

**ENTROPİ OPTİMİZASYON METODLARIYLA
RASSAL DEĞİŞKENLERİN
DAĞILIMLARININ İNCELENMESİ**

Yeliz MERT KANTAR
Doktora Tezi

İstatistik Bölümü Anabilim Dalı
Mayıs- 2006

ÖZET
Doktora Tezi
ENTROPİ OPTİMİZASYON METODLARIYLA
RASSAL DEĞİŞKENLERİN
DAĞILIMLARININ İNCELENMESİ

Yeliz MERT KANTAR

Anadolu Üniversitesi

Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman: Prof. Dr. Aladdin ŞAMILOV

2006, 79 sayfa

Bu tezde, MaxEnt metodu yardımıyla moment fonksiyonlarına bağlı yeni bir fonksiyonel tanımlanarak, bu fonksiyonele minimum değer veren ve informasyonu maksimum yapan bir dağılım belirleme süreci geliştirilmiştir. Bu süreç sonucu elde edilmiş dağılım MinMaxEnt olarak adlandırılmıştır. Bu dağılımın anakütlenin dağılımına MaxEnt ölçümüne göre en yakın dağılım olduğu gösterilmiştir. Geliştirilen bu süreç MaxEnt metodunun bir genelleşmesi olarak düşünülebilir.

Benzer şekilde Kullback-Leibler (K-L) ölçüsü yardımıyla moment fonksiyonlarına bağlı yeni bir fonksiyonel tanımlanmış ve bu fonksiyonele maksimum değer veren ve Kullback-Leibler ölçüsüne göre ana kütlelin dağılımına en yakın dağılım bulma süreci geliştirilmiştir. Bu süreç sonucu elde edilmiş dağılım MaxMinxEnt olarak adlandırılmıştır.

Uygulama olarak, MinMaxEnt ve MaxMinxEnt dağılımları kullanılarak, Eskisehir bölgesinde ölçülen rüzgâr hızının dağılımları bulunmuş ve klasik istatistiğin bir modeli olan, rüzgâr alanında uygulanabilirliği kabul edilen 2-parametrelili Weibull dağılımıyla karşılaştırılmıştır. Sonuç olarak, MinMaxEnt ve MaxMinxEnt dağılımları rüzgar hızını modellemede Weibull'a göre daha iyi sonuçlar verdiği gösterilmiştir.

Anahtar Kelimeler: Maximum Entropi, Kullback-Leibler ölçüsü, MinMaxEnt ve MaxMinxEnt dağılımları, Model belirleme kriterleri

ABSTRACT
PhD Dissertation
THE DETERMINATION OF DISTRIBUTIONS OF RANDOM
VARIABLE BY USING ENTROPY OPTIMIZATION METHODS

Yeliz MERT KANTAR

Anadolu University
Graduate School of Sciences
Statistics Program

Supervisor: Prof. Dr. Aladdin ŞAMILOV

2006, 79 pages

In this thesis, a new functional which depends on moment functions is defined by means of MaxEnt measure. By using mentioned functional, the process of determining a distribution is developed. This distribution has the greatest information and is closest to the distribution of the population in the sense of MaxEnt measure. It is natural to call the determined distribution as MinMaxEnt distribution. This process can be seen as the generalization of MaxEnt method.

Moreover, a functional which depends on moment functions is defined by means of Kullback-Leibler (K-L) measure. By using mentioned functional, the process of determining a distribution is developed. This distribution is closest to the distribution of the population in the sense of K-L measure. It is natural to call the determined distribution as MaxMinxEnt distribution. This process can be seen as the generalization of MinxEnt method

In application, using MinMaxEnt and MaxMinxEnt distributions, the distribution of the wind speed measured in Eskisehir are found and compared with 2-parameter Weibull distribution which is classical statistical model, accepted in wind study. Consequently, it is showed that MinMaxEnt and MaxMinxEnt distributions give better results for wind speed distributions, compared to the Weibull which is widely used in wind studies.

Keywords: Maximum Entropy, Kullback-Leibler measure, MinMaxEnt and MaxMinxEnt distributions, Model selection criteria

TEŞEKKÜR

Hazırlamış olduğum Doktora tezine hazırlık aşamasında, bölümümüzde açmış olduğu “Ölçüm Teorisi ve Olasılık”, “Entropi ve İnfomasyon Teorisi”, “İstatistiğin Matematiksel Yöntemleri” gibi çok önemli konuları kapsayan dersleriyle bilgilerini bizlere aktaran, ayrıca tez konusunun belirlenmesi, çalışmaların planlanması, içeriğin oluşturulması ve geliştirilmesinde bana gereken yardımını esirgemeyen saygı değer tez danışmanım Prof. Dr. Aladdin ŞAMİLOV’ a en derin minnet ve şükranlarımı sunarım. Ayrıca yapılan seminerler ve çalışmalar süresince değerli düşünceleriyle, engin bilgisiyle bizi yönlendirdiği ve zaman ayırdığı için özellikle teşekkür ederim.

