EDİTÖRLER

Doç. Dr. Güzide ŞENEL
Doç. Dr. Emre EKİN

ARALIK
2024

İmtiyaz Sahibi / Yaşar Hız
Yayına Hazırlayan / Gece Kitaplığı
Birinci Basım / Aralık 2024 - Ankara
ISBN / 978-625-430-284-8

Gece Kitaplığı
Kızılay Mah. Fevzi Çakmak 1. Sokak
Ümit Apt No: 22/A Çankaya/ANKARA
0312 384 80 40
www.gecekitapligi.com / gecekitapligi@gmail.com

Baskı & Cilt
Bizim Büro
Sertifika No: 42488

İSTATİSTİK ALANINDA ARAŞTIRMALAR VE DEĞERLENDİRMELER

EDİTÖRLER

Doç. Dr. Güzide ŞENEL
Doç. Dr. Emre EKİN

gece kitaplığı

İÇİNDEKİLER / CONTENTS

BÖLÜM 1

CHAPTER 2

BÖLÜM 1

HİSSE SENEDİ FİYAT TAHMİNİ: ARIMA VE ÜSTEL DÜZELTME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Mehmet Talha COŞGUN[1], Demet SEZER[2]

1 Yüksek Lisans Öğrencisi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik ABD, Konya, Türkiye.

ORCID ID: 0009-0005-6999-5805

2 Dr. Öğr. Üyesi, Selçuk Üniversitesi, Fen Fakültesi, Aktüerya Bilimleri Bölümü, Konya, Türkiye.

ORCID ID: 0000-0002-0680-948X

1. GİRİŞ

Tarih boyunca insanlar varlıklarını korumak ve yatırım yapmak için çeşitli alanlarda değerlendirmeler yapmaktadır. Bu alanlardan bir tanesi ise borsa yatırım araçlarından olan hisse senetleridir. Risk oranı fazla olan hisse senetlerinin tahmin ve değerlendirme aşaması bu yatırım araçlarının korku derecesini yükseltmiştir. Matematik ve istatistiksel alanda çalışmalar neticesinde hisse senetlerinin fiyatlanmasına etken olan birçok yöntem geliştirilmiştir. Daha çok ileri dönem tahminleri üzerine yoğunlaşan yatırımcılar bu alanda bir çok tahmin yöntemlerini ve algoritmalarını kullanmışlardır.

2. TAHMİN YÖNTEMLERİ

İstatistiksel çalışmalarda tahmin yöntemleri yatırım araçları, veri analizleri, bütçe tahminleri gibi bir çok alanda kullanılmaktadır. Bu yöntemler işveren, çalışan veya bir bireyin hedeflediği beklenti ile karşılaşılabilecek olasılıkların belirlenmesinde önem arz etmektedir. Yatırım araçlarının tahminlemesi için ARIMA ve Üstel Düzeltme önemli tahmin yöntemlerindendir. Üstel Düzeltme yönteminin basitliği, trend ve mevsimselliğe uyumu, uygulama kolaylığı, düşük maliyeti, veriye uygun hareket etme kabiliyeti başlıca avantajlarındandır. ARIMA ise esneklik, uzun vadeli tahminleme, geniş uygulama alanları ve teorik temel gibi birçok avantaja sahiptir (Mashadihasanli, 2022). Bu iki yöntem ile kullanıcılar düşük maliyet ve uygulama kolaylığı ile çalışmalarında doğruluk olasılığı yüksek bilgiler elde etmektedir. Bu çalışmada bu iki yöntem ile hisse senedi fiyat tahmini yapılmıştır.

2.1. ARIMA

ARIMA (Otoreğresif Entegre Hareketli Ortalama) modeli, zaman serisi verilerinin geçmiş değerleri arasındaki ilişkiyi modelleyerek gelecekteki değerleri tahmin etmek için kullanılan güçlü bir istatistiksel yöntemdir(Box ve Jenkins, 2015). Bu modelin temelini oluşturan üç bileşen vardır. Bunlar:

OtoreGresif (AR): Bir gözlemin, önceki gözlemlerinin doğrusal bir fonksiyonu olduğunu ifade eder. Yani, bir veri noktasının değeri, kendinden önceki değerlerle belirli bir ilişki içindedir.

Entegre (I): Verinin durağan olmaması durumunda, durağan hale getirmek için fark alma işlemi uygulanır. Fark alma işlemi, bir gözlem ile bir önceki gözlem arasındaki farkı alarak yapılır.

Hareketli Ortalama (MA): Bir gözlemin, beyaz gürültü hatalarının doğrusal bir fonksiyonu olduğunu ifade eder. Yani, bir veri noktasının değeri, geçmişteki hataların ağırlıklı ortalamasıdır.

Bir ARIMA modelini genel olarak aşağıda verilen eşitlikteki gibi verebiliriz:

$$Yt = \alpha + \beta 1 Yt - 1 + \beta 2 Yt - 2 + \cdots + \beta p Yt - p + \varphi 1 \varepsilon t - 1 + \varphi 2 \varepsilon t - 2 + \cdots + \varphi q \varepsilon q - 1 \qquad (1)$$

Eşitlik (1)'de,

α: sabit değeri,

β: gözlem değerleri için katsayılarını,

Yt: zaman serisinin t anındaki değerini,

φ: hata terimlerinin katsayılarını,

ε: hata terimlerini,

p: oto regresif katsayısını başka bir ifade ile gecikme değerlerini,

göstermektedir. q ise hareketli ortalama parametresidir. Sonuç olarak, ARIMA (p, d, q) ifadesi ile otoregresif sıralama(p), fark alma işleminin kaç kez yapılacağı (d) ve hareketli ortalama parametresi(q) bilgileri ortaya konulur (Aslan ve Erdur, 2020).

