, Statistical Inference, *California*, Duxbury Press, p.

Chakraborty, S. ve Chakravarty, D., 2012, Discrete gamma distributions: properties and parameter estimation., *Communications in Statistics-Theory and Methods*, 41, 3301–3324.

Chakraborty, S. ve Chakravarty, D., 2014, A Discrete Gumbel Distribution, *arXiv:1410.7568.*

Delgarm, L. ve Mohammad, R., Z., 2015, A new generalization of lifetime distributions, *Computational Statistics & Data Analysis*, 30 (4), 1185-1198.

Déniz, E. G., 2013, A new discrete distribution: Properties and applications in medical care. , *Journal of Applied Statistics*, 40 (12), 2760-2770.

Gradshteyn, I. S. ve Ryzhik, I. M., 2007, Table of Integrals, Series, and Products, *San Diego, CA.*, Academic Press, p.

Hu, Y., Peng, X., Li, T. ve Guo, H., 2007, On the Poisson approximation to photon distribution for faint lasers, *Phys. Lett*, 367, 173-176.

Hussain, T. ve Ahmad, M., 2014, Discrete inverse Rayleigh distribution, *Pakistan Journal of Statistics*, 30, 230-222.

Inusah, S. ve Kozubowski, T. J., 2006, A discrete analogue of the Laplace distribution, *Journal of Statistical Planning and Inference*, 136, 1090-1102.

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

Keilson, J. ve Gerber, H., 1971, Some results for discrete unimodality, *Journal of the American Statistical Association*, 66, 386-389.

Khan, M., S.,A.,, Khalique, A. ve Abouammoh, A. M., 1989, On estimating parameters in a discrete Weibull distribution, *IEEE Transactions on Reliability*, 38, 348-350.

Krishna, H. ve Pundir, P. S., 2009, Discrete Burr and discrete Pareto distributions., *Statistical Methodology*, 6, 177-188.

Kulasekera, K. B., 1994, Approximate mles of the parameters of a discrete Weibull distribution with type-I censored-data, *Microelectron. Reliab.*, 34, 1185-1188.

Kuş, C., 2007, A new lifetime distribution, *Computational Statistics & Data Analysis*, 51 (9), 4497-4509.

Leemis, L., M., 1986, Lifetime Distribution identities, *IEEE Trans. Reliability*, 35, 170-174.

Lu, W. ve Daimin, S., 2012, A new compounding life distribution: the Weibull–Poisson distribution, *Journal of Applied Statistics*, 39 (1), 21-38.

Mahmoudi, E. ve Mitra, S., 2012, Exponentiated Weibull-geometric distribution and its applications, *arXiv preprint arXiv:1206.4008*.

Mahmoudi, E. ve Afsaneh, S., 2013, Exponentiated Weibull–Poisson distribution: Model, properties and applications, *Mathematics and Computers in Simulation*, 92, 76-97.

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

Nakagawa, T. ve Osaki, S., 1975, The discrete Weibull distribution, *IEEE Transactions on Reliability*, 24, 300-301.

Nekoukhou, V., Alamatsaz, M. H. ve Bidram, H., 2012, A discrete analog of the generalized exponential distribution, *Communications in Statistics-Theory and Methods*, 41, 2000–2013.

Nekoukhou, V., Alamatsaz, M. H., Bidram, H. ve Aghajani, A. M., 2013a, Discrete beta-exponential distribution *Communications in Statistics-Theory and Methods*, DOI: 10.1080/03610926.2013.773348.

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

Nekoukhou, V. ve Bidram, H., 2015, The exponentiated discrete Weibull distribution, *SORT*, 39, 127-146.

Noughabi, M., S.,, Roknabadi, A., H.,R., ve Borzadaran, G., R.,M., 2013, Some Discrete Lifetime Distributions with Bathtub-Shaped Hazard Rate Functions, *Quality Engineering*, 25, 225-236.

Noughabi, M. S., Roknabadi, A. H. R. ve Borzadaran, G. M., 2011, Discrete modified Weibull distribution, *Metron*, LXIX, 207–222.

Percontini, A., Betsabé, B. ve Gauss, M. C., 2013, The beta Weibull Poisson distribution, *Chilean Journal of statistics*, 4 (2), 3-26.

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

Rezaei, S., Saralees, N. ve Nahid, T., 2013, A new three-parameter lifetime distribution, *Statistics*, 47 (4), 835-860.

