T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

REGRESYON MODELLLERİNDE DEĞİŞKENLERDEKİ EĞRİSELLİĞİN KARŞILIKLI BİLGİ KULLANILARAK BELİRLENMESİ

Zehra Zeynep ŞAHİNBAŞOĞLU

DOKTORA TEZİ

İstatistik Anabilim Dalı

İstatistik Programı

Danışman

Doç. Dr. Atıf Ahmet EVREN

Temmuz, 2023

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

REGRESYON MODELLLERİNDE DEĞİŞKENLERDEKİ EĞRİSELLİĞİN KARŞILIKLI BİLGİ KULLANILARAK BELİRLENMESİ

Zehra Zeynep ŞAHİNBAŞOĞLU tarafından hazırlanan tez çalışması 10.07.2023 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstatistik Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Doç. Dr. Ahmet Atıf EVREN
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Doç. Dr. Atıf Ahmet EVREN, Danışman
Yıldız Teknik Üniversitesi _____

Doç. Dr. Serpil KILIÇ DEPREN, Üye
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üye. Elif TUNA, Üye
Yıldız Teknik Üniversitesi _____

Prof. Dr. Deniz İNAN, Üye
Marmara Üniversitesi _____

Dr. Öğr. Üye. Erhan USTAOĞLU, Üye
Marmara Üniversitesi _____

Danışmanım Doç. Dr. Atıf Ahmet EVREN sorumluluğunda tarafımca hazırlanan Regresyon Modellerinde Değişkenlerdeki Eğriselliğin Karşılıklı Bilgi Kullanılarak Belirlenmesi başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Zehra Zeynep ŞAHİNBAŞOĞLU

İmza

Anneme

Ve

Babama

TEŞEKKÜR

Tez çalışmamda bilgi ve tecrübelerinden yararlandığım sayın hocam Doç. Dr. Atıf Ahmet EVREN'e, her zaman yanımda olan ve bana doktora sürecini ve mesleğimi sevdiren ve desteğini hiç esirgemeyen değerli hocalarım Dr. Öğr. Üyesi Elif TUNA ve Dr. Öğr. Üyesi Doğan YILDIZ olmak üzere, Yıldız Teknik Üniversitesi İstatistik Bölümü'nde bulunan tüm hocalarıma ve arkadaşlarıma, tezden üretilmiş yayında önemli katkıları bulunan değerli arkadaşlarım Arş. Gör. Dr. Mehmet Şamil GÜNEŞ'e ve Mert VEZNİKLİ'ye, tez süresince beni hep destekleyen Arş. Gör. Dr. Coşkun PARİM'e ve Arş. Gör. Egemen ÖZKAN'a teşekkürü bir borç bilirim. Ayrıca bu sürece başlamamda beni cesaretlendiren ve her zaman fikirlerinden yararlandığım değerli hocalarım Prof. Dr. Sevinç ERTÜRK'e ve Doç.Dr. Feyza KOLCU'ya teşekkür ederim.

Üzerimde çok büyük emekleri olan ve haklarını asla ödeyemeyeceğim saygıdeğer annem Şehma ŞAHİNBAŞOĞLU ile saygıdeğer babam Mahmut Fazıl ŞAHİNBAŞOĞLU'na, kardeşlerim Niyazi ŞAHİNBAŞOĞLU, Emir ŞAHİNBAŞOĞLU, Nil ŞAHİNBAŞOĞLU ve birtanem Ela ŞAHİNBAŞOĞLU'na müteşekkirim.

Zehra Zeynep ŞAHİNBAŞOĞLU

İÇİNDEKİLER

SİMGE LİSTESİ vii

KISALTMA LİSTESİ ix

ŞEKİL LİSTESİ x

TABLO LİSTESİ xi

ÖZET xii

ABSTRACT xiv

1 GİRİŞ 1

1.1 Literatür Özeti ... 1

1.2 Tezin Amacı ... 3

1.3 Hipotez ... 3

2 ENTROPİNİN TARİHÇESİ VE KULLANILDIĞI ALANLAR 5

2.1 Entropinin Tarihçesi ... 5

2.2 Entropinin Kullanıldığı Alanlar ... 6

3 ENTROPİNİN TANIMI VE ÇEŞİTLERİ 9

3.1 Entropinin Tanımı ... 9

3.2 Termodinamikte Entropi Kavramı ... 9

3.3 Bazı Entropi Çeşitleri ... 11

3.4 Bilgi Entropisi (Information Theory) ... 12

3.5 İstatistiksel Entropi ... 13

3.6 Karşılıklı Bilgi (Mutual Information) ... 26

3.7 Entropi ve Varyans Arasındaki İlişki ... 30

3.8 Karşılıklı Bilgi ve Entropi Arasındaki İlişki ... 31

4 İSTATİSTİKSEL ANALİZLER 34

4.1 Kutu Boyutu (Bin Size) Belirleme ... 34

4.2 Akaike Bilgi Kriteri (AIC) ... 35

4.3 Bayes Bilgi Kriteri (BIC) ... 36

4.4 AIC ve BIC Arasındaki Farklar ... 36

4.5 Bilgi Kazanım Fonksiyonu ... 37

4.6 Regresyon Analizinde Doğrusal Olamama Durumunu Belirlemeye Yarayan Bazı Yöntemler ... 38

5 UYGULAMA 48

5.1 Simülasyon Verileri Kullanılarak Eğrisellik Belirleme ... 48

5.2 Eğrisellik Belirleme Ölçülerinin Çevre Verileri Üzerine Uygulanması 56
5 SONUÇ VE ÖNERİLER **67**
KAYNAKÇA **69**
TEZDEN ÜRETİLMİŞ YAYINLAR **75**

SİMGE LİSTESİ

$H_\alpha(X)$	Alfa Dereceli Rényi Entropisi
Y_{ij}	Bağımsız Değişken
$E(.)$	Beklenen Değer
$G(p)$	Bilgi Miktarı Kazancı
k_B	Boltzmann Sabiti
α	Derece Değeri
F	F testi (Model Uyum Testi)
S_{top}	Farklı Makro Durumların Toplamı
$D_{KL}(p\|\|q)$	Göreli Entropi (Kullback – Leibler Mesafesi)
ξ	Hata terimi
S	İstatistiksel Entropi
I_i	Karşılıklı Bilgi Miktarı
$I(X,Y)$	Karşılıklı Bilgi Miktarı
$M(X,Y)$	Karşılıklı Bilgi Miktarı
D_S	Kesikli Durum için Shannon Entropisi Değeri
D_R	Kesikli Durum için Rényi Entropisi Değeri
D_T	Kesikli Durum için Tsallis Entropisi Değeri
$H(X\|Y)$	Koşullu Entropi
Δ_{bin}	Kutu (Optimum Boyut) Sayısı
S_i	Makro Durumları Oluşturan Mikro Durum Entropisi
$H_{\max}(X)$	Maximum Entropi
μ_j	Model Parametre Sayısı ($j=1,\ldots,c$)
$p_1,\ldots,p_N$	N tane Olayın Meydana Gelme Olasılıkları

$f(x_i)$	Olasılık Yoğunluk Fonksiyonu
$p(.)$	Olasılık Yoğunluk Fonksiyonu
$q(.)$	Olasılık Yoğunluk Fonksiyonu
S_{micro}	Ölçülemeyen Entropi Miktarı
$\beta_0, \ldots, \beta_p$	Parametreler
X	Rasgele Değişken
$H_R(X)$	Rényi Entropisi
df	Serbestlik Derecesi
λ_S	Shannon Doğrusallık Ölçüsü
H_S	Shannon Entropisi
$H_S(X)$	Shannon Entropisi
N	Sistemdeki Mikro Durum Sayısı (Olay Sayısı)
p_i	Sistemdeki i. durumda Bulunma Olasılığı
λ_R	Rényi Doğrusallık Ölçüsü
H_T	Tsallis Entropisi
$H_T(X)$	Tsallis Entropisi
λ_T	Tsallis Doğrusallık Ölçüsü

KISALTMA LİSTESİ

AIC	Akaike Bilgi Kriteri (Akakike Information Criteria)
SSE(R)	Azaltılmış Modelin Hata Kareleri Toplamı (Square Sum of Reduced Model)
BIC	Bayesyen Bilgi Kriteri (Bayesian/Schwarz Information Criteria)
BP	Breusch Pagan Testi
EKC	Çevresel Kuznets Eğrisi (Environmental Kuznets Curve)
IQR	Çeyrekler Arası Aralık (Interquartile Range)
LIC	Düşük Gelirli Ülkeler (Low Income Countries)
EKK	En Küçük Kareler
FD	Freedman – Diaconis Kuralı
GSYİH	Gayri Safi Yurtiçi Hasıla
GDP	Gayri Safi Yurtiçi Hasıla (Gross Domestic Product)
HKO	Hata Kareler Ortalaması
CO_2	Karbon Dioksit
RMSE	Karekök Ortalaması (Root Mean Square)
MAE	Ortalama Mutlak Hata (Mean Absolute Error)
MI	Karşılıklı Bilgi (Mutual Information)
KL	Kullback – Leibler Mesafesi
NAFTA	Kuzey Amerika Serbest Ticaret Antlaşması (North American Free Trade Agreement)
F	Model Uyum Analizi Testi
PDF	Olasılık Yoğunluk Fonksiyonu (Probability Densitiy Function)
LMC	Orta Gelirli Ülkeler (Middle Income Countries)
SSTO	Regresyon Modelindeki Toplam Değişim
SD	Serbestlik Derecesi
SSPE	Saf Hata Kareleri Toplamı(Square Sum Of Pure Error)
SSE(F)	Tam Modelin Hata Kareleri Toplamı (Square Sum of Full Model)
ANOVA	Varyans Analizi (Analysis of Variance)

ŞEKİL LİSTESİ

Şekil 3.1 Entropi ile Karşılıklı Bilgi Arasındaki İlişki 18
Şekil 3.2 Entropi ile Karşılıklı Bilginin Venn Diyagramı 32
Şekil 3.3 Entropi ile Karşılıklı Bilgi Arasındaki Fark 33
Şekil 5.1 Doğrusallık Durumunda λ ile α Değerleri Arasındaki İlişki 53
Şekil 5.2 Eğrisellik durumunda λ ile α Değerleri Arasındaki İlişki 53
Şekil 5.4 Çevresel Kuznets Eğrisi(EKC) 56
Şekil 5.5 Kalıntı Grafikleri 59
Şekil 5.6 Kuadratik Regresyon Modelinin Grafiği 60
Şekil 5.7 Çevre Verilerine Göre λ ve α Arasındaki İlişki 64

TABLO LİSTESİ

Tablo 3.1 Kesikli ve Sürekli Değişkenler için Entropi Hesaplamaları.....33
Tablo 4.1 Varyans Analizi Tablosu (ANOVA).....42
Tablo 5.1 Doğrusal Model için Parametre Değerleri.....48
Tablo 5.2 Doğrusal Model için ANOVA Tablosu.....49
Tablo 5.3 Kuadratik Model için Parametre Değerleri 50
Tablo 5.4 Kuadratik Model için ANOVA Tablosu 51
Tablo 5.5 Farklı α Değerlerine Göre λ_S, λ_R ve λ_T Değerleri.....51
Tablo 5.6 Simülasyon Verileri için Model Belirleme Kriterleri.....54
Tablo 5.7 Çevre Verileri için Kuadratik model Parametre Değerleri.....60
Tablo 5.8 ANOVA Tablosu.....61
Tablo 5.9 Uygun Kutu (Bölme) Sayısı.....62
Tablo 5.10 Farklı α Değerleri için EKC Verisinin λ_S, λ_R ve λ_T Değerleri.....62
Tablo 5.11 Çevre Verileri için Model belirleme Kriterleri.....65

ÖZET

Regresyon Modellerinde Değişkenlerdeki Eğriselliğin Karşılıklı Bilgi Kullanılarak Belirlenmesi

Zehra Zeynep ŞAHİNBAŞOĞLU

İsatistik Anabilim Dalı

Doktora Tezi

Danışman: Doç. Dr. Atıf Ahmet EVREN

Ekonomik büyüme ve çevre kirliliği arasındaki ilişkiyi analiz eden ampirik çalışmalar son 30 yıl içinde artış göstermiştir. Literatürde yoğun bir şekilde ekonomik büyüme ve çevresel kirliliğinin ilişkilendirilmesinde Çevresel Kuznets Eğrisi (ÇKE) hipotezinin geçerliliğinin test edilmektedir. Bu hipoteze göre gelir artışı ve çevre kirliliği arasındaki ilişki ters-U şeklindedir. Yani ekonomik büyümenin ilk aşamasında çevresel kirlilik ve tahribat artarken gelirin belli bir seviyeye ulaşmasından sonra çevresel kirlilik ve tahribat azalmaya başlamaktadır.

Bu tezin amacı, regresyon modellerinde kullanılan eğrisellik analizlerinin farklı bir yöntem olan karşılıklı bilgi kullanılarak tahmin edilmeye çalışılmasıdır. Çalışmanın örneklemi için 46 yıllık bir dönemde, az gelişmiş ve orta gelişmiş ülkelerin CO_2 emisyonları ile GSYİH değişkenleri kullanılmıştır. Bu verilerin ters U şeklinde bir dağılım gösterdiği görülmüş ve kuadratik olarak modellenerek, modelin eğriselliği karşılıklı bilgi ölçüleri kullanılarak açıklanmaya çalışılmıştır. Yapılan analizler sonucunda Rényi ve Tsallis Karşılıklı bilgi ölçülerinin Shannon

karşılıklı bilgi ölçülerinden verilerdeki eğriselliği daha iyi açıkladığı sonucuna varılmıştır.

Anahtar Kelimeler: Karşılıklı bilgi miktarı, EKC hipotezi, shannon entropisi, rényi entropisi, tsallis entropisi.

ABSTRACT

Determining The Curvature Of Variables In Regression Models Using Mutual Information

Zehra Zeynep ŞAHİNBAŞOĞLU

Department of Statistics

Doctor of Philosophy Thesis

Supervisor: Doç.Dr. Atıf Ahmet EVREN

Empirical studies analyzing the relationship between economic growth and environmental pollution have shown an increasing trend over the past 30 years. In the literature, there is a substantial focus on the association between economic growth and environmental pollution, specifically testing the validity of the Environmental Kuznets Curve (EKC) hypothesis. According to this hypothesis, the relationship between income growth and environmental pollution follows an inverted U-shaped pattern. In other words, environmental pollution and degradation increase in the early stages of economic growth, but start to decrease after income reaches a certain level.

The aim of this thesis is to attempt to predict the curvilinearity in regression models using a different method, Mutual Information. The study employs data from a 46-year period, utilizing CO_2 emissions and GDP variables from both underdeveloped and moderately developed countries as the sample. It is observed that these data exhibit a reversed U-shaped distribution, and the curvature of the model is explained using Mutual Information measures. The analysis results

indicate that Rényi and Tsallis Mutual Information measures provide a better explanation of the curvature in the data compared to Shannon Mutual Information measures.

Keywords: Mutual information, EKC hypothesis, shannon entropy, rényi entropy, tsallis entropy.

1
GİRİŞ

1.1 Literatür Özeti

Entropi, zaman ilerledikçe hayatın neden daha az değil ama daha karmaşık hale gelmesini açıklayan bir düzensizlik ölçüsü olarak tanımlanır. Entropi kavramı ilk olarak Alman fizikçi Rudolf Julius Emanuel Clausius tarafından ele alınmıştır [1]. Entropi, termodinamik yasasının ilk temel 4 kuralından birisidir ve her ne kadar literatürlerde “düzensizliğin bir ölçüsü” olarak tanımlansa da değişik disiplinlerde veri analizi ve ilgili konu hakkında yeterli bilgi sahibi olabilmek için kullanılan çok önemli ve gerekli bir yöntem olmuştur[2].

Murphy Yasası der ki, "Bir işi tamamlamak için birden fazla olasılık mevcutsa ve bu olasılıklardan biri istenmeyen sonuçlar veya felaketle sonuçlanacaksa; bu olasılık mutlaka gerçekleşecektir”. Bu söz Amerikalı mühendis Edward A. Murphy, jr. tarafından 1949 yılında, hayatın sorun çıkarma ve işleri zorlaştırma eğilimine göndermede bulunmak için söylenmiştir. Problemler kendiliğinden ortaya çıkarmış gibi görünse de çözümler her zaman insanların dikkatini, enerjisini ve çabasını gerektirir. Hayat, kendiliğinden sorunları çözmez gibi görünür. Aksine, insanların hayatları daha karmaşık hale gelir ve düzen yerine bozulmaya doğru yavaşça geriler. Peki bunun nedeni nedir? Murphy Yasası, sadece bir konuşma içinde kullanılan yaygın bir deyim olmasına rağmen, yaşanılan evrenin büyük güçlerinden biriyle ilişkilidir. Bu güç, dünyanın çalışma şekline o kadar temel bir şekilde etki eder ki, neredeyse insanoğlunun harcadığı her çabaya nüfuz eder. Karşılaşılan pek çok sorunu yönlendirir ve düzensizliğe yol açar. Çoğu literatürde bu etkinin tanımı “Entropi” olarak ifade edilir.

Entropi, insanoğlu ile birebir ilişkilidir. Örneğin insanoğlunun vücudundaki hücreler zamanla ölüyor ve bozuluyor, bir çalışan hata yapıyor, çevre, hava kirleniyor, işletmeler başarısız oluyor, suçlar işleniyor, düzen bozuluyor ve ilişkiler sona eriyor gibi. Her şey düzensizliğe doğru eğilim gösterir. Özellikle, termodinamiğin ikinci yasası, "herhangi bir izole veya kapalı sistemdeki entropinin

(düzensizlik derecesinin) zamanla artacağı (veya en azından aynı kalacağı)" şeklinde ifade edilir. Entropi, sadece düzensizlik seviyesini ölçen bir kavram gibi görünse de, günlük hayatın her yönünü etkiler [3-5].

Enerji dağılır ve sistemler kaosa doğru giderler. Düzensizlik arttıkça, bir şey o kadar entropik olarak kabul edilir. Başka bir deyişle, entropi evrenin hem makro düzeyde hem de mikroskobik düzeydeki düzensizlik seviyesinin bir ölçüsü olarak tanımlanabilir. Dolayısıyla insanlığın önemli sorunların yaşandığı bu zaman diliminin kritik bir süreç olduğu ve çevresel bozulmalar, kirlenmeler göz önüne alındığında diğer tüm yapılardaki bozulmalarla birlikte, tüm evrenin bir çöküşe doğru ilerlediği görülmektedir[6, 7]. Dolayısıyla entropinin hayatın her yönüyle, insanoğlunun yaşamıyla, evrenin sürekliliği ile yakından ilgisi olduğu görülmektedir.

1990'lardan bu yana çevre ile ekonomik büyüme arasındaki ilişki, ekonomistler arasında yoğun bir şekilde tartışılan bir konu haline gelmiştir. Bu çalışmalar genellikle ekonomik gelişme düzeyleri ile gelir seviyesi arasındaki ilişkiyi ve çevresel kirlilik üzerindeki etkisini incelemektedir. Çevresel Kuznets eğrisi hipotezi, çevresel kirliliğin ekonomik büyümenin erken aşamalarında arttığını göstermektedir. Ancak belli bir gelir seviyesini aştığında, ekonomik büyüme çevresel iyileştirmelere olanak sağlar [8, 9]. Dolayısıyla gelir seviyesi ile çevresel kirlilik arasındaki ilişkinin ters-U şeklinde olduğu söylemektedir. Bu ilişkinin şekillenmesinde üç faktör etkilidir: ölçek etkisi, bileşim etkisi ve teknoloji etkisi. Doğru, ölçek etkisi kavramı, üretimin artmasıyla birlikte daha fazla doğal kaynağın kullanılacağını ve çevrenin daha fazla zarar göreceğini ifade eder. Bu nedenle, ölçek etkisi ilişkideki pozitif kısmı temsil eder. Yani, ekonomik büyüme ve çevresel kirlilik arasındaki ilişkide, ölçek etkisi doğrultusunda bir artış görülür. Bu durum, üretim ölçeğinin büyümesinin çevresel etkilerini artırabileceğini gösterir. Bileşim etkisi, ekonomide imalat sektöründen hizmetler sektörüne geçişi sağlar ve bu da ekonomik büyümenin bir sonucu olarak çevresel etkileri değiştirir. Birincil fosil yakıtların yanması, atmosfere sera gazı salınımı ile bağlantılı en önemli çevresel etkiyi oluşturur. Kömür, petrol ve doğalgaz gibi birincil fosil yakıtların tamamı karbon içerir. Bu yakıtların yakılması sonucunda, karbon dioksit (CO_2) gibi sera gazları atmosfere salınır. Sera gazları, atmosferde birikerek güneş

ışınlarının yeryüzüne geri yansımasını engeller ve dünyanın ısınmasına yol açabilir, bu da iklim değişikliği ve çevresel etkilerine neden olabilir. Bu nedenle, fosil yakıtların kullanımının azaltılması ve yenilenebilir enerji kaynaklarının teşvik edilmesi, sera gazı salımının kontrol edilmesi ve çevre üzerindeki olumsuz etkilerin azaltılması açısından önemlidir. Bu yakıtların yanması sırasında karbon oksijenle birleşir ve başlıca sera gazı olan CO_2 oluşur[10, 11]. CO_2 atmosferde birikir ve bunun sonucunda dünya iklimi üzerinde olumsuz etkiler meydana gelmesi beklenir. Bu etkiler arasında biyolojik çeşitlilik kaybı, tropikal fırtına yoğunluğunun artması, deniz seviyesinin yükselmesi ve küresel ısınma yer almaktadır[12].

1.2 Tezin Amacı

Rasgele değişkenler arasındaki doğrusal olmayan bağımlılığın derecesini belirlemek, uzun süredir birçok istatistiksel problemin konusu olmuştur. Bununla birlikte, günümüzde değişkenler arasındaki doğrusal bağımlılığı araştırmak için çeşitli çalışmalar yapılmıştır ve yapılmaktadır. Toplam bağımlılığı değerlendirmek için C.Shannon karşılıklı bilgiyi önermiştir. Bu sayede bağımlılığın lineer ve doğrusal olmayan bileşenlerini ayırt etmek için birkaç yöntem geliştirilmiş ve bağımlılık ölçüsü çeşitleri arasında en kapsamlı ölçü olarak kabul edilmiştir.

Bu tezin amacı, parametrik doğası nedeniyle toplam bağımlılığı ölçmek için Rényi ve Tsallis karşılıklı bilgi ölçümlerinin kullanarak modelde ortaya çıkabilecek değişkenlerindeki doğrusal olamama durumunun tespitine çalışmaktır. İlk olarak, değişkenler arasındaki lineer ilişkiyi incelemek ve gidermek için bir kalıntı analizi kullanılmıştır. Daha sonra orijinal verinin Rényi ve Tsallis karşılıklı bilgi ölçüleri hesaplanmış, ardından doğrusal modelden elde edilen kalıntılar ile açıklayıcı değişken arasındaki ortak bilgi karşılaştırılarak, bu çalışmada geliştirilen ölçüye dayalı doğrusallık derecesi belirlenmeye çalışılmıştır.

1.3 Hipotez

Çevresel Kuznets Eğrisi (ÇKE) varsayımı, bireysel düzeydeki milli gelir (BMG) ile çevresel tahribat göstergeleri arasında ters U şeklinde bir ilişkinin mevcut olduğunu ifade etmektedir. Karbondioksit (CO_2), sera gazı emisyonlarının ana

faktörü olduğu için çevresel bozulmanın başlıca nedeni olarak kabul edilir. Dolayısıyla, GSYİH ile CO_2 arasında aynı ilişki varsayılır. Bu nedenle, EKC, ekonomilerin tarım temelli olarak başlayıp endüstriyel ve ardından hizmet temelli post-endüstriyel ekonomilere geçiş yaparken kullandıkları "ekonomik büyüme aşamalarının" bir göstergesidir.