Analizler genellikle orta çaplı büyüklükteki veri setlerinin analizlerinde kullanılır. Büyük çaplı veriler için daha çok makine öğrenmesi ve derin öğrenme, bulanık mantık gibi yöntemler kullanılmaktadır. BIST endeksinde süreçlerin karmaşıklığı, fiyatları etkileyen faktörlerin çok olması sebebiyle yatırımcılar, ekonomistler, matematikçiler birden fazla tekniğin uygulaması, kıyaslaması ve hedef belirlenmesini amaçlamıştır. Optimal sonuç başarısı için karşılaştırmalar göz önünde bulundurulmuştur. Bu çalışmada ise mevsimsellik faktörü ile öğrenme tekniği olan ARIMA yönteminin kıyaslanması sonucu elde edilen çıktıların doğruluğu kontrol edilecektir. Bu sayede zaman faktörü ve bağıl hareketlerin piyasa fiyatlanmasındaki baskınlık test edilmiş olacaktır (Bengio, 2009).

2.2. ÜSTEL DÜZELTME

Üstel Düzeltme, zaman serisi verilerindeki eğilimleri ve düzgün olmayan değişimleri tahmin etmek için kullanılan bir istatistiksel yöntemdir. Bu yöntemde, daha eski verilere göre daha yeni verilere daha fazla ağırlık verilir. Yani, zaman serisi boyunca ilerledikçe, yeni gözlemler daha fazla önem kazanır ve eski gözlemler daha az önem kazanır. Bu sayede, verideki en güncel eğilimler daha iyi yakalanabilir. Üstel Düzeltme Yöntemleri genel hatları ile 3 yöntem üzerinden değerlendirilebilir. Bunlar Single, Double ve Triple Exponential Smoothing olarak adlandırılır.

2.2.1. Single Exponential Smoothing (SES/Single HWES):

Basit Üstel Düzeltme Yöntemi (SES), yalnızca durağan zaman serilerinde etkili olup, trend ve mevsimsellik gibi yapıları modelleyemez. Bu yöntem, geçmiş verilere daha fazla ağırlık vererek üstel bir düzeltme gerçekleştirmektedir. Geleceğin, yakın geçmişten daha fazla etkilendiği varsayımıyla geçmiş veriler ağırlıklandırılarak kullanılır. Single Exponential Smoothing formülü Denklem 2 de verilmiştir.

$$\hat{y}_t = a \cdot y_{t-1} + (1-a) \cdot (\hat{y}_{t-1}) \quad (2)$$

α, 0 ile 1 arasında bir değer alarak düzeltme faktörü olarak işlev görür ve ağırlıklandırma düzeyini belirler. Tahminler, geçmiş gerçek değerler (learning) ile önceki tahminlerin (remembering) üstel ağırlıklandırılmasıyla oluşturulur.

2.2.2.Doble Exponential Smoothing (DES/Double HWES):

Çift Üstel Düzeltme yöntemi (DES), Basit Üstel Düzeltme (SES) yöntemine ek olarak trend etkisini de dikkate alarak üstel düzeltme yapar. Ancak, bu yöntem mevsimsellik etkilerini modelleyemez.

$$l_t = ay_t + (1-a)(l_{t-1} + b_{t-1}) \quad (3)$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \quad (4)$$

$$\hat{y}_{(t+1)} = l_t + b_t \quad (5)$$

Denklem 3 level'i, Denklem 4 trendi ifade eder. α parametresi level'a ilişkin geçmiş değer ve tahmin edilen değeri ne kadar ağırlıklandıracağımızla ilgilidir. DES, α parametresine ek olarak β parametresini de kullanır.

β; Trend bileşenine ilişkin optimize edilmesi gereken parametredir.

2.2.3.Triple Exponential Smoothing (TES/Triple HWES):

Winters Üstel Düzeltme Yöntemi, Çift Üstel Düzeltme (DES) yöntemine ek olarak mevsimsellik etkilerini de modelleyerek en gelişmiş düzeltme yöntemlerinden biri olarak kabul edilir.

$$l_t = a(y_t - s_{t-p}) + (1-a)(l_{t-1} + b_{t-1}) \quad (6)$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \quad (7)$$

$$s_t = \gamma(y_t - l_t) + (1-\gamma)s_{t-p} \quad (8)$$

$$\hat{y}_{t+m} = l_t + mb_t + s_{t-p+1+(m-1)modp} \quad (9)$$

Level ve trend' e ek olarak 3.formül mevsimselliği ifade eder.

γ; Mevsimsellik bileşenine ilişkin optimize edilmesi gereken parametredir.

Bu yöntem, dinamik bir şekilde düzey (level), trend ve mevsimsellik etkilerini değerlendirerek tahminler üretir.