Kendisini tanımış olduğum zamandan itibaren, geçirmiş olduğum tüm süreçlerde desteğini benden hiç esirgemeyen, en ufak bir problemimde dahi beni yalnız bırakmayan, güven, sevgi ve saygı duyduğum değerli hocam Prof. Dr. Embiya AĞAOĞLU’ na yürekten teşekkürü bir borç bilirim.

Çalışmalarım süresince bana oldukça huzurlu bir çalışma ortamı sağlayıp, manevi desteğini hiç esirgemeyen ve sürekli beni cesaretlendiren değerli hocam Prof. Dr. Ali Fuat YÜZER’ e teşekkür ederim.

Tez çalışmalarım süresince bana zaman ayıran, destek olan ve yardım eden, her an yanımda olduklarını hissettiğim sevgili hocalarım Haluk HÜSEYİN, Mammadagha MAMMADOV, Berna YAZICI, çok sevdiğim arkadaşlarım R.Ece OMAV, Şenay YOLAÇAN, İlhan USTA, Çiğdem GİRİFTİNOĞLU, Betül KAN’a teşekkür ederim.

Tezin içeriğine ve uygulamalarına yaptığı katkılardan dolayı arkadaşım İlhan USTA’ ya ayrıca teşekkür etmek isterim.

Beni yetiştiren aileme, bu çalışma süresince gösterdiği sabır ve anlayış için eşim Hakan KANTAR’ a teşekkür ederim.

Yeliz MERT KANTAR

Mayıs-2006

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
2 MAKSİMUM ENTROPİ METODU	6
2.1. Kesikli Rassal Değişkenler İçin Maksimum Entropi Metodu.....	6
2.1.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı.....	9
2.1.2. Varyans-Covaryans matrisi	11
2.1.3. Koşullu Ekstremumun varlığı	11
2.1.4. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği.....	13
2.1.5. Entropi Fonksiyonun maksimum değeri.....	14
2.1.6. H_{\max} 'ın a_1, a_2, \dots, a_m 'a göre Konkavlığı.....	14
2.2. Sürekli Rassal Değişkenler İçin Maksimum Entropi Metodu.....	16
2.2.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı.....	18
2.2.2. Koşullu Ekstremumun Varlığı.....	20
2.2.3. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği.....	21
3. MİNİMUM ÇAPRAZ ENTROPİ METODU	23
3.1. Kesikli Rassal Değişkenler İçin MinxEnt Metodu.....	23
3.1.1. Koşullu Ekstremumun varlığı.....	24
3.1.2. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı.....	25
3.1.3. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre	

Konveksliđi.....	27
3.1.4. D_{\min} 'in a_1, \dots, a_m 'e gre Konveksliđi.....	29
3.2. Srekli Rassal Deđiřken İin MinxEnt Metodu.....	30
3.2.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange arpanlarının Varlıđı.....	32
3.2.2. Kořullu Ekstremumun Varlıđı.....	34
3.2.3. λ_0 arpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange arpanlarına gre Konveksliđi.....	35
4. İNFORMASYON TEORİSİNE DAYALI MODEL BELİRLEME KRİTERLERİ.....	37
4.1. Akaike bilgiyon kriteri.....	37
4.2. Bayes bilgiyon kriteri.....	38
4.3. İstatistiksel kriterler.....	38
5. GENELLEŐTİRİLMİŐ MAKSİMUM ENTROPİ METODU VE GELİŐTİRİLMİŐ DAĐILIMLAR.....	40
5.1. Kesikli Rassal Deđiřkenler İin Maksimum Entropi Fonksiyonu ve Moment Fonksiyonları Yardımıyla TanımlanmıŐ Bir Fonksiyonel...	40
5.2. Kesikli Rassal Deđiřkenler iin MinMaxEnt Dađılımı.....	42
5.3. Srekli Rassal Deđiřkenler İin Maksimum Entropi Fonksiyoneli ve Moment Fonksiyonları Yardımıyla TanımlanmıŐ Bir Fonksiyonel.....	43
5.4. Srekli Rassal Deđiřkenler iin MinMaxEnt Dađılımı.....	44
5.5. MinMaxEnt ve MaxMaxEnt dađılım dizisi.....	45
5.6. Rzgar Hızınının MinMaxEnt dađılımı: Weibull dađılımına Alternatif.....	46
6. GENELLEŐTİRİLMİŐ MİNİMUM APRAZ ENTROPİ METODU VE GELİŐTİRİLMİŐ DAĐILIMLAR.....	58
6.1. Kesikli Rassal Deđiřkenler iin Kullback-Leibler Fonksiyonu ve Moment Fonksiyonları Yardımıyla TanımlanmıŐ Bir Fonksiyonel...	58