Bu tezin amacı; EKC hipotezini desteklemeye yarayacak şekilde değişkenlerinde doğrusal olamama durumunun tespitinin Karşılıklı Bilgi ölçüsü kullanılarak belirlenmesidir. Kuznets eğrisinin kesin şeklinin belirlenmesi önemli olsa da, doğrusal olmama durumunun tespiti, EKC hipotezini desteklemeye yardımcı olacaktır. Rényi ve Tsallis karşılıklı bilgi ölçüleri hesaplanacak ve elde edilen sonuçlar Shannon Karşılıklı Bilgi ölçüleri ile karşılaştırılacaktır. EKC hipotezinin doğruluğunu göstererek, hipotezin ortaya koyduğu "şimdi büyü ve kirlet, sonra temizle" stratejisinin büyük çevresel maliyetlere yol açıp açmadığı sonucu araştırılacaktır.

Bölüm 2, Entropinin tarihçesi ve kullanıldığı alanlardan oluşmaktadır. Bölüm 3'te Entropi çeşitleri ve teorileri anlatılmış, Bölüm 4'te İstatistiksel Analizlere yer verilmiştir. Bölüm 5'te Uygulama gerçekleştirilmiştir. Uygulama iki bölümden oluşmaktadır. Birinci bölümünde simülasyon verileri kullanılarak karşılıklı bilgi ölçüleri hesaplanmıştır. Simülasyon ile elde edilen verilerin değişkenlerindeki doğrusallığı karşılıklı bilgi ölçüleri kullanılarak araştırılmıştır. Daha sonra EKC hipotezi uygulanması için, Doğu Asya ve Asya-Pasifik ülkelerinin (yüksek gelirli ülkeler hariç) 46 yıllık bir dönemde CO_2 emisyonları ile GSYİH değişkenleri analizde kullanılmıştır. Bu veriler ile Karşılıklı bilgi ölçüleri elde edilmiş ve modelin yapısı araştırılmıştır. Bölüm 6'da Sonuçlar ve Öneriler belirtilmiştir. Model ve veri arasındaki yapının değişkenlerinde doğrusal mı eğrisel mi olduğu karşılıklı bilgi ölçüleri ile elde edilmeye çalışılmıştır. Bu bulgular sayesinde analizi yapılacak olan verilerin ve modelin birbiri ile uyumu karşılıklı bilgi ölçüleri kullanılarak araştırılabilir ve eğrisellik araştırmaları için bundan sonraki çalışmalarda literatüre katkı sağlayabilir.

2
ENTROPİNİN TARİHÇESİ VE KULLANILDIĞI ALANLAR

2.1 Entropinin Tarihçesi

Entropi tarihi oldukça geniştir ve çeşitli disiplinlerde ele alınmıştır. Entropi kelimesi yunanca bir kelime olup, anlamı “belirsizlik” olarak ifade edilmiştir. İlk olarak, termodinamikte kullanılan entropi terimi, 19. yüzyılın başlarında Nicolas Léonard Sadi Carnot tarafından tanımlanmıştır [13].

Carnot, ısı motorları aracılığıyla enerjinin dönüşümünü inceleyerek, enerjinin sürekli bir şekilde değiştiğini ve bu değişimin sadece bir yönde olabileceğini öne sürmüştür. Bu değişimin termodinamik entropi kavramı ile tanımlandığı düşünülmektedir.

Daha sonra, 19. yüzyılın ortalarında, Clausius ve Kelvin gibi bilim insanları, entropiyi daha da geliştirerek, termodinamik sistemlerin davranışını anlamak için önemli bir araç haline getirdiler. Bu çalışmaların sonucunda, termodinamiğin ikinci yasası ortaya çıktı ve entropi artışının evrensel bir yasa olduğu keşfedildi[14].

Ancak, termodinamik entropi, belirli bir sistemdeki düzensizliği ölçebilirken, termodinamik dışındaki sistemlerdeki düzensizliği ölçemiyordu. Bu nedenle, 19. yüzyılın sonlarında, Ludwig Boltzmann tarafından geliştirilen istatistiksel mekanik ile entropi kavramı daha da geliştirildi. Boltzmann, termodinamik entropi kavramını istatistiksel entropi ile bağlantılı hale getirdi ve entropiyi moleküler düzeydeki hareketlerle ilişkilendirdi[15].

Boltzmann'ın entropi konusundaki çalışmaları, başlangıçta birçok bilim insanı tarafından eleştirildi. Ancak, Boltzmann'ın ölümünden sonra, çalışmaları daha fazla kabul gördü ve istatistiksel mekaniğin gelişmesine yol açtı. Günümüzde, istatistiksel mekanik ve entropi, birçok farklı alanda, özellikle fizik, kimya,

matematik, bilgisayar bilimi ve mühendislikte önemli bir rol oynamaktadır[16, 17].

Entropinin gelişimi aşağıdaki gibi özetlenebilir:

i. Rudolf Julius Emanuel Clausius, bir Alman fizikçi ve matematikçidir. "Mechanical Theory of Heat" adlı kitabında, termodinamik bir sistemin enerji miktarını ölçmek için kullanılan bir ifade olan "entropi" terimini tanımlamıştır. Clausius, belirli süreçlerin geri döndürülemez olduğunu tahmin eden termodinamik bir özellik olarak da bahsetmiştir.

ii. Ludwig Boltzmann, 1844 ile 1906 yılları arasında yaşamış bir bilim insanıdır ve bilim camiasında istatistiksel mekaniğin babası olarak anılmaktadır. Boltzmann ilkesi ve denklemi gibi konuları da içeren istatistiksel mekanik alanında yaptığı çalışmalar, bilim dünyasına büyük katkılar sağlamıştır. Özellikle Maxwell ile geliştirdiği Maxwell-Boltzmann istatistiği, yaptığı araştırmalar arasında en dikkat çekicilerinden biridir. Ayrıca Boltzmann dağılımı, Boltzmann olasılığı ve mikro durumlar da bir sistemi doğru bir şekilde hesaplamak için kullanılan temel fikirler arasındadır.

iii. Claude E. Shannon, 1948'de ortaya koyduğu teorik çalışmasıyla, bilgi transferi sırasında gürültü içeren iletişim kanallarında meydana gelen bilgi kaybındaki belirsizliği ölçmek amacıyla "entropi" terimini kullanmıştır [18, 19].

iv. Matematikte ters alma problemlerinin tahmini sırasında, entropi kullanımı 1957 yılında Jaynes tarafından keşfedilmiştir. Jaynes, ters alma problemlerinin çözümlerinin entropiyi maksimum yaparak elde edilebileceğini, bunu yaparken de Shannon'un entropi ölçüsünden yararlanılabileceğini göstermiştir. Jaynes'in bu yaklaşımı literatürde "maksimum entropi yöntemi" olarak adlandırılmaktadır [20]

v. Entropi, ilerleyen senelerde Golan ve meslektaşları tarafından istatistik ve ekonometrik modellerin tahmin edilmesinde kullanılmıştır [21].

2.2 Entropinin Kullanıldığı Alanlar

Shannon 1948 yılında “Bilgi Entropisi Teorisi” üzerine çalışmalarda bulunmuştur. 1951 yılında Kullback ve Leibler’ın bilgi teorisine katkıları ile 1956 yılında Lindley “Minimum Çapraz Entropi” ilkesini geliştirmiştir [18].

Karşılıklı bilgi teorisinin geliştirilmesi ile günümüzde entropi kavramı farklı alanlarda kullanılmaya başlanmıştır. Bu alanlardan bazıları aşağıda verilmiştir [13, 22-24]:

- Biyoloji
- Genetik
- Kimya
- Fizik
- Matematik
- Kuantum mekaniği
- İstatistiksel Mekanik
- Güvenilirlik Analizi
- Termodinamik
- Olasılık teorisi
- Elektronik ve İletişim Mühendisliği
- Görüntü İşleme
- Fotogrametri
- Harita yapımı
- Yöneylem araştırması
- Ekonomi
- Psikoloji
- Sosyal bilimler
- Veri toplama ve depolama
- Jeoloji
- Coğrafya
- Jeofizik
- Hidrolik

- Ulaşım
- Mühendislik

3
ENTROPİNİN TANIMI VE ÇEŞİTLERİ

3.1 Entropinin Tanımı

İstatistiksel entropi, James Clerk Maxwell ve Ludwig Boltzmann gibi fizikçiler tarafından istatistiksel mekaniğin bir dalı olarak ortaya çıkmıştır. Boltzmann, denge halindeki bir sistemin düzensizliği veya rastlantısallığının bir ölçüsü olarak entropi kavramını geliştirmiştir. Ancak, Boltzmann'ın çalışmaları bilimsel topluluk tarafından hemen kabul edilmemiş ve Ernst Mach ve Wilhelm Ostwald dahil diğer fizikçiler tarafından ciddi bir muhalefetle karşı karşıya kalmıştır. Boltzmann'ın fikirleri etrafındaki tartışmalar, 1906'daki trajik intiharına neden olmuştur. Bu tarihten sonra istatistiksel mekanik alanı gelişmeye devam etmiş ve Boltzmann'ın entropi hakkındaki fikirleri daha geniş bir çapta kabul görmeye başlamıştır. 20. yüzyılın ortalarına gelindiğinde, entropi kavramı termodinamiğin temel prensiplerinden biri haline gelmiştir ve 2. Yasası kullanılarak ortaya çıkmıştır.

3.2 Termodinamikte Entropi Kavramı

Entropi kavramını açıklamak için öncelikle termodinamik kavramından bahsetmek gerekmektedir. Termodinamik, mantıksal ve matematiksel bilimlerden biri olarak kabul edilir. Enerji ve enerjinin dönüşümleri ile ısı etkilerini içeren sistemlerdeki değişim ve denge eğilimlerini inceleyen bir bilim dalıdır. Fiziğin temel yapı taşlarından biri olan termodinamiğin dört yasası bulunmaktadır.

Termodinamikte “entropi” kavramı; zaman simetrisinin bozulduğu, tersine çevrilebilir ve tersine çevrilemez süreçler arasındaki farklıları tanımlamak için kullanılmıştır [5]. Sonlu bir değişim sırasında, termodinamik değişkenler sistemin her noktasında homojen bir şekilde değişiyorsa ve denge durumunda değişiyorsa (ya da sistem ara durumlarında çevreyle ısısal denge halinde ilerliyorsa), bu sürece *“tersinir (reversible) süreç”* denir. Diğer süreçler ise *“tersinemez (irreversible) süreçler”* olarak adlandırılır. Entropi kavramı, ikinci yasa tarafından tanımlanmıştır. Hacim, basınç, sıcaklık, enerji, entropi gibi kavramlar termodinamiğin temel kavramları arasında yer almaktadır. Termodinamik

kavramları, istatistik ile birleştirilerek kullanılabilmektedir. Bu yöntemlerden birisi de "entropi yasası"dır. Entropi, termodinamik sisteminin düzensizliği veya belirsizliğini ölçen bir termodinamik kavramdır. Entropi arttıkça, sistemdeki düzensizlik artar ve enerjinin kullanılabilirliği azalır. Bu nedenle, entropi genellikle bir sistemin bozulma derecesi olarak da düşünülebilir [25].

Termodinamiğin birinci yasası, enerjinin tür değiştirme işlemine tabi olabileceğini ancak yaratılamayacağını ve yok edilemeyeceğini belirtir. Bir sistemin iç enerjisindeki değişim, sisteme dışarıdan verilen ısı eksi sistemin çevresi üzerinde yaptığı iş miktarına eşittir. Bu yasa, "her şey iş yapmak için enerjiye ihtiyaç duyar" anlamını taşımaktadır [7].

Termodinamiğin ikinci yasası, her şeyin zamanla yıpranacağını, eskiyeceğini ve çürüyeceğini ifade eder. Bu yasa, sıcak su içindeki şekerin neden çözündüğünü ve neden asla tekrar birleşip şeker oluşturmadığını açıklar. Bu yasanın anlaşılabilmesi için, "entropi" kavramının tanımı gerekmektedir. Entropi, dağınıklık ve düzensizliğin bir ölçüsüdür. Düzensizlik arttıkça sistemin entropisi artar. Aynı zamanda, "entropi", bir işi yerine getirmek için gereken enerji harcama potansiyeli olarak da tanımlanabilir. Bu durumda, sistemin enerji harcama potansiyeli ne kadar yüksekse, entropisi o kadar düşüktür. Bu yasaya göre, dışarıdan enerji verilmediği sürece entropi her zaman azalır [26].

Termodinamiğin üçüncü prensibi, bir sistem sıcaklığının mutlak sıfıra yaklaştıkça entropisinin istikrar kazandığı veya entropideki değişimin sıfır olduğunu vurgular. Bu noktaya "mutlak entropi" denir. Eğer bir sistem tamamen düzenli ise, entropisi sıfırdır.

Termodinamiğin dördüncü yasasına göre eğer iki sistem birbiriyle etkileşim içindeyken aralarında ısı veya madde alışverişi olmuyorsa bu sistemler dengededir.

Bir sistem izole edildiğinde enerji artışı sıfır olacaktır. Bu durumda sistem entropisi geri dönüşümsüz süreçler nedeniyle artmaya devam eder ve ulaşabileceği maksimum değere, yani termodinamik denge durumuna ulaşır. Denge durumunda, tüm geri dönüşümsüz süreçler durur. Bir sistemin dışarıyla entropi değişimi başladığında, genel olarak denge durumundan uzaklaşır ve entropi

üreten geri dönüşümsüz süreçler devreye girer. Entropi değişimi ısı ve madde alışverişi nedeniyle gerçekleşir [5].

3.3 Bazı Entropi Çeşitleri

Bir sistemin entropisi, kitle, enerji, hacim, momentum, yük veya kimyasal türlerin atom sayısı gibi yaygın termodinamik özelliklerden biridir. Yaygın termodinamik özellikler, bu özelliklere sahip bir denge halindeki bir sistemin, ikiye eşit parçaya bölündüğünde yarıya inerken, yoğun özellikler değişmeden kalır. Örneğin, hacim, kütle, molekül sayısı ve entropi gibi değişkenler, yaygın özelliklere örnek olarak gösterilebilirken, sıcaklık ve basınç yoğun özelliklere örnek olarak verilebilir. Bir sistem denge halinde olduğunda, birim kütle başına düşen entropi üretim hızı, dış kısıtlamalarla uyumlu olarak minimum olmalıdır. Termodinamikte, entropi, sistemdeki düzensizlik derecesini ölçen ve sistem durumunu tanımlayan bir ölçü olarak kullanılır.

Kapalı ve izole bir sistemde, entropi her zaman en büyük değere doğru artma eğilimindedir. Enerji kaybı ve diğer nedenler, sistemi düzensiz ve kaotik hale getirir. Bu nedenle entropi, bir sistemdeki karmaşa veya düzensizlik miktarının bir ölçüsü olarak yorumlanabilir. Enerji dönüşümünde genellikle enerji kaybı olarak adlandırılan bir durum meydana gelir. Geri dönüşü olmayan enerji akışının miktarı, korunmayan ve her zaman artan, yani geri döndürülemez olan entropi değişiminin bir ölçüsüdür. Entropi artışı, düzensizliğin artması anlamına gelir[17].

Termodinamikte, bir sistem evrenin herhangi bir bölümünü temsil eder ve bu bölüme büyük miktarda parçacık dahildir. Geri kalan kısım ise çevre olarak adlandırılır. Termodinamik, uygulanan kısıtlamalara bağlı olarak dört farklı sınıfa ayrılan sistemlere sahiptir. Sistemlerin sınıflandırılması madde, basınç, enerji ve ısı transferi gibi faktörlere dayanır. Termodinamiğin ikinci ilkesi ise entropiyi ifade eder. Zaman içinde, entropinin sabit kalacağı ya da artacağı ancak azalmayacağı belirtilir. Bu ilke, sadece izole veya adiabatik (ısı veya kütle kaybı veya kazancının olmadığı) sistemlere uygulanır [3].

Bir sistem mikro durum ve makro durum olmak üzere iki durumdan oluşur. Bir sistem ve çevresi birbirinden izole edilebilir ve böyle bir sistemle çevresi arasında

ısı veya madde değişimi olmaz. Böyle bir sistem sonunda termodinamik anlamda bir denge durumuna ulaşır, bu da sistemin durumunda önemli bir değişiklik olmayacağı anlamına gelir. Makro durum, basınç, hacim, sıcaklık gibi termodinamik değişkenleri gözlemleyerek tamamen tanımlanabilen bir termodinamik durumdur. Dolayısıyla, klasik termodinamikte, bir sistem, deneysel olarak gözlemlenebilen özellikleri ve sistem ile çevresi arasındaki etkileri içeren makroskobik durumuyla tanımlanır. Termodinamik, sistemde var olabilecek çeşitli mikro durumlar arasında ayrım yapmaz ve bu nedenle atomik ölçekte çalışan mekanizmalarla ilgilenmez [27].

3.4 Bilgi Teorisi (Information Theory)

Bilimsel ve insanlık tarihinin en büyük icatları ve keşiflerinin çoğu, enerjinin dönüşümü, transferi ve depolanması veya bilginin iletimi ve depolanmasıyla ilişkilidir. Örneğin, ateşin ve suyun kullanımı, kaldıraç ve tekerlek gibi basit makinelerin icadı, buhar motorunun icadı, elektrik ve atom enerjisinin keşfi, enerjiyle bağlantılıyken, tarih öncesi dönemde konuşmanın ortaya çıkışı ve insanlık tarihinin başlangıcında yazının icadı, ardından kâğıt, baskı, telgraf, fotoğraf, telefon, radyo, televizyon ve nihayetinde bilgisayar ve bilgisayar ağı gibi bilgiyle ilgilidir. Birçok icat ve keşif her iki amaç için de kullanılabilir. Ateş, örneğin, pişirme, ısıtma ve sinyal iletimi için kullanılabilir. Elektrik ise enerji ve bilgi iletimi için kullanılabilir.

Enerji ve bilgi çeşitli formlarda mevcuttur. Enerjinin gerçek formunu (mekanik, termal, kimyasal, elektriksel ve atomik vb.) ve bilginin gerçek içeriğini göz ardı edersek, geriye saf bir miktar kalır. Enerji korunumu prensibi, geçtiğimiz yüzyılın ortalarında formüle edilmiş ve geliştirilmiştir, oysa bilginin miktarı 1940'lardan sonra incelenmiştir. Bilginin saf miktarıyla, optik kablo normal elektrik telefon hattından çok daha fazla bilgi iletebilir ve genel olarak iletişim kanalının kapasitesi, bilgi miktarı hızı açısından belirtilebilir. Bilginin nicel ölçümü iletişim çalışmalarından kaynaklanmasına rağmen, o kadar temel bir kavram ve yöntemdir ki istatistik, fizik, kimya, biyoloji, yaşam bilimleri, psikoloji, bilişsel bilim, sinirbilim, bilgisayar bilimleri, ekonomi, işletme araştırmaları, dilbilim, felsefe gibi birçok alanda yaygın olarak uygulanmıştır.

İletişim sistemlerinde bilginin nicel ölçümünün incelenmesi 1920'lerde başlamıştır. 1924'te Nyquist, sabit bir hattın hızıyla bir telgraf devresi üzerinden iletişim hızının, mesajı kodlamak için kullanılan $W = k \log m$ akım değeri sayısının logaritmasına bağlı olduğunu göstermiştir: k , burada bir sabittir. 1928'de Hartley, iletişimin tüm formları için bunu genelleştirmiş ve her bir sembol seçimi için kullanılabilir sembol sayısını belirtmiştir. Hartley açıkça bilgi için nicel ölçümün sorununa değinmiş ve bunun psikolojik faktörlerden bağımsız (veya objektif) olması gerektiğini belirtmiştir. 1948 yılında, Shannon ünlü "Bir İletişim Kuramı" isimli makalesini yayımlamıştır. Bu makalede mesajların istatistiksel yapısını incelemiştir. Daha sonra Nyquist ve Hartley'in logaritmik bilgi ölçeğini olasılıksal bir logaritmaya genişletmiştir.

3.5 İstatistiksel Entropi

Sistemdeki parçacıkların hareketi istatistiksel mekaniğin konusudur. Bu nedenle, sistem mikro durumu ile ilgilidir. Sistemdeki parçacık sayısının fazlalığından dolayı, her bir parçanın kendine has durumunu tek tek incelemek mümkün değildir. Bu nedenle, mikro durumların dağılım fonksiyonlarını karakterize etmek için istatistiksel yöntemlere başvurulması daha uygun bir yaklaşım olarak görülmüştür.

Termodinamik durumların birçok gerçekleşme olasılığı vardır. Sistemdeki mikro durum sayısı N ile gösterildiğinde, istatistiksel entropi şu şekilde ifade edilir:

$$S = k_B \ln N \tag{3.1}$$

(3.1) denkleminde k_B, Boltzmann sabiti olarak adlandırılır ve bu denkleme "Boltzmann entropisi" adı verilir. Boltzmann sabiti genellikle J/K (joule / kelvin) biriminde ifade edilir. Yaklaşık değeri ise $k_B \approx 1.380649x10^{-23} \approx 1.380649x10^{-23}$ J/K 'dir.

Boltzmann entropisi, termodinamiğin ikinci yasasının temelindeki entropi artışı prensibini, sistemin mikro durumlarındaki belirsizlik ve rastgeleliğin artması olarak yorumlamaktadır. Bu prensip, termodinamiğin temel yasalarından biri olarak kabul edilir ve birçok alanda uygulama alanı vardır. Örneğin, istatistiksel

mekaniğin birçok alanında Boltzmann entropisi kullanılır ve moleküler dinamik simülasyonları gibi birçok uygulama için temel bir kavramdır [6, 7].

Denklem (3.1)'de sistem tüm mikro durumların aynı gerçekleşme olasılığına sahip olduğunu belirtir. Buna göre N arttıkça S'de artar ve en olası maksimum durumu temsil eder. Gerçekleşme durumunun maksimum sayısını verir. Bu ölçü termodinamik durumun olasılığı olarak kabul edilmiştir.

Boltzmann, termodinamik alanında yaptığı çalışmalar kapsamında, gazların mikro durumlarına ilişkin olasılık dağılımlarının tahmin edilmesinde entropi ölçüsünü kullanmıştır. Bu bağlamda, gaz moleküllerinin yönlü hızlarındaki değişimlere bağlı olarak ortaya çıkan olasılık dağılımının belirlenmesi problemine odaklanarak çalışmalarını sürdürmüştür [28].