Ristić, M. M. ve Saralees, N., 2014, A new lifetime distribution, *Journal of Statistical Computation and Simulation*, 84 (1), 135-150.

Roussas, G. G., 1973, A First Course in Mathematical Statistics, *U.S.A.*, Addison-Wesley Publishing Company, p.

Roy, D. ve P., G. R., 1992, Classifications of discrete lives, *Microelectronics and Reliability*, 32 (10), 1450-1459.

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

Roy, D., 2004, Discrete Rayleigh distribution, *IEEE Trans. Reliability*, 53, 255-260.

Stein, W. E. ve Dattero, R., 1984, A new discrete Weibull distribution, *IEEE Transactions on Reliability*, R-33, 196-197.

Tiku, M. ve Akkaya, A., 2004, Robust Estimation and Hypothesis Testing, *New Delhi*, New Age International, p.

Torabi, H. ve Narges, M. H., 2012, The gamma-uniform distribution and its applications, *Kybernetika*, 48 (1), 16-30.

Wang, M. ve Ibrahim, E., 2015, The modified Weibull geometric distribution, *Metron*, 73 (3), 303-315.

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

Wu, S.-J. ve Kuş, C., 2009, On the Estimation Based on Progressive First Failure-Censored Sample, *Computational Statistics and Data Analysis*, 53 (10), 3659-3670.

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.

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

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Yunus AKDOĞAN
Uyruğu : T.C
Doğum Yeri ve Tarihi : Nizip / 14.08.1986
Telefon : 05382029587
Faks :
e-mail : yakdogan@selcuk.edu.tr

EĞİTİM

Derece		Adı, İlçe, İl	Bitirme Yılı
Lise	:	**Gaziantep Lisesi, Şahinbey, Gaziantep**	**2005**
Üniversite	:	**Selçuk Üniversitesi, Selçuklu, Konya**	**2009**
Yüksek Lisans	:	**Selçuk Üniversitesi, Selçuklu, Konya**	**2011**
Doktora	:	**Selçuk Üniversitesi, Selçuklu, Konya**	**2011-**

İŞ DENEYİMLERİ

Yıl	Kurum	Görevi
2010	**Selçuk Üniversitesi**	**Araştırma Görevlisi**

UZMANLIK ALANI

İstatistik teorisi
Uygulamalı İstatistik

YABANCI DİLLER

İngilizce

YAYINLAR*

A-2	**SCI, SCI-expanded, SSCI, AHCI'de yer alan ve Impact faktörü 0.1 - 0.9 arası olan uluslararası dergilerde yayımlanan araştırma makalesi**
1	Kuş C., **Akdoğan Y.,** Wu S.J., 2013. Optimal progressive group censoring scheme under cost considerations for pareto distribution, Journal of Applied Statistics, Vol. 40, No. 11, 2437–2450.(5 years Impact Factor=0,656) (Bu yayın B-1.1-12' te verilen bildirinin yayına dönüşmüş halidir.) [Yüksek Lisans Tezinden Elde Edilen Makale]
2	**Akdoğan Y.**, Kuş C., Wu S.J., 2015. Optimal Progressive Group-Censoring Plans for Chen Distribution under Cost Constraint., Pakistan Journal of Statistics, Vol. 31, No.1, 107-120. (Impact Factor=0,336) (Bu yayın B-1.1-12' te verilen bildirinin yayına dönüşmüş halidir.) [Yüksek Lisans Tezinden Elde Edilen Makale]
3	**Akdoğan Y.,** Kuş C., Asgharzadeh A., Kınacı I., S. Fatemeh., 2016., Uniform-Geometric distribution., Journal of Statistical Computation and Simulation, Vol. 86(9), 1754-1770. [DoktoraTezinden Elde Edilen Makale]

