ISSN 2757-5543 GÜFFD 5. Cilt (2): 108-121 (2024)

Gazi Üniversitesi Fen Fakültesi Dergisi

https://dergipark.org.tr/tr/pub/guffd

GÜFFD GAZİ ÜNİVERSİTESİ FEN FAKÜLTESİ DERGİSİ

Hisse Senedi Fiyatlarinin Benford Kanunu ve Uç Değer Dağılımlar ile İncelenmesi

Hatice Nur Karakavak*, Gamze Özel

Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06800, Ankara, Türkiye

Öne Çıkanlar

• Bu çalışmada hisse senedi fiyatlarının analizinde Benford Kanunu ve uç değer dağılımları kullanılarak, hisse senedi işlem hacimlerinin zamansal değişimleri incelenmiştir.

• Çalışmanın sonuçları, hisse senedi işlem hacimlerinin anormal hareketlerini ortaya koyarak, bu yöntemlerin finansal analizlerdeki etkinliğini göstermiştir.

Makale Bilgileri

Geliş: 15/05/2024
Kabul: 23/07/2024

Anahtar Kelimeler

Benford Kanunu,
Uç Değer Dağılım,
Blok Maksima,
Finans

Öz

Bu çalışmada hisse senedi fiyatlarının analizinde Benford Kanunu ve uç değer dağılımları kullanılarak, hisse senedi işlem hacimlerinin zamansal değişimleri incelenmiştir. Benford Kanunu'nun veri setlerindeki ilk rakamların frekansını tahmin ederek anormallikleri saptama kabiliyeti, finansal verilerin doğruluğunu test etmede kullanılmıştır. Ayrıca, uç değer teorisiyle finansal piyasalardaki büyük dalgalanmalar ve nadir olaylar analiz edilerek risk yönetimi ve finansal modelleme için önemli bulgular sunulmuştur. Çalışmanın sonuçları, hisse senedi işlem hacimlerinin anormal hareketlerini ortaya koyarak, bu yöntemlerin finansal analizlerdeki etkinliğini göstermiştir.

Investigation of Stock Prices with Benford's Law and Extreme Value Distributions Highlights

Highlights

• In this study, the temporal changes in stock transaction volumes were examined using Benford's Law and extreme value distributions to analyze stock prices.

• The results of the study have highlighted the abnormal movements in stock transaction volumes, demonstrating the effectiveness of these methods in financial analyses.

Article Info

Received: 15/05/2024
Accepted: 23/07/2024

Keywords

Benford's Law,
Extreme Value,
Distribution,
Block Maxima,
Finance

Abstract

In this study, the temporal changes in stock transaction volumes were examined using Benford's Law and extreme value distributions to analyze stock prices. Benford's Law's ability to predict the frequency of the first digits in datasets and detect anomalies has been utilized to test the accuracy of financial data. Furthermore, extreme value theory was applied to analyze significant fluctuations and rare events in financial markets, providing important findings for risk management and financial modeling. The results of the study have highlighted the abnormal movements in stock transaction volumes, demonstrating the effectiveness of these methods in financial analyses.

Araştırma Makalesi

* Sorumlu Yazar/Corresponding Author: Hatice Nur Karakavak, hatice.karakavak@hacettepe.edu.tr

1. GİRİŞ

Benford Kanunu, büyük veri kümeleri üzerinde anlamlı bir analiz aracı olarak kabul edilmektedir. Bu kanun, veri setlerindeki ilk rakamların dağılımını açıklamak için kullanılmaktadır. Genel olarak, Benford Kanunu, büyük ve gerçek dünya veri setlerinde ilk rakamların dağılımını tanımlayan bir olasılık dağılımıdır. Bu kanun, verilerdeki ilk rakamların 1'den 9'a doğru azalan bir frekansta olacağını öngörür; yani, 1 rakamı en sık görülen ilk rakamken, 9 en seyrek olanıdır. Modern teknolojinin olanaklarıyla birleştiğinde, Benford Kanunu, büyük ve karmaşık veri kümelerinde hileli işlemleri tespit etmede etkili bir yöntem hâline gelmiştir.

Benford Kanunu, çeşitli alanlarda veri setlerinin doğruluğunu değerlendirme ve anormallikleri saptama aracı olarak kullanılır. Demografik veriler ve nüfus sayımlarını incelemek için kullanılır [1]. Seçim verilerinde oy sayılarının ilk rakamlarının Benford dağılımına uyması beklenir; sapmalar manipülasyon belirtisi olabilir [2]. Bilimsel verilerde, özellikle çevre bilimleri ve astrofizikte, veri manipülasyonu tespitinde kullanılmıştır [3]. Sağlık bilimlerinde, epidemiolojik verilerdeki hataları belirlemek amacıyla uygulanmıştır [2]. İnternet trafiği ve sosyal medya analizlerinde, bot trafiğini veya anormal aktiviteleri belirlemek için de yararlanılmıştır [4]. Çevresel bilimlerde, iklim değişikliği verileri veya çevresel kirlilik düzeyleri gibi çeşitli veri setlerinin analizinde Benford Kanunu kullanılmıştır [5]. Özkan [6], Benford Kanunu'nu kullanarak ekosistemlerin doğallığını ölçmeyi amaçlamıştır. Çalışmasında Benford Kanunu, ekosistem verilerindeki sayısal desenleri analiz ederek, insan müdahalesi veya doğal olmayan değişikliklerin belirlenmesine yardımcı olmuştur. Bu geniş uygulama yelpazesi, Benford Kanunu'nun çeşitli alanlar için bir analitik araç olduğunu göstermektedir.

Benford Kanunu, finans üzerine yapılan çalışmalarda da önemli bir yere sahiptir. Denetim süreçlerinin planlanması aşamasında da önemli bir rol oynar. Denetçilere, hile veya hata içerme olasılığı olan işlem ve kalemleri ortaya çıkararak rehberlik eder. Muhasebe verileri, beklenen dağılımlara uyması beklenir. Benford Kanunu'na uygunluğun analizi, veri setlerinin bu teorik dağılımı ne ölçüde takip ettiğini belirleyerek herhangi bir anormalliği veya sapmayı ortaya çıkarmak için etkili bir yöntem olarak kullanılabilir. Bu analiz, satışlar, satış fiyatları, birim fiyatlar, stok girişleri, giderler, alacak hesapları ve satın alma işlemleri gibi muhasebe kalemleri üzerinde yapılabilir [7]. Ekonomik veriler üzerinde Benford Kanunu'nun nasıl kullanılabileceğini inceleyen bu çalışma, kanunun ekonomi alanında nasıl bir etki yaratabileceğini ortaya koymuştur. Nigrini ve Mittermaier [5] veri denetimi ve adli analizlerde Benford Kanunu'nun rolünü ve etkilerini araştırmıştır. Durtschi vd. [1] tarafından muhasebe verilerinde hile tespiti için Benford Kanunu'nun nasıl kullanılabileceğini ayrıntılı bir şekilde incelenmiştir.