3. HİSSE SENEDİ GELECEK DÖNEM FİYAT TAHMİNİ

Küreselleşen dünyada artan rekabetle birlikte hisse senedi fiyatları, arz-talep dengesine bağlı olarak şekillenmekte ve çeşitli değişkenlerden etkilenmektedir. Başarılı bir hisse senedi fiyat tahmini için en az veri ve en basit modellerle en doğru sonuçlara ulaşmak büyük önem taşır. Politik, ekonomik, toplumsal ve ticari birçok faktör hisse senedi endekslerini etkilediğinden ve endeksler kompleks ve doğrusal olmayan bir yapıya sahip olduğundan tahminler güçlükle yapılır. (Aktas vd., 2022) hisse senetlerinin bu şekilde karmaşık yapıda olması ise daha çok parametre içeren istatistiksel yöntemlerin önemini destekler niteliktedir. Hisse senedi fiyatları şiddetli volatilite nedeniyle diğer yatırım şekillerine göre riski fazladır (Çalışkan ve Deniz, 2015). Hisse senedi fiyat tahminlerinde en önemli veri setlerinden biri günlük, haftalık ve aylık kapanış fiyatlarıdır. Yatırımın devamlılığı, sermayenin korunması ve yatırımcı psikolojisinin yönetimi açısından kapanış

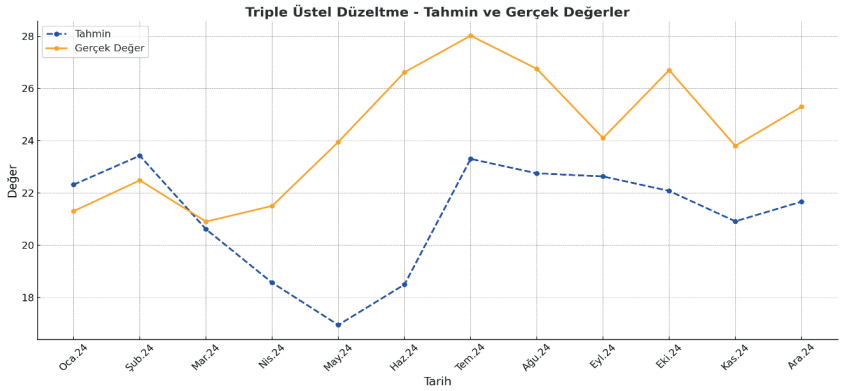
fiyatlarına dayalı analizler güven sağlar. BİST endeksinin karmaşık yapısı nedeniyle uzun vadeli tahminlerin genellikle gerçekçi olmadığı görülmüş, buna karşın orta vadeli analizlerde tahminlerin gerçekleşme oranının belirgin olduğu gözlenmiştir.

4. ÜSTEL DÜZELTME VE ARIMA YÖNTEMLERİ İLE TAHMİN UYGULAMASI

Uygulama aşamasında BİST100'de işlem gören EREGL hisse senedinin 2017 ve 2023 verileri kullanılarak Triple Üstel Düzeltme Yöntemi ile 2024 yılı tahmini yapılmıştır. Ay kapanış verileri kullanılarak oluşturulan model, gerçek değerler ile tahmin değerleri karşılaştırılarak kıyaslanmıştır. ARIMA Yöntemi ile yapılan uygulamada ise EREGL hisse senedinin 2017-2023 verileri kullanılmıştır. Elde edilen sonuçlar Tablo 1-2 ve Şekil 1-2'de verilmiştir.

Tablo 1: *Triple Üstel Düzeltme ile elde edilen tahmin ve gerçek değer veri tablosu*

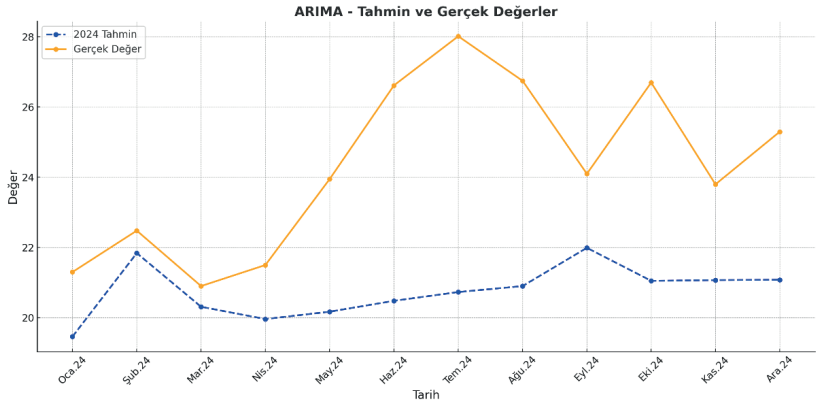
TARİH	TAHMİN	GERÇEK DEĞER
Oca.24	22,31	21,3
Şub.24	23,42	22,48
Mar.24	20,61	20,9
Nis.24	18,56	21,5
May.24	16,94	23,95
Haz.24	18,49	26,62
Tem.24	23,30	28,02
Ağu.24	22,75	26,75
Eyl.24	22,63	24,1
Eki.24	22,07	26,7
Kas.24	20,91	23,8
Ara.24	21,66	25,3



Şekil 1. *Triple Üstel Düzeltme ile elde edilen tahmin ve gerçek değer karşılaştırma grafiği*