Denklem (3.1) çeşitli sistem toplulukları için genelleştirilebilir. Sistem farklı mikro durumlar için kullanılabilir. Bir sistemin N tane, farklı, eşit olasılığa sahip mikro durumu olabileceği varsayıldığında, bu mikro durumların ayrıntılarını doğrudan ölçmek zordur. Bu mikro durumlar farklı gruplara ayrılır (bu gruplar makro durumlar olarak adlandırılır) ve i. makro durumunda n_i mikro durum bulunur. Makro durumlar deneysel olarak daha kolay ayırt edilebilir çünkü bazı makroskobik, ölçülebilir bir özelliğe karşılık gelirler [29]. Her makro durumdaki tüm mikro durumların toplamının, mikro durumların toplam sayısına eşit olması gerekmektedir, bu durum Denklem (3.2)'de gösterilmiştir:

$$\sum_i n_i = N \tag{3.2}$$

N, tüm mikro durumların oluşturduğu olay sayısıdır. Sistemin i. makro durumda bulunma olasılığı aşağıdaki gibi ifade edilir:

$$p_i = n_i/N \tag{3.3}$$

$\sum p_i = 1$ olduğu bilindiğinden toplam entropi $S_{top} = k_B \ln N$ olarak ifade edilir. Ancak sistemdeki mikro durumlar hakkında olası tüm bilgiler elde edilemeyeceği için ve kolayca ölçülemeyeceği için bu sonuca ulaşmak zordur. S_{top}; farklı makro

durumların toplamının entropi miktarını ifade eder. S ölçülebilen entropi miktarı olmak üzere;

$$S_{top} = S + S_{micro} \tag{3.4}$$

Farklı mikro durumlarda meydana gelen, ölçülemeyen entropi miktarı S_{micro} Denklem (3.5) ile ifade edilir:

$$S_{micro} = \sum_i p_i S_i \tag{3.5}$$

Burada S_i makro durumları oluşturan mikro durum entropisidir. $S_i = k_B \ln n_i$ olarak ifade edilir. p_i belirli bir makro durumun olasılığı olmak üzere;

$$S = S_{top} - S_{micro} \tag{3.6}$$

$$= k_B(\ln N - \sum_i p_i \ln n_i) \tag{3.7}$$

$$= k_B \sum_i p_i(\ln N - \ln n_i) \tag{3.8}$$

Logaritmik özelliklerden yararlanıldığında $\ln N - \ln n_i = -\ln(n_i/N) = -\ln p_i$ elde edilir. Bu durumda Gibbs entropisi aşağıdaki gibi ifade edilir [29]:

$$S = -k_B \sum_i p_i \ln p_i \tag{3.9}$$

İstatistiksel entropiyi oluşturmak için tüm olası durumları N ile bölerek istatistiksel entropi oluşturulur ve S'in ölçümü mikro durumların sayısına bağlıdır.

Entropi, deney öncesi olaylarla ilgili belirsizliğin bir ölçüsü olarak kabul edilir. Deney tamamlandığında belirsizlik giderilir. Bu, deneyin olaylar hakkında bilgi sağladığı anlamına gelir ve bu bilgi, olayların entropisiyle eşdeğer olan belirsizlik anlamına gelir.

3.6 Shannon Entropisi

1948 yılında, Shannon teorik bilgi ve basit bilgilendirici entropisi kavramlarını ortaya attı [18]. Bu kavramlar "Shannon entropisi" olarak adlandırılmaktadır. Shannon, bilgi miktarı belirli olduğu zaman, belirsizliğin azaldığını ya da ortadan kalktığını ifade etmiştir. Bu şekilde "belirsizlik ölçüsü" elde edilir. Bu durumda entropi "belirsizlik miktarı" olarak da isimlendirilir. Bilgi miktarı ise belirsizliğin tersi olarak ifade edilir [30]. Shannon entropi teorisini bilgi miktarını ya da belirsizlik miktarını ölçebilmek için geliştirmiştir. Bir kesikli rasgele X değişkeninin N tane olayın $p_1, p_2, \ldots, p_N$ olasılıkları ile meydana gelmesi Shannon entropisinin temelini inşa eder.

$x_i,\ i = 1, 2, \ldots N$ tane kesikli durumun olasılıkları $p_1, p_2, \ldots, p_N$ olsun. Burada $P = \{p_1, p_2, \ldots, p_N\}$ her değer bir olayı temsil eder. Bu durumda;

$$-\log\left(\frac{1}{p_1 p_2 \cdots p_N}\right) = -\log(\frac{1}{p_1}) - \log(\frac{1}{p_2}) - \ldots - \log(\frac{1}{p_N}) \tag{3.10}$$

Denklem (3.10), N tane olayın birlikte gerçekleşmesini gözlemleyerek elde edilen bilgiyi ifade eder. Denklem (3.10) da verilen serinin ortalama bilgi miktarı, bu serinin beklenen değeri olarak ifade edilir ve aşağıdaki gibi gösterilir:

$$H_S = -\sum_{i=1}^{N} p_i \log p_i \tag{3.11}$$

Denklem (3.11)'de ifade edilen "H_S" değeri kesikli değerler alan x_i değişkeninin Shannon tarafından 1948 yılında açıklanan entropi değeridir [31]. Shannon'ın 1948'te "bilgi entropisi" olarak kullandığı Denklem (3.11)' te verilen ifade ile denklem (3.9)'da verilen termodinamik entropi ifadesi aynı formdadır ve bu ilişki "Boltzmann ve Gibbs" tarafından geliştirilmiştir. Dolayısıyla bazı araştırmacılar "Shannon entropisi" olarak kullanılan bu ifadeyi "Boltzmann – Gibbs entropisi" ya da "Shannon-Boltzmann-Gibbs entropisi" olarak da kullanmaktadırlar [32].

H_S entropi denkleminde $X = (x_1, x_2, \ldots, x_N)$ rastgele değişken olmak üzere, $P = \{p_1, p_2, \ldots, p_N\}$; N boyutlu X örnekleminin olasılık dağılımıdır.

$P(X = x_1) = p_1,\ P(X = x_2) = p_2,\ \ldots,\ P(X = x_N) = p_N$ olarak gösterildiğinde $\sum_{i=1}^{N} p(x_i) = 1$ olmak üzere, $H(X)$'e veya $H(P)$'ye *"Shannon Entropisi"* veya *"Bilgi Entropisi"* denir [33]. Kesikli değerler alan x_i değişkeni için, olasılık yoğunluk fonksiyonu olmak $p(x_i)$ olmak üzere entropi değeri aşağıdaki gibi hesaplanır:

$$H_S(X) = H(P) = -K\sum_{i=1}^{N} p(x_i)\log([\,p(x_i)\,]) \tag{3.12}$$

Sürekli değerler alan x_i değişkeni için, yoğunluk fonksiyonu $f(x_i)$ olmak üzere Shannon entropi değeri aşağıdaki gibi hesaplanır:

$$H_S(X) = H(P) = -\int_{-\infty}^{+\infty} f(x_i)\log f(x_i)dx \tag{3.13}$$

Denklem (3.12)'ye göre aşağıdaki ifadeler yazılabilir:

- $0\log 0 = 0$ olduğunda ve $x \to 0$ olduğu durumda $x\log x \to 0$ olacağından dolayı sıfır olasılıklı terimler eklemek entropi değerini değiştirmez.
- Entropi sadece X değişkeninin olasılık dağılımına bağlıdır, değişkenin kendi değerine bağlı değildir.

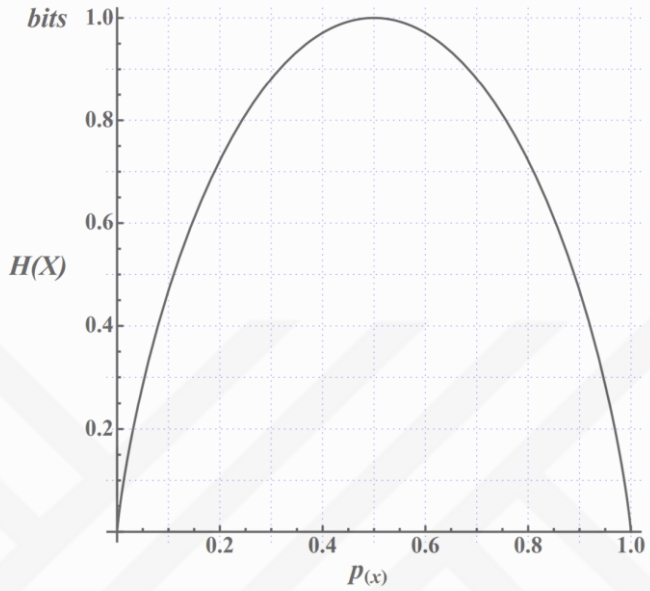
k, Boltzmann sabiti kullanılan logaritma tabanına bağlı bir değerdir. Bu değere *"entropi birimi"* de denilebilir. Eğer farklı entropi birimleri kullanılırsa logaritma tabanları da değişir. Örneğin taban;

- 2 olarak kullanılırsa birim değeri "bits (binary digits)" olarak isimlendirilir [4, 34].
- e olarak kullanılırsa birim değeri "nat (natural)" olarak isimlendirilir.
- 10 olarak kullanılırsa "decibels", "logit" ya da "docit" olarak kullanılır [34].

Genel olarak $k = 1$ olarak alınır ve $H(x)$ ifadesi aşağıdaki gibi yazılır:

$$H_S(X) = H(P) = -\sum_{i=1}^{N} p(x_i)\log[\,p(x_i)\,] \tag{3.14}$$

Bu ifade rastgele X değişkenindeki bilgi miktarını ya da $P(x)$ olasılık dağılımındaki bilgi miktarını ölçer. Bu ölçüm, bir değişkenin belirsizliğinin bir ölçüsüdür ve tek bir değerdeki ortalama bilgi miktarını gösterir. Bu sayede, Shannon entropisiyle bilgi miktarı arasındaki ilişki açığa çıkar.



Şekil 3.1 Entropi ile Karşılıklı Bilgi Arasındaki İlişki[35]

Entropi fonksiyonun özellikleri aşağıda belirtilmiştir:

i. $p_i = 1$ ise $H_S(X) \geq 0$ olur. Entropi değeri hiçbir zaman negatif olmaz, her zaman pozitif değerler alır:

$$0 \leq H(x) \leq \log N \tag{3.15}$$

Birçok olasılık dağılımının ve entropilerin bu iki limit değer arasında yer aldığı söylenebilir: Entropi minimum olduğunda $H_{\min} = 0$ olur ve sistem tamamen düzenlidir, herhangi bir belirsizlik söz konusu değildir. H'nin alabileceği en büyük değer $\log N$'dir. Diğer yandan maksimum entropi ($H_{\max}(x)$) maksimum belirsizlik ölçüsü olarak adlandırılır ve sistemdeki bütün olaylar aynı olasılık ile gerçekleştiğinde düzensizlik maksimum olur.

ii. Eğer dağılım tekdüze (uniform) ise, entropi en yüksek seviyede olur ($\sum_{i=1}^{n} p_i = 1$).

iii. $p = 1/2$ olduğunda Shannon entropisi maksimum değerini alır, en fazla bilgi değeri elde edilir.

iv. $p = 1$ ya da $p = 0$ olduğunda entropi sıfır değerini alır. İlgilenilen durum ile ilgili bilgi miktarı minimum düzeyde elde edilir. Bir sonuç neredeyse kesin olduğunda, yani olasılığı (p) neredeyse 1 olduğunda, durum çok az bilgi sağladığından sonuç tahmin edilebilir. Bilgi teoride tanımlandığı gibi, bilgi kavramı bir olayla ilişkili şaşkınlık veya belirsizlik derecesiyle yakından ilişkilidir.

Bilgi teorisine göre, elde edilen bilgi miktarı bir olayın olasılığıyla ters orantılıdır. Bir olay çok olası olduğunda ($p \cong 1$), bu olayın beklenildiği ve çok fazla yeni veya beklenmedik bilgi sağlamadığı anlamına gelir. Başka bir deyişle, bir sonucun ne kadar kesin olduğu, taşıdığı bilginin miktarını azaltır.

Bunu anlatmak için, adil bir şekilde atılan bir para örneğini düşünsün. Eğer para tarafsız ise yazı (Y) veya tura (T) olmak üzere iki olası sonucu vardır. Her bir sonucun olasılığı $p = 0,5$ olur. Eğer sonuç yazı ise, elde edilen bilgi miktarı $log_2(1/p) = log_2(1/0.5) = 1$ bit olur. Ancak, sonuç kesin ise, örneğin yazıysa ve $p = 1$ ise, elde edilen bilgi $log_2(1/p) = log_2(1/1) = log_2(1) = 0$ bit olur. Bu, bir sonuç neredeyse kesin olduğunda ($p \approx 1$), ortalama bilgiye çok az katkıda bulunduğunu gösterir. Özetlemek gerekirse, yüksek bir olasılığa sahip bir sonuç $p \approx 1$), şaşkınlık veya belirsizlik eksikliği nedeniyle az bilgi sağlar [34, 36].

Eğer X değişkeni rastgele bir değişken değilse, olasılık kesin bir değer olan “1” değerini alacaktır ve diğer alternatif değerlerin olasılıkları “0” olacaktır. Bu durumda Denklem (3.13) ile hesaplanan entropi fonksiyonu minimum değerini alır. Bu değere göre sistem tertiplidir. Diğer taraftan bütün x_i'lerin eşit olduğu, değişkenin tekdüze(uniform) dağılım gösterdiği ($p_i = 1/N$, $i = 1, 2, \ldots, N$) varsayıldığında, tüm olasılıklar her değer için eşittir. Bu durumda Denklem (3.13) aşağıdaki gibi ifade edilebilir ve entropi fonksiyonunun maksimuma ulaştığı durumu gösterir ve maksimum entropiyi ifade eder:

$$H(X) = H_{\max}(X) = \log N \tag{3.16}$$

Herhangi bir sürecin/değişkenin entropisinin Denklem (3.16)'te tanımlanan sınırlar dahilinde her zaman pozitif değerler aldığı söylenir.

Olayların gerçekleşme olasılıkları sıfıra yaklaştığında, bu süreçten daha fazla bilgi elde edilir. Belirsizlik veya düzensizlik en yüksek olduğunda, maksimum entropi ortaya çıkar.

3.6.1 Rényi Entropisi

Parametrik entropi ailesi, Alman fizikçi Alfred Rényi tarafından 1961 yılında Shannon entropisinin matematiksel bir genelleştirmesi olarak tanıtılmıştır. Rényi, Shannon Entropisi, Kapur entropisi ve diğer entropi çeşitlerini genelleştirerek Rényi entropisini oluşturmuştur.

Rényi'nin entropisinin geliştirilmesiyle birlikte bilgi ölçülerinin parametrik genelleştirilmesi ortaya çıkmıştır. Shannon'ın entropisinin tanımında ele alınan aksiyomatik yaklaşım, Rényi tarafından 1961 yılında yeniden incelenmiştir. Bu yeniden incelenmede, zincir kuralı yerine iki ayrı kesikli dağılımın çarpımının ilgili rastgelelik ölçülerinin toplamına eşit olduğu daha kırılgan bir rastgelelik ölçüm gerekliliğiyle değiştirilmiştir. Başka bir deyişle, iki bağımsız olayın birlikte gözlemlenmesinin, toplamlarına ilişkin bilgi kadar bilgi sağlaması gerektiği ifade edilmiştir [37].

Belirsizlik miktarı ya da $p_i \geq 0$ ve $\sum_{i=1}^{N} p_i = 1$ olan $P = (p_1, p_2, \ldots, p_n)$ olasılık dağılımının entropisi, $H(p) = H(p_1, p_2, \ldots, p_n)$ ile ölçülür ve $p_1, p_2, \ldots, p_n$ deney sonuçlarına sahiptir. Burada $H(p,1-p)$, $0 \leq p_i \leq 1$ aralığında p'nin sürekli bir halidir.

1961 yılında Rényi kesikli değerler alan x_i değişkeni aşağıdaki entropi formülünü tanımlamıştır:

$$H_\alpha(p_1, p_2, \ldots, p_n) = -\frac{1}{1-\alpha} log\left(\sum_{i=1}^{n} p_i^\alpha\right) \tag{3.17}$$

Denklem (3.17)'de $\alpha > 0$ ve $\alpha \neq 1$'dir. Bu denkleme "P *dağılımının* α *dereceli Rényi entropisi*" denir.

x_i değişkeninin sürekli bir değişken olması durumunda Rényi entropisi aşağıdaki gibi hesaplanır:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \int_x f^\alpha(x) dx \tag{3.18}$$

Sürekli değerler alan x_i değişkeni için, yoğunluk fonksiyonu $f(x)$ olmak üzere Denklem (3.17) ile Denklem (3.11) aynı olduğu görülmektedir [38]. Dolayısıyla Shannon entropisi $\alpha \to 1$ Rényi entropisi Shannon entropisine yakınsar. Rényi ve Shannon entropilerine bakıldığında, Shannon entropisinin Rényi entropisinin özel bir hali olduğu görülür, Shannon entropisi Rényi entropisinin limit alınmış halidir [39].

$$H_R(X) = \frac{1}{1-\alpha} \log \frac{\sum_{i=1}^{N} p_i^a}{\sum_{i=1}^{N} p_i} \tag{3.19}$$

Shannon ve Rényi entropileri arasındaki temel ayrım, logaritmanın ifade içindeki konumudur. Shannon entropisinde, olasılık yoğunluk fonksiyonunun ($\log[p(x_i)]$) ağırlıklandırılması teriminde yer alırken, Rényi entropisinde ise logaritma, olasılık yoğunluk fonksiyonunun içermiş olduğu α kuvvetine sahip bir terimin dışında bulunur. Ayrıca Rényi 'nin entropi ölçüsü, α parametresi nedeniyle çok daha esnektir [40].

3.6.1.1 Rényi Entropisinin Özellikleri

- Rényi entropi ölçüsü $H_\alpha(p_1, p_2, \ldots, p_n)$, p_K tane tüm olasılığın sürekli bir fonksiyonudur. Olasılık dağılımındaki bir değişiklik entropi değerinde de değişiklik meydana gelmesine neden olacaktır [41].
- Rényi entropisi negatif değer alamaz; $H_R(X) \geq 0$.
- α'nın farklı değerleri için Rényi entropisi korelasyonludur.

- Rényi entropisi sürekli değerler alır ve sınırlıdır, $H_R(X) = H_\alpha(0,\ 1) = H_\alpha(1,\ 0) = 0$.
- $\alpha < 1$ için Rényi entropisi konkavdır. $\alpha > 1$ için entropi konkav ya da konveks değildir.
- $H(1/n,\ \ldots.,\ 1/n)$, n'nin artan bir fonksiyonudur. Seçenekler arttığında belirsizlik ve düzensizlik artar, dolayısıyla entropi de artar.
- Rényi entropisi toplanabilirdir. Eğer $p = (p\ ,\ldots,p_N\ )$ ve $q = (q_1,\ldots,q_N)$ iki bağımsız olasılık dağılımı ise ve bunların birleşik olasılık dağılımları $p \cdot q$ ile temsil ediliyorsa, toplanabilirlik özelliği yandaki gibi sağlanmış olur: $H(p \cdot q)\ =\ H(p)\ +\ H(q)$

3.6.2 Tsallis Entropisi

Constantino Tsallis 1988 yılında p olasılıkla gerçekleşen bir olayın bilgi miktarındaki kazancı aşağıdaki formül ile açıklamıştır [42]:

$$G(p) = \Delta I = \frac{k}{q-1}\left[\left(1 - p_i^{q-1}\right)\right] \tag{3.20}$$

k, pozitif bir tamsayı temsil eder; q ise herhangi bir sayıdır. Tsallis entropisi kazanç fonksiyonunun beklenen değeri olarak ifade edilebilir ve kesikli değerler alan x_i değişkeni aşağıdaki gibi yazılır [43]:

$$H_T = E(\Delta I) = \frac{k}{q-1}\sum_{i=1}^{N} p_i\left[\left(1 - p_i^{q-1}\right)\right] \tag{3.21}$$

Sürekli değerler alan x_i değişkeni aşağıdaki gibi hesaplanır [38]:

$$H_T = E(\Delta I) = \frac{k}{q-1}\int_x f^{\alpha}(x)dx - 1 \tag{3.22}$$

H_T değeri her durum için sıfıra eşit veya sıfırdan büyüktür. Bu durum Shannon entropisinin (ya da bazı kaynaklara göre Boltzmann- Gibbs Entropisinin) genelleştirilmiş halidir [44].

Tsallis entropisine q entropisi ismi de verilmektedir ve aşağıdaki özelliklere sahiptir [45]:

• Her durum için tüm olasılıklar eşit olduğu zaman Tsallis entropisi maksimum değerini alır [46].

• $N=1$ olduğu durumda entropi sıfır olur.

• q değeri "1" olduğu zaman Tsallis entropisi Shannon Entropisine yakınsar ($\alpha>0, \alpha\neq 1$).

3.6.3 Bileşik (Joint) Entropi

Bilginin kavramı, tek bir tanımlamayla tam olarak açıklanması mümkün bir kavram değildir. Bununla birlikte, herhangi bir olasılık dağılımı için, bilgi ölçüsü olarak kabul edilen bir nicelik olan entropi tanımlanabilir. Bu kavram, bir rasgele değişkenin diğer bir rasgele değişken hakkında ne kadar bilgi içerdiğini ölçen karşılıklı bilgiyi tanımlamak için genişletilir. Entropi, daha sonra bir rasgele değişkenin kendi bilgisi haline gelir.

Aynı olasılık uzayında tanımlanmış *X* ve *Y* iki bağımsız rasgele değişken olsun. X , $\{x_1, x_2, \ldots, x_n\}$ değerlerini alırken, Y değişkeni $\{y_1, y_2, \ldots, y_m\}$ değerlerini alsın. Denklem (3.23) ve (3.24)'te X ve Y rasgele değişkenlerinin birleşik olasılık dağılımı ($p(x_i, y_j)$) :

$$p_i = \sum_{j=1}^{m} p_{ij} = P(X = x_i), \qquad i = 1,\ldots,n,\ p_1 + \ldots p_n = 1 \tag{3.23}$$

$$q_j = \sum_{i=1}^{n} p_{ij} = P(Y = y_i), \qquad i = 1,\ldots,m,\ q_1 + \ldots q_m = 1 \tag{3.24}$$

X ve Y rasgele değişkenlerinin bileşik entropisi ise aşağıdaki gibi hesaplanır:

$$H_{mn}(X,Y) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(x_i, y_j) \log p(x_i, y_j) = -\sum_{j=1}^{m}\sum_{i=1}^{n} p_{ij} \log(p_{ij}) \tag{3.25}$$

Eğer X ve Y değişkenleri bağımsız ise bileşik entropi, değişkenlerin entropilerinin toplamına eşit olur. Bağımsızlık durumunda $p(x,y) = p(x)q(y)$ yazılabilir ve bileşik entropi aşağıdaki şekilde ifade edilebilir:

$$H(X,Y) = H(X) + H(Y) \tag{3.26}$$

Birleşik entropi aynı zamanda aşağıdaki gibi yazılabilir:

$$H(X,Y) = -E\log p(X,Y) \tag{3.27}$$

Denklem (3.27) yandaki gibi açıklanabilir: "Bir rastgele değişken verildiğinde, aynı zamanda bir başka rastgele değişkenin koşullu entropisi, koşullu dağılımların entropilerinin beklenen değeri olarak tanımlanır ve bu beklenen değer, koşullandırma rastgele değişkeni üzerinden ortalama alınarak hesaplanır.

3.6.4 Koşullu (Conditional) Entropi

Rastgele X değişkenin bir başka rastgele değişken olan Y'ye göre entropisine *"koşullu entropi"* denir. $p(x|y)$ koşullu olasılığı ifade etmektedir ve $p(x|y) = \frac{p(x,y)}{p(y)}$ ile hesaplanmaktadır. Denklem (3.28)'da koşullu entropi ifade edilmiştir:

$$H(X|Y) = -\sum_{y} p(y) \sum_{x} p(x|y) \log p(x|y) \tag{3.28}$$

Diğer bir ifadeyle, "bir değişkenin diğerine göre koşullu olasılığının logaritmasının beklenen değeri, negatiftir" denir. Zincir kuralı, koşullu entropi ile bileşik entropi arasında ilişkilendirmek açısından önemlidir [13].

- X ve Y gibi rastlantısal değişkenlerin bileşik entropisi, zincir kuralına göre, bir değişkenin entropisi ile diğer değişkenin bu değişkene göre koşullu entropisinin toplamına eşittir. Bu teorem *"zincir kuralı teoremi"* olarak adlandırılır ve aşağıdaki gibi ifade edilir:

$$H(X,Y) = H(X) + H(Y|X) \tag{3.29}$$

$$H(X,Y) = H(Y) + H(X|Y) \tag{3.30}$$

- X ve Y gibi rastlantısal değişkenler bağımsız ise:

$$H(Y|X) = H(Y) \tag{3.31}$$

$$H(X|Y) = H(X) \tag{3.32}$$

- X ve Y gibi rastlantısal değişkenlerin koşullu entropisi, değişkenlerin kendi entropileri toplamına eşittir ya da bu toplamdan küçüktür.

$$H(X,Y) \leq H(X) + H(Y) \tag{3.33}$$

3.6.5 Göreli (Relative) Entropi ve Kullback – Leibler Mesafesi

Relative entropi ilk olarak Kullback ve Leibler tarafından tanımlanmıştır [47, 48]. Kullback Leibler uzaklığı, çapraz entropi, bilgi ayrışımı ve ayrımcılık için bilgi gibi çeşitli isimlerle bilinir ve Csiszar [2] ve Amari [49] tarafından detaylı bir şekilde incelenmiştir.