Benford Kanunu'nun etkin kullanımı, veri setinin özelliklerine bağlıdır. Özellikle muhasebe departmanlarında oluşturulan yapay sayı dizileri gibi düzensizlikler bu yöntemle tespit edilebilir. Ancak, yüksek kazanç sağlayan hesaplardan yapılan zimmete geçirmeler gibi bazı durumlar bu yöntemle tespit edilemeyebilir. Etkili bir Benford analizi için, veri seti homojen birimlerden oluşmalı ve belirli alt veya üst sınırlara sahip olmamalıdır. Ayrıca, veri seti çok küçükse Benford Kanunu uygulanmayabilir. Benford Kanunu'nun uygulanamayacağı durumlar arasında tekdüze dağılım gösteren veriler, rastgele oluşturulan sayılar, kişisel düşüncelerle etkilenen sayılar, veri setindeki maksimum ve minimum değerler, kesik noktaları olan sayılar ve hatalı verilerin bol olduğu durumlar yer almaktadır. Ayrıca, eğer veri sayısı çok azsa, bu kanunun uygulanması uygun olmayabilir.

Uç değer dağılımı, Benford Kanunu'nun uygun olmadığı durumlar için alternatif bir yöntem olarak kullanılabilir. Özellikle finansal verilerdeki büyük dalgalanmalar ve anormallikler, uç değer teorisinin önemli bir uygulama alanıdır [8]. Finansal piyasalarda beklenmedik büyük çıkışlar veya düşüşler, risk yönetimi ve finansal modellemeye önemli katkılar sağlar. Sigorta şirketleri ve yatırım bankaları, bu teoriyi kullanarak büyük kayıpları veya kazançları öngörmekte ve finansal krizleri daha etkili bir şekilde yönetebilmektedir [9]. Bu yöntem, veri setlerindeki uç noktaları inceleyerek sıra dışı durumları belirlemenin yanı sıra, risk değerlendirmelerini iyileştirmeye yardımcı olur. Beirlant vd. [10], uç değerleri analiz etmek için pratik yöntemler sunarak, finansal dâhil çeşitli gerçek dünya verilerine uygulanabilecek teknikleri tartışmaktadır.

Embrechts vd. [8], finans ve sigortacılıkta aşırı olayları modelleme üzerine kapsamlı bir çalışma yaparak aşırı risk ve kuyruk bağımlılıkları için istatistiksel yöntemleri incelemiştir. Longin [11], finansal piyasalarda nadir ve büyük fiyat hareketlerinin risklerini anlamak ve ölçmek için uç değer teorisini kullanmanın önemini tartışarak, daha sağlam risk değerlendirme araçlarına olan ihtiyacı savunmuştur. Reiss ve Thomas [12], uç değer analizinin istatistiksel yöntemlerini araştırarak çevresel ve finansal verilere uygulamalarıyla aşırı olayların uygun şekilde sınıflandırılması ve tahmin edilmesini vurgulamıştır. McNeil vd. [9], kantitatif risk yönetimi ve piyasa, kredi ve operasyonel riskleri modellemek için uç değer teorisinden yararlanmıştır.

Bu çalışmada finansal piyasalarda işlem gören hisse senetlerinin fiyat davranışlarının analizi için Benford Kanunu ve uç değer dağılım teorisinden yararlanılmıştır. Bölüm 2'de ilk olarak Benford Kanunu ve daha sonra uç değer dağılım teorisi açıklanmıştır. Bölüm 3'te elde edilen bulgular üzerinde durulmuş ve sonuçlar yorumlanmıştır. Bölüm 4'te ise, tartışma ve önerilere yer verilmiştir.

2. YÖNTEM

2.1. Benford Kanunu

Benford Kanunu, 1881'de astronom ve matematikçi Simon Newcomb [13] tarafından keşfedilmiş ve "American Journal of Mathematics" dergisinde yayımlanan iki sayfalık bir makale ile literatüre girmiştir. Newcomb, o dönemlerde hesap makinelerinin olmaması nedeniyle hesaplamaların logaritmik çizelgeler kullanılarak yapıldığını gözlemlemiştir. Bu çizelgelerin ilk sayfalarının, son sayfalara göre daha kirli ve yıpranmış olduğunu fark etmiştir. Bu gözlem, Newcomb'un 1 ile başlayan sayıların, 2 ile başlayanlara göre; 2 ile başlayanların ise 3 ile başlayanlara göre daha sık kullanıldığını tespit etmesine yol açmıştır. Newcomb bu tespitleri matematiksel bir formülle ifade ederek, sayıların baş rakamlarının görülme sıklığının logaritmik bir dağılım izlediğini ortaya koymuştur. İlk basamağın d'ye eşit olma olasılığı:

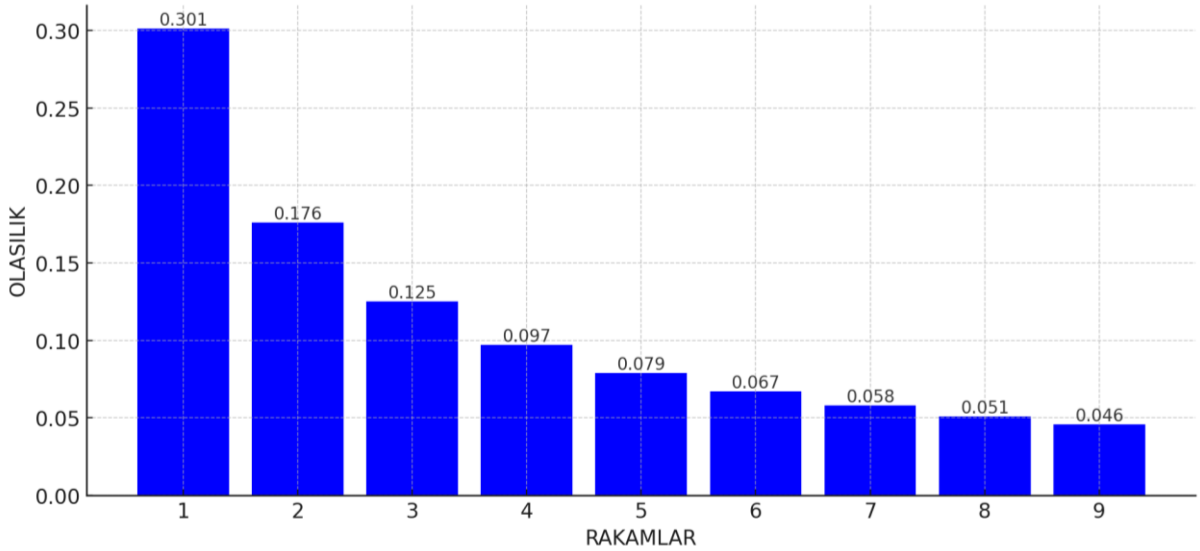
$$P(d) = \log_{10}\big(1 + (1/d)\big),\ d = 1, 2, \ldots, 9.$$

Newcomb'un çalışmasında yer alan, rakamların bir sayının ilk ve ikinci basamaklarında görülme olasılıkları Çizelge 1 olarak sunulmuştur.

***Çizelge 1.** Rakamların birinci ve ikinci basamakta bulunma olasılıkları*

Rakam	İlk Basamak İçin Beklenen Oran	İkinci Basamak İçin Beklenen Oran
0		0,1197
1	0,3010	0,1139
2	0,1761	0,1088
3	0,1249	0,1043
4	0,0969	0,1003
5	0,0792	0,0967
6	0,0669	0,0934
7	0,0580	0,0904
8	0,0512	0,0876
9	0,0458	0,0850

Benford [14], General Electric'teki laboratuvarında yapılan bir çalışmada 20229 veri seti incelenmiş ve bulgulara göre, sayıların ilk basamağının '1' olma olasılığı %30.6 olarak tespit edilmiştir. Bu sonuç, sayıların beklenen eşit dağılımının aksine, logaritmik bir dağılıma sahip olduğunu göstermektedir.