Tablo 2: *ARIMA ile elde edilen tahmin ve gerçek değer veri tablosu*

TARİH	2024 TAHMİN	GERÇEK DEĞER
Oca.24	19,46	21,3
Şub.24	21,84	22,48
Mar.24	20,31	20,9
Nis.24	19,96	21,5
May.24	20,17	23,95
Haz.24	20,48	26,62
Tem.24	20,73	28,02
Ağu.24	20,9	26,75
Eyl.24	21,99	24,1
Eki.24	21,05	26,7
Kas.24	21,07	23,8
Ara.24	21,08	25,3



Şekil 2. *ARIMA ile elde edilen tahmin ve gerçek değer karşılaştırma grafiği*

5. SONUÇ

Çalışma sonucu elde edilen bulgular aşağıda verilmiştir:

- Triple Üstel Düzeltme modeli, stabil ve düşük dalgalanmalı dönemlerde (örneğin Mart ve Nisan 2024) daha tutarlı tahminler üretmiştir. ARIMA modeli ise genel olarak gerçek değerlere daha yakın sonuçlar elde etmiş, ancak ani değişimlerde sapma göstermiştir.

- Triple Üstel Düzeltme modeli, yüksek volatilite dönemlerinde (Mayıs ve Haziran 2024) ciddi tahmin hataları yapmıştır. ARIMA modeli, dalgalanmaları daha iyi takip etmesine rağmen, ani artış ve düşüşlerde performansı yetersiz kalmıştır.

- Triple Üstel Düzeltme modeli, daha düz ve yumuşak bir tahmin eğilimi göstermiş, keskin iniş-çıkışları öngörmekte zorlanmıştır. ARIMA modeli ise keskin geçişlere duyarlı olsa da yüksek volatilite dönemlerinde tutarsızlık göstermiştir.

- Triple Üstel Düzeltme modelinde tahminler, özellikle Mayıs ve Haziran aylarında gerçek değerlerden oldukça uzaklaşmıştır. ARIMA modelinde aynı dönemlerde sapmalar daha küçüktür, ancak hâlâ kayda değer bir fark bulunmaktadır.

- Her iki model de Mart, Nisan ve Eylül aylarında gerçek değerlere oldukça yakın sonuçlar üretmiştir, bu da düşük dalgalanma dönemlerinde iyi performans sergilediklerini göstermektedir.

- Her iki model de ani değişimler ve volatiliteyi daha iyi yakalayacak şekilde iyileştirilmelidir. Sezonluk etkiler ve dışsal faktörler modele dahil edilerek performans artırılabilir.

- Triple Üstel Düzeltme modeli, dalgalanması az olan dönemlerde güvenilir sonuçlar verirken, ARIMA modeli genel doğruluk açısından daha üstündür. Ancak, her iki modelin de yüksek volatilite dönemlerinde performansı artırılması gerekmektedir.

REFERANSLAR

1) Aktas, O. U., Kryzanowski, L., & Zhang, J., 2022. Price-limit effectiveness: Evidence from the Borsa Istanbul (BIST). International Journal of Islamic and Middle Eastern Finance and Management, 15(3), 527-568. https://doi.org/10.1108/IMEFM-04-2020-0151

2) Aslan, B., & Erdur, R. C., 2020. Stock Market Prediction with Deep Learning Using Public Disclosure Platform Data. 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), 1-5.

3) Bengio, Y., 2009. Learning deep architectures for AI. Foundations and trends® in Machine Learning, 2(1), 1-127. https://doi.org/10.1561/2200000006

4) Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M., 2015. Time series analysis: Forecasting and control. John Wiley & Sons.

5) Çalışkan, M. M. T., Deniz, D., 2015. Yapay Sinir Ağlarıyla Hisse Senedi Fiyatları ve Yönlerinin Tahmini, Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 10(3), 177- 194.

6) Mashadihasanli, T., 2022. Stock Market Price Forecasting Using the Arima Model: an Application to Istanbul, Turkiye, Journal of Economic Policy Researches, 9(2), 439-454.

CHAPTER 2

PARAMETER ESTIMATION AND COMPARISON CRITERIA

Ecem DEMİR YURTSEVEN[1]

1 Lecturer Dr., Sivas Cumhuriyet University Health Services Application and Research Hospital, Statistics Department, Sivas, Türkiye, ecemdemir@cumhuriyet.edu.tr, ORCID ID:0000-0001-9714-0672

INTRODUCTION

Estimation of unknown parameters of distributions used to model real data is one of the main problems of statistical science. Parameter estimation plays a central role in statistical analysis and modeling processes. The determination of parameters representing the basic characteristics of a population or process based on observed data has a multifaceted importance in both theoretical and applied studies. In this context, parameter estimation forms a critical leg for the development of generalizable models in various disciplines ranging from linear regression analysis to machine learning algorithms (Aster, Borchers, Thurber 2019; Kutz 2023).

Contemporary statistics employs various parameter estimation techniques, ranging from traditional methods like Least Squares Estimation (LSE) and Maximum Likelihood Estimation (MLE) to more advanced approaches, including Bayesian methods and strategies focused on robustness. Various metrics are used to compare the performance of these methods. Measures such as mean squared error (MSE), information criteria (AIC, BIC) and error rates are the main tools used to reveal the advantages and disadvantages of different methods.