X, kesikli bir dağılıma sahip rastgele bir değişken olsun. Bu değişkenin olasılık yoğunluk fonksiyonu $p(x) = P\{X = x\}$ olsun.

Göreli entropi (relative entropy), $D_{KL}(p\|q)$ ifadesi "Göreli (Relative) Entropiyi ifade eder. İki dağılım arasındaki mesafe Göreli Entropi kullanarak ölçülebilir. Entropi, Denklem (3.34)'te ifade edildiği gibi, olasılık oranının beklenen logaritmasıdır.

Gerçek dağılımın p olduğu durumda, dağılımın q olduğunu varsaymanın etkinliğinin ölçüsüdür.

İki olasılık kütle fonksiyonu $p(x)$ ve $q(x)$ arasındaki *"Göreli entropi (Relative Entropy) veya Kullback-Leibler uzaklığı ($D_{KL}(.)$)"* aşağıdaki şekilde tanımlanır:

$$\mathrm{D}_{KL}(p\|q) = \sum p(x) \log \frac{p(x)}{q(x)} = E_p \log \frac{p(X)}{q(X)} \tag{3.34}$$

Entropi sadece $p = q$ olduğu durumlarda sıfır değerini alır. Entropi miktarı ne kadar az ise, kullanılan değişkenlerin dağılımları birbirlerine yakındır fakat entropi miktarı fazla ise dağılımlar birbirlerinden o kadar farklıdır. Matematiksel olarak bu ifade aşağıdaki gibi ifade edilir:

$$D_{KL}(p \| q) \neq D_{KL}(q \| p) \tag{3.35}$$

Göreli entropi dağılımlar arasındaki uzaklık ölçüsü olarak tanımlanabilir [50].

Göreli entropi aynı zamanda entropi değerinin maksimum entropi değerine oranı olarak da ifade edilebilir:

$$H_R = \frac{H(X)}{H_{\max}(X)} = \frac{H(X)}{\log N} \tag{3.36}$$

Göreli (Relative) entropi bazı kaynaklarda "boyutsuz entropi" olarak da adlandırılmaktadır [17].

X, sürekli bir dağılıma sahip rastgele bir değişken olduğunda $p(x)$ olasılık yoğunluk fonksiyonunun entropisi denklemin integrali alınarak aşağıdaki gibi hesaplanır:

$$H(X) = -\int \log p(x) p(x) dx = -E_X\left[\log(p(x))\right] \tag{3.37}$$

$$I(X) = -\int\int I(x,y) p(x,y) dx dy = -E_{XY}\left[I(x,y)\right] \tag{3.38}$$

$$D_{KL}(p(x) \| q(x)) = \int p(x) \log \frac{p(x)}{q(x)} dx = E_P\left[\log \frac{p(x)}{q(x)}\right] \tag{3.39}$$

Yukarıda verilen Denklem (3.37),(3.38),(3.39)'ta kullanılan $p(.)$ ve $q(.)$ ifadeleri olasılık yoğunluk fonksiyonlarını ifade ederken, $I(x,y)$ karşılıklı bilgiyi (MI), $E(.)$ ise beklenen değeri gösterir [40].

3.7 Karşılıklı Bilgi (Mutual Information- MI)

Bilgi teorisinin gelişimine bakıldığında Nyquist (1924) ve Hartley (1928), iletişim kanalları için matematiksel bir ölçüm geliştirmek için çalışmışlardır [51, 52]. Daha

sonra, Shannon (1948) ve Wiener (1948), İkinci Dünya Savaşı'ndan kısa bir süre sonra bilgi teorisinin doğmasını sağlayan önemli katkılarda bulunmuşlardır [53].

Bilgi teorisi alanında temel bir kavram olan entropi, veri sıkıştırma ve iletiminin temel yasasını olasılıksal bir yaklaşımla düşünme konusunda devrim niteliğinde bir yenilik olarak Claude E. Shannon tarafından tanıtılmıştır. Shannon'ın entropisi, bilgi teorisi literatüründe son yıllarda gerçekleşen en önemli ilerlemelerden biri olarak kabul edilmektedir. Fikirleri büyük bir heyecan yaratmış ve iletişim, veri depolama ve işleme, bilgisayarlı görüşme, tıbbi görüntüleme, güvenilirlik teorisi ve diğer bilgi teknolojilerinin gelişimi için temel oluşturmaya devam etmiştir, iletişim veya kanallar üzerinden sinyal iletimi problemlerini incelemek ve çözmek için bir araç olarak tanıtılmıştır [37, 54].

Rastlantı değişkeni farklı değerler aldığında bir olayın gerçekleşmesi ile ilgili belirsizlik söz konusu olur. Bir durum ile ilgili belirsizlik söz konusu ise "bilgi miktarı" bu olay gözlemlenerek belirlenebilir. Bir olay yüksek ihtimalle meydana gelirse, olaydan elde edilecek bilgi miktarı minimum olur. Diğer taraftan, bir olayın geçekleşme olasılığı az olduğunda elde edilen bilgi miktarı artar ya da bir olayın gerçekleşmesi belirsizlik miktarını azaltır. Aynı şekilde eğer bir olayın gerçekleşmesi kesin ise bu olayın gerçekleşmesi ile daha az miktarda bilgi ortaya çıkar ya da bu süreci tanımlayabilmek için daha fazla miktarda bilgiye ihtiyaç duyulur. Bu durum entropi, bilgi miktarı, belirsizlik arasında bir bağlantı olduğunu ortaya koyar.

Karşılıklı bilgi, iki olasılık dağılımı arasındaki mesafeyi ölçen daha genel bir nicelik olan *"Göreli entropinin (Relative Entropy)"* özel bir durumudur. Tüm bu tanımlamalar birbirleriyle yakından ilişkilidir ve bir dizi basit özelliği paylaşırlar [48].

$P(x)=\{p_1,\ p_2,\ \ldots,\ p_N\}$ olasılık dağılımına sahip, $X:\{x_1,\ \mathrm{x}_2,\ \ldots,\ x_N\}$ bir rasgele değişken düşünülsün. Değişken gözlemlendiğinde x_i değerini alıyorsa, bu süreçten bir bilgi elde edilir. Elde edilen bu bilgi miktarı I_i ile gösterilir ve olasılığın logaritmik büyüklüğü olarak ifade edilir:

$$I_i = -\log p_i = |\log p_i| \tag{3.40}$$

Gözlem yapılmadan önce değişken hakkındaki belirsizlik miktarını ölçebilmek için belirsizlik miktarı ile bilgi miktarı arasında bir ilişki kurulabilir. Belirsizlik miktarı gözlem yoluyla elde edilmesi beklenen bilginin ortalama değeridir. Dağılımın entropisi Denklem (3.41) ile verilmiştir. Bilgi miktarı olarak tanımlanır ve aşağıdaki gibi ifade edilir:

$$H = \sum_{i=1}^{N} p_i I_i = -\sum_{i=1}^{N} p_i \log(p_i) = \sum_{i=1}^{N} p_i |\log p_i| \tag{3.41}$$

Bu olasılık dağılımın entropisi, gözlemden elde edilen beklenen bilgi miktarını gösterir. $0 \le p_i \le 1$ olduğu için, H her zaman pozitif değer alır.

Eşit olasılıklara sahip iki olasılık dağılımı olduğu düşünülsün. Bunlar $p_i = p$ ($i = 1, 2, \ldots, N$) ve $q_i = q$ ($i = 1, 2, \ldots, M$) ile ifade edilsin. Bu iki dağılımdan elde edilen bilgi miktarı arasındaki fark aşağıdaki şekilde ifade edilir:

$$I = \Delta H = H_p - H_q = \log_2 p - \log_2 q = \log_2 (p/q) \tag{3.42}$$

Burada H_p bilgi miktarını ya da $\{p_i \ , \ i = 1, 2, \ldots , N\}$ olasılığının entropisini, H_q ise aynı şekilde bilgi miktarını ya da $\{q_i \ , \ i = 1, 2, \ldots , M\}$ olasılığının entropisini gösterir. İki rastgele değişken arasındaki karşılıklı bilgi, birbirleri hakkında iletilen karşılıklı bilgi miktarını ölçer. Başka bir ifade ile Y hakkında bilgi edinmekten kaynaklanan X hakkındaki belirsizlikteki ortalama azalmayı ölçer ve "I" ile ifade edilir:

- Eğer $q > p$ ya da $M < N$ ise $I = \Delta H > 0$ ise entropi artar. Belirsizlik ortadan kalkar, olası sonuçların artması ile bilgi kaybolur.
- Diğer taraftan, bilgi miktarının artması, $q < p$ ya da $M < N$ olması durumunda ve $I = \Delta H < 0$ ise gerçekleşir. Elde edilecek olası sonuçların azalması ile elde edilecek bilgi miktarı artar.

Maksimum entropi ($H_{\max}(X)$) "maksimum belirsizlik ölçüsü" olarak adlandırılır ve sistemdeki bütün olaylar aynı olasılık ile gerçekleştiğinde düzensizlik maksimum olur. $H_{\max}$ ile karşılaştırıldığında bilgi miktarı aşağıdaki gibi hesaplanabilir:

$$I = H_{\max} - H = log\ n + \sum_{i=1}^{n} p_i\ log\ p_i = \sum_{i=1}^{n} p_i\ log\left(\frac{p_i}{1/n}\right) = \sum_{i=1}^{n} p_i\ log\left(\frac{p_i}{q_i}\right) \quad (3.43)$$

Denklem (3.43) de $q_i = 1/n$ ile ifade edilir. $\{q_i\}$ prior dağılım, $\{p_i\}$ ise posterior dağılım olarak kabul edilir. Her olay farklı bir olasılıkla gerçekleşir.

X ve Y gibi iki rasgele değişken tanımlansın. Bu değişkenlerin olasılık yoğunluk fonksiyonları $p(x)$ ve $p(y)$ ile birleşik olasılık yoğunluk fonksiyonu ise $p(x,y)$ ile gösterilsin. İki değişken arasındaki karşılıklı bilgi hesaplaması değişkenlerin kesikli veya sürekli olması durumunda aşağıdaki gibi hesaplanır [13]:

- x_i değişkeni kesikli ise aşağıdaki gibi hesaplanır:

$$I(X,Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (3.44)$$

- x_i değişkeni sürekli ise karşılıklı bilgi miktarı aşağıdaki gibi hesaplanır:

$$I(X,Y) = \int_{x,y} f(x,y) \log \frac{f(x,y)}{f(x)f(y)} \quad (3.45)$$

Bilginin ölçüsü entropidir, entropi rastgelelik veya karışıklığın bir ölçüsüdür. Fiziksel sistemler giderek daha az düzenli hale gelmeye eğilimlidir. Eğer bir sistem son derece düzenli ve rastgelelikle açıklanamıyorsa, bu sistemin entropisi (bilgi miktarı) düşüktür.

Bir olayla ilgili belirsizlik ne kadar yüksek ise o olay ile ilgili bilgi elde etme miktarı artacaktır. Belirsizlik bir olayın farklı değerler alabildiğini ifade eder. Büyük olasılık ile meydana gelmesi beklenen değer daha az bilgi taşır. Bir olayın kesin bir değer ile ortaya çıkması "belirsizliğin ölçüsü (the measure of uncertainty)" olarak ifade edilir. Shannon (1948) entropiyi bir olayın alternatif değerlerinin

olasılıklarının beklenen değeri olarak belirtmiştir. Elde edilen bilgi miktarı dolaylı olarak ölçülen belirsizlik ya da entropi miktarındaki azalma olarak ifade edilmiştir. Bu durumda belirsizliğin bir bilgi miktarı olarak tanımlanabileceği belirtilir [55].

Olasılıksal belirsizlik sonuçların olasılığı ile ilgili olduğu gibi, geçmişten günümüze ve geleceğe doğru artması sonucu ile zaman ile de bağlantılıdır. Ayrıca verilerin, değişkenlerin durumları, tahmin, planlama, simülasyon gibi konularla da ilgilidir. Bazı değişkenlerin bilinmeyen durumları kullanılarak diğer değişkenlerin durumları ifade edilmeye çalışılabilir. Eğer değişkenler arasında tek bir ilişki varsa model deterministiktir(kesindir). Aksi durumda ise model kesin değildir ve belirsizlik söz konusu olur. Belirsizlik doğru modeli elde etmek ile ilgilidir (doğru modelin kurulması ile ilgilidir- veriyi en iyi temsil eden modelin elde edilmesi ile ilgilidir).

3.8 Entropi ve Varyans Arasındaki İlişki

Deneysel verilerin analizinde, belirsizlik miktarının ölçümü ya da bilgi kaybı veya kazanımı ölçümü için genellikle varyans kullanılır. Entropi, olasılık yoğunluk fonksiyonunun daha yüksek düzeyli istatistikleri hakkında bilgi içererek Gauss dağılımlarındaki varyansın rolünü genelleştiren bir belirsizlik ölçüsüdür [56] [57]. Dolayısıyla entropi ölçüsünü belirlemek mümkün olan her durumda varyans hesaplaması da gerçekleştirilebilir ama bunun tersi söz konusu değildir. Ayrıca varyans örnek boyutu küçük olduğunda kullanılması çok uygun bir ölçü değildir. Bu durumlarda entropi ölçüsünün hesaplanması daha uygun olabilir.

Varyans ve entropi arasında bazı sezgisel benzerlikler ve farklılıklar bulunmaktadır. İkisi de bir durum altında yoğunlaşma veya yayılma ölçütleridir. İkisi arasındaki önemli bir fark, varyansın yalnızca yoğunluk ortalaması etrafında merkezlenen bir ölçü olduğu, entropinin ise yoğunluk içindeki konumlarından bağımsız bir ölçü olduğudur. Bir diğer fark ise entropinin yoğunluğun eşit dağılımdan sapmasını ölçmesi (maksimum belirsizlik sağlayan dağılım), varyansın ise kendi dağılımının ortalamasından uzaklıkların bir ortalaması olduğudur. Bu, dağılımın entropisinin her zaman dağılıma bağlı olmayan tutarlı bir temel çizgiye sahip olmasıyla açıklanabilirken, varyansın kendi dağılımına bağlı olduğu ifade edilebilir [58].

Sürekli rasgele değişkenler için, $E(X^2) < \infty \Rightarrow H(X) < \infty$, entropi ve varyans arasında bir karşılaştırma yapılabilir, ancak bu iki özellik her zaman birbiriyle eşit olmayabilir. Shannon, sürekli bir rasgele değişkenin entropisi ile varyansı arasında aşağıdaki karşılaştırmayı önermiştir:

$$\frac{e^{2H(X)}}{2\pi e} \leq Var(X) \tag{3.46}$$

Denklem (3.46) ile verilen eşitlik yalnızca rasgele değişken X normal dağılıma sahipse geçerlidir. Bu, *"entropi güç oranı"* olarak bilinir ve varyans ile entropi arasındaki ilişkiye vurgu yapar [58].

3.9 Karşılıklı Bilgi ve Entropi Arasındaki İlişki

Karşılıklı bilgi $I(X;Y)$; ortak dağılım ve çarpım dağılımı ($p(x)p(y)$) arasındaki göreli entropidir.

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{3.47}$$

$$= \sum_{x,y} p(x,y) \log \frac{p(x|y)}{p(x)} \tag{3.48}$$

$$= \sum_{x,y} p(x,y) \log p(x) + \sum_{x,y} p(x,y) \log p(x|y) \tag{3.49}$$

$$= -\sum_{x,y} p(x) \log p(x) - (-\sum_{x,y} p(x,y) \log p(x|y)) \tag{3.50}$$

$$= H(X) - H(X|Y) \tag{3.51}$$

Bu nedenle, karşılıklı bilgi $I(X;Y)$, Y'nin bilgisine bağlı olarak X'in belirsizliğinde meydana gelen azalmayı ifade eder[59].

Simetriden dolayı Denklem (3.51) aşağıdaki gibi yazılabilir:

$$I(X;Y) = H(Y) - H(Y|X) \tag{3.52}$$

Denklem (3.52) ile simetriden dolayı, X değişkeninin Y değişkeni hakkında verdiği bilgi kadar, Y değişkeninin de X değişkeni kadar bilgi verdiği görülür. Bu nedenle, bir rasgele değişkenin kendiyle olan karşılıklı bilgisi, rasgele değişkenin entropisiyle aynıdır.

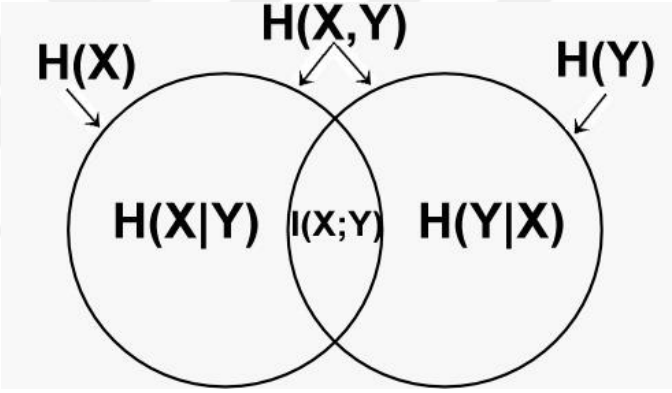
$$I(X;Y) = H(X) - H(X|Y) \tag{3.53}$$

$$I(X;Y) = H(Y) - H(Y|X) \tag{3.54}$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{3.55}$$

$$I(X;X) = H(X) \tag{3.56}$$

Karşılıklı bilgi ve entropi arasındaki ilişki aşağıdaki tabloda gösterilmiştir:

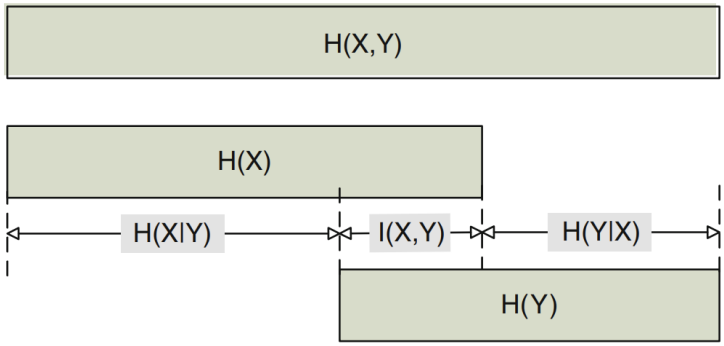


Şekil 3.2 Entropi ile Karşılıklı Bilginin Venn Diyagramı

$H(X)$, $H(Y)$, $H(X,Y)$, $H(X|Y)$, $H(Y|X)$ arasındaki ilişki Şekil 3.2'de gösterilmiştir. Karşılıklı Bilgi $I(X;Y)$, X'teki bilgi ile Y'deki bilginin kesişimi olarak görülmektedir. Bilgi miktarındaki artış, aynı miktarda belirsizliği kaybetmek anlamına gelir, bu nedenle karşılıklı bilgi (I) ve entropi değeri (H), hesaplamalarında sadece işaret açısından farklılık gösterir ve aşağıdaki şekilde ifade edilebilir:

$$H = -I = -\sum_i p_i \log p_i \tag{3.57}$$

Entropi, rasgele değişkenlerin olasılıklarına bağlıdır, rastgele değişkenlerin gerçek değerlerine bağlı değildir. Dikkat edilmesi gereken bir nokta, I hiçbir zaman pozitif olmaz, değeri 0 ile $-\infty$ arasında değişirken p değeri 0 ile 1 arasında değer alır.



Şekil 3.3 Entropi ile Karşılıklı Bilgi Arasındaki Fark [40]

Tablo 3.1 Kesikli ve Sürekli Değişkenler için Entropi Hesaplamaları

Bilgi Ölçüsü	Kesikli Değişkenler	Sürekli Değişkenler
Shannon Entropisi	$H(X) = -\sum_x p_x log p_x$	$H(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx$
Kullback – Leibler Mesafesi	$D_{KL}(p\|q) = \sum_{x,y} p(x,y)\log\frac{p(x,y)}{q(x,y)}$	$D_{KL}(p\|q) = \int_{x,y} f(x,y)\log\frac{f(x,y)}{f(x)f(y)}$
Karşılıklı Bilgi	$I(X;Y) = \sum_{x,y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$	$I(X;Y) = \int_{x,y} f(x,y)\log\frac{f(x,y)}{f(x)f(y)}dxdy$
Rényi Entropisi	$H_{R-\alpha}(X) = \frac{1}{1-\alpha}\log\sum_x p_x^\alpha$	$H_{R-\alpha}(X) = \frac{1}{1-\alpha}\int_x f^\alpha(x)dx$
Tsallis Entropisi	$H_{T-\alpha}(X) = \frac{1}{\alpha-1}\sum_x p_x^\alpha - 1$	$H_{T-\alpha}(X) = \frac{1}{\alpha-1}\int_x f^\alpha(x)dx - 1$

4
İSTATİSTİKSEL ANALİZLER

4.1 Kutu Boyutu (Bin Size) Belirleme

Karşılıklı bilgi, öncelikle hem kutu(bin) boyutuna hem de örneklem büyüklüğüne bağlıdır. Bu nedenle, en iyi parametre değeri seçimi sorunu ortaya çıkar. En iyi "kutu boyutu(bin)" sayısını elde etmek için çeşitli yöntemler vardır. Bu yöntemlerden bir tanesi "Freedman - Diaconis" kuralıdır. Bu kurala göre, optimal bin sayısı, $IQR = (Q_3 - Q_1)$ (çeyrekler arası aralık) hesaplanarak elde edilir ve veri noktalarının sayısı "n" temel alır.

""Freedman-Diaconis", verilerin yayılımını değerlendirmek için standart sapma yerine verilerin "çeyrek aralığını (IQR)" kullanır. Bu nedenle, bu yöntem diğer bazı yöntemlere göre daha güvenilir olarak kabul edilir. Yöntem aşağıdaki gibi hesaplanır:

$$\Delta_{bin} = 2 \times \frac{IQR(X)}{\sqrt[3]{n}} \tag{4.1}$$

Freedman-Diaconis kuralı, verinin simetrisizliğini göz önünde bulundurarak ve kutu (bin) boyutunu veri aralığına orantılı olarak ayarlar[60]. Verilerinin temelindeki dağılımın bilinmediği durumlarda, entropi tahminlerini hesaplamak için veri noktalarını boyutlandırma (ayarlanabilir hale getirme-discretization) işlemi önemlidir. Çünkü kesin olarak en iyi kutu (bin) boyutu tahmini bulunmamaktadır ve farklı kutu boyutları verinin farklı özelliklerini ortaya çıkarabilir. Çok küçük bir kutu genişliği, aşırı pürüzlü bir histograma neden olabilirken, çok büyük bir kutu genişliği gerçek olasılık yoğunluk fonksiyonuna (PDF) kıyasla daha düzgün bir histograma yol açar. Pürüzlülük arttıkça tahmini PDF'nin varyansı çok yüksek olurken, düzgünlük arttıkça önemli miktarda yanlılık ortaya çıkar. Uygulama bölümünde Freedman – Diaconis yöntemi kullanılmıştır.