A-3	**Diğer uluslararası hakemli dergilerdeki yayınlar**
1	**Akdoğan Y.**, Kuş C., Wu S.J., 2012. Planning life tests for Burr XII distributed products under progressive group-censoring with cost considerations, Gazi University Journal of Science, Vol 25, no 1, (Bu yayın B-1.1-14' te verilen bildirinin yayına dönüşmüş halidir.) [Yüksek Lisans Tezinden Elde Edilen Makale]
2	**Akdoğan Y.**, Kuş C., Kınacı İ., 2015. Point Estimation of Parameters in Discrete Burr Distribution based on Type I censored sample, Journal of the Turkish Statistical Association, Vol 8, No. 1, 112-123. [DoktoraTezinden Elde Edilen Makale]
3	Kınacı İ., **Akdoğan Y.**, Kuş C., Ng H.K.T., 2014. Statistical Inference for Weibull Distribution Based on a Modified Progressive Type-II Censoring Scheme, Sri Lankan Journal of Applied Statistics. (Special Issue, 2014), 95-116.
4	Kınacı İ., Kuş C., Karakaya K., **Akdoğan Y.**, On the number of l-overlapping success runs of length k-under q-sequence of binary trials, Journal of the Turkish Statistical Association, (Accepted)
5	Kınacı İ., Karakaya K., **Akdoğan Y.**, Kuş C., 2016. Bayesian Estimation for Discrete Chen Distribution, Selçuk Üniversitesi Fen Fakültesi Fen Dergisi, (Accepted)
6	Kınacı İ., Karakaya K., **Akdoğan Y.,** Kuş C., 2016. Interval Estimation Based on Progressively Censored Data. Selçuk Üniversitesi Fen Fakültesi Fen Dergisi, (Accepted)
B. Uluslararası Bildiriler	
B-1	**Yabancı dilde sözlü sunulan ve tam metni yayınlananlar**
1	
B-1.1	**Yabancı dilde sözlü sunulan ve özeti yayınlananlar**
1	**Akdoğan Y.,** Kuş C., Asgharzadeh A., Kınacı I., S. Fatemeh., 2016., 7th European Congress of Mathematics. Berlin, July 18-22. pp. 67. [Doktora Tezinden Elde Edilen Bildiri]
2	Karakaya K., Kınacı İ., Kuş C., **Akdoğan Y.,** 2016. A New Familiy of Distributions 2th International Research Science Young Statistician Conference, Ankara,Turkey, May 4-8. pp. 28.
3	Kınacı İ., Kuş C., Karakaya K., **Akdoğan Y.,** 2016. APT-Pareto Distribution and its Properties, 2th International Researchers Science Young Statistician Conference, Ankara,Turkey, May 4-8. pp. 56.
4	Akdoğan Y., Ng T., **Kuş C.**, Kınacı İ., 2014. A Simulation Study to Compare the Estimators for Several Discrete Distributions Under Type I Censoring, 20th Conference of the International Federation of Operational Research Societies. Barcelona, July 13-18. pp. 131. [Doktora Tezinden Elde Edilen Bildiri]
5	Karakaya K., Akdoğan Y., Kınacı İ., **Kuş C.**, 2014. Bayesian Estimation for Discrete Chen Distribution, 9th International Statistics Day Symposium ISDS 2014 (İGS 2014), May 10-14, Antalya, pp. 1.
6	Akdoğan Y., **Kuş C.**, Kınacı İ., 2014. Statistical Inference for Several Discrete Distribution based on Type I Censored Sample, 9th International Statistics Day Symposium ISDS 2014 (İGS 2014), May 10-14, Antalya, pp. 43.
7	Şahin F., Akdoğan Y., **Kuş C.**, Kınacı İ., 2014. On the Kolmogorov-Smirnov Test for Normality, 9th International Statistics Day Symposium ISDS 2014 (İGS 2014), May 10-14, Antalya, pp. 100.
8	Akdogan Y., Çalık A., Altındağ İ., **Kuş C**., Kınacı İ., 2013. Interval Estimation for some Life Distributions Based on Progressively Censored Sample, XXIX-th European Meeting of Statisticians, Budapest, pp. 29.
9	Çalık A., Altındağ İ., **Kuş C.,** Akdoğan Y., Kınacı İ., 2013. Parameter Estimation for EEG Distribution based on Progressively Censored Sample, XXIX-th European Meeting of Statisticians, 2013, Budapest, pp. 27.
10	Altındağ, İ., Akdoğan, Y., Çalık, A., **Kuş C.**, Kınacı, İ., 2013. On the Uniqueness of MLEs based on Censored Data for Some Life time Distributions, XXIX-th European Meeting of Statisticians, 2013, Budapest, pp. 30-31.
11	**Kuş C.**, Akdoğan Y., Wu, S-J., 2012. Optimal Progressive Group Censoring Scheme under Cost Considerations for Pareto Distribution, 25th European Conference on Operational Research, 8-11 July, Vilnius, Lithuania., pp. 248. [Yüksek Lisans Tezinden Elde Edilen Bildiri]