6.2. Kesikli Rassal Değişkenler için MaxMinxEnt Dağılımı.....	62
6.3. Sürekli Rassal Değişkenler için Kullback-Leibler Fonksiyonu ve Moment Fonksiyonları Yardımıyla Tanımlanmış Bir Fonksiyonel...	63
6.4. Sürekli Rassal Değişkenler için MaxMinxEnt Dağılımı.....	64
6.5. Rüzgâr Hızınının MaxMinxEnt dağılımı: Weibull dağılımına Alternatif.....	66
7. SONUÇLAR VE ÖNERİLER.....	75
KAYNAKLAR.....	75

ŞEKİLLER DİZİNİ

5.1.	Aralık ve Mayıs ayında ölçülen rüzgâr hızının MinMaxEnt dağılımları ve Weibull dağılımı.....	55
5.2.	(Celik 2004) alınan rüzgar hızı verisi için MinMaxEnt dağılımları ve Weibull dağılımı.....	57
6.1.	Aralık ayında ölçülen rüzgâr hızının MaxMinxEnt dağılımları ve Weibull dağılımı.....	72

ÇİZELGELER DİZİNİ

5.1.	Aralık ayında ölçülen rüzgâr hızı için iki bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımlarının entropileri.....	48
5.2.	Aralık ayında ölçülen rüzgâr hızı için üç bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımlarının entropileri.....	48
5.3.	Aralık ayında ölçülen rüzgâr hızı için MinMaxEnt dağılımları ve Weibull dağılımından elde edilen teorik frekanslar.....	49
5.4.	Aralık ayında ölçülen rüzgâr hızı için MinMaxEnt dağılımları ve Weibull dağılımından elde edilen teorik frekanslar.....	51
5.5.	2005 yılı tüm aylarında ölçülen rüzgâr hızı için MinMaxEnt dağılımlarının ve Weibull dağılımının karşılaştırılması.....	55
5.6.	(Celik 2004) alınan rüzgar hızı verisi için MinMaxEnt dağılımlarının Weibull dağılımları ile karşılaştırılması.....	56
6.1	Aralık ayında ölçülen rüzgâr hızı için iki bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımlarından hesaplanan K-L ölçümü.....	67
6.2.	Aralık ayında ölçülen rüzgâr hızı için üç bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımlarından hesaplanan K-L ölçümü.....	67
6.3.	Aralık ayında ölçülen rüzgâr hızı için MaxMinxEnt dağılımları ve Weibull dağılımından elde edilen teorik frekanslar.....	69
6.4.	2005 yılı tüm aylarında ölçülen rüzgâr hızı için MaxMinxEnt dağılımlarının ve Weibull dağılımının karşılaştırılması.....	69

1. GİRİŞ

Temelde entropi termodinamiğin bir kavramı olmakla beraber istatistiksel mekanikte entropinin yeri ve metodlarının geliřtirmesinde rolü büyüktür. Shannon (1948) “Mathematical theory of communication” adlı çalıřmasında, entropi kavramını informasyonla iliřkilendirmiş ve ardından Jaynes (1957), ilgili sistemin moment kısıtlarına göre Shannon’un entropi ölçümünü maksimize ederek olasılık dağılımlarının elde edilebildiđi MaxEnt (Maksimum entropi) olarak adlandırdıđı bir metot geliřtirmiřtir. Daha sonraki yıllarda, Kullback, Leibler, Rengi rassallıđın söz konusu olduđu ortamlarda, ilgilenilen olasılıkların hesaplanmasında entropi ve MaxEnt metodunun başarıyla uygulanabileceđini göstermişlerdir (Kullback ve Leibler 1951; Kullback 1959; Rengi 1961).

Hem istatistik hem de MaxEnt kısmi veya tamamlanmamış informasyondan yola çıkarak sonuç çıkarsama ile ilgilenir. MaxEnt’de mümkün olan sonsuz momentler içerisinden sadece birkaç momentin bilinmesi ile olasılık yoğunluk fonksiyonunu tahmin etmeye çalıřılırken, istatistikte ise sonsuz sayıda mümkün rassal örneklem içerisinden sadece bir rassal örneklem mümkün olduđunda yoğunluk fonksiyonunu tahmin etmeye çalıřılır. Böylece bu iki disiplinin birbiriyle içten iliřkili oldukları ve birbiri üzerinde etkileri olduđu şaşırtıcı deđildir.