The methods used in parameter estimation are characterized by the bias, consistency and efficiency of the estimators (Kosmidis 2013; Sultana, Muhammad, Aslam 2019). Bias refers to the systematic distance of an estimator from the true parameter value, while consistency describes the tendency of the estimator to converge to the true parameter value as the sample grows. Efficiency refers to the minimal variance of an estimator under a given level of bias (Edge, 2019). These properties are important criteria for highlighting different parameter estimation methods and determining their areas of use.

In this section, we will discuss the theoretical basis and comparison criteria of parameter estimation methods. The aim is to provide the reader with a comprehensive overview of the underlying principles and applications of both traditional and modern parameter estimation approaches.

1. Maximum Likelihood Estimation

The core principle of the Maximum Likelihood Method is to identify the parameter value that maximizes the likelihood of observing the given sample data, thereby offering an estimate for the population parameter.

Consider $X_1, X_2, \ldots, X_n$ as a sample drawn from a population with a probability density function $f(x;\theta)$. When the probability density function of this sample is viewed as a function of the parameter θ alone, it

is referred to as the likelihood function of θ. The likelihood function is expressed as:

$$L(\theta|x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

The value of θ that maximizes the likelihood function (if such a value exists) is considered the maximum likelihood estimate of θ. In other terms, the maximum likelihood estimator of θ is:

$$\hat{\theta} = \underset{\theta \in \Theta}{argmax} L(\theta|x_1, x_2, \dots, x_n)$$

To simplify the process, the natural logarithm of the likelihood function $L(\theta|x_1, x_2, \dots, x_n)$ is often used instead of the likelihood function itself when finding the maximum likelihood estimator of θ. Since the natural logarithm is monotonically increasing, it can be expressed as (Akdi, 2014: 329).

$$\underset{\theta \in \Theta}{max} L(\theta|x_1, x_2, \dots, x_n) = \underset{\theta \in \Theta}{max} logL(\theta|x_1, x_2, \dots, x_n)$$

Here, the function $logL(\theta|x_1, x_2, \dots, x_n)$ is called the *logL* likelihood function and is abbreviated as.

Example: The maximum likelihood estimator is widely used for estimating the parameters of the Weibull distribution. When assessed using goodness-of-fit statistics, it tends to offer more reliable results in comparison to other estimation methods (Lei, 2008).

Consider *X1, X2,…, Xn* as a sample of *n* units drawn from the Weibull distribution. The likelihood function is:

$$L(\mu, \alpha, \beta|x) = \frac{\beta^n}{\alpha^n} \prod_{i=1}^{n} \left(\frac{x_i - \mu}{\alpha}\right)^{\beta - 1} exp\left(-\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\alpha}\right)^{\beta}\right)$$

if the logarithm of the likelihood function is taken;

$$\ln L(\mu, \alpha, \beta|x) = n \ln \beta - n \ln \alpha + (\beta - 1) \sum_{i=1}^{n} \ln\left(\frac{x_i - \mu}{\alpha}\right) - \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\alpha}\right)^{\beta}$$

is obtained. When the first order partial derivatives of $ln\, L(\mu, \alpha, \beta|x)$ concerning μ, α, β are taken and equated to zero;

$$\frac{d \ln L(\mu, \alpha, \beta|x)}{d\mu} = -(\beta - 1) \sum_{i=1}^{n} \left(\frac{1}{x_i - \mu}\right) + \frac{\beta}{\alpha} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\alpha}\right)^{\beta - 1}$$

$$\frac{d \ln L(\mu, \alpha, \beta|x)}{d\alpha} = -\frac{n}{\alpha} - \frac{n(\beta - 1)}{\alpha} + \frac{\beta}{\alpha} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\alpha}\right)^{\beta}$$

$$\frac{d \ln L(\mu, \alpha, \beta||x)}{d\beta} = \frac{n}{\beta} + \sum_{i=1}^{n} \ln\left(\frac{x_i - \mu}{\alpha}\right) - \sum_{i=1}^{n} ln\left(\frac{x_i - \mu}{\alpha}\right)\left(\left(\frac{x_i - \mu}{\alpha}\right)^{\beta}\right)$$

equations are obtained. Since the likelihood equations are nonlinear functions, their solutions are obtained using numerical methods. Second derivatives are examined to show that the functions are minimum or maximum at the point or points that make their first derivatives zero.

1.2. Least Squares Method

It is particularly one of the methods used for estimating the parameters of a regression equation. Let there be any two variables, *X* and *Y*. If there is a relationship of the form *Y = f(X)* between these variables, the value of Y is determined for *X = x*. When any experiment is repeated under the same conditions and *X = x* is held constant, different outcomes may be observed for *Y*. Therefore, a relationship of the form $Y = f(x) + e$ is more meaningful. Here, it will be considered as a function in the form $f(x) = a + b\,x$. It is one of the methods used, especially for estimating the parameters of a regression equation.