Optimal kutu genişliğini belirlemek için kullanılan bazı yöntemler aşağıda belirtilmiştir:

i. Freedman–Diaconis' Kuralı
ii. Sturges Formülü
iii. Scott's Kuralı
iv. Rice Kuralı
v. Doane's Formülü
vi. Bendat ve Piersol Kuralı
vii. Cochran Formülü
viii. Knuth Kuralı
ix. Shimazaki and Shinomoto's Kuralı

4.2 Akaike Bilgi Kriteri (AIC)

Çoğu zaman, uygun olabilecek ve veriyi iyi bir şekilde temsil edebilecek birden fazla model bulunabilir. 1973 yılında Hirotsugu Akaike, bu durumda en iyi modeli bulabilmek için Akaike Bilgi Kriteri (AIC - Akaike Information Criterion) adı verilen bir seçim yöntemini geliştirmiştir [61]. AIC, herhangi bir tahmin edilen istatistiksel modelin uyum iyiliğinin bir göstergesi olarak da adlandırılabilir. AIC hesabı Denklem (4.2)' de gösterilmiştir:

$$AIC = -2\log(\text{maximized Likelihood}) + 2k \tag{4.2}$$

$$AIC = -2\log(L(\hat{\theta}|y)) + 2k \tag{4.3}$$

$L(\hat{\theta} \mid y)$, log -olabilirlik fonksiyonunun en yüksek noktasındaki sayısal değeridir. θ, k boyutlu bilinmeyen parametreler vektörü olmak üzere, $\hat{\theta}$, θ'nın en çok olabilirlik tahmincisidir.

Log-olabilirlik fonksiyonundaki bu en yüksek nokta, maksimum olabilirlik tahminlerinin değerlerine karşılık gelir. k, modeldeki parametre sayısını ifade eder.

Akaike Bilgi Ölçütü (AIC) bilgi teorisinde bir temel ile geliştirilmiştir. Birden fazla model olduğunda minimum AIC kriterine sahip model en iyi model olarak seçilir. AIC, ekonometrik modellerin belirli bir veri seti için göreceli kalitesini ölçmeye

çalışırken kullanılan bir ölçüdür. Bu nedenle, AIC, belirli bir modelin karmaşıklığı ile uyumluluğun iyiliği arasındaki dengeyi sağlamak için çalışır. AIC, modelin veri veya gözlemler kümesine ne kadar iyi uyduğunu açıklayan istatistiksel bir terimdir [62] [63].

4.3 Bayes Bilgi Kriteri (BIC)

Bayes bilgi kriteri ile Akaike bilgi kriteri arasındaki en büyük fark örneklem büyüklüğüdür. Bu yaklaşıma göre tasarlanan hesaplamalar verilere göre en iyi modeli bulmak için BIC kriterini kullanır. L olasılığı ifade etmek üzere, BIC kriteri aşağıdaki gibi hesaplanır [64]:

$$BIC = -2LogL(\hat{\theta}) + klogn \tag{4.4}$$

BIC değerinin $\log n$ oranında büyümesine modeldeki parametre sayısının artması neden olmaktadır. Dolayısıyla model seçiminde en az değere sahip BIC modelleri tercih edilir.

Bayesian Bilgi Kriteri (BIC), doğru bir model seçiminde daha kullanışlıyken, AIC gelecekteki gözlemleri tahmin etmek için ve en iyi modeli bulmak için daha uygun bir seçenektir.

4.4 AIC ve BIC Arasındaki Farklar

- BIC, model seçiminde tutarlıdır. Örneklem büyüdükçe, gerçek modelin aday modeller arasında olduğu ve diğer varsayımlar sağlandığında, gerçek modeli seçme olasılığı oldukça yüksektir. AIC ise asimptotik olarak tahmin ortalama karesel hatayı en aza indiren modeli seçer. AIC ayrıca sonlu örneklem büyüklüklerinde maksimum olası riski en aza indirir.
- AIC, farklı modeller arasında karşılaştırma yaparken, en küçük AIC değerine sahip modeli seçmeyi önerir. AIC, daha iyi bir uyum için daha fazla parametreyi kabul edebilir.

BIC, farklı modeller arasında karşılaştırma yaparken en küçük BIC değerine sahip modeli seçmeyi önerir. BIC, daha düzenlidir ve daha az parametreli modelleri tercih eder [61].

4.5 Bilgi Kazanım Fonksiyonu (Information Gain Function)

Bir olayda elde edilen bilgi miktarındaki kazanç, bu olayın gerçekleşme olasılığı ile ters orantılıdır. Bu kazanç $G(p)$ veya ΔI ile ifade edilsin. Shannon'a göre kazanç fonksiyonu aşağıdaki gibi ifade edilir:

$$G(p) = \Delta I = log\left(\frac{1}{p_i}\right) = - log(p_i) \tag{4.5}$$

Denklem (4.5)'deki $G(p)$ *"kazanç fonksiyonu (gain function)"* olarak ifade edilir ve bilgi miktarındaki kazancı ölçer. Diğer taraftan, $i.$ olay gerçekleştiğinde belirsizlik ortadan kalkar. Bu tarz denklemlerde, ilgilenilen olaylar birbirlerinden bağımsız ise, bağımsız olayların çarpımsal kuralından logaritma kullanmak hem kolay hem de uygundur.

Örneğin $P(A \cap B) = P_A P_B$ ise, $H(AB) = - logP_A - logP_B = H(A) + H(B)$ gerçekleşir. Eğer bir olayın olasılığı oldukça düşükse, örneğin $p_i = 0.01$ ise, olay tarafından aktarılan kısmi bilgi miktarı logaritma tabanı 10 kullanılarak $\Delta I = 2$ oldukça büyük bir sayı olur.

$$I = \sum_{i=1}^{N} \Delta I_i = -\sum_{i=1}^{N} \log(p_i) \tag{4.6}$$

Standart entropi ile yakından ilgili önemli bir nicelik, Kullback bilgi kazancıdır. Belirli bir ölçüm kümesinin, j indeksiyle gösterilen sonuçlarını elde etme olasılıkları P_j olarak verilsin. Bilgi miktarındaki değişim $\Delta\sigma_j$ ile gösterilir ve $\Delta\sigma_j = -(\ln p'_j - \ln p_j)$ ile hesaplanır.

Bu bilgi kazancı iki olasılık dağılımı (P_i ve P_j) ile ilişkili bilgi değişiminin ortalama değişimini ölçer ve Kullback - Leibler bilgi kazancı olarak ifade edilir:

$$K(p, p') \equiv \sum_j p_j \Delta\sigma_j = \sum_j p_j \ln \frac{p_j}{p'_j} \tag{4.7}$$

- İlk olarak, rastgele koordinat dönüşümlerine karşı değişim göstermezler, bu da fiziksel teorilerin beraber formüle edilmesine olanak tanır.
- İkinci olarak, istatistiksel sonuçlar için bir minimum entropi formülasyonu sağlar ve aynı olasılığa sahip olmayan bir önsel dağılımı otomatik olarak dahil eder.
- Kullback entropisi, herhangi iki çözüm arasındaki göreceli entropiye dayalı olarak H teoreminin doğru bir şekilde formüle edilmesinde son derece önemlidir.
- Bu entropi aynı zamanda iki rasgele değişken arasındaki istatistiksel bağımlılık derecesini de ölçer.

4.6 Regresyon Analizinde Doğrusal Olmama Durumunu Belirlemeye Yarayan Bazı Yöntemler

İki veya daha fazla rasgele değişken arasındaki bağımlılığın analizi, Gauss ve Laplace gibi matematikçilerin çalışmalarıyla 19. yüzyılın sonlarına kadar takip edilebilir. Sonrasında, Galton tarafından korelasyon kavramı önerildi ve bu, Pearson korelasyon katsayısının türetilmesini sağladı. Bu korelasyon katsayısı o zamandan beri çeşitli istatistiksel analizlerde yaygın bir şekilde kullanılmaktadır.

Eğer ilişki doğrusal veya yaklaşık olarak doğrusal ise, korelasyon katsayısı rasgele değişkenler arasındaki ilişkinin en güçlü göstergesidir. Ayrıca, ilişkinin yönünü, pozitif veya negatif olup olmadığını basit bir şekilde açıklar.

Ancak, bağımlılık doğrusallıktan saparsa, doğrusal korelasyon katsayısı kullanışsız hale gelebilir. Bu nedenle, doğrusallık olmayan durumları değerlendirmek için Spearman korelasyon katsayısı gibi çeşitli yöntemler önerilmiştir. Bu katsayı, parametrik olmayan bir yöntemdir ve sıralanmış değerleri kullanarak iki rasgele değişken arasındaki monoton olmayan durumu değerlendirmek için kullanılır [65].

Doğrusal olmayan bağımlılığın başka bir ölçüsü, korelasyon oranıdır. Bu ölçüm, rasgele değişkenler arasındaki ilişkiyi tek bir ölçüde ifade eder. Lineer olmayan ilişkilerde, korelasyon oranı korelasyon katsayısından daha büyük bir değere

sahiptir. Bu nedenle, korelasyon oranı ile korelasyon katsayısı arasındaki fark, bağımlılığın doğrusallık derecesini gösterir [66].

4.7 F Testi (Model Uyum Testi)

Veri toplandığında ve bir regresyon modeli uygulandığında, uyarlanmış regresyon modelinin yanıt değişkenini (Y) yeterince iyi açıklayıp açıklanmadığı öğrenilmek istenir. Bu durumda "Model uyum testinden" yararlanılabilir. Model uyum testi, tek bir X değeri için Y üzerindeki gözlemlerin bağımsız ve normal dağıldığını araştırdığı gibi, bir veya daha fazla X düzeyinde tekrarlanan gözlemleri ($X_1, X_2, \ldots X_c$) araştırabilir ve Y'nin koşullu dağılımlarının aynı varyansa (σ^2) sahip olduğunu varsayar [67]. Model uyum testi literatürde "Uyumsuzluk testi" olarak da adlandırılmaktadır. Uyumsuzluk, veriye uygun şekilde uymayan bir başlangıç modelini içerir. Genellikle tam bir model ile araştırmaya başlanır ve daha sonra azaltılmış modellerle devam edilir. Uyumsuzluk araştırılırken, başlangıç model ile azaltılmış model arasındaki değerler kontrol edilir. Ve azaltılmış modelden önemli ölçüde daha iyi uyan modeller aranır [68]. Varyans analizi yapılarak, bağımlı değişkenin bağımsız değişkenler tarafından açıklanıp açıklanmadığı, yani bağımlı değişkenle bağımsız değişkenler arasında doğrusal bir ilişki olup olmadığı "F testi" ile test edilir.

Model uyum testi için hipotezler aşağıdaki gibi kurulur:

H_0 : *Modelde varsayılan ilişki doğrusaldır ve modelde uyumsuzluk yoktur.*
H_a : *Modelde varsayılan ilişki doğrusal değildir ve modelde uyumsuzluk vardır.*

Regresyon fonksiyonunun doğrusallığını test etmek için hipotezler aşağıdaki şekilde kurulur:

$$H_0 : E(Y) = \beta_0 + \beta_1 X_1 + \ldots \beta_{p-1} X_{p-1}$$
$$H_a : E(Y) \neq \beta_0 + \beta_1 X_1 + \ldots \beta_{p-1} X_{p-1} \quad (4.8)$$

Y_{ij}, X'in j.düzeyindeki tekrar sayısı n_j ile gösterilecek olursa, X'in (j. düzeyi için) i.tekrarı gözlemlenen yanıt değişkeninin değeridir $(i = 1, \ldots, n_j, j = 1, \ldots, c)$.

$X = X_J$ noktasındaki bağımsız değişkenin ortalaması $\overline{Y}_j$ ile gösterilir. Bu durumda genel doğrusal test yaklaşımı, uyumsuzluk testi için kullanılan tam modelin belirlenmesiyle başlar. Tam model aşağıdaki gibidir:

$$Y_{ij} = \mu_j + \varepsilon_{ij} \tag{4.9}$$

Bu denklemde μ_j ($j = 1,\ldots,c$) modeldeki parametreleri temsil eder. Bu parametreler bağımsız ve normal dağılıma sahip oldukları ($N(0,\sigma^2)$), hata terimlerinin ise beklenen değerinin sıfır olduğu varsayılır ($E(Y_{ij}) = \mu_j$)).

μ_j parametresi için en küçük kareler (EKK) tahmincileri örnek ortalamalarıdır ($\hat{\mu}_j = \hat{Y}_j$). Y_{ij} gözlemi için tahmini beklenen değer $\hat{\mu}_j = \overline{Y}_j$ şeklindedir. Tam modelin hata kareleri ortalamaları aşağıdaki gibi hesaplanır ve $SSE(F)$ olarak gösterilir:

$$SSE(F) = \sum_j \sum_i (Y_{ij} - \overline{Y}_j)^2 = SSPE \tag{4.10}$$

Tam modelin hata kareleri toplamı (SSPE=SSE(F)) "saf hata kareleri toplamı" olarak adlandırılır. $SSPE$'nin serbestlik derecesi aşağıdaki gibidir:

$$df = \sum_j (n_j - 1) = \sum_j n_j - c = n - c \tag{4.11}$$

H_0 hipotezi μ_j parametrelerinin $p-1$ tane açıklayıcı değişken ile doğrusal bağlantılı olduğunu ifade eder:

$$\mu_j = \beta_0 + \beta_1 X_1 + \ldots + \beta_P X_{P-1} \tag{4.12}$$

Y_{ij}'nin tahmin değerleri ise aşağıdaki gibidir:

$$\hat{Y}_j = b_0 + b_1 X_1 + \ldots + b_P X_{P-1} \tag{4.13}$$

Azaltılmış modelin hata kareleri toplamı ise Denklem (4.14) ile ifade edilir:

$$SSE(R) = \sum\sum (Y_{ij} - \hat{Y}_{ij})^2 = SSE \tag{4.14}$$

$SSE(R)$'nin serbestlik derecesi $df_R = n - p$'dir. Bu durumda doğrusal test istatistiği ile ifade edilen F^* aşağıdaki gibidir:

$$F^* = \frac{SSLF}{(n-p)-(n-c)} \div \frac{SSPE}{n-c} \tag{4.15}$$

F^* denkleminde ifade edilen $SSLF$, SSE ile $SSPE$ arasındaki farka eşittir. Eğer $F^* < F(1-\alpha; c-p, n-c)$ ise doğrusal regresyon modelinin verilerle açıklanabilir olduğunu ifade eden H_0 kabul edilir [69].

$SSTO = \sum (Y_i - \bar{Y})^2$, tahmin değişkeni olan X_i dikkate alınmadığında, Y_i gözlemindeki değişimi ölçer.

SSE, X açıklayıcı değişkeni kullanılan model için Y_i gözlemlerindeki değişimi gösterir.

R^2, çoklu belirlilik katsayısı olarak isimlendirilir. Regresyon modelinde, bağımlı değişkendeki toplam varyasyonun ne kadarının bağımsız değişken tarafından açıklandığını belirtir. Diğer bir ifadeyle; Y_i gözlemindeki değişkenliğin yüzde kaçının doğrusal regresyon tarafından açıklandığını belirtir. Regresyon modelinin veriye uyumunun ölçüsü olarak da kullanılan R^2, regresyon modeli ile açıklanabilen değişimin(SSR), toplam değişim içindeki (SSTO) yüzdesidir [70].

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \tag{4.16}$$

$0 \le R^2 \le 1$ olmalıdır. R^2 sıfır değerini $b_k = 0$ $(k = 1, ..., p-1)$ olduğu zaman alır. 1 değerini ise, tahmin değerleri tahmini regresyon doğrusuna tam uyum gösterdiğinde yani $Y_i = \hat{Y}_i$ (bütün i'ler için) olduğunda geçerlidir.

Hipotez testlerinin doğruluğunu araştırmak için Tablo 4.1'de verilen formüller kullanılır ve elde edilen sonuçlar aşağıdaki gibi değerlendirilir:

- F^* istatistiği hesaplanır [71]. Kullanılan yazılım programlarından p değeri de elde edilebilir ve bu değere göre bir karar da verilebilir. F^* değeri, $(c-2, n-c)$ serbestlik derecesine sahiptir. Elde edilen sonuç F^* ile karşılaştırılarak modelin doğrusallığı ile ilgili bir karara varılır. F^* değeri yeterince büyük ise modelin doğrusal olmadığı kanısına varılır.
- Eğer X ile Y arasında doğrusal bir ilişki varsa, o zaman $\mu_j = \beta_0 + \beta_1 X_1 + \ldots \beta_{p-1} X_{p-1}$ olur. Yani doğrusal regresyon modelinde uyumsuzluk yoktur. Eğer p değeri anlamlılık düzeyinden (α) daha küçükse, H_0 reddedilir. Sonuç olarak, "doğrusal regresyon modelinde uyumsuzluk olduğunu göstermek için yeterli kanıtın, α düzeyinde olduğunu" belirtiriz.
- X ile Y arasında doğrusal bir ilişki yoksa, o zaman $\mu_j \neq \beta_0 + \beta_1 X_1 + \ldots \beta_{p-1} X_{p-1}$ olur. Yani doğrusal regresyon modelinde uyumsuzluk vardır. Eğer p değeri anlamlılık düzeyinden (α) daha büyükse, H_0 kabul edilir. Sonuç olarak, "doğrusal regresyon modelinde uyumsuzluk olduğunu sonuçlandırmak için yeterli kanıtın, α düzeyinde olmadığı" belirtilir.

Tablo 4.1 Varyans Analizi Tablosu (ANOVA Tablosu)

Kaynak	**df (Degrees of Freedom)**	**SS (Sum of Squares)**	**MS (Mean Square)**	**F**
Regresyon	1	$SSR = \sum_{i=1}^{c}\sum_{j=1}^{n_i}(\hat{y}_{ij} - \bar{y})^2$	$MSR = \frac{SSR}{1}$	$F = \frac{MSR}{MSE}$
Kalıntılar (Residual Error)	$n-2$	$SSE = \sum_{i=1}^{c}\sum_{j=1}^{n_i}(y_{ij} - \hat{y}_{ij})^2$	$MSE = \frac{SSE}{n-2}$	$F^* = \frac{MSLF}{MSPE}$
Uyum Testi (Lack of Fit)	$c-2$	$SSLF = \sum_{i=1}^{c}\sum_{j=1}^{n_i}(\bar{y}_i - \hat{y}_{ij})^2$	$MSLF = \frac{SSLF}{c-2}$	

Tablo 4.1 Varyans Analizi Tablosu (ANOVA Tablosu) (devamı)

Saf Hata	$n-c$	$SSPE=\sum_{i=1}^{c}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2$	$MSPE=\frac{SSPE}{n-c}$	
Toplam	$n-1$	$SSTO=\sum_{i=1}^{c}\sum_{j=1}^{n_i}(y_{ij}-\bar{y})^2$		

4.8 Entropi ve Karşılıklı Bilgi (MI) Yöntemi

Bir istatistiksel deneyin entropisi belirsizliğin bir ölçüsüdür. Olasılık dağılımlarının çeşitli entropi formülasyonları ve çeşitli entropi ölçüleri Tablo 3.1'de gösterilmiştir [72]. Shannon entropisi, bilgi kuramında, bilgi seçimi ve belirsizlik ölçüsü olarak önemli bir rol oynar. Koşullu entropi veya karşılıklı bilgi (MI), herhangi bir temel dağılımı önceden varsaymadan ve rastgele değişkenler arasındaki stokastik ilişkiyi, doğrusal veya doğrusal olmayan özellikleriyle tam olarak yansıtır. Bu özellikler, karşılıklı bilginin bağımlılıkları analiz etmek için iyi bir seçenek olduğunu göstermektedir. Özellikler finans [73-75] ve genetik [76-78] alanında bağımlılık analizi için genellikle karşılıklı bilgi kullanılmaktadır. Karşılıklı bilgi, rastgele değişkenler arasındaki ilişkiyi ölçmek için kullanılan bir yöntem olmakla birlikte, bağımlılığın doğrusal veya doğrusal olmayan bir ilişki olup olmadığı hakkında herhangi bir bilgi sunmaz. Literatürde Shannon'ın karşılıklı bilgisinin doğrusal bileşenini çıkarmak suretiyle bağımlılığın doğasını belirleyebilmek için çok az araştırma yapılmıştır [1, 17].

Shannon, Rényi ve Tsallis karşılıklı entropileri, kesikli(ayrık) durumlar için aşağıdaki gibi tanımlanır:

$$D_S=(p\|q)=\sum P(x)\log\frac{P(x)}{q(x)} \tag{4.17}$$

$$D_{\mathrm{R}}=(p\|q)=\frac{1}{\alpha-1}\log\sum P(x)\left(\frac{P(x)}{q(x)}\right)^{\alpha-1} \tag{4.18}$$

$$D_T = (p\|q) = \frac{1}{\alpha - 1} \sum P(x) \left[\left(\frac{P(x)}{q(x)} \right)^{\alpha - 1} - 1 \right] \tag{4.19}$$

İki değişkenli kesikli durumlar için ise hesaplamalarda aşağıdaki denklemler kullanılır:

$$D_S(p(x,y)\|q(x,y)) = \sum\sum p(x,y) \log \frac{p(x,y)}{q(x,y)} \tag{4.20}$$

$$D_R(p(x,y)\|q(x,y)) = \frac{1}{\alpha - 1} \log \sum\sum p(x,y) \left(\frac{p(x,y)}{q(x,y)} \right)^{\alpha - 1} \tag{4.21}$$

$$D_T((p(x,y)\|q(x,y)) = \frac{1}{\alpha - 1} \sum\sum p(x,y) \left[\left(\frac{p(x,y)}{q(x,y)} \right)^{\alpha - 1} - 1 \right] \tag{4.22}$$

Değişkenlerin bağımsızlığını kontrol etmek için sıfır ve alternatif hipotezler aşağıdaki gibi ifade edilebilir:

$$H_0 : p_{X,Y}(x,y) = q_{X,Y}(x,y) \tag{4.23}$$

$$H_A : p_{X,Y}(x,y) \neq q_{X,Y}(x,y) \tag{4.24}$$

Karşılıklı bilginin üç alternatif denklemi Shannon, Rényi ve Tsallis tarafından yapılmıştır. Shannon karşılıklı bilgi (veya Kullback-Leibler uzaklığı (KL)), aşağıdaki gibi tanımlanır:

$$M(X,Y) = D_S\left(p_{X,Y}(x,y) \| p_X(x) p_Y(y) \right) = \sum\sum p(x,y) \log \frac{p(x,y)}{p_X(x) p_Y(y)} \tag{4.25}$$

Denklem (4.25)'de ifade edilen karşılıklı bilgi aynı zamanda *"çapraz entropi (cross entropy)"* olarak da adlandırılır.

α dereceli Rényi karşılıklı bilgisi şu şekilde yazılır:

$$D_R\left(p_{X,Y}(x,y) \| p_X(x) p_Y(y) \right) = \frac{1}{\alpha - 1} \log \sum\sum \frac{p_{X,Y}(x,y)^{\alpha}}{\left(p_X(x) p_Y(y) \right)^{\alpha - 1}} \tag{4.26}$$

α dereceli Tsallis karşılıklı bilgisi ise aşağıdaki gibi ifade edilir:

$$D_T\left(p_{X,Y}(x,y)\|p_X(x)p_Y(y)\right)=\frac{1-\sum\sum\frac{p_{X,Y}(x,y)^{\alpha}}{\left(p_X(x)p_Y(y)\right)^{\alpha-1}}}{1-\alpha} \tag{4.27}$$

Bağımsızlık durumunda, Rényi ve Tsallis karşılıklı bilgileri de Shannon karşılıklı bilgisi gibi "0" olur. $\alpha\rightarrow 1$ olduğunda, Rényi ve Tsallis karşılıklı bilgileri Shannon karşılıklı bilgiye yaklaşır [79]. İki değişkenin karşılıklı bilgisi, bir değişkenin diğerini bilerek değişkenliğinin azalmasını yansıtır. Karşılıklı bilgi (MI), yalnızca rastgele değişkenler bağımsız olduğunda 0 olur. Ayrıca vurgulanmalıdır ki karşılıklı bilgi genel bağımlılığı ölçerken, korelasyon katsayısı doğrusal bağımlılığı ölçer [77].