Şekil 1. *Benford Kanunu'na göre ilk basamaktaki rakamların olasılık dağılımı*

Benford Kanunu'na göre rakamların ortaya çıkma olasılıkları Şekil 1'de gösterilmiştir ve sayıların ilk basamağı için olasılık dağılımları birinci, ikinci ve ilk iki basamak için sırasıyla aşağıdaki gibi verilmiştir:

$$P(d_1) = \log\big(1 + (1/d_1)\big),\ d_1 \in \{1,2, \dots ,9\}$$

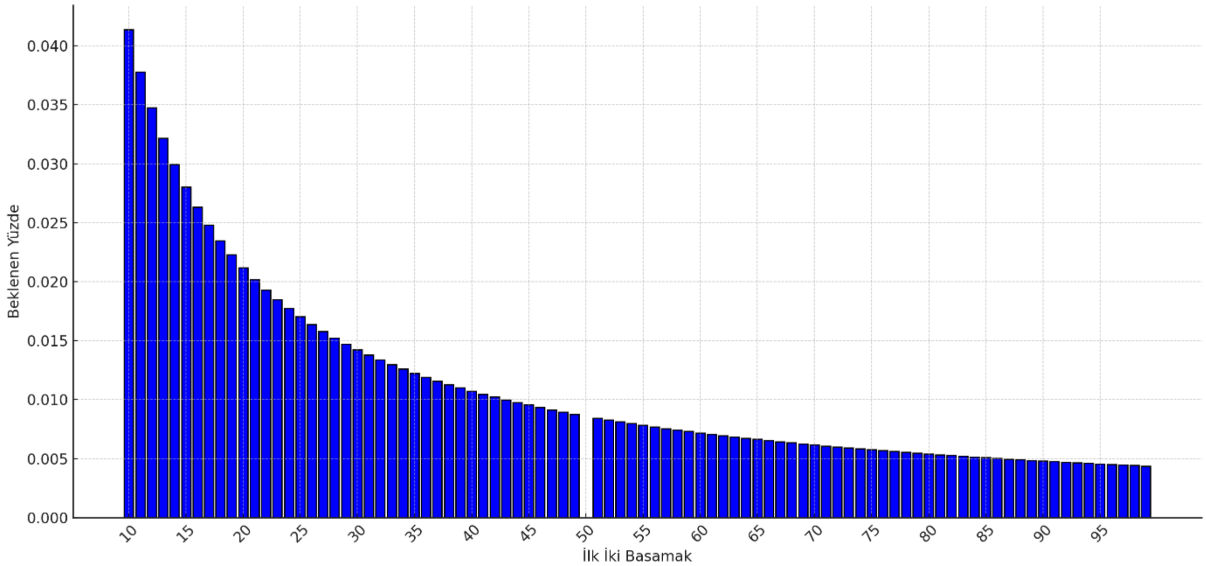
$$P(d_2) = \log\big(1 + (1/d_2)\big),\ d_2 \in \{0,1, \dots ,9\}$$

$$P(d_1 d_2) = \log\big(1 + (1/d_1 d_2)\big),\ d_1 d_2 \in \{10,11, \dots ,99\}.$$

2.1.1. Benford Kanunu'na uyumun analizinde kullanılan testler

Birinci basamak testi, sayıların soldan ilk rakamlarının frekanslarını hesaplayarak Benford Kanunu'nun beklenen oranlarıyla karşılaştırır. Bu test, veri kümesinin Benford Kanunu'na uygunluğunu belirlemek için kullanılır. Birinci basamak testi, veri kümesinde anormallikler olduğuna dair ipuçları sağlar ve genellikle veri setinin mantıklı olup olmadığını değerlendirmede kullanılır. İkinci basamak testi, veri setindeki sayıların ikinci rakamlarının (0'dan 9'a kadar) görülme sıklıklarının hesaplanarak bir çizelgeye dökülmesi ve ardından bu gözlemlenen frekansların Benford Kanunu tarafından öngörülen beklenen frekanslarla karşılaştırılmasını içermektedir. Bu test, olasılıkların %8.5 ile %11.9 arasında değiştiği genel bir uygunluk testidir ve birinci basamak testiyle benzer şekilde verinin Benford Kanunu'na ne derece uyduğunu değerlendirmek için kullanılmaktadır.

İlk iki basamak testi, sayıların ilk iki rakamının (10'dan 99'a kadar) varoluş frekanslarını Benford Kanunu'nun beklenen oranlarıyla karşılaştırır. Bu test, daha büyük bir veri seti içinde daha detaylı bir analiz sağlar ve olasılık aralığı %0.4 ile %4.1 arasındadır. Genellikle 10000'den az veri içeren küçük veri kümelerinde kullanılır ve hem grafiksel farklılıkları düzenlemek hem de anormal durumları belirlemek için tercih edilir. Bu testle, ilk iki basamakta 90 farklı kombinasyon analiz edilir. İlk iki basamak testi, veri setlerindeki anormal durumları belirlemek amacıyla kullanılır. Bu test, sayıların ilk iki basamağındaki rakam kombinasyonlarının frekanslarını analiz ederek en fazla sapma gösteren kombinasyonları saptamaktadır. Böylece belirlenen şüpheli rakam kombinasyonları daha detaylı incelenebilir. Şekil 2'de ilk iki basamağa göre Benford Kanunu'na uyan dağılımın yüzde değerleri gösterilmiştir..



Şekil 2. *Benford Kanunu'na göre ilk iki basamaktaki rakamların olasılık dağılımı*

2.1.2 Benford Kanunu analizinde kullanılan testlerin yorumlanması

Benford Kanunu'na göre yapılan analizlerde, sayıların her basamağındaki rakamların dağılımı hesaplandıktan sonra, elde edilen sonuçlar çeşitli istatistiksel testlerle değerlendirilir. Bu değerlendirme sürecinde, gözlemlenen ve beklenen oranlar arasındaki farklar incelenir. Bu sapmalar incelenirken, ki-kare testi, z-istatistiği, ortalama mutlak sapma ve Kolmogorov-Smirnov testi kullanılmaktadır.

Ki-kare testi

Ki-kare testi, verilerin beklenen dağılımlara uygunluğunu test etmek için kullanılır ve testin adımları gözlemlenen ve beklenen oranların farkının karesinin alınması ve beklenen oranlara bölünmesi şeklinde ilerler. Ki-kare testi aşağıdaki gibi hesaplanır:

$$X^2 = \sum \frac{(p_0 - p_e)^2}{p_e}.$$

Bu formülde p_0, gözlemlenen oranı ve p_e, beklenen oranı temsil eder. Bu istatistik, örnek verilerin altında yatan dağılımın beklenen dağılım hipotezine ne kadar uygun olduğunu ölçer. Bu testte elde edilen ki-kare değeri, belirlenen bir kritik değerle karşılaştırılır. Bu değer, kritik değeri aşarsa, bu durum gözlemlenen oranların beklenen oranlardan istatistiksel olarak önemli bir farklılık gösterdiğini ve hipotezin reddedilmesi gerektiğini gösterir. Bu sonuç, verilerde olası yolsuzluk veya hata olabileceğine işaret eder ve daha detaylı bir inceleme gerektirir. Eğer ki-kare değeri kritik değerin altında kalırsa, verilerin beklenen dağılıma uyduğu kabul edilir.