Let there be *X* and *Y*. If there is a relationship between these variables in the form of *Y=f(X)*, then for *X=x*, the value of *Y* is known. When an experiment is conducted repeatedly under identical conditions, varying outcomes for *Y* may be observed while *X=x* remains fixed. Therefore, a relationship of the form Y=*f(X)*+e is more meaningful. Here, it will be considered as a function of *f(X)=a+bx*

The simple linear regression equation model involves dependent random variables Y_i, where $x_i, i = 1,2,3, \dots, n,$ and independent error terms e_i. These error terms have an expected value of zero and a variance of σ^2. The parameters α_0 and α_1 represent the intercept and slope, respectively.

$$Y_i = \alpha_0 + \alpha_1 x_i + e_i, \quad i = 1,2,3, \dots, n$$

One objective is to estimate the model parameters α_0 α_1 so that the sum of squared errors is the smallest.

$$minQ(\alpha_0, \alpha_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \alpha_1 X_i - \alpha_0)^2$$

For this purpose, by setting the first derivatives of $Q(\alpha_0, \alpha_1)$ with respect to α_0 and $\alpha_1$equal to zero, the values that make the function $Q(\alpha_0, \alpha_1)$ minimum or maximum are found. By setting these derivatives equal to zero

$$\frac{dQ(\alpha_0,\alpha_1)}{d\alpha_0} = -2\sum_{i=1}^{n}(Y_i - \alpha_1 X_i - \alpha_0)$$

$$\frac{dQ(\alpha_0, \alpha_1)}{d\alpha_1} = -2\sum_{i=1}^{n} x_i(Y_i - \alpha_1 X_i - \alpha_0)$$

equations are obtained. $\hat{\alpha}_0$ and $\hat{\alpha}_1$

$$n\ \hat{\alpha}_0\ + \hat{\alpha}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} Y_i \qquad \text{and}$$

$$\hat{\alpha}_0 \sum_{i=1}^{n} x_i\ + \hat{\alpha}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i Y_i$$

equations are obtained. These equations are known as *normal equations* in literature. Solutions of normal equations are

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_n)(Y_i - \bar{Y}_n)}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}, \quad \hat{\alpha}_0 = \bar{Y}_n - \hat{\alpha}_1 \bar{x}_n$$

is in the form of. These solutions are least squares estimators of the parameters α_0 and α_1. In order to show that they are minima, the second derivatives of the function $Q(\alpha_0, \alpha_1)$ must also be analyzed. (Akdi, 2014: 342).

Example: Obtaining Weibull distribution parameters with least squares estimator

Consider $X_1, X_2, \ldots, X_n$ as a sample of n units selected from a Weibull distribution

$$F(x; b, c) = 1 - exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)$$

The following function is obtained by taking the logarithm of both sides of the cumulative distribution function,

$$F(x) = 1 - exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)$$

$$1 - F(x) = exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)$$

$$-\ln(1 - F(x)) = \left(\frac{x}{\alpha}\right)^{\beta}$$

$$ln\left(-ln\left(1 - F(x)\right)\right) = \beta \ln x - \beta \ln \alpha$$

To obtain the least squares estimator of the parameters β_0 and β_1

$$Y = ln\left(-ln\left(1 - F(x)\right)\right); X = \ln x; \beta_1 = \beta; \beta_0 = -\beta \ln \alpha$$

can be written as a linear equation using the above transformations.

$$Y = \beta_1 X + \beta_0$$

For *i=1,2,...,n* the (i) th smallest rank statistic of $X_1, X_2, \ldots, X_n$ $X_{(i)}$. The range of means is used to estimate the values of the cumulative distribution function $F(X)$,

$$\hat{F}\left(x_{(i)}\right) = \frac{i}{n+1}$$

The regression coefficients β_0 and $\beta_1$are determined by minimizing the sum of error squares.

$$minQ(\beta_0,\beta_1)=\sum_{i=1}^{n}(Y_i-\beta_1X_i-\beta_0)^2$$

To obtain the least squares, an estimator of β_0 and β_1, the partial derivative of Q with respect to β_0 ve β_1 is taken and set equal to zero

$$\frac{dQ}{d\beta_0}=\bar{Y}-\widehat{\beta_1}\bar{X}=\frac{\sum_{i=1}^{n}\ln\left(-\ln\left(1-F(x_i)\right)\right)-\widehat{\beta_1}\sum_{i=1}^{n}\ln x_i}{n}$$

$$\frac{dQ}{d\beta_1}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{Y}\bar{X}}{\sum_{i=1}^{n}x_i^2-n\bar{X}^2}$$

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

Accordingly, the estimators $\hat{\alpha}$ and $\hat{\beta}$ can be written as follows.

$$\hat{\beta}=\widehat{\beta_1}=\frac{n\sum_{i=1}^{n}\ln x_i\ln\left(-\ln\left(1-F(x_i)\right)\right)-\sum_{i=1}^{n}\ln\left(-\ln\left(1-F(x_i)\right)\right)\sum_{i=1}^{n}\ln x_i}{n\sum_{i=1}^{n}\ln^2 x_i-\left(\sum_{i=1}^{n}\ln x_i\right)^2}$$

$$\widehat{\beta_0}=-\hat{\beta}\ln\hat{\alpha}=\frac{1}{n}\sum_{i=1}^{n}\ln\left(-\ln\left(1-F(x_i)\right)\right)-\hat{\beta}\sum_{i=1}^{n}\ln x_i$$

$$\ln\hat{\alpha}=-\frac{\sum_{i=1}^{n}\ln\left(-\ln\left(1-F(x_i)\right)\right)-\widehat{\beta_1}\sum_{i=1}^{n}\ln x_i}{\widehat{\beta}n}$$

$$\hat{\alpha}=exp\left(-\frac{\sum_{i=1}^{n}\ln\left(-\ln\left(1-F(x_i)\right)\right)-\widehat{\beta_1}\sum_{i=1}^{n}\ln x_i}{\widehat{\beta}n}\right)$$

1.3 Weighted Least Square Estimation

The weighted least squares estimators of β_0 and β_1 are $\hat{\beta}_0$ ve $\hat{\beta}_1$;

$Q(\beta_0,\beta_1)=\sum_{i=1}^{n}w_i(Y_i-\beta_1X_i-\beta_0)^2$

are the smallest values of the function. Weight factor w_i proposed by Bergman (1986).

$$w_i=\left[\left(1-\hat{F}\left(x_{(i)}\right)\right)ln\left(1-\hat{F}\left(x_{(i)}\right)\right)\right]^2,\ i=1,2,\ldots,n$$

is defined as.