12	Akdoğan Y., Oral E., Körez M. K., **Kuş C.**, 2012. Simulation Study for the Comparison of Maximum Likelihood and Modified Maximum Likelihood Estimators in Logistic Regression, 25th European Conference on Operational Research, 8-11 July, Vilnius, Lithuania., pp. 227.
13	Karadayı N., Akdoğan Y., **Kuş C.**, Wu, S-J., 2012. Optimal Warranty Length For a Pareto Distributed Product Wıth Progressive Censorıng, 25th European Conference on Operational Research, 8-11 July, Vilnius, Lithuania., pp. 247.
14	Körez M.K., Oral E., Akdoğan Y., **Kuş C.**, 2012. Simulation Study for Modified Maximum Likelihood Estimations in Poisson Regression, 25th European Conference on Operational Research, 8-11 July, Vilnius, Lithuania., pp. 206.
15	**Kuş C.**, Akdoğan Y., 2011. Kısıtlı Maliyet Durumunda Bathtube-Shape Dağılımı için Optimal İlerleyen Tür Grup Sansürleme Planı, Uluslararası 7. İstatistik Kongresi, Antalya, 28 Nisan-01 Mayıs 2011., Sayı 4, 2011, No:1, pp. 214-215. [Yüksek Lisans Tezinden Elde Edilen Bildiri]
16	Akdoğan Y., **Kuş C.**, 2011. Kısıtlı Maliyet Durumunda Burr Tip III Dağılımı için Optimal İlerleyen Tür Grup Sansürleme Planı, Uluslararası 7. İstatistik Kongresi, Antalya, 28 Nisan-01 Mayıs, Sayı 4, No:1, pp. 210-211. [Yüksek Lisans Tezinden Elde Edilen Bildiri]
17	**Kuş C.**, Akdoğan Y., Wu S-J., 2011. Optimal progressive group-censoring plans for Burr XII distribution under cost constraint, 5th Annual International Conference on Mathematics, Statistics & Mathematical Education, 13-16 June Athens, Greece. pp. 45. [Yüksek Lisans Tezinden Elde Edilen Bildiri]
18	Akdoğan Y., **Kuş C.**, Wu S-J., 2011. Optimal progressive group-censoring plans for Logistic distribution under cost constraint, 5th Annual International Conference on Mathematics, Statistics & Mathematical Education, 13-16 June, Athens, Greece. pp. 30. [Yüksek Lisans Tezinden Elde Edilen Bildiri]
D. Ulusal Çalışmalar	
E. Ulusal Bildiriler	
E-1	**Sözlü sunulan ve tam metin yayınlananlar**
1	**Kuş C.**, Akdoğan Y., 2010. Burr XII Dağılımının Parametrelerinin İlerleyen Tür İlk Bozulma Sansürlemeye Dayalı Güven Aralıkları ve Güven bölgeleri, Orta Doğu Teknik Üniversitesi İGS 2010 7. İstatistik Günleri Sempozyumu, Ankara, s.78. [Doçentlik sonrası bildiri]
2	Yaman Ü., Akdoğan Y., Pekgör A., **Kuş C.**, 2010. Ölçüm yeterlilik Parametrelerinin Tahmin edicileri ve Güven Aralıkları, Orta Doğu Teknik Üniversitesi İGS 2010 7. İstatistik Günleri Sempozyumu, Ankara, s.78. [Doçentlik sonrası bildiri]
E-1.1	**Sözlü sunulan ve özeti yayınlananlar**
1	Akdoğan Y., **Kuş C.,** 2011. Likert Tipi Verilere Varyans Analizi Uygulanabilir mi?, III. Ulusal Konya Ereğli Kemal Akman Meslek Yüksek Okulu Tebliğ Günleri, Sayı 3, No:1, 374-387.
2	Akdoğan Y., Pekgör A., Karakoca A., **Kuş C.**, 2011. Bir Örneklem Hotelling T^2 Testi ve Sağlamlığı Üzerine Bir Çalışma, III. Ulusal Konya Ereğli Kemal Akman Meslek Yüksek Okulu Tebliğ Günleri, Sayı 3, No:1, 374-387.
3	Akdoğan Y., **Kuş C.**, 2010. İlk Bozulma Sansürlü Örnekleme Planına Dayalı Burr XII Dağılımının Parametrelerinin Tahmini, Güven Aralıkları, Güven bölgeleri ve Beklenen Test Süresi, II. Ulusal Konya Ereğli Kemal Akman Meslek Yüksek Okulu Tebliğ Günleri, Sayı 2, No:1, 374-387.