Z istatistiği testi

Z istatistiği testi, bir rakam veya rakam kombinasyonlarının Benford Kanunu'na göre beklenen oranından sapmasını ölçer. Bu test, belirli bir kategoriye odaklanırken, ki-kare testi tüm veri kümesini değerlendirir. Z istatistiği formülü aşağıdaki gibidir:

$$Z = \left[|p_0 - p_e| - \frac{1}{2n} \right] / \sqrt{\left(p_e * \frac{1-p_e}{n} \right)}.$$

Burada p_0, Benford Kanunu'na göre beklenen değeri; p_e, veri kümesindeki gözlemlenen oranı ve n, gözlem sayısını ifade etmektedir. Z istatistik testinin hesaplanan değeri, belirlenen güvenilirlik derecesine göre olasılık çizelgelerdeki kritik değerlerle karşılaştırılır. Örneğin, %95 güvenilirlik için 1.96 ve %90 güvenilirlik için 1.64 olan Z değerleri, önemli bir sapmanın göstergesidir. Testin hassasiyeti, incelenen veri miktarı arttıkça yükselir, bu yüzden tüm verilerin analize dâhil edilmesi önerilir. Bu şekilde, örnekleme yapılmadan direkt tüm veri seti üzerinden analiz yapmak daha doğru sonuçlar verebilir.

Kolmogorov-Simirnov testi

Kolmogorov-Smirnov testi de Benford Kanunu'na verinin uyumunun incelenmesinde kullanılmaktadır. Bir veri kümesinin normal dağılıma uygun olup olmadığını belirlemek için kullanılır. Bu testte, gözlemlenen dağılım ile bir referans dağılım arasındaki farkı ölçen bir test istatistiği hesaplanır. Test istatistiği hesaplandıktan sonra, belirli bir anlamlılık düzeyi için kritik değerlere bakılır. Test istatistiği, belirlenen anlamlılık düzeyi için kritik değerden büyükse, yokluk hipotezi reddedilir ve veri setinin normal dağılıma uymadığı sonucuna varılır.

Ortalama mutlak sapma

Mutlak sapma, bir rakamın ya da rakam kombinasyonlarının gözlemlenen frekanslarının Benford Kanunu tarafından öngörülen beklenen frekanslardan olan sapmasının mutlak değeri olarak tanımlanmaktadır. Ortalama mutlak sapma ise, bu mutlak sapmaların her biri için ayrı ayrı hesaplandıktan sonra bunların ortalamasının alınmasıyla elde edilen bir değerdir. Bu ölçüm, veri setinin büyüklüğünden bağımsızdır ve hesaplaması nispeten basittir. Denetimde, ortalama mutlak sapma en iyi uyum derecesi testi olarak kabul edilir. Ortalama mutlak sapma, üç adımdan oluşur:

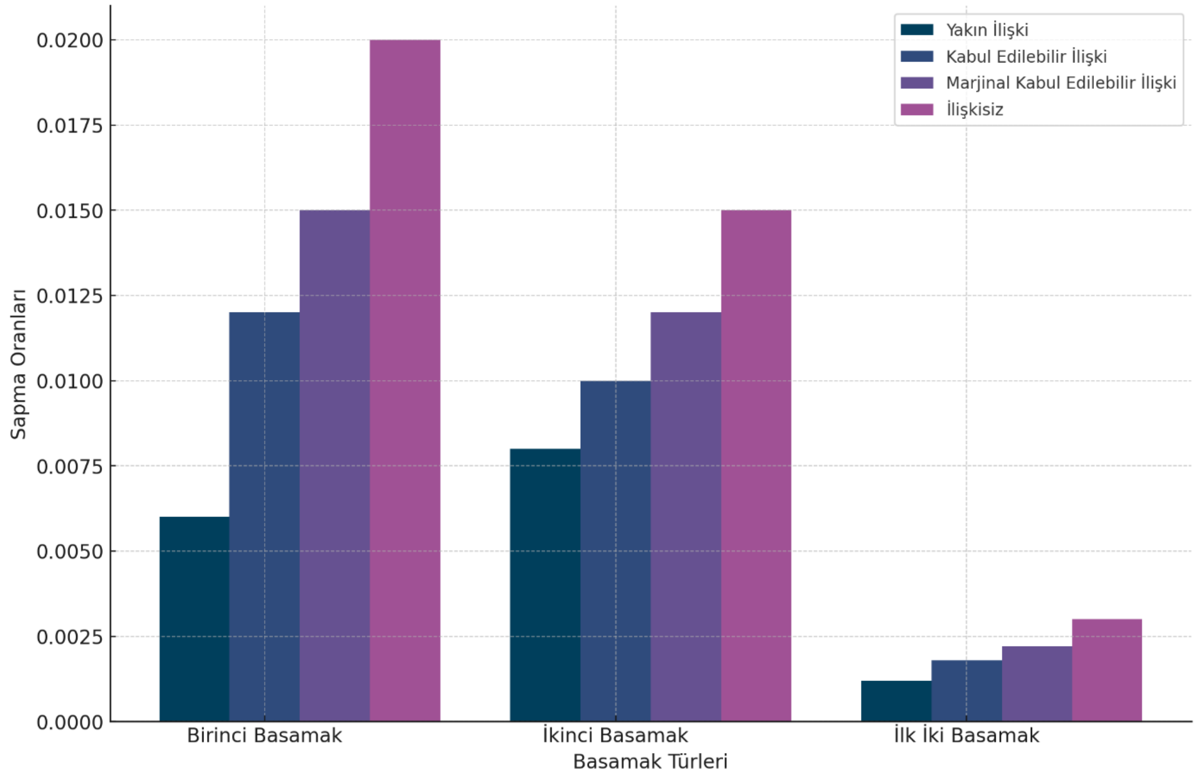
- Beklenen oranla gerçekleşen oran arasındaki farklar hesaplanır. Bu hesaplama sırasında farkların pozitif ya da negatif olmasına dikkat edilmez çünkü sayıların mutlak değeri alınır.
- Elde edilen farkların toplamı hesaplanır.
- Toplam farklar, gözlem sayısına bölünerek ortalama mutlak sapma değeri bulunur.

Bu adımlar, veri setinin Benford Kanunu'na ne kadar uyumlu olduğunu değerlendirmek için kullanılır. Ortalama mutlak sapma aşağıdaki biçimde hesaplanır:

$$\sum_{i=1}^{K} \frac{|AP-EP|}{K}.$$

Burada K, basamak sayısını; AP, veri kümesindeki gözlemlenen oranı; EP, Benford Yasası'nın beklenen değerini göstermektedir.

Mutlak sapmaların toplanması ve daha sonra 9'a bölünmesiyle (9 tane birinci basamak olduğundan) hesaplanan ortalama mutlak sapma, Benford Kanunu'nda sık kullanılan testlerden biridir. Örneğin birinci basamak testinde 1 rakamının veri kümesi içerisinde gözlemlenen oranı 0.320 olsun. Benford Kanunu'nda 1 rakamının beklenen oranı 0.301'dir. Bu durumda sapma eksi yönlü 0.019 olur. Sapmanın pozitif yönde olması için gözlemlenen oranın beklenen orandan küçük olması gerekir. Şekil 3'te verilen ortalama mutlak sapma değerleriyle veri kümesinin Benford Kanunu'na uyup uymadığı değerlendirilir.