Example: Obtaining Weibull distribution parameters with weighted least squares estimator

Consider $X_1, X_2, \ldots, X_n$ as a sample of n units selected from a Weibull distribution.

$Y=ln\left(-ln\left(1-F(x)\right)\right)$; $X=\ln x$; $\beta_1=\beta$; $\beta_0=-\beta\ln\alpha$ can be written as a linear equation using transformations.

$$minQ(\beta_0,\beta_1)=\sum_{i=1}^{n}w_i\left(ln\left(-ln\left(1-F(x)\right)\right)-\beta_1\ln x_i+\beta ln\alpha\right)^2$$

2. STATISTICAL CRITERIA USED IN THE COMPARISON OF PARAMETER ESTIMATES

After obtaining the estimators of the unknown parameters, the next important problem is determining which estimator is the best. Below, the statistical criteria used in the application areas of the study are explained

2.1. Mean

It is obtained by summing all the observed values and dividing the total by the number of observations. This value represents the mean of the parameter estimates obtained via Monte Carlo simulation, and is mathematically expressed as:

$\bar{\hat{\theta}} = \frac{1}{n}\sum_{i=1}^{n} \hat{\theta}_i$

2.2. Bias

Bias reflects the difference between the expected value of an estimator and the actual value, and is mathematically expressed as:

$Bias(\hat{\theta}) = E(\hat{\theta}) - \hat{\theta}$

If the *Bias=0*, the estimator is called an 'unbiased estimator'.

2.3. Variance

Variance provides information about the concentration of the estimators around the mean and is mathematically defined as;

$Var(\hat{\theta}) = \frac{1}{n-1}\sum_{i=1}^{n} \left(\hat{\theta}_i - \bar{\hat{\theta}}\right)^2$

Variance values can be used to compare the efficiency of unbiased estimators. An unbiased estimator with a smaller variance is considered more efficient according to the variance criterion.

2.4. Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) is a second-order metric that quantifies the error magnitude between predicted and actual values in a model. It is represented as the standard deviation of the prediction errors, providing a numerical measure of the model's ability to accurately align with the observed data.

Let $X_1, X_2, \ldots, X_n$ denote random variables, with their corresponding order statistics given by $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ and the observed ordered values as $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$. The following average rank is used to estimate the values of the cumulative distribution function.

$$\hat{F}\left(x_{(i)}\right) = \hat{y}_{\mathrm{i}} = \frac{i}{n+1}, \qquad i = 1,2,\ldots,n$$

Here i is the minimum index value of the observed variables $x_{(1)}, x_{(2)}, \ldots, and\ x_{(n)}$. The RMSE minimization function is given below.

$$\psi(\mu,\sigma) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

Example: Let $X_1, X_2, \ldots, X_n$ represent a sample of size n drawn from a Gumbel distribution with mean μ and standard deviation σ. The RMSE function of this distribution is as follows;

$$min\psi(\mu,\sigma) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i-\mu}{\sigma}\right)\right) - F\left(x_{(i)}\right)\right)^2}$$

2.5. Specification coefficient (R^2)

The specification coefficient indicates how much the independent variables explain the dependent variable. It is the best indicator of the goodness of fit and explanatory power of a linear model. As explained in the Least Squares method, let there be a relationship between the X and Y variables in the form of $Y = f(x) + e$:

The most general definition of the specification coefficient is as follows;

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - F(x_i))^2}{\sum_{i=1}^{n}(y_i - \bar{y}))^2}$$

The specification coefficient tends to increase as more independent variables (X_i) are added to the model. Since the degrees of freedom decrease, the prediction errors increase, creating a disadvantage when comparing the added X variables to the model.

Example: The R^2 function of the Gumbel distribution

Let $X_1, X_2, \ldots, X_n$ be a sample of size n drawn from a Gumbel distribution with mean μ and standard deviation σ. The R^2 function of this distribution is as follows;

$$\frac{\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i-\mu}{\sigma}\right)\right) - F\left(x_{(i)}\right)\right)^2}{\left(\sum_{i=1}^{n} exp - \left(exp - \left(\frac{x_i-\mu}{\sigma}\right)\right) - \frac{1}{n}\sum_{i=1}^{n} exp - \left(exp - \left(\frac{x_i-\mu}{\sigma}\right)\right)\right)^2}$$

2.6. Theil's inequality coefficient

Theil's inequality coefficient is one method used to measure the accuracy of predictions generated by certain models. Known as Theil's U coefficient, it indicates the degree of difference between a model's

predicted values and the corresponding observed values. The first inequality coefficient used by Theil, U_1 , is given by the equation *4.33*.

$$TIC(U_1) = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}y_i^2}}$$

The difference between the U_1 and U_2 coefficients is that U_1 has the prediction values in the denominator. The presence of the prediction term in the denominator causes U_1 to be bounded between 0 and 1 ($0< U_1<1$). In contrast, the U_2 coefficient, which does not include prediction values, has no finite upper limit. When U_1 approaches zero, it indicates a good fit, while approaching one indicates a poor fit. When $\hat{y}_i = y_i$, the U coefficient reaches zero, indicating perfect prediction. If there is an inverse relationship between the predicted and observed values or if one of the variables is zero, the TIC reaches its maximum value of 1. Therefore, the U_1 value alone does not provide sufficient information about the reliability of the applied prediction method (Bliemel, 1973).