Shannon karşılıklı bilgi ölçüsünün doğrusal olmama durumunun tespit edilmesi problemine uygulanması Tanaka, Okamoto ve Naito [80] ve Smith [81] tarafından önerilmiştir.

Bu yöntem, normal lineer regresyon modeli tarafından elde edilen kalıntıları kullanır. Uyum iyiliği sağlanmış bir lineer regresyon modeli, değişkenler arasında doğrusal bir ilişkinin iyi bir göstergesi olduğundan, lineer bir modele ait kalıntıların bağımsız değişkenlere göre doğrusal bağımlılık içermediği kabul edilir.

$$\xi_i = Y_i - b_0 - \sum_{j=1}^{p} b_j X_j \tag{4.28}$$

ξ_i kalıntıları temsil etmektedir. Kalıntıların ve bağımsız değişkenin gözlemlenen değerleri (X_j) arasındaki karşılıklı bilgi hesaplanabilir. Hesaplama adımları aşağıdaki gibi yazılabilir:

i. Bağımsız ve bağımlı değişkenler arasındaki karşılıklı bilgi $M(X,Y)$ hesaplanır.
ii. Lineer regresyon sonucu elde edilen kalıntılar ile bağımsız değişken arasındaki karşılıklı bilgi $M(X,\xi)$ hesaplanır.
iii. Karşılıklı bilgi hesaplanırken, bağımsız değişken ile kalıntılar arasındaki miktar, bağımlı ve bağımsız değişkenler arasındaki karşılıklı bilgi

miktarından çok farklı değilse, o zaman ilişki lineer olmayan bir ilişkidir ($M(X,Y) = M(X,\xi)$).

iv. $M(X,\xi)$ ile $M(X,Y)$ karşılaştırarak, bağımlılıkta yer alan doğrusallık derecesini değerlendirilebilir.

Rényi ve Tsallis karşılıklı bilgi ölçüleri, parametrik doğaları nedeniyle doğrusal olmamanın tespitinde daha iyi sonuçlar verebilir. Özellikle Tsallis karşılıklı bilgi ölçümü, α'nın gücüne dayanarak hesaplandığı için, Tsallis karşılıklı bilgi ölçüsü α değerine bağlıdır. α ne kadar büyükse Tsallis karşılıklı bilgi de o kadar büyür. Bu nedenle, bu iki karşılıklı bilgi arasındaki fark açıklanabilmesi zor bir hale gelir.

α parametresi doğrusal olmayan ilişkilere duyarlılığı kontrol eder. α 1'e yaklaştıkça, Tsallis karşılıklı bilgisi Shannon karşılıklı bilgiye yakınsar ve bu genellikle doğrusal bağımlılıkları yakalar. α 1'den sapmaya başladıkça, Tsallis karşılıklı bilgisi doğrusallık dışındaki ilişkilere daha duyarlı hale gelir.

Farklı α değerleri için $M(X,\xi)$ değerlerini inceleyerek, ilişkideki doğrusallık derecesini değerlendirilebilir. $M(X,\xi)$ değerlerinin daha yüksek olması, daha güçlü doğrusal olmayan bağımlılıkları gösterirken, daha düşük değerler bağımsız değişken ve kalıntılar arasında daha lineer bir ilişkiyi işaret eder.

Bu nedenle, Tsallis karşılıklı bilgi ölçümünü kullanarak ve farklı α değerlerini keşfederek, doğrusallık derecesi nicel olarak ölçülebilir ve değişkenler arasındaki ilişkinin doğasına dair bilgi elde edilebilir.

α'ya bağlı olarak ilişkinin doğrusallığını yorumlamayı kolaylaştıracak formül Denklem (4.29) aşağıda verilmiştir. Bu denklem kullanılarak bir doğrusallık ölçüsü elde edilmiş olur:

$$\lambda_{S,R,T} = \left| 1 - \frac{M(X,\xi)}{M(X,Y)} \right| \tag{4.29}$$

İndeksteki S, R, T harfleri Shannon, Rényi ve Tsallis karşılıklı bilgilerinin hesaplanmasında kullanıldığını gösterir. Eğer $M(X,\xi)$ ve $M(X,Y)$ birbirine yakınsa, ilişki doğrusal olmayan bir ilişkidir ve λ 0'a yakınsar. Tersine, doğrusal olmayan bağımlılığın bir ölçüsü olan $M(X,\xi)$ 0'a yakın ise, λ da 1'e yakınsar.

Bir değerine yakın λ değerleri doğrusal bir ilişkiyi gösterdiği için, λ_T, λ_S ve λ_R doğrusallık varsayımını temsil eden H_0 destekler.

5
UYGULAMA

Uygulama bölümünde iki farklı uygulama gerçekleştirilmiş ve bu uygulamalar ile karşılıklı bilgi kullanılarak modeldeki eğrisellik tanımlanmaya çalışılmıştır. Hipotezin gerçekliğini test edebilmek için 2 farklı uygulama gerçekleştirilmiştir. Birinci uygulama için biri doğrusal diğeri eğriselliği yansıtan iki veri seti seçilmiştir. Bu veri setleri için simülasyon ile elde edilen X ve Y değişkenlerinin sayısı 1000'dir. İkinci veri seti için ise Karbon Emisyonu ve GSYİH verileri kullanılmıştır.

5.1 Simülasyon Verileri ile Eğrisellik Analizi

Hipotezin gerçekliğini test edebilmek için 2 farklı uygulama gerçekleştirilmiştir. Birinci uygulama için biri doğrusal diğeri eğriselliği yansıtan polinom iki veri seti seçilmiştir. Bu veri setleri için simülasyon ile sayısı 1000 olan X ve Y değişkenleri elde edilmiştir. Doğrusal regresyon modeli temsil eden simüle edilmiş veriler Denklem (5.1) tarafından üretilmiştir.

$$Y = \beta_0 + \beta_1 X + e \tag{5.1}$$

Eğrisel regresyon modelini temsil eden simüle edilmiş veriler Denklem (5.2) tarafından üretilmiştir.

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + e \tag{5.2}$$

Tablo 5.1 Doğrusal Model için Parametre Değerleri

Analiz	**Tahmin**	**St. Hata**	**t Değeri**	$P(>\|t\|)$
Sabit	2.916	0.1857	15.71	0.0000
x	250	0.003207	77956.20	0.0000

- Oluşturulan doğrusal regresyon modeli aşağıdaki gibidir:

$$Y = 2.916 + 250X$$

- Model parametrelerinin anlamlılığı incelendiğinde aşağıdaki sonuçlar elde edilmiştir:
  - β_0 için aşağıdaki hipotezler test edilmiştir:

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

    %95' lik güven aralıkları incelendiğinde $\beta_0 = 0$'ın bu aralıkların içinde olmadığı görülür. Dolayısıyla H_0 hipotezi reddedilir. β_0 katsayısı anlamlıdır.
  - β_1 için aşağıdaki hipotezler test edilmiştir:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

    %95' lik güven aralıkları incelendiğinde $\beta_1 = 0$'ın bu aralıkların içinde olmadığı görülür. Dolayısıyla H_0 hipotezi reddedilir. β_1 katsayısı anlamlıdır.
- P değerlerine bakılarak da hipotezler test edilebilir. Görüldüğü gibi p değerleri $\alpha = 0,05$ düzeyinden az olduğu için hem β_0 için hem de β_1 için H_0 hipotezleri reddedilir.
- Varyans analizi tablosu incelendiğinde aşağıdaki sonuçlar elde edilmiştir:

Tablo 5.2 Doğrusal Model ANOVA Tablosu

	sd	**Kareler Toplamı**	**F Değeri**	$P(> F)$
Sabit	1	2137	246.66	0.0000
X	1	5.2644e+10	6.0772e+09	0.0000
Kalıntılar	998	8645		

Kuadratik model için elde edilen analizler aşağıdaki gibidir:

Tablo 5.3 Kuadratik Model için Parametre Değerleri

Analiz	**Tahmin**	**St. Hata**	**t Değeri**	$P(>\|t\|)$
Sabit	2.803	0.2783	10.074	0.0000
x	250	0.01274	19626.921	0.0000
x_1	-2.50007	0.0001224	-20422.16789	0.0000

- Oluşturulan doğrusal regresyon modeli aşağıdaki gibidir:

$$Y = 2.803 + 250X - 2.50007X^2$$

- Model parametrelerinin anlamlılığı incelendiğinde aşağıdaki sonuçlar elde edilmiştir:
  - β_0 için aşağıdaki hipotezler test edilmiştir:

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

    %95' lik güven aralığında H_0 hipotezi reddedilir. β_0 katsayısı anlamlıdır.
  - β_1 için aşağıdaki hipotezler test edilmiştir:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

    %95' lik güven aralığında H_0 hipotezi reddedilir. β_1 katsayısı anlamlıdır.
  - β_2 için aşağıdaki hipotezler test edilmiştir:

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

%95' lik güven aralığında H_0 hipotezi reddedilir. β_2 katsayısı anlamlıdır.

- P değerlerine bakılarak da hipotezler test edilebilir. Görüldüğü gibi p değerleri $\alpha = 0,05$ düzeyinden az olduğu için hem β_0 için hem de β_1 hem β_2 için H_0 hipotezleri reddedilir. Model parametreleri anlamlı bulunmuştur.

- Varyans analizi tablosu incelendiğinde aşağıdaki sonuçlar elde edilmiştir:

Tablo 5.4 Kuadratik Model ANOVA Tablosu

	sd	**Kareler Toplamı**	**F Değeri**	$P(>F)$
Sabit	1	2137	246.66	0.0000
X	1	5.2644e+10	6.0772e+09	0.0000
Kalıntılar	998	8645		

Farklı α değerleri için α'nın doğrusallık ölçüleri üzerindeki etkisini değerlendirmek için 0 ile 5 arasında çeşitli α değerleri rastgele olarak seçilmiştir. Tablo 5.1, farklı α değerleri için 50 rastgele gözlemi ve bunlara karşılık gelen Rényi ve Tsallis ölçüleri için λ değerlerinin uniform dağılımdan elde edilmiştir.

Tablo 5.5 Farklı α Değerlerine Göre λ_s, λ_R ve λ_T Değerleri

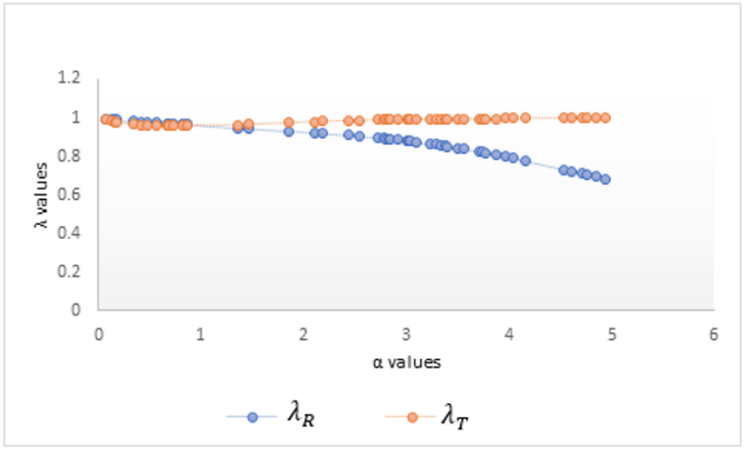
	Doğrusal İlişki		**Eğrisel İlişki**			**Doğrusal İlişki**		**Eğrisel İlişki**	
α	λ_R	λ_T	λ_R	λ_T	α	λ_R	λ_T	λ_R	λ_T
0.07	0.9956	0.9906	0.0489	0.0308	3.01	0.8801	0.9914	0.0133	0.0436
0.13	0.9917	0.983	0.0457	0.0263	3.02	0.8795	0.9915	0.0133	0.0437
0.17	0.9894	0.9788	0.0447	0.0247	3.04	0.8782	0.9917	0.0133	0.0441
0.18	0.9888	0.9778	0.0445	0.0244	3.09	0.8749	0.9921	0.0132	0.045
0.35	0.9808	0.9664	0.0433	0.0228	3.23	0.8652	0.9931	0.0131	0.0476
0.41	0.9785	0.964	0.0431	0.0232	3.29	0.8608	0.9934	0.0131	0.0487

Tablo 5.5 Farklı α Değerlerine Göre λ_R ve λ_T Değerleri (Devamı)

α	λ_R	λ_T	λ_R	λ_T	α	λ_R	λ_T	λ_R	λ_T
0.48	0.976	0.9619	0.0429	0.024	3.34	0.857	0.9937	0.013	0.0497
0.56	0.9734	0.9604	0.0429	0.0254	3.38	0.8538	0.994	0.013	0.0504
0.67	0.9702	0.9595	0.0435	0.0281	3.4	0.8522	0.9941	0.013	0.0508
0.69	0.9696	0.9595	0.0437	0.0287	3.5	0.844	0.9946	0.0129	0.0528
0.74	0.9684	0.9596	0.0444	0.0304	3.57	0.8379	0.9949	0.0128	0.0542
0.82	0.9668	0.9605	0.0464	0.0339	3.71	0.8249	0.9954	0.0128	0.057
0.87	0.9663	0.9618	0.0483	0.0369	3.74	0.822	0.9956	0.0128	0.0576
1.36	0.9473	0.964	0.0087	0.0099	3.77	0.8191	0.9957	0.0127	0.0583
1.46	0.9446	0.9662	0.003	0.0037	3.88	0.808	0.996	0.0127	0.0607
1.86	0.9323	0.9749	0.0129	0.0214	3.97	0.7985	0.9963	0.0127	0.0627
2.11	0.9236	0.9797	0.0138	0.0271	4.04	0.7908	0.9964	0.0127	0.0643
2.18	0.9209	0.981	0.0139	0.0285	4.17	0.7762	0.9967	0.0127	0.0673
2.44	0.9102	0.9851	0.0139	0.0334	4.54	0.7318	0.9974	0.0128	0.0769
2.54	0.9056	0.9865	0.0138	0.0352	4.62	0.7219	0.9975	0.0128	0.0792
2.73	0.8962	0.9888	0.0136	0.0386	4.71	0.7108	0.9976	0.0129	0.0818
2.78	0.8935	0.9893	0.0136	0.0395	4.76	0.7046	0.9976	0.0129	0.0833
2.8	0.8924	0.9895	0.0136	0.0398	4.85	0.6934	0.9977	0.013	0.0861
2.83	0.8908	0.9898	0.0135	0.0403	4.94	0.6822	0.9978	0.0131	0.089
2.92	0.8856	0.9907	0.0134	0.0419	Mean	0.8783	0.9848	0.0211	0.0443
					Std.Dev.	0.0869	0.0134	0.0143	0.0199
					λ_s	0.9589		0.0121	

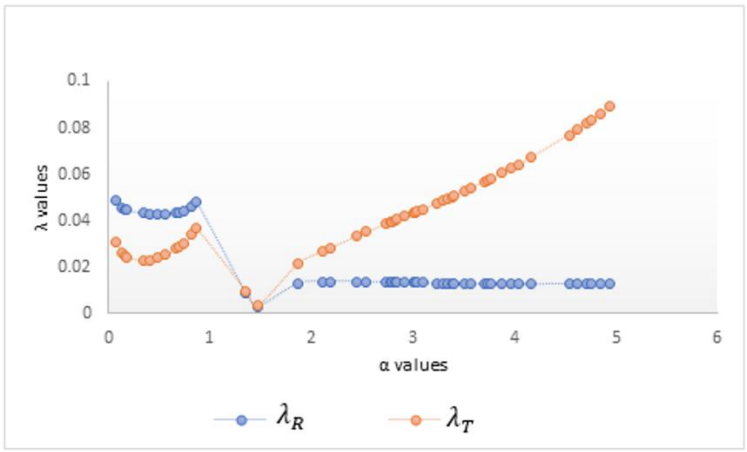
Karşılıklı bilgi ölçüsünün ölçeğini temsil eden α parametresini değiştirmek, bu ölçünün duyarlılığını doğal olarak değiştirir. λ değerlerinin ölçeğe bağlı olarak α'ya karşı duyarlılık değişimi görsel olarak görülebilir. Sonuçları görsel olarak

yorumlamak için farklı α değerlerine göre λ_R *ve* λ_T değerleri aşağıdaki grafiklerle gösterilmiştir.



Şekil 5.1 Doğrusallık durumunda λ değerleri ile α arasındaki ilişki

Şekil 5.1'den görüldüğü gibi, doğrusal ilişki için $\alpha > 1$ durumunda, λ_T λ_R'ye göre daha başarılı ve kararlıdır. λ_T ölçeği sürekli olarak 1'e yakın değerler alırken, λ_R α değerleri arttıkça daha küçük değerler alır.



Şekil 5.2 Eğrisellik durumunda λ değerleri ile α arasındaki ilişki

Eğrisel ilişkide, Şekil 5.2'de görüldüğü gibi, λ_T'nin α 1.4'ten sonra artmaya başladığı görülmektedir. λ_R'nin ise tüm α değerlerinde sıfıra yakın değerler aldığı ortadadır. Ancak dikkat edilmesi gereken nokta, λ_T'nin maksimuma ulaştığı değer 0.09101'dir. Bu durumda hesaplanan iki ortak bilgi ölçüsü de "eğrisellik" ölçüsü olarak kullanılabilir. Ancak λ_R, doğrusal olmama durumunu daha tutarlı bir

şekilde göstermektedir. Tsallis karşılıklı bilgi formülünde logaritmik bir fonksiyon olmadığı için, α değeri 1'den büyük olduğunda, Tsallis karşılıklı bilgi, Rényi karşılıklı bilgiden daha az doğrusallıktan sapmalara önem verir. Bu nedenle, λ_T, λ_R'ye kıyasla doğrusallıktan sapmalara daha az duyarlı olacaktır. Dolayısıyla λ_R, λ_S ve λ_T'ye kıyasla doğrusallığa daha az tepki verecektir. Aynı nedenle, λ_T, doğrusal ilişkide Rényi'ye kıyasla doğrusallığı daha iyi temsil edecektir.

Belirli bir olasılık dağılımı için Rényi entropisinin α'ya (1'den farklı olan herhangi bir gerçel sayı) göre tekdüze azalan bir fonksiyondur. Ve bu ifade Rényi entropisinin önemli bir genel özelliğidir. Şekil 5.1'de görüldüğü gibi, α değerleri 5 ile sınırlandırılmıştır. Çünkü α değerlerini arttırmak ek bilgi sağlamayacağı görülmüştür.

5.1.1 Sonuç 1

Kuadratik ilişki gösteren simülasyon verileri için, uygulanan model belirleme kriterleri doğrusal ve kuadratik regresyon için çalışılmıştır:

Tablo 5.6 Simülasyon Verileri için Model Belirleme Kriterleri

Model Belirleme Kriterleri	AIC	BIC	RMSE	MAE	R^2
Doğrusal Regresyon	17944.59	17959.3	1901.41	1634.72	0.0007085
Kuadratik Regresyon	5002.582	5022.21	2.93983	2.34643	0.9999976

Uygulamaya bakıldığında aşağıdaki sonuçlar elde edilmiştir:

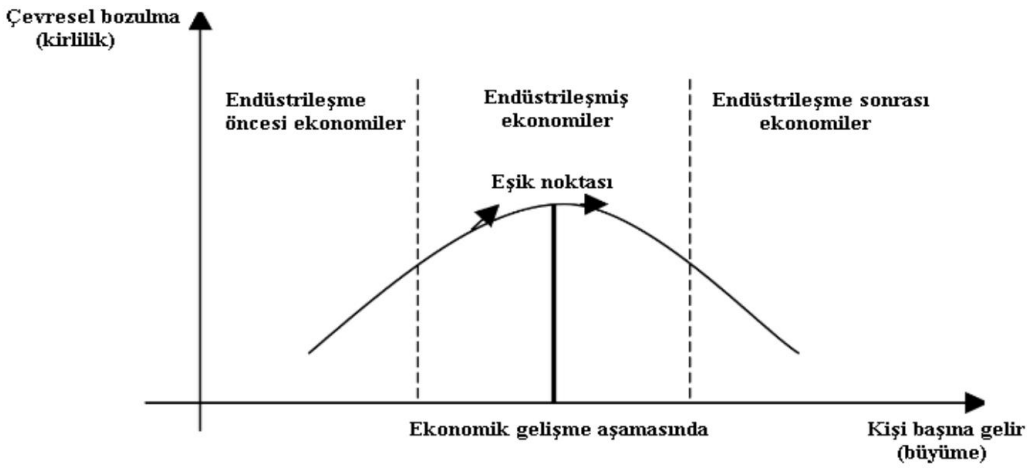
- R^2, bağımlı değişkenin değişkenliğini açıklama açısından doğrusal regresyon modelindeki bağımsız değişkenlerin ne kadar iyi olduğunu açıklamak için kullanılır. R Kare değeri bağımsız değişkenlerin eklenmesiyle her zaman artar, bu da modele gereksiz değişkenlerin eklenmesine yol açabilir. Simülasyon verileri için doğrusal ve eğrisel modeller karşılaştırıldığında iki modelin R^2'leri arasında önemli bir

farklılık olduğu, kuadratik ilişkiyi doğrusal regresyon modeliyle tahmin etmenin ilişkiyi açıklamada ne kadar eksik olduğu görülmektedir.

- İstatistikte, ortalama mutlak hata (MAE), gözlemler arasındaki hataların bir ölçüsüdür. Y'ye karşı X'in örnekleri, tahmin edilenin gözlemlenenle karşılaştırılması, ardışık zamanın başlangıç zamanına karşı gelmesi ve bir ölçüm tekniğinin alternatif bir ölçüm tekniğiyle karşılaştırılması gibi durumları içerir. MAE, orijinal veriyle aynı birime sahiptir ve hataları aynı birimlerde ölçülen modeller arasında karşılaştırılabilir.
- Hem RMSE hem de R Kare, bir doğrusal regresyon modelinin bir veri kümesine ne kadar iyi uyduğunu ölçer. RMSE, bir regresyon modelinin yanıt değişkeninin değerini mutlak olarak ne kadar iyi tahmin edebildiğini söylerken, R Kare, tahmin edici değişkenlerin yanıt değişkenindeki değişkenliği ne kadar iyi açıklayabildiğini gösterir. MAE, MSE ve RMSE'nin daha düşük değerleri, bir regresyon modelinin daha yüksek doğruluğunu ifade eder.
- Farklı doğrusal regresyon modellerinin doğruluğunu karşılaştırırken, RMSE (Kök Ortalama Kare Hata) gerçekten R-kare'ye göre daha iyi bir seçimdir. R-kare bağımsız değişkenler tarafından açıklanan değişkenlik oranını ölçer, ancak tahmin hatalarının büyüklüğünü doğrudan yakalamaz. Öte yandan, RMSE hem hataların yönünü hem de büyüklüğünü içeren ortalama tahmin hatası ölçüsünü sağlar. Bu nedenle, RMSE regresyon modellerinin doğruluğunu değerlendirmek için daha kapsamlı bir ölçüdür. Bu değerin karşılaştırması sonucu kuadratik regresyon modeli daha uygun görülmektedir.
- λ_T denkleminde logaritmik bir ifade olmadığı için, α 1'den büyük değerler aldığı zaman, λ_R'ye göre doğrusal olmama durumuna karşı daha az hassastır.
- Yapılan karşılıklı bilgi analizinde ise; λ_T ve λ_R ölçümlerinin doğrusal olmamanın bir göstergesi olabileceği sonucuna varılmıştır. Tablo 5.1'de sonuçlar ayrıntılı olarak gösterilmiştir.