Şekil 3. *Basamaklara göre sapma oranları*

2.2. Uç Değer Teorisi

Çalışmalarda genellikle verilerin ortalaması alınarak istatistiksel analizler yapılmakta ve bu analizlerde uç değerler özellikle dikkate alınmaktadır. Uç değerler, genellikle veri grubundan ayrıştırılmak veya ayrı bir şekilde incelenmek üzere analiz edilir. Bu değerlerin analiz edilmesi gerekiyorsa, Uç Değerler Teorisi (UDT) gibi yöntemler kullanılır. UDT, dağılımın merkezindeki gözlem değerlerinden ziyade, dağılımın kuyruğundaki verilere odaklanılmaktadır. Rasgele değişkenlerin toplamı için merkezi limit teoreminin işlevini üstlenen UDT, rasgele değişkenlerin uç değerleri için modeller oluşturur ve limit dağılımlarının nasıl olması gerektiğini ifade etmektedir.

UDT, hidroloji, sigortacılık, finans, telekomünikasyon ve diğer birçok alanda uygulamaya sahiptir. UDT, olasılık dağılımlarının kuyruklarına odaklanarak nadir olaylara teorik bir çerçeve sunmaktadır. Bu teorinin önemli bir yönü, ön varsayım yapılmadan uygulanabilmesidir. UDT geçmiş uç olaylarla geleceği tahmin etmeye çalışan bir teoridir ve bu nedenle geçmiş veriler önemlidir. Hidrolojide, sel, kuraklık ve diğer hidrolojik olayların incelenmesinde UDT önemli bir rol oynar. Örneğin, sel olaylarının frekansı ve büyüklüğü gibi hidrolojik olayların modellenmesinde uç değerler teorisi kullanılır. UDT'nin hidrolojiye uygulanmasına dair çalışmalar arasında Smith [15], Davison ve Smith [16] gibi önemli referanslar bulunmaktadır. Sigorta sektöründe, özellikle aşırı risklerin modellenmesi ve tahmin edilmesinde UDT'nin kullanımı yaygındır. Beirlant [17], Mikosch [18] ve McNeil [19], UDT'nin sigortacılık alanındaki uygulamalarını incelemişlerdir.

Finansal piyasalarda risk yönetimi, varlık fiyatlandırması ve portföy optimizasyonu gibi konularda uç değerler teorisi önemli bir rol oynamaktadır. Özellikle aşırı olayların ve krizlerin modellenmesi için UDT'nin kullanımı yaygındır. Danielsson ve de Vries [20], McNeil [21, 22] finansal alandaki UDT uygulamalarını araştırmışlardır. Özellikle nadir ve aşırı olayların analizi için UDT'nin kullanımı önemlidir. Bu alanlarda UDT'nin kullanımına dair çeşitli çalışmalar bulunmaktadır.

Uç değerler, nadir olayların sıklığını ve büyüklüğünü ölçen teknikler ile analiz edilmektedir. Nadir olayların sayımı, belirlenen bir zaman dilimi veya deneysel seriler boyunca meydana gelen nadir olayların sayısını hesaplamaktadır. Bu olaylar genellikle kesikli dağılımlar kullanılarak modellenir; Binom, Poisson, Geometrik ve Hipergeometrik dağılımlar bu tür analizler için sıklıkla tercih edilir. Özellikle Poisson dağılımı, nadir olayların modellemesinde yaygın olarak kullanıldığından, uç değerlerin dağılımı genellikle bu tür olayların dağılımı olarak kabul edilir.

UDT'de iki temel yaklaşım öne çıkmaktadır. İlk yöntem, veri serisinin maksimum ve minimum değerlerinin dağılımını modellemektir ve bu "Blok Maksima" olarak adlandırılır. Belirli şartlar altında, serilerin maksimum değerlerinin dağılımı, Gumbel (Fisher-Tippett Tip I), Frechet (Fisher-Tippett Tip II) veya Weibull (Fisher-Tippett Tip III) dağılımlarına yakınsar. Fisher-Tippett teorisi, normal koşullar altında uç değerler dağılımları için belirli bir sınırlı dağılım formunu tanımlar. Bu üç dağılım, "Genelleştirilmiş Uç Değerler (GUD) Dağılımı" olarak birleştirilir ve Fisher-Tippett-Gnedenko teorisi, n adet gözlemin maksimum değerlerinin genellikle bağımsız ve özdeş dağıldığını varsayar. Bu durumda, gözlemler serisi $(x_1, x_2, x_3, \dots, x_n)$, genel koşullar altında yaklaşık olarak GUD olarak modellenebilir. GUD dağılımının birikimli olasılık dağılım fonksiyonu, uç değer analizlerinde kritik bir rol oynar. GUD dağılımının birikimli olasılık dağılım fonksiyonu aşağıdaki gibidir:

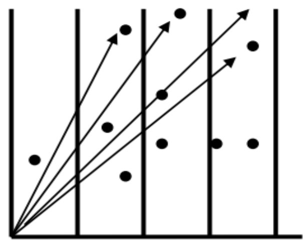
$$F(x) = \begin{cases} exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, & \xi \neq 0 \\ exp\left\{-\left[exp\left(\frac{x-\mu}{\sigma}\right)\right]\right\}, & \xi = 0. \end{cases}$$

Burada μ, konum parametresi dağılımın başlangıç noktasını belirler. σ, ölçek parametresi dağılımın yayılımını, yani değerlerin ne kadar geniş bir aralıkta dağıldığını gösterir. ξ (biçim) ise kuyruk indeks parametresidir. Dağılımın kuyruğunun ne kadar "ağır" veya "hafif" olduğunu ifade eder. GUD dağılımı, ξ > 0 ise Frechet dağılımına, ξ < 0 ise Weibull dağılımına, ξ = 0 ise Gumbel dağılımına dönüşür. Bu parametreler, belirlenen eşik değerinin üzerindeki veriler kullanılarak tahmin edilir.

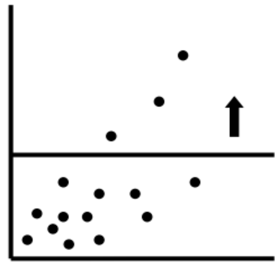
Eşiği Aşan Değer (EAD) yöntemi, belirli bir eşik değeri aşan uç değerlerin olasılıklarını hesaplamak için kullanılır. Bu yöntem, sadece belirlenen bir eşik değerin üzerindeki verilere odaklanarak, nadir ve aşırı olayların analizinde önemli bir araçtır. Eşik aşım modeli, genellikle uç değer teorisinde sıkça başvurulan Genelleştirilmiş Pareto Dağılımı (GPD) kullanılarak modelleme yapar. GPD, özellikle finans, meteoroloji ve mühendislik gibi alanlarda aşırı riskleri değerlendirmede kritik öneme sahiptir. GPD'nin birikimli dağılım fonksiyonu aşağıdaki gibidir:

$$G_{\mu,\varepsilon\,\varphi}(y) = \begin{cases} 1-e^{-z}, & \varepsilon = 0 \\ 1-(1+\varepsilon z)^{-\frac{1}{\varepsilon}}, & \varepsilon \neq 0 \\ z \geq 0, & \varepsilon \geq 0 \\ 0 \leq z \leq -\frac{1}{\varepsilon}, & \varepsilon < 0. \end{cases}$$

Şekil 4.a'da Blok Maksima ile kullanılacak verilerin seçimi ve Şekil 4.b'de EAD yönteminde eşik seviyesi tespiti sonrasında belirlenen örneklem seçimi gösterilmiştir.