Example: Theil's Inequality Coefficient for the Gumbel Distribution

Let $X_1, X_2, \ldots, X_n$ be a sample of size nnn drawn from a Gumbel distribution with mean μ and standard deviation σ. The Theil's inequality coefficient for this distribution is as follows;

$$\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right) - F\left(x_{(i)}\right)\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(exp - \left(exp - \left(\frac{x_i - \mu}{\sigma}\right)\right)^2\right)} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}F\left(x_{(i)}\right)^2}}$$

2.7. Mean Squared Error (MSE)

It is a criterion used to compare the efficiency of biased estimators and is mathematically expressed as;

$$MSE\left(\hat{\theta}\right) = Var\left(\hat{\theta}\right) + \left(Bias\left(\hat{\theta}\right)\right)^2$$

According to the MSE criterion, an estimator with a smaller MSE value is more efficient. For unbiased estimators, the variance and MSE values are equal.

2.8. Defect criterion (Def)

It is an important measure used to test the effectiveness of various methods for estimating a set of parameters. When multiple parameters are

to be estimated simultaneously, a criterion is used to compare the efficiency of estimators and is mathematically expressed as follows.

$$Def(\hat{\alpha},\hat{\beta},\hat{\Gamma}) = MSE(\hat{\alpha}) + MSE(\hat{\beta}) + MSE(\hat{\Gamma})$$

Since estimators with smaller MSE values are more efficient than others, estimators with smaller Def values are also more efficient.

2.9. Akaike Information Criterion (AIC)

It is used to compare the proposed models for a dataset, and the model with the lowest AIC value is always preferred. The AIC values (Akaike, Petrov, and Csaki, 1973) are mathematically expressed as:

$$AIC = -2ln\mathrm{L} + 2\mathrm{k}$$

The AIC is valid within the selected sample size and for future predictions.

2.10. E^2 criterion

The metric most commonly used for parameter estimation is the minimum mean squared error. In all cases, parameters are chosen that minimize the sum of the squared differences between the actual and predicted values (Nisbet, Elder, and Miner, 2009).

Abbasi, Niaki, Khalife, and Faize (2011) used the E^2 value, which is the sum of the squared differences between the real parameter values and the predicted parameter values, to measure the performance of their new algorithm and to compare the prediction results of the Levy Flight Artificial Bee Colony (LABC), Artificial Bee Colony (ABC), Genetic Algorithm, and LPSO algorithm used by Yonar (2020) for parameter estimation.

For the proposed models, the best solution among the Pareto points in the parameter space was obtained by selecting the point corresponding to the prediction points' lowest (min) E^2 value.

$$E^2 = \sum_{i=1}^{n} (\gamma - \hat{\gamma})^2$$

REFERENCES

Abbasi, B., Niaki, S., Khalife, M. and Faize, Y. (2011). Hybrid variable neighborhood search and simulated annealing algorithm to estimate the three parameters of the Weibull distribution, *Expert Systems with Applications*, 38 (1), 700-708.

Akaike, H., Petrov, B. N. and Csaki, F. (1973). Information theory and an extension of the maximum likelihood principle, Second International Symposium on Information Theory, Budapest, 267-281.

Akdi, Y. (2014). Matematiksel istatistiğe giriş (4. basım). Ankara: Gazi Kitabevi, 326-342.

Aster, R.C., Borchers, B., Thurber, C.H. (2019). Chapter Tw: Linear Regression. Parameter Estimation and Inverse Problems. 25-53.

Bliemel, F. (1973). Theil's Forecast Accuracy Coefficient: A Clarification. Journal of Marketing Research, 10(4), 444-446.

Kutz, J.N. (2023). Machine learning for parameter estimation. Applied Mathematics. 120 (12).

Kosmidis, I. (2014). Bias in parametric estimation: reduction and useful side-effects. Wiley Interdisciplinary Reviews: Computational Statistics, 6(3), 185-196.

Lei, Y. (2008). Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (Pinus tabulaeformis), Journal of Forest Science, 54(12), 566-571.

Nisbet, R., Elder, J. and Miner, G. (2009). Chapter 13- model evaluation and enhancement, handbook of statistical analysis and data mining applications (1st edition). New York: Academic Press, 285-312.

Sultana, T., Muhammad, F., & Aslam, M. (2019). Estimation of Parameters for the Lifetime Distributions. Journal of Reliability and Statistical Studies, 77-92.

Edge, M. D. (2019). Statistical thinking from scratch: A primer for scientists. Oxford University Press.92-109.

Yonar, A. (2020). Tek değişkenli ve çok değişkenli dağılımların parametre tahmini için metasezgisel yaklaşımlar, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, 81.