5.2 Eğrisellik Belirleme Ölçülerinin Çevre Verileri Üzerine Uygulanması

Çevresel kalite ile ekonomik büyüme arasındaki ilişki, 1970'lerden sonra, Club of Rome üyelerinin Gelecek Sınırları [82] adlı çalışmasında dramatik çevresel senaryoları önlemek için sıfır büyüme veya istikrarlı bir ekonomi önermelerinden sonra kapsamlı bir şekilde araştırılmıştır. Bu raporda; ekonomik büyümenin önemli çevresel etkilere sahip olduğunu ve gelecekte dünya ekonomisinin yenilenemez kaynaklar, tarımsal üretim ve aşırı kirlilik açısından fiziksel sınırlarına ulaşacağı için dünyanın çökmesi ile sonuçlanacağı belirtilmiştir. Beckerman [10], teknolojik ilerlemenin doğal kaynaklara olan bağımlılığı azaltacağını ve sürekli bir büyümeye olanak tanıyacağını iddia etmiştir. Uzun yıllar boyunca mevcut çevresel veri eksikliği nedeniyle bu tartışma yalnızca teorik bir temelde kalmıştır. Çevresel kalitenin tek bir kriter ile ifade edilememesinden dolayı çevresel bozulmanın birkaç göstergesi önerilmiştir. Bu kriterler arasında kükürt dioksit, azot oksit emisyonları ve karbondioksit gibi çevresel göstergelerin uygulandığı deneysel çalışmalardaki kanıtlar, gelirle ters-U şeklinde bir ilişki olduğunu göstermiştir. Bu durumda, çevresel bozulmanın büyümenin erken aşamalarında kötüleştiği, ancak gelirin bir "dönüm noktası" olarak adlandırılan bir eşik seviyesini geçtikten sonra zirveye ulaştığı ve azalmaya başladığı görülmüştür (Şekil 5.3).



Şekil 5.3 Çevresel Kuznets Eğrisi (EKC) [83]

Kuznets (1955) ekonomik ilerlemeyle beraber bireylerin gelir düzeylerinin arttığını, fakat gelişimin erken evrelerinde gelir eşitsizliğinin de arttığını ifade etmiştir. Ek olarak, yükselen gelir eşitsizliğinin ekonomik ilerlemenin devam etmesiyle birlikte belirli bir noktadan sonra azalmaya başladığını öne sürmüştür. Kişi başına düşen gelir miktarı ile gelir eşitsizliği arasındaki ilişkiyi açıklayan ve "Kuznets eğrisi" olarak bilinen bu şekil, çan eğrisi veya ters U şeklinde tasvir edilmektedir. Bu eğri, gelir ve eşitsizlik arasında benzer bir ilişkiyi ilk kez 1950'lerde gözlemleyen ve öne süren Simon Kuznets'e ithafen *"Çevresel Kuznets Eğrisi (EKC)"* olarak tanımlanmıştır. 1991 yılında, ekonomistler Gene M. Grossman ve Alan B. Krueger, Kuznets Eğrisi'nin çevre verilerine uygulanabileceğini önermiştir [84].

Grossman ve Kruger [84] EKC hipotezini test etmiş ve ekonomik büyüme ile çevresel bozulma arasında ters-U şeklinde bir ilişki olduğunu bulmuşlardır. EKC hipotezi, genellikle "Kişi Başına Gelir" olarak ölçülen ekonomik faaliyetler ile çevresel kalite arasında, kişi başına CO_2 emisyonu gibi çevresel göstergelerle ölçülen, ters-U şeklinde bir ilişkinin var olduğunu varsayar. Yani ekonomik büyümenin ilk aşamasında, çevresel bozulma kişi başına gelirdeki artışla birlikte artar, ancak kişi başına gelirin gelir dönüm noktasını aşmasıyla azalmaya başlar.

EKC hipotezi, bir ülkenin çevresel zararın göreceli seviyesini tahmin etmek için kişi başına düşen GSYİH'ya odaklanır. EKC'nin formuyla ilgili olarak iki temel açıklama bulunmaktadır.

i. EKC hipotezi, gelişmiş ülkelerin tarihini ve kalkınmayı gerçekleştirmek için izledikleri yolları yansıtır,
ii. Gelir arttıkça çevresel kaliteye yönelik tercihlerin değişmesini ifade eder.

Tarihsel olarak, gelişmiş ülkelerin ekonomik yapıları tarıma dayanıyordu ve bu dönemde çevresel zararlar genellikle düşük düzeydeydi. Ancak, sanayileşme süreciyle birlikte kirlilik genellikle artmış, kaynakların tükenmesi ve atık oluşumu hızlanmıştır. Gelişmişlik düzeyi yükseldikçe, bilgiye dayalı endüstrilerdeki ve hizmetlerdeki yapısal değişim, daha verimli teknolojiler ve çevresel kaliteye olan talepteki artış sonucunda, çevresel bozulmanın düzelmeye başlamasına ve sürekli bir azalmayla çevresel kirliliğin azalmasına sebep olmuştur [83]. Yani, yüksek

gelir seviyesi, bireyleri ekonomik büyümeden önce çevresel kaliteyi ön planda tutmaya yönlendirmiştir.

Ancak, Shafik ve Bandyopadhyay [9], Panayotou [85], ve Selden ve Song [8], ekonomik büyüme ile çevresel kalite arasındaki ilişkinin, olumlu veya olumsuz olsun, bir ülkenin gelişim yolunda sabit olmadığını öne sürmüşlerdir. Bu tahmine göre her ülkenin kendi EKC'si vardır ve ülkelerin kendi GSYİH'sına göre; kaynak zenginliği, ekonomideki faaliyetlerin bileşimi ve toplumsal adetler gibi faktörlere dayanmaktadır. Bu konudaki literatür 1992 yılında başlamış olup, 1991 yılında Grossman ve Krueger [84] çevresel kalite ile ekonomik büyüme arasındaki ilişkiyi modelleyen ilk çalışmayı yapmışlardır. Onlar, EKC ilişkisini Kuzey Amerika Serbest Ticaret Anlaşması (NAFTA) bağlamında analiz etmişlerdir. Ardından birçok deneysel çalışma yapılmış ve EKC yoğun bir şekilde araştırılmıştır. Panayotou [85], Selden ve Song [8], Shafik [9] ve Grossman ve Krueger [84], bazı çevreyi kirleten maddelerin gelire bağlı olarak ters-U şeklinde bir eğri izlediğine dair ilk kanıtları sunmuşlardır. Diğer çalışmalar ise EKC'nin var olmadığını, uygulanan modellere göre varsayımlardan sapmalar meydana geldiği için sonuçlarına güvenilemeyeceğini ifade etmişlerdir [86].

EKC hipotezi, çevresel bozulma ile ekonomik gelişme düzeyi arasındaki uzun vadeli ilişkiyi temsil etmeyi amaçlar. EKC genellikle belirli bir ülke veya ülkeler grubu için, kişi başına GSYİH ve seçilen çevresel göstergeler arasındaki ilişki olarak tanımlanır ve her bir ilişki ayrı ayrı ele alınır. Çevresel Kuznets Eğrisi (EKC), karbondioksit (CO_2) emisyonu ile kişi başına düşen Gayri Safi Yurtiçi Hasıla (GSYİH) arasındaki ilişkiyi araştırmak için uygulanır. Bu çalışmalar genellikle çevresel bozulma ve gelir arasında ters U şeklinde bir ilişki olduğunu bulmaktadır, yani gelirin düşük seviyelerinde kirlilik gelirle birlikte artarken, gelirin yüksek seviyelerinde kirlilik azalmaktadır [84].

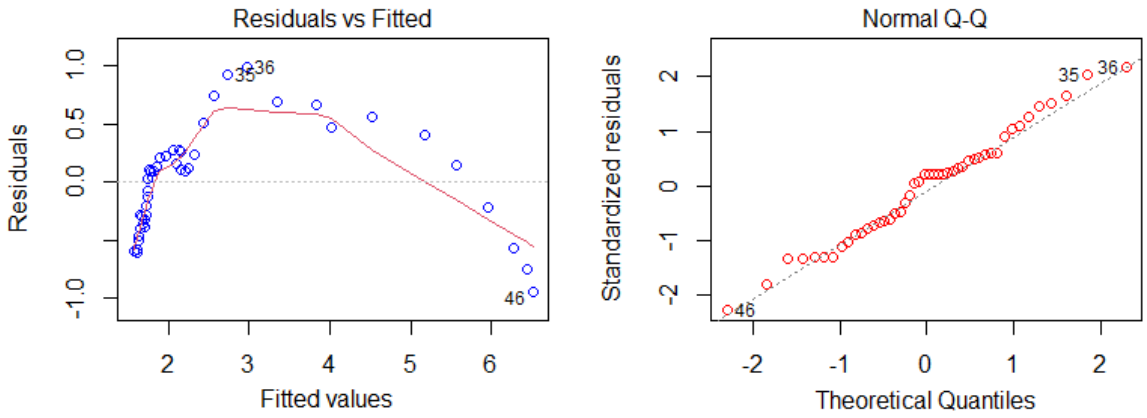
EKC hipotezinin test edilmesi aşamasında literatürde çok farklı değişkenler kullanılmıştır. Örneğin; Stern and Common [87], Culas [88], Managi and Jena [89] gibi çalışmalar kükürt dioksit (SO2), su kirliliği ve ormansızlaşma gibi yerel kirlilik meydana getiren ögeleri kullanmıştır. Zanin and Marra [90], Lean [11] gibi çalışmalar ise çevresel bozulmanın bir ölçüsü olarak karbondioksit (CO_2) gibi

küresel bir kirlilik ögesini kullanmıştır. EKC ile ilgili deneysel çalışmaların sonuçları, incelenen kirliliğin türüne bağlı olarak değişmektedir.

5.3 Çevresel Kuznets Eğrisi (EKC) Uygulaması

EKC hipotezini test etmek için basit bir doğrusal regresyon modeli uygulanmıştır. Bu hipotezi test etmek için veriler Dünya Bankasından elde edilmiştir [91]. Basit en küçük kareler (EKK) yöntemini kullanarak, Doğu Asya ve Asya-Pasifik ülkelerinin (yüksek gelirli ülkeler hariç) 46 yıllık bir dönemde CO_2 emisyonları ile GSYİH değişkenleri analizde kullanılmıştır. CO_2 emisyonları (kişi başına metrik ton) ile GSYİH (güncel ABD doları) arasında kareli bir ilişki ("ters U-hipotezi") bulunmuştur. Başlangıçta, Doğu Asya ve Asya-Pasifik ülkeleri 1990'ların düşük gelirli ülkeleri (LIC) olarak sınıflandırıldı, ancak 2010 yılında düşük orta gelirli ülkeler (LMC) kategorisine dahil edildi. Aslında, CO_2 emisyonlarının en yüksek büyüme oranı %5.6 (1990-2008), en yüksek GSYİH büyüme oranlarının %7.2 (1990-2000) ve %9.4 (2000-2010)) gerçekleştiği Doğu Asya ve Asya-Pasifik bölgesinde gözlemlenmiştir.

Öncelikle kalıntıların diyagramlarını inceleyerek varsayımlardan ciddi sapmalar olmadığı belirlenmiştir:



(a) (b)

Şekil 5.4 Kalıntı grafikleri (a) Kalıntı- Tahmin Değerleri Grafiği (b) Standart Kalıntıların Q-Q grafiği

Lineer ilişkinin incelenmesi sırasında, artıklar ve tahmin edilen değerler bir araya getirilerek dağılım grafikleri elde edilebilir. Eğer dağılım grafiğinde noktalar bir eğriyi temsil ediyorsa, doğrusal ilişki varsayımı geçerli olmamaktadır. Şekil 5.6 incelendiğinde kuadratik bir modelin EKC veri setine daha uygun olduğu saptanmıştır.

Tablo 5.7 Çevre Verileri için Kuadratik Model Parametre Değerleri

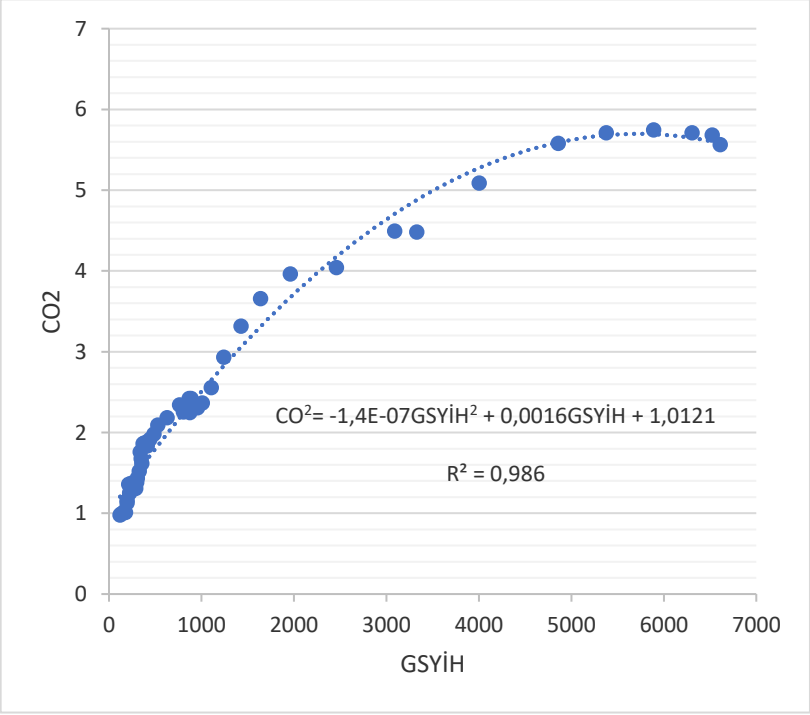
Kaynak	a	b	c	F	R^2
Parametre Değerleri	1,0121	0,0016	1,4E-0,7	1581,224	0,986585
Standard Hata	0,04599	5,9E-05	9,35E-09		
P Değeri	4,99E-25	4,39E-29	5,08E-19		

a: sabit, b and c: model katsayıları, F: F test, R^2: Belirlilik Katsayısı

Elde edilen parametrelere göre kuadratik model aşağıdaki gibi ifade edilir:

$$CO^2 = 1,0121 + 0,0016GDP - (1,4E - 0,7)GDP^2 \qquad (5.3)$$

Denklem (5.3)'de ifade edilen kuadratik modelin grafiği ise aşağıdaki gibidir:



Şekil 5.5 Kuadratik Regresyon Modelinin Grafiği

Modelin varyans analizi (ANOVA) incelendiğinde ise Tablo 5.8 elde edilmiştir:

Tablo 5.8 ANOVA Tablosu

Source of Variation	Df	Sum of Squares	Mean Squares
Explained variation by linear regression	1	SSR= 98,65863	98,65863
Explained variation by nonlinear regression	3	SSLF=4,925404	1,64180133
Unexplained variation	41	SSPE =4,425963	0,107950317
Total	45	SST=108,01	

Basit doğrusal regresyon fonksiyonunun uygunluğunu test etmek için kullanılan sıfır ve alternatif hipotezler şunlardır:

$$H_0 : E(Y) = \beta_0 + \beta_1 X \tag{5.4}$$

$$H_a : E(Y) \neq \beta_0 + \beta_1 X \tag{5.5}$$

Basit regresyon modeli için F istatistiği uygulanmıştır. F istatistiği, regresyon modelinin verilere ne derece uyumlu olduğunu göstermektedir. Eğer F istatistiği istatistiksel olarak anlamlı değilse, iki değişken arasındaki ilişki doğrusal değildir ve regresyon doğrusuyla ifade edilemez.

$$F^* = \frac{SSLF}{(c-2)} \div \frac{SSPE}{(n-c)} = \frac{1.641}{0.107} = 15.209$$

F istatistiği $F(0,05;3,41)$ serbestlik derecesine sahiptir. Bu durumda $F^* > F(0,05;3,41) = 2,833$ olduğundan, "doğrusal regresyon modelinin verilere iyi bir uyum sağladığı" anlamına gelen H_0 hipotezi reddedilir. Veriler ile model arasında doğrusal bir ilişki olmadığı sonucuna varılır.

5.3.1 Karşılıklı Bilgi (MI) Kullanılarak Eğrisellik Analizi

Tanaka, Okamoto ve Naito [80] ile Smith [81] yöntemi, orijinal veri serisi ile lineer bağımlılığın kaldırıldığı yeni veri serisi arasındaki Shannon karşılıklı bilgi ölçüsünü karşılaştırmaya dayanır.

Entropi ve karşılıklı bilgi hesaplamaları bir kontenjans tablosuna dayanır. EKC hipotezinin bir nedeni, fakir ülkelerde üretimin büyük bir kısmının tarım

sektöründe üretiliyor olmasıdır. Bu nedenle, CO_2 emisyonları diğer ülkelere göre daha düşüktür. Orta gelirli ülkelerde ise kirlilik artmaya başlar. Ülke büyüdükçe, daha temiz bir teknolojiye geçme eğiliminde olması beklenir. Freedman- Diaconis kuralına göre, optimum kutu (bölme) sayısı hesaplanmış ve Tablo 5.9'da gösterilmiştir:

Tablo 5.9 Uygun Kutu(Bölme) Sayısı

Variables	n_{bins}
CO_2	7
GSYİH	14
Kalıntılar	9

Doğrusallık tespiti için Shannon, Rényi ve Tsallis karşılıklı bilgi ölçüleri hesaplanmış ve farklı α değerleri için karşılaştırmalar yapılmıştır. Doğrusal olmama durumunu değerlendirmek için iki karşılıklı bilgi ölçüsü karşılaştırılmıştır. $M(X,\xi) = M(X,Y)$ ise değişkenler arasındaki ilişkinin doğrusal olmama durumuna dayandığı sonucu elde edilir. Bu durumda $\lambda_S, \lambda_R, \lambda_T$ değerleri doğrusal olmama durumunun birer kriteri olarak görülebilir. Tablo 5.10'da farklı α değerleri için EKC verisinin $\lambda_S, \lambda_R, \lambda_T$ değerleri görülmektedir.

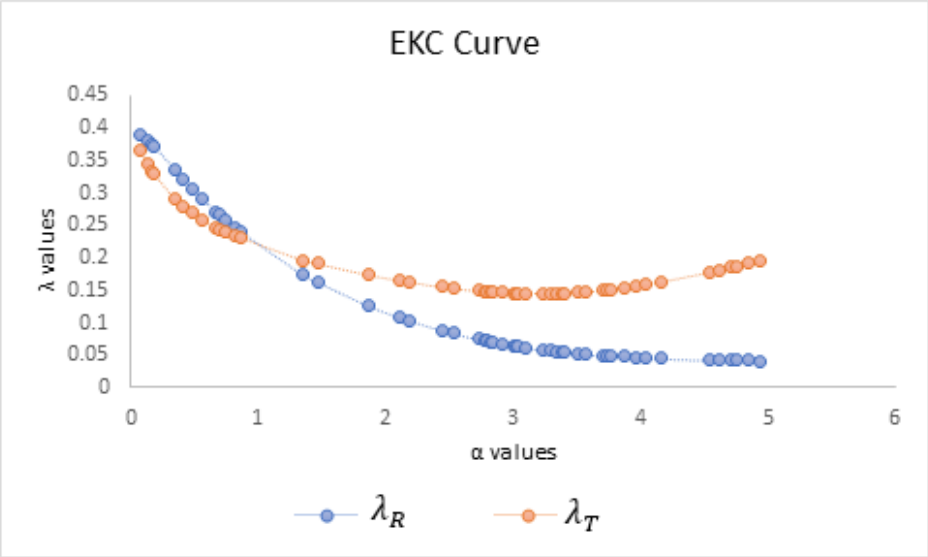
Tablo 5.10 Farklı α değerleri için EKC verisinin $\lambda_S, \lambda_R, \lambda_T$ değerleri

α	λ_R	λ_T	α	λ_R	λ_T
0.07	0.3892	0.3655	3.01	0.0649	0.1460
0.13	0.3809	0.3444	3.02	0.0646	0.1459
0.17	0.3741	0.3323	3.04	0.0640	0.1458
0.18	0.3722	0.3295	3.09	0.0625	0.1455
0.35	0.3355	0.2902	3.23	0.0587	0.1453
0.41	0.3221	0.2797	3.29	0.0573	0.1455
0.48	0.3071	0.2690	3.34	0.0562	0.1457
0.56	0.2910	0.2586	3.38	0.0553	0.1460

Tablo 5.10 Farklı α değerleri için EKC verisinin λ_S, λ_R, λ_T değerleri (devamı)

α	λ_R	λ_T	α	λ_R	λ_T
0.67	0.2708	0.2466	3.4	0.0549	0.1461
0.69	0.2673	0.2447	3.5	0.0530	0.1471
0.74	0.2590	0.2402	3.57	0.0518	0.1481
0.82	0.2468	0.2339	3.71	0.0497	0.1505
0.87	0.2400	0.2307	3.74	0.0493	0.1511
1.36	0.1735	0.1961	3.77	0.0489	0.1518
1.46	0.1631	0.1913	3.88	0.0476	0.1545
1.86	0.1271	0.1743	3.97	0.0467	0.1570
2.11	0.1087	0.1653	4.04	0.0461	0.1591
2.18	0.1040	0.1630	4.17	0.0451	0.1635
2.44	0.0888	0.1556	4.54	0.0430	0.1782
2.54	0.0837	0.1532	4.62	0.0427	0.1817
2.73	0.0751	0.1494	4.71	0.0423	0.1858
2.78	0.0731	0.1486	4.76	0.0421	0.1881
2.8	0.0723	0.1483	4.85	0.0418	0.1923
2.83	0.0711	0.1479	4.94	0.0415	0.1965
2.84	0.0708	0.1477	λ_S	0.2181	0.2181
2.92	0.0679	0.1468	Mean	0.1313	0.1914

Tablo 5.10' da görüldüğü gibi λ_S, λ_R, λ_T değerleri sıfıra yakın elde edilmiştir. Bu durumda ilişkinin doğrusal olmadığı ifade edilir. Ayrıca λ_R, λ_T'nin standart sapmaları ise sırasıyla 0,1147 ve 0.0607 bulunmuştur.



Şekil 5.6 Çevre Verilerine Göre λ ve α İlişkisi

Şekil 5.6'da EKC verilerine göre λ ve α ilişkisi görülmektedir. Simülasyon datasından elde edilen α değerlerine göre $\alpha<1$ için λ_T ve $\alpha>1$ için λ_R eğriselliği başarılı bir şekilde açıklamaktadır. Simülasyon sonucunda elde edilen bu bulguları EKC verisi ile elde edilen sonuçlarda desteklemektedir. Sonuç olarak, λ_S, λ_R, λ_T parametreleri EKC hipotezini destekler şekilde sıfıra yakın çıkmışlardır, bu durumda eğrisel ilişkiyi desteklemektedir [92].

5.3.2 Sonuç 2

Akaike Bilgi Kriteri veya kısaca AIC, bir modeli seçme yöntemidir. Bu yöntemin geliştiricisi olan Hirotugu Akaike'nin adını taşır ve bilgi teorisine dayanır. En düşük AIC değerine sahip model en iyi model olarak seçilir. AIC, herhangi bir tahmin edilen istatistiksel modelin uygunluk derecesi olarak adlandırılabilir. AIC, karşılaştırılabilir olmayan modelleri karşılaştırmak için kullanılabilir, yani modeller doğrudan karşılaştırılabilir olmasa bile kullanılabilir. Örneğin, AIC'yi bir doğrusal regresyon modelini bir lojistik regresyon modeli ile karşılaştırmak için hesaplanabilir.

BIC ise farklı parametre sayılarına sahip bir parametrik model sınıfı arasında model seçimi yapmanın bir türüdür. BIC veya Bayesian Bilgi Kriteri, iki modelin uyum iyiliğini karşılaştırmak için kullanılan istatistiksel bir araçtır. BIC, en iyi

modelin verilerin olasılığını en üst düzeye çıkaran, aynı zamanda modele eklenen parametre sayısına da ceza veren fikrine dayanır. Bu, modelin veriye aşırı uyumlu olmamasını sağlar. BIC, regresyon analizi ve makine öğrenimi gibi çeşitli alanlarda kullanılabilir. Genel olarak, BIC, farklı modelleri karşılaştırmak zorunda olan istatistikçiler ve veri bilimciler için değerli bir araçtır.