Şekil 4.a. Blok Maksima Yöntemi ile Örneklem Seçimi



Şekil 4.b. EAD Yöntemi ile Örneklem Seçimi

3. BULGULAR

01.01.2022-31.12.2023 tarihleri arasındaki Aksigorta hisse senedinin TL cinsinden işlem hacmi incelenmiş ve toplam 501 adet veri değerlendirilmiştir. Veriler İş Bankasının İş Yatırım portalından [23] alınmıştır. Bu analizde, veri setindeki sayısal değerlerin ilk iki basamağının frekans dağılımını incelemek amacıyla İlk İki Basamak Testi (First Two Digits Test) kullanılmıştır. Verilerin frekansları ile Benford olasılıkları arasındaki farkın uygunluğunu belirlemek için ortalama mutlak sapma ve Ki-kare uygunluk testi uygulanmıştır. Ki-kare testinde hipotezler, $\alpha = 0.05$ anlamlılık düzeyinde test edilmiştir. Hipotezler aşağıdaki gibidir:

H_0: İşlem hacmi sayıları ile Benford Kanunu tarafından öngörülen beklenen sayılar arasında bir fark yoktur.

H_1: İşlem hacmi sayıları ile Benford Kanunu tarafından öngörülen beklenen sayılar arasında bir fark vardır.

Hisse senedi hacim verilerinin Benford Kanunu'na uygunluğunu değerlendirirken gözlemlenen en büyük beş sapmanın olduğu basamakları ve mutlak fark değerleri Çizelge 2'de özetlenmiştir.

Çizelge 2. Hisse senedi işlem hacim verilerinin Benford Kanunu'na uymayan basamak değerleri ve mutlak farkları

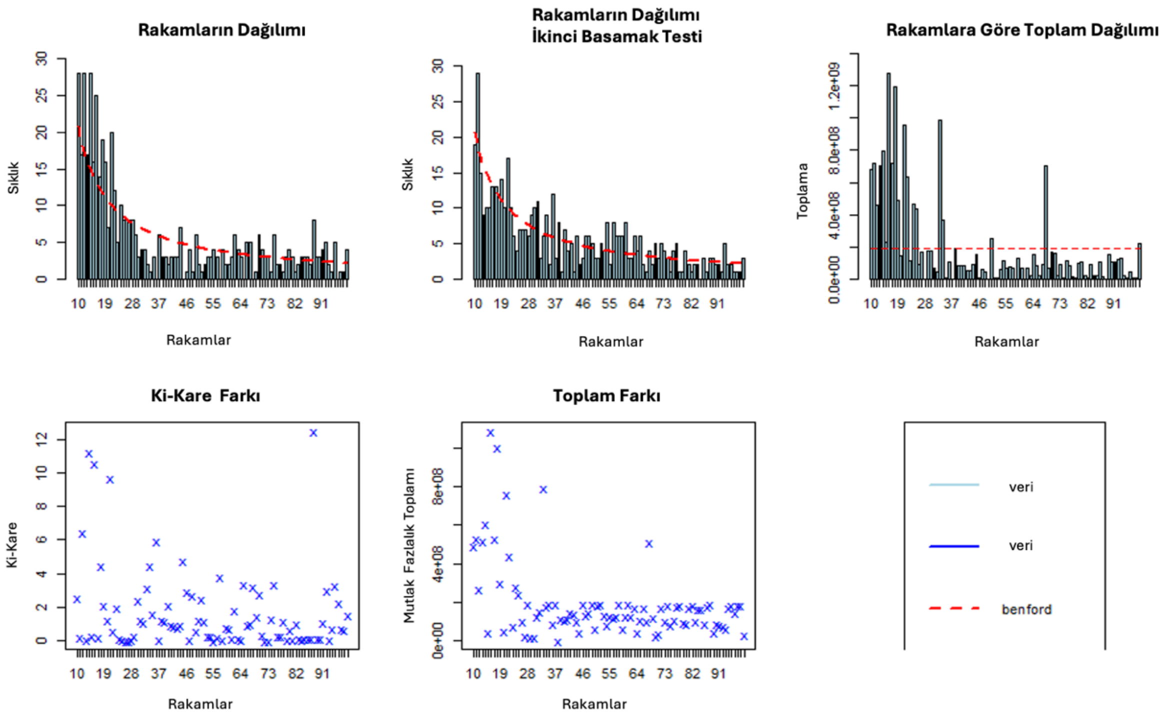
Sıra	Basamaklar	Mutlak Fark
1	14	12,99
2	16	11,81
3	12	10,58
4	21	9,88
5	10	7,26

Basamaklar sütunu, sayıların başlangıç basamaklarını göstermekte ve mutlak fark sütunu, bu basamakların gerçek verilerdeki frekansı ile Benford Kanunu tarafından öngörülen frekans arasındaki mutlak farkı ifade etmektedir. Yüksek mutlak farklar, belirli basamakların Benford'un beklediği dağılımdan önemli ölçüde sapmasını işaret etmektedir. Çizelge 3'te Ki-kare testi ve Mantissa Arc test sonuçları sunulmuştur.

Çizelge 3. *Ki-kare ve Mantissa Arc test sonuçları*

Ki-kare Testi		Mantissa Arc Testi	
İstatistik	**Değer**	**İstatistik**	**Değer**
Ki-kare	150,99	L^2	0,050961
sd	89	sd	2
p-değeri	0,00004542	p-değeri	0,000000008161

Çizelge 3'te verilen Ki-kare testi sonuçları, hisse senedi hacim verilerinin Benford Kanunu'na ne derece uyduğunu değerlendirmek için kullanılmıştır. Ki-kare test istatistiği değeri 150.99 olup, gözlemlenen frekanslar ile Benford Kanunu tarafından öngörülen frekanslar arasında önemli bir fark olduğunu göstermektedir. p-değeri çok düşük (0.00004542) olduğundan, hisse senedi hacim verilerinin Benford Kanunu'na uymadığı sonucuna varılmıştır. Bu, verilerde beklenmedik bir dağılım olduğunu ve potansiyel olarak anormal hareketlerin bulunabileceğini işaret etmektedir. Benzer olarak, Mantissa Arc testinin p-değeri 0.000000008161 olduğundan veri kümesindeki ilk basamaktan sonraki rakamların tekdüze (uniform) bir dağılıma uymadığını göstermekte ve Benford Kanunu'na uyumsuzluğu teyit etmektedir. Ortalama mutlak sapma değeri de 0.004859391 olarak elde edilmiştir. Bu değer, gözlemlenen ve teorik Benford dağılımları arasındaki ortalama mutlak sapmayı ifade etmektedir. Çizelge 4'te verilen ilk iki basamak test değerleri ile karşılaştırıldığında Benford Kanunu'ndan sapmalar olduğu görülmektedir. Şekil 5'te Benford Kanunu'na uyumu incelemek amacıyla elde edilen grafiklere yer verilmiştir.



***Şekil 5.** İşlem hacmi verisine ait benford kanunu'na uyum grafikleri*

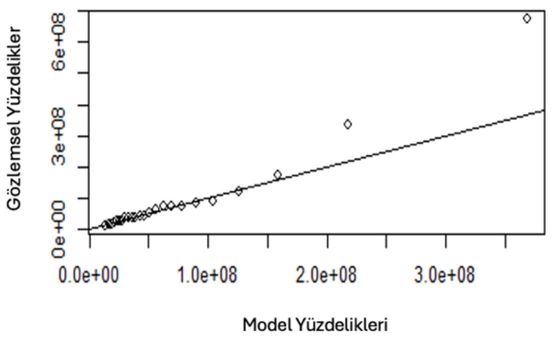
Şekil 5'te, Benford Kanunu'nun beklenen değerleri kırmızı çizgiyle gösterilmiştir ve işlem hacmi değerleri ile Benford Kanunu'nun kuramsal değerleri arasında fark olduğu söylenebilir. RStudio programında "Benford" paketi kullanılarak yapılan analizler sonucunda veri setindeki 501 gözlemden 53'ünün Benford Kanunu'na uymadığı ve dolayısıyla şüpheli olduğu belirlenmiştir. Bu gözlemler, Benford yasasına uymayan ve dikkate alınması gereken gözlemler olarak işaretlenmiştir.

Bir sonraki aşamada uç değer dağılım teorisinden yararlanarak Benford Kanunu'na uymayan hisse senedi verisinin ekstrem değerleri dikkate alınarak Blok Maksima yöntemi ile analiz yapılmıştır. İncelenen hisse senedine ait işlem hacmi değerleri öncelikle bir zaman serisine dönüştürülmüştür. Böylece, Blok Maksima yönteminin uygulanabilmesi için aylar blok olarak kabul edilerek, aylık bazda maksimum değerlerin hesaplanması sağlanmıştır. Böylece, uç değer dağılımı için uygun bir model oluşturulması amaçlanmıştır. Modele ait parametre tahminleri en çok olabilirlik (Maximum Likelihood) yöntemi kullanılarak Çizelge 4'te verilmiştir.

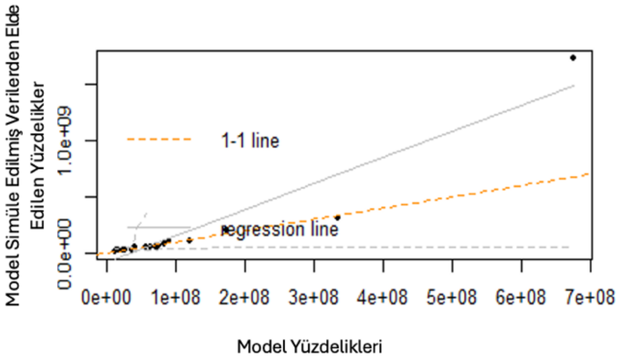
Çizelge 4. *Blok Maksima yöntemi ile elde edilen en çok Olabilirlik parametre tahmin değerleri*

Parametre	**Tahmin**	**Standart Hata**
Konum	33,766,540	6,139,631
Ölçek	25,900,590	6,839,863
Şekil	0,736	0,255

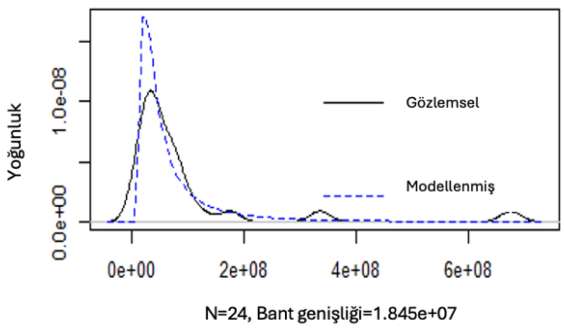
Çizelge 4'te verilen parametreler, uç değer dağılımının karakteristiklerini belirtmektedir. Şekil parametresi pozitif olduğu için, dağılımın kuyruğu sağa doğru ağırlıklı olduğu söylenebilir. Ayrıca Blok Maksima yönteminde verinin GUD dağılımına uyumunun testi gereklidir. Bu amaçla, Kolmogorov-Smirnov testi ile verinin GUD dağılımına uygunluğu değerlendirilmiştir. Kolmogorov-Smirnov test istatistiği değeri 0.1026684 ve p-değeri 0.9396785 olarak bulunmuştur. Buna göre, %5 anlamlılık düzeyinde, verilerin GUD dağılımına uygun olduğu söylenebilir. Ayrıca model için Akaike değeri (AIC) 921.1607 ve Bayesci Bilgi Kriteri (BIC) değeri 924.6948 olarak elde edilmiştir. Modele ait uyum ölçütlerinin ve geri dönüş periyodu (return period) grafikleri Şekil 6'da gösterilmiştir.



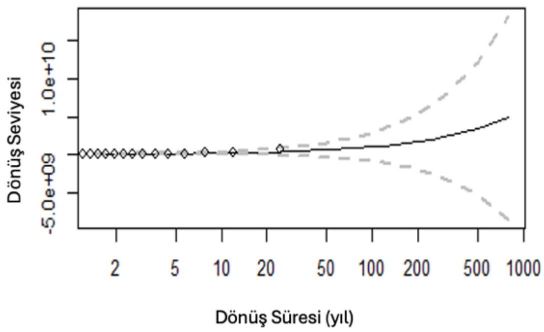
a. Deneysel ve kuramsal kantil grafiği



b. Simulasyon verisinin kantil grafiği



c. Yoğunluk grafiği



d. Geri dönüş periyodu grafiği

Şekil 6. *Blok Maksima yöntemine ait uyum iyiliği ve geri dönüş periyodu grafikleri*

Şekil 6.a'da verilen Q-Q grafiğinde gözlemlenen verilerin kantillerini (quantiles) model tahminlerinin kantilleriyle karşılaştırılmıştır. İdeal olarak, noktaların çoğu y=x çizgisine (45 derecelik açıyla çizilen doğru) yakın olmalıdır. Grafikteki noktaların çoğunun doğru etrafında yoğunlaşması modelin uç değerleri iyi modellediğini göstermektedir. Şekil 6.b'de, model tahminlerinin gerçekte ne kadar iyi performans gösterdiğini anlamak için simüle edilmiş veri kantilleri ile model tahminleri karşılaştırılmıştır. 1-1 çizgisi (kesikli turuncu çizgi), ideal uyumu gösterir. Gri "regresyon çizgisi" ise gerçek model uyumunu gösterir. Bu çizgi, modelin düşük ve yüksek kantillerde verileri biraz daha düşük tahmin ettiğini göstermesine rağmen genel olarak modelin uyumu iyidir. Şekil 6.c'deki grafik, gözlemlenen verilerin (kesikli siyah çizgi) ve model tarafından üretilen verilerin (kesikli mavi çizgi) yoğunluk dağılımlarını karşılaştırmaktadır.

Modelin yoğunluğu, gözlemlenen verilerin yoğunluğunu genel olarak iyi takip etmektedir. Şekil 6.d'de farklı dönüş periyotları için tahmini dönüş seviyeleri gösterilmiştir. Yatay eksen dönüş periyodunu, dikey eksen ise bu periyotta beklenen maksimum değeri ifade etmektedir. Noktalı çizgiler güven aralıklarını göstermektedir. Modelin dönüş seviyeleri, uzun dönemlerde hızlı bir artış göstermesi nadir olayların çok yüksek etkilerinin olabileceğini işaret etmektedir. Çizelge 5'te geri dönüş seviyesi, alt ve üst güven sınırları verilmiştir.