Hem AIC hem de BIC, bir modelin veriye uygunluğunu tahmin eder, ancak AIC bunu olabilirlik fonksiyonunu maksimize ederek yapar, BIC ise Bayesian bilgi kaybı fonksiyonunu minimize ederek yapar.

AIC genellikle BIC'den daha güvenilir olarak kabul edilir, ancak BIC daha hızlı hesaplanabilir. AIC aşırı uyumdan etkilenebilirken, BIC etkilenmez. Sonuç olarak, hem AIC hem de BIC model seçimi için kullanışlı araçlar olabilir, ancak genellikle AIC daha doğru kabul edilir.

Uygulamaya bakıldığında Kuadratik model daha iyi bir model olarak ortaya çıkar. Tablo 5.11'de model belirleme kriterleri sonuçları görülmektedir:

Tablo 5.11 Çevre Verileri için Model Belirleme Kriterleri

Model Belirleme Kriterleri	AIC	BIC	RMSE	MAE	R^2
Doğrusal Regresyon	63.2588	68.7448	0.4508776	0.3724944	0.9134213
Kuadratik Regresyon	-20.51759	-13.203	0.1774773	0.1501555	0.9865853

Negatif bir AIC, pozitif bir AIC'den daha az bilgi kaybı gösterir ve dolayısıyla daha iyi bir modeli ifade eder.

R^2, gözlemlenen sonuç değerleri ile model tarafından tahmin edilen değerler arasındaki korelasyonun karesini ifade eder. R-kare ne kadar büyükse, modelin başarısı o kadar yüksektir.

Uygulamaya bakıldığında aşağıdaki sonuçlar elde edilmiştir:

- Doğrusal ve eğrisel modellerin R^2' leri karşılaştırıldığında kuadratik modelin R^2'si daha yüksek çıkmıştır. Dolayısıyla kuadratik model daha iyi bir modeldir.
- Uygulamaya bakıldığında iki modelinde MAE değerleri birbirlerinden farklı çıkmıştır. Kuadratik modelin MAE değeri doğrusal modelden daha düşük çıktığı için kullanılan veriler için kuadratik model daha iyi bir model olarak belirlenmiştir.
- Kuadratik modelin RMSE'si daha düşük çıktığı için, veriler için daha doğru, daha uygun bir model olduğu sonucuna varılmıştır.
- Yapılan karşılıklı bilgi analizinde ise; λ_S, λ_R, λ_T parametreleri EKC hipotezini destekler şekilde sıfıra yakın çıkmışlardır. Bu durumda bu parametreler eğrisel ilişkinin birer göstergesi olarak kullanılabilirler.

6
SONUÇ VE ÖNERİLER

EKC hipotezinin etkinliği gelişmekte olan ülkelerin özel durumlarını inceleyerek, sürdürülebilir kalkınma için bir çerçeve olarak analiz edilmelidir. Bu çalışmada ekonomik büyümenin çevresel endişeleri göz ardı etmesinin uzun vadeli çevresel bozulmaya ve beraberinde gelen sonuçlara yol açıp açmadığını incelemiştir.

EKC'nin temel mesajı "önce büyü ve kirlet, sonra temizle" dir. EKC hipotezi, gelişmekte olan ve gelişmiş ülkelerin ekonomi politikaları üzerinde önemli bir etki yaratmıştır. Bu hipotez aynı zamanda IMF ve Dünya Bankası'nın politikalarını ve önceliklerini etkilemiştir ki bu da büyümeye yönelik politikalarından anlaşılmaktadır. Hızlı ekonomik büyüme, çevre sorunlarına yeterli özen göstermeksizin yoksulluğu ve işsizliği azaltmanın temel hedefi haline gelmiştir.

EKC'nin "önce büyü ve kirlet, sonra temizle" yaklaşımı, dünyayı küresel ısınma ve iklim değişikliği gibi çevresel değişikliklerle karşı karşıya getirmiştir. Yükselen ekonomilerdeki yüksek gelir büyümesi, su ve hava kirliliği, ormanlardaki azalmalar, bozulmuş hava kalitesi, kentsel ve endüstriyel atıkların birikmesi ve biyolojik çeşitlilik kaybı gibi çok yüksek çevre maliyetlerine neden olduğu görülmüştür. Bu çevre sorunları, insan hayatı için çok tehlikeli sonuçlar doğurmaya başlamıştır.

Bununla birlikte bilimsel araştırmalar ekonomik büyümenin çevre üzerindeki olumsuz etkilerine dikkat çekmektedir. Bu etkiler arasında sera gazlarının aşırı birikimi, hava kirliliğinin artması ve atık üretiminin önemli ölçüde artması gibi durumlar yer almaktadır.

Bu tez çalışmasında doğrusal en küçük kareler tahmin yöntemini kullanarak, Doğu Asya ve Asya-Pasifik ülkelerinin 46 yıllık bir zaman serisinde CO_2 emisyonu ve GSYİH arasında kuadratik bir ilişki olduğu görülmüştür. İlk olarak, EKC hipotezini kontrol etmek için F testi kullanılmış ve bu sayede doğrusal modelin verilere iyi bir uyum sağlayıp sağlamadığı sonucu araştırılmıştır. İkinci bir teknik olarak, doğrusal belirleme katsayısı ile korelasyon oranını karşılaştırmak faydalı olacağı

düşünülerek bu ölçüler karşılaştırılmıştır. EKC verileri için bu iki ilişki ölçüsü arasındaki fark anlamlı bulunmuştur ve eğrisellikle ilgili bilgi sahibi olunmasını sağlamıştır.

Alternatif olarak, Karşılıklı Bilgi Miktarına dayanan iki bağımlılık ölçüsü arasındaki fark hesaplanmıştır. Literatürde daha sık olarak Shannon karşılıklı bilgisi kullanılmış olsa da, Rényi ve Tsallis karşılıklı bilgi ölçülerinin parametrik esneklikleri nedeniyle değişkenler arasındaki ilişkinin doğasını daha iyi yakaladığı tespit edilmiştir.

Bu çalışmada, ilk olarak bağımlı ve bağımsız değişkenler arasındaki karşılıklı bilgi miktarı ($M(X,Y)$) bulunmuştur. İkinci olarak, basit bir doğrusal regresyon modeli kullanılarak kalıntılar (ξ) hesaplanmıştır. Ardından, bağımsız değişken ile kalıntılar arasındaki karşılıklı bilgi miktarı $M(X,\xi)$ elde edilmiştir. Son olarak, bu iki karşılıklı bilgi ölçüsünün karşılaştırılmasıyla, bağımlılığa dahil edilen eğrisellik derecesi belirlenmiştir. Ayrıca, gelecek çalışmalar için literatüre önemli katkı sağlayacağı düşünülen alternatif (değişkenlerinde) eğrisellik ölçüsü olan λ önerilmiştir. Bu sayede Rényi ve Tsallis karşılıklı bilgi ölçülerinin, belirli α değerleri aralıklarında Shannon karşılıklı bilgi ölçüsüne göre eğrisellikleri daha iyi belirlediği tespit edilmiştir.

Bu ölçümlerin CO_2 emisyonları ve gayri safi yurt içi hasıla (GSYİH) verilerine uygulanması, eğrisellik durumunu tespit edilmesine yardımcı olmuş ve dolayısıyla EKC hipotezinin öngörülen modelinin gerçekçi ve tutarlı olduğunu ortaya koymuştur. Sonuçlar, "şimdi büyü ve kirlet, sonra temizle" stratejisinin çok fazla kaynağı israf ettiğini ve büyük çevresel maliyetlere sahip olduğunu göstermektedir. Bu nedenle, ülkelere çevre kirliliklerine neden olmayacak alternatif büyüme stratejileri aramaları önerilmektedir.

KAYNAKÇA

[1] UNKNOWN, "OBITUARY. RUDOLF JULIUS EMANUEL CLAUSIUS, 1822-1888," in *Minutes of the Proceedings of the Institution of Civil Engineers*, 1889, vol. 96, no. 1889: Thomas Telford-ICE Virtual Library, pp. 307-316.

[2] I. Csiszár and P. C. Shields, "Information theory and statistics: A tutorial," *Foundations and Trends® in Communications and Information Theory,* vol. 1, no. 4, pp. 417-528, 2004.

[3] K. G. Denbigh, "Note on entropy, disorder and disorganization," *The British Journal for the Philosophy of Science,* vol. 40, no. 3, pp. 323-332, 1989.

[4] R. M. Gray, *Entropy and information theory*. Springer Science & Business Media, 2011.

[5] I. Dincer and Y. A. Cengel, "Energy, entropy and exergy concepts and their roles in thermal engineering," *Entropy,* vol. 3, no. 3, pp. 116-149, 2001.

[6] R. Pathria and P. D. Beale, "1-The Statistical Basis of Thermodynamics," *Statistical Mechanics,* pp. 1-23, 2011.

[7] H. Callen, "Introduction to Thermodynamics and Thermostatistics," ed: John Wiley and Sons, 1985.

[8] T. M. Selden and D. Song, "Environmental quality and development: is there a Kuznets curve for air pollution emissions?," *Journal of Environmental Economics and management,* vol. 27, no. 2, pp. 147-162, 1994.

[9] N. Shafik and S. Bandyopadhyay, *Economic growth and environmental quality: time-series and cross-country evidence*. World Bank Publications, 1992.

[10] W. Beckerman, "Economic growth and the environment: Whose growth? Whose environment?," *World development,* vol. 20, no. 4, pp. 481-496, 1992.

[11] F. Tan, H. H. Lean, and H. Khan, "Growth and environmental quality in Singapore: is there any trade-off?," *Ecological indicators,* vol. 47, pp. 149-155, 2014.

[12] Ş. SARAÇ and A. YAĞLIKARA, "Environmental Kuznets Curve: The Evidence from BSEC Countries," *Ege Academic Review,* vol. 17, no. 2, pp. 255-264, 2017.

[13] T. M. Cover and J. A. Thomas, "Entropy, relative entropy and mutual information," *Elements of information theory,* vol. 2, no. 1, pp. 12-13, 1991.

[14] E. A. Guggenheim, "The history of the entropy concept," *Journal of Chemical Education,* vol. 35, pp. 584-591, 1958, Art no. 12.

[15] R. Swendsen, *An introduction to statistical mechanics and thermodynamics*. Oxford University Press, USA, 2020.

[16] L. Boltzmann, "Uber die Beziehung zwischen des zweiten Hauptsatze der mechanischen der Warmetheorie," *On the Relation of a General Mechanical Theorem to the Second Law of Thermodynamics"), Sitzungsberichte, K. Akademie der Wissenschaften in Wien, Math.-Naturwiss,* vol. 75, pp. 67-73, 1877.

[17] V. P. Singh, *Introduction to tsallis entropy theory in water engineering*. CRC Press, 2016.

[18] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal,* vol. 27, no. 3, pp. 379-423, 1948.

[19] A. D. Al-Nasser, "The Journey from Entropy to Generalized Maximum Entropy," *Journal of Quantitative Methods,* vol. 3, no. 1, pp. 1-7, 2019.

[20] E. T. Jaynes, "Information theory and statistical mechanics. II," *Physical review,* vol. 108, no. 2, p. 171, 1957.

[21] A. Takada, R. Conradt, and P. Richet, "Residual entropy and structural disorder in glass: A review of history and an attempt to resolve two apparently conflicting views," *Journal of Non-Crystalline Solids,* vol. 429, pp. 33-44, 2015.

[22] F. Abraham, *Homogeneous nucleation theory: the pretransition theory of vapor condensation*. Elsevier, 2012.

[23] V. Berdichevskii, "The connection between thermodynamic entropy and probability," *Journal of Applied Mathematics and Mechanics,* vol. 52, no. 6, pp. 738-746, 1988.

[24] L. M. Martyushev and V. D. Seleznev, "Maximum entropy production principle in physics, chemistry and biology," *Physics reports,* vol. 426, no. 1, pp. 1-45, 2006.

[25] J. Bensinger, "Introduction to Thermodynamics and Statistical Mechanics," *Instructor,* 2016.

[26] A. Wehrl, "General properties of entropy," *Reviews of Modern Physics,* vol. 50, no. 2, p. 221, 1978.

[27] J. D. Fast, *Entropy: the significance of the concept of entropy and its applications in science and technology*. Macmillan, 1982.

[28] C. De Fraiture, *The use of entropy optimization principles in parameter estimation: Applications to global water demand modeling*. University of Colorado at Boulder, 2003.

[29] S. J. Blundell and K. M. Blundell, *Concepts in thermal physics*. Oxford University Press on Demand, 2010.

[30] D. Ellerman, "An introduction to logical entropy and its relation to Shannon entropy," vol. 7, ed: World Scientific, 2013, pp. 121-145.

[31] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin, "Inequalities for Shannon entropy and Kolmogorov complexity," *Journal of Computer and System Sciences,* vol. 60, no. 2, pp. 442-464, 2000.

[32] A. Ben-Naim, "Entropy, Shannon's measure of information and Boltzmann's H-theorem," *Entropy,* vol. 19, no. 2, p. 48, 2017.

[33] N. Eshima, *Statistical data analysis and entropy*. Springer, 2020.

[34] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[35] W. Dong, "Mutual information: inferring tie strength and proximity in bipartite social network data with non-metric associations," 2012.

[36] J. V. Stone, "Information theory: a tutorial introduction," 2015.

[37] S. SINGH, "ON SOME GENERALIZED ENTROPY MEASURES FOR DOUBLY TRUNCATED RANDOM VARIABLE," RAJIV GANDHI INSTITUTE OF PETROLEUM TECHNOLOGY, 2021.

[38] M. S. Tabass, G. M. Borzadaran, and M. Amini, "Renyi entropy in continuous case is not the limit of discrete case," *Mathematical Sciences and Applications E-Notes,* vol. 4, no. 1, pp. 113-117, 2016.

[39] T. Villmann, J. C. Principe, and A. Cichocki, "Information theory related learning," in *ESANN*, 2011: Citeseer.

[40] J. C. Principe, *Information theoretic learning: Renyi's entropy and kernel perspectives*. Springer Science & Business Media, 2010.

[41] J. C. Principe, D. Xu, and D. Erdogmuns, "Renyi's entropy, divergence and their nonparametric estimators," *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives,* pp. 47-102, 2010.

[42] C. Tsallis, "Generalized entropy-based criterion for consistent testing," *Physical Review E,* vol. 58, no. 2, p. 1442, 1998.

[43] S. Furuichi, K. Yanagi, and K. Kuriyama, "Fundamental properties of Tsallis relative entropy," *Journal of Mathematical Physics,* vol. 45, no. 12, pp. 4868-4877, 2004.

[44] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *Journal of statistical physics,* vol. 52, pp. 479-487, 1988.

[45] S. Furuichi, F.-C. Mitroi-Symeonidis, and E. Symeonidis, "On some properties of Tsallis hypoentropies and hypodivergences," *Entropy,* vol. 16, no. 10, pp. 5377-5399, 2014.

[46] F. Nielsen and R. Nock, "On R\'enyi and Tsallis entropies and divergences for exponential families," *arXiv preprint arXiv:1105.3259,* 2011.

[47] S. Kullback, *Information theory and statistics*. Courier Corporation, 1997.

[48] P. Réfrégier and F. Goudail, "Kullback relative entropy and characterization of partially polarized optical waves," *JOSA A,* vol. 23, no. 3, pp. 671-678, 2006.

[49] S.-i. Amari, *Differential-geometrical methods in statistics*. Springer Science & Business Media, 2012.

[50] L. Pardo, *Statistical inference based on divergence measures*. CRC press, 2018.

[51] H. Nyquist, "Certain factors affecting telegraph speed," *Transactions of the American Institute of Electrical Engineers,* vol. 43, pp. 412-422, 1924.

[52] R. V. Hartley, "Transmission of information 1," *Bell System technical journal,* vol. 7, no. 3, pp. 535-563, 1928.

[53] N. Wiener, "Cybernetics Wiley," *New York,* 1948.

[54] F. Bavaud, J. Chappelier, and J. Kohlas, "An introduction to information theory and applications," *UniFr course, version,* vol. 2, pp. 9-11, 2005.

[55] S. Roman, *Introduction to coding and information theory*. Springer Science & Business Media, 1996.

[56] A. Toomaj and A. Di Crescenzo, "Generalized entropies, variance and applications," *Entropy,* vol. 22, no. 6, p. 709, 2020.

[57] D. Mukher jee and M. V. Ratnaparkhi, "On the functional relationship between entropy and variance with related applications," *Communications in Statistics-Theory and Methods,* vol. 15, no. 1, pp. 291-311, 1986.

[58] A. M. Peccarelli, *A comparison of variance and renyi's entropy with application to machine learning*. Northern Illinois University, 2017.

[59] M. Thomas and A. T. Joy, *Elements of information theory*. Wiley-Interscience, 2006.

[60] D. Freedman and P. Diaconis, "On the histogram as a density estimator: L 2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete,* vol. 57, no. 4, pp. 453-476, 1981.

[61] H. Akaike, "Information theory and an extension of the maximum likelihood principle," *Selected papers of hirotugu akaike,* pp. 199-213, 1998.

[62] K. P. Burnham, D. R. Anderson, and K. P. Huyvaert, "AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons," *Behavioral ecology and sociobiology,* vol. 65, pp. 23-35, 2011.

[63] K. P. Burnham and D. R. Anderson, "Multimodel inference: understanding AIC and BIC in model selection," *Sociological methods & research,* vol. 33, no. 2, pp. 261-304, 2004.

[64] R. E. Kass and L. Wasserman, "A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion," *Journal of the american statistical association,* vol. 90, no. 431, pp. 928-934, 1995.

[65] Y. Wang, Y. Li, H. Cao, M. Xiong, Y. Y. Shugart, and L. Jin, "Efficient test for nonlinear dependence of two continuous variables," *BMC bioinformatics,* vol. 16, pp. 1-8, 2015.

[66] C. E. Weatherburn, *A first course mathematical statistics*. CUP Archive, 1949.

[67] J. Neter, W. Wasserman, and M. H. Kutner, "Applied linear statistical models: Regression," *Analysis of variance, and experimental designs,* vol. 1, no. 985, pp. 382-393, 1985.

[68] A. Hector, S. Von Felten, and B. Schmid, "Analysis of variance with unbalanced data: an update for ecology & evolution," *Journal of animal ecology,* vol. 79, no. 2, pp. 308-316, 2010.

[69] J. Neter, M. H. Kutner, C. J. Nachtsheim, and W. Wasserman, "Applied linear statistical models," 1996.

[70] Z. Z. Şahinbaşoğlu, "Doğrusal olmayan regresyonda bazı eğrisellik ölçüleri," 2005.

[71] J. Q. Su and L. Wei, "A lack-of-fit test for the mean function in a generalized linear model," *Journal of the American Statistical Association,* vol. 86, no. 414, pp. 420-426, 1991.

[72] M. D. Esteban and D. Morales, "A summary on entropy statistics," *Kybernetika,* vol. 31, no. 4, pp. 337-346, 1995.

[73] A. Dionisio, R. Menezes, and D. A. Mendes, "Mutual information: a measure of dependency for nonlinear time series," *Physica A: Statistical Mechanics and its Applications,* vol. 344, no. 1-2, pp. 326-329, 2004.

[74] A. Q. Barbi and G. A. Prataviera, "Nonlinear dependencies on Brazilian equity network from mutual information minimum spanning trees," *Physica A: Statistical Mechanics and its Applications,* vol. 523, pp. 876-885, 2019.

[75] W. Mohti, A. Dionísio, P. Ferreira, and I. Vieira, "Frontier markets' efficiency: mutual information and detrended fluctuation analyses," *Journal of Economic Interaction and Coordination,* vol. 14, no. 3, pp. 551-572, 2019.

[76] X. Wu, L. Jin, and M. Xiong, "Mutual information for testing gene-environment interaction," *PloS one,* vol. 4, no. 2, p. e4578, 2009.

[77] J. Li, W. Dong, and D. Meng, "Grouped gene selection of cancer via adaptive sparse group lasso based on conditional mutual information," *IEEE/ACM transactions on computational biology and bioinformatics,* vol. 15, no. 6, pp. 2028-2038, 2017.

[78] F. E. Rosas, P. A. Mediano, M. Gastpar, and H. J. Jensen, "Quantifying high-order interdependencies via multivariate extensions of the mutual information," *Physical Review E,* vol. 100, no. 3, p. 032305, 2019.

[79] A. Ullah, "Entropy, divergence and distance measures with econometric applications," *Journal of Statistical Planning and Inference,* vol. 49, no. 1, pp. 137-162, 1996.

[80] N. Tanaka, H. Okamoto, and M. Naito, "Detecting and evaluating intrinsic nonlinearity present in the mutual dependence between two variables," *Physica D: Nonlinear Phenomena,* vol. 147, no. 1-2, pp. 1-11, 2000.

[81] R. Smith, "A mutual information approach to calculating nonlinearity," *Stat,* vol. 4, no. 1, pp. 291-303, 2015.

[82] D. H. Meadows, D. H. Meadows, J. Randers, and W. W. Behrens III, "The limits to growth: a report to the club of Rome (1972)," *Google Scholar,* vol. 91, p. 2, 1972.

[83] C. Cialani, "Economic growth and environmental quality: An econometric and a decomposition analysis," *Management of Environmental Quality: An International Journal,* 2007.

[84] G. M. Grossman and A. B. Krueger, "Economic growth and the environment," *The quarterly journal of economics,* vol. 110, no. 2, pp. 353-377, 1995.

[85] T. Panayotou, "Empirical tests and policy analysis of environmental degradation at different stages of economic development," 1993.

[86] D. I. Stern, "Beyond the environmental Kuznets curve: diffusion of sulfur-emissions-abating technology," *The Journal of Environment & Development,* vol. 14, no. 1, pp. 101-124, 2005.

[87] D. I. Stern and M. S. Common, "Is there an environmental Kuznets curve for sulfur?," *Journal of environmental economics and management,* vol. 41, no. 2, pp. 162-178, 2001.

[88] R. J. Culas, "Deforestation and the environmental Kuznets curve: An institutional perspective," *Ecological Economics,* vol. 61, no. 2-3, pp. 429-437, 2007.

[89] S. Managi and P. R. Jena, "Environmental productivity and Kuznets curve in India," *Ecological Economics,* vol. 65, no. 2, pp. 432-440, 2008.

[90] L. Zanin and G. Marra, "Assessing the functional relationship between CO2 emissions and economic development using an additive mixed model approach," *Economic Modelling,* vol. 29, no. 4, pp. 1328-1337, 2012.

[91] "World Bank." https://databank.worldbank.org/source/world-development-indicators (accessed 2022, 2022).

[92] E. Tuna, A. Evren, E. Ustaoğlu, B. Şahin, and Z. Z. Şahinbaşoğlu, "Testing Nonlinearity with Rényi and Tsallis Mutual Information with an Application in the EKC Hypothesis," *Entropy,* vol. 25, no. 1, p. 79, 2022.

TEZDEN ÜRETİLMİŞ YAYINLAR

Makale

1. Tuna; E.; Evren;A.; Ustaoğlu;E.; Şahin;B.;Şahinbaşoğlu;Z.Z; “Testing Nonlinearity with Rényi and Tsallis Mutual Information with an Application in the EKC Hypothesis” Journal of Entropy Volume 25 Issue 1;2022. DOI: 10.3390/e25010079 (SCI Expanded)

Konferans Bildirisi

1. Şahinbaşoğlu; Z, Z; Tuna; E.; Evren; A; “Comparision of Shannon; Renyi and Tsallis Mutual Information in Detecting Nonlinearity” International Conference on Global Practice of Multidisciplinary Scientific Studies - IV, Kuzey Kıbrıs Türk Cumhuriyeti, 28 - 30 Nisan 2023, ss.277