Çizelge 5. *Geri Dönüş Seviyesi, Alt ve Üst Güven Sınırları*

Dönüş Seviyeleri	%95 Alt güven sınırı	Tahmin	%95 Üst güven sınırı
3-yıllık dönüş seviyesi	38014979	66956567	95898156
6-yıllık dönüş seviyesi	47518008	121695272	195872536
9-yıllık dönüş seviyesi	42219070	168372576	294526083
12-yıllık dönüş seviyesi	29756738	210741711	391726685

Çizelge 5'te GUD modeli kullanılarak hesaplanan, farklı dönemler için dönüş seviyelerini ve bunların güven aralıklarını göstermektedir. Bu değerler, belirli bir süre içinde (yıl bazında) beklenen maksimum değerleri ifade etmektedir. 3 yıllık dönüş seviyesi işlem hacmi seviyesi tahmini 66,956,567 ile en düşük 38,014,979 ve en yüksek 95,898,156 arasında değişmektedir. 6 yıllık dönüş seviyesi tahmini 121,695,272 olup, 47,518,008 ile 195,872,536 arasında bir güven aralığına sahiptir. 9 yıllık dönüş seviyesi tahmini 168,372,576 tahmini değeri, 42,219,070 ile 294,526,083 arasında bir güven aralığıyla belirlenmiştir. 12 yıllık dönüş seviyesi tahmini en uzun süreli tahmin olan 210,741,711, 29,756,738 ile 391,726,685 arasında geniş bir güven aralığına sahiptir.

3. SONUÇ

Bu çalışmada, hisse senedi fiyatlarının analizi için Benford Kanunu ve uç değer teorisi kullanılmıştır. Benford Kanunu, hisse senedi işlem hacimlerindeki ilk rakam frekanslarını analiz ederek, bu verilerde beklenen dağılımdan sapmaları başarılı bir şekilde tespit etmiştir. Bu sapmalar, potansiyel anormalliklerin veya manipülasyonların göstergesi olabilir. Özellikle, hisse senedi verileri üzerinde yapılan Benford Kanunu analizleri, belirgin anormal hareketleri ortaya çıkarmıştır. Örneğin, Durtschi vd. [1] tarafından yapılan çalışmalar, muhasebe verilerinde Benford Kanunu'nun nasıl kullanılabileceğini detaylı bir şekilde incelerken, bu çalışma da hisse senedi piyasalarında benzer bir analitik yaklaşımın uygulanabilirliğini test etmiştir. UDT, finansal piyasalardaki aşırı dalgalanmaları ve nadir olayları incelerken kullanılmıştır. Bu yöntem, özellikle risk yönetimi ve finansal modelleme bağlamında önemli sonuçlar üretmiştir. Hisse senedi fiyatlarında gözlemlenen aşırı maksimum ve minimum değerler, piyasa risklerinin daha iyi anlaşılmasını sağlamıştır. Bu bulgular, Embrechts vd. [8] tarafından sigortacılık ve finans alanlarında uç değerlerin nasıl modellenebileceğini incelediği çalışmalarla paralellik göstermektedir. Embrechts vd. [8]'nin çalışmaları, aşırı değerlerin tahmin edilmesi ve risk değerlendirmelerinin geliştirilmesinde kritik öneme sahipken, bu çalışma da benzer metodolojilerin hisse senedi piyasası verileri üzerinde uygulanabilirliğini ortaya koymaktadır.

Sonuç olarak, bu çalışma, Benford Kanunu ve uç değer teorisinin hisse senedi işlem verilerinin analizinde ne kadar etkili olduğunu göstermiştir. Bu yöntemler, finansal denetimlerde ve piyasa analizlerinde kullanılarak, daha sağlam finansal kararlar alınmasına yardımcı olabilir. Ancak, her iki metodun da kendi sınırlılıkları olduğunu ve en iyi sonuçlar için birlikte kullanılmaları gerektiğini belirtmek önemlidir. Gelecekte bu yöntemlerin, daha geniş veri kümeleri ve farklı piyasa koşulları altında nasıl performans gösterdiğini araştırmak, bu alanlardaki bilgi birikimini artırabilir. Bu şekilde, finansal piyasaların daha derinlemesine anlaşılması ve piyasa dalgalanmalarının etkin bir şekilde yönetilmesi mümkün olabilir.

TEŞEKKÜR

Bu makalenin hazırlanmasında kullanılan yöntem ve analizlerin seçiminde TÜBİTAK 4005 122B264 No'lu "Eğitimde Yenilikçi Bir Yaklaşım: Veri Görselleştirme Teknikleri ve Uygulamaları" adlı etkinlikten yararlanılmıştır. İlgili projeye teşekkür ederiz.

YAZAR KATKI ORANI

Hatice Nur Karakavak: Kaynaklar, Araştırma, Metodoloji, Yazma. **Gamze Özel**: Biçimsel analiz, Görselleştirme, Metodoloji, Yazma

ÇIKAR ÇATIŞMASI BEYANI

Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

KAYNAKLAR

[1] Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting*, 5(1), 17-34.
[2] Deckert, J., Myagkov, M., & Ordeshook, P. C. (2011). Benford's Law and the detection of election fraud. *Political Analysis*, 19(3), 245-268.
[3] Diekmann, A. (2007). Not the first digit! Using benford's law to detect fraudulent scientif ic data. *Journal of Applied Statistics*, 34(3), 321-329.
[4] Giles, D. E. (2007). Benford's law and naturally occurring prices in certain ebaY auctions. *Applied Economics Letters*, 14(3), 157-161.
[5] Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's Law as an Aid in Analytical Procedures. *Auditing: A journal of Practice & Theory*, 16(2).
[6] Özkan, K. (2021). Estimating ecosystem naturalness using Benford's law and generalized Benford's law. *Turkish Journal of Forestry*, 22(2), 73-82.
[7] Carslaw, C. A., & Kaplan, S. E. (1991). An examination of audit delay: Further evidence from New Zealand. *Accounting and Business Research*, 22(85), 21-32.
[8] Embrechts, P., Klüppelberg, C., Mikosch, T., Emberchts, P., Klüppelberg, C., & Mikosch, T. (1997). Risk theory. *Modelling Extremal Events: for Insurance and Finance*, 21-57.
[9] McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: concepts, techniques and tools-revised edition. *Princeton University Press*.
[10] Beirlant, J., Vynckier, P., & Teugels, J. L. (1996). Excess functions and estimation of the extreme-value index. *Bernoulli*, 293-318.
[11] Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7), 1097-1130.
[12] Reiss, R. D., Thomas, M., & Reiss, R. D. (1997). Statistical analysis of extreme values (Vol. 2). *Basel: Birkhäuser*.
[13] Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, 4(1), 39-40.
[14] Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 551-572.
[15] Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 367-377.
[16] Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 52(3), 393-425.
[17] Beirlant, J., Teugels, J. L., & Vynckier, P. (1994). Extremes in non-life insurance. *In Extreme Value Theory and Applications: Proceedings of the Conference on Extreme Value Theory and Applications*, Volume 1 Gaithersburg Maryland 1993 (pp. 489-510). Springer US.
[18] Mikosch, T. (1997). Heavy-tailed modelling in insurance. *Communications in statistics. Stochastic models*, 13(4), 799-815.
[19] McNeil, A. J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA*, 27(1), 117-137.
[20] Danielsson, J., & de Vries, C. G. (1997). Extreme returns, tail estimation, and value-at-risk (Vol. 11). University of Iceland, Institute of Economic Studies.

[21] McNeil, A. J. (1998). Calculating quantile risk measures for financial return series using extreme value theory. *ETH Zurich*, 4(3), 1-18.
[22] McNeil, A. J. (1999). Extreme value theory for risk managers. *Departement Mathematik ETH Zentrum*, 12(5), 217-37.
[23] İş Bankası, https://www.isyatirim.com.tr/tr-tr/analiz/hisse/Sayfalar/Tarihsel-Fiyat-Bilgileri.aspx [17.03.2024 Tarihli Erişim]