İstatistik biliminde, olasılık dağılımları temel yapı taşlarından birisidir. (Shamilov ve ark. 2005), (Shamilov ve Mert 2005). Yüzyıllar boyunca, çok sayıda olasılık dağılımları; kesikli veya sürekli, tek deđişkenli ya da çok deđişkenli, deđişik düşüncelerden özdeřleştirilmiştir. Burada, MaxEnt güçlü bir birleřtirici etki sağlar çünkü pek çok olasılık dağılımı bazı basit momentlerin bilgisi altında, entropi maksimize edilerek elde edilebilir. Bazı dağılımların sadece birkaç momentiyile karakterize edildiđini bulmak memnun edici bir kolaylıktır. MaxEnt metodu, çeřitli dağılımların karakterize edici momentlerini bulmak için bize bir araç sağlamaktadır.

İstatistiksel çıkarsamanın temel hedeflerinden biri, sadece bir rassal örneklem verildiđinde yoğunluk fonksiyonunu tahmin etmektir. Yoğunluk fonksiyonunun formu bilinir ve yoğunluk fonksiyonundaki parametreler rassal örneklem deđerlerini kullanarak tahmin edilir. Benzer yolla, olasılık

yoğunluğunun karakterize edici momentleri görüntü ile verilir ve bu momentlerin değerleri rassal örneklemede ilgili momentlerin değerlerinde aritmetik ortalaması gibi elde edilir. Bu tahminler Fisher'in maksimum benzerlik metodu (Maximum Likelihood) kullanıldığında elde ettiği tahminleri ile aynıdır.

Hem MaxEnt hem İstatistik kendilerini olasılık yoğunluk fonksiyonu hakkında bilgi elde etmek için yönlendirmiştir. İstatistikte bunu rassal örneklemeden elde ederken, MaxEnt'de informasyon momentler hakkındaki bilgiden elde edilmektedir. Fisher, parametrelerde bulunan informasyonu ele alırken, Shannon ve Jaynes momentlerde bulunan informasyonu ele almıştır.

Günümüzde ise entropi optimizasyon metodları, istatistik, matematik, coğrafya, uzay bilimleri, ekonomi, finans, pazarlama, sistem analizi, görüntü işleme, model belirleme gibi alanlarda önemli uygulamalara sahiptir.

Özellikle istatistikte pek çok konu maksimum entropi (MaxEnt) ve minimum çapraz entropi (MinxEnt) metodları uygulanarak incelenmiştir. (Kullback 1951, 1959; Kapur, Kesevan 1992; Shamilov 2004; Shamilov ve Mert Kantar 2005; Shamilov ve Yolacan 2005). En önemli konulardan bir tanesi MaxEnt metoduyla non-parametrik bir yoğunluk fonksiyonunun bulunabilmesidir. Bu yoğunluk fonksiyonuna MaxEnt yoğunluk fonksiyonu denilmektedir. Fonksiyon oldukça esnek güçlü olmasının yanı sıra MaxEnt metodunda eğer moment kısıtları özel bir dağılımın karakterize eden momentleri ise bu durumda metodun bulduğu sonuç dağılımın kendisidir. Örneğin moment kısıtı x ve x^2 alındığında metod MaxEnt dağılımı olarak normal dağılımı, x ve $\ln x$ moment kısıtı alındığında MaxEnt dağılımı olarak Gama dağılımını bulmaktadır. Literatürde, MaxEnt dağılımlarının uygulandığı pek çok çalışma vardır. Özellikle Ebrahimi (1996, 1997, 2000) yaşam dağılımları için MaxEnt metodunu farklı şekillerde uygulamıştır. Zellner (1997) ve Zellner ve Tobias (2001) de Bayesian metottaki önsel dağılımın tahmininde MaxEnt metodları uygulamışlardır. Ximing (2003)'de MaxEnt metoduyla Amerika'nın gelir dağılımını elde etmiştir ve klasik istatistiksel modeller olan Log-normal, gamma dağılımlarıyla karşılaştırmıştır. Ayrıca yine aynı yazar makalesinde, çeşitli kriterlere göre, elde edilen MaxEnt dağılımları ile Log-normal, gamma dağılımlarını karşılaştırarak MaxEnt dağılımlarının hatasının en küçük olduğu bu dağılımın oldukça esnek olduğunu

göstermiştir. (Buchen ve Kelly 1996; Cozzolino ve Zahner1973) de MaxEnt'i finans alanında uygulamıştır. Enerji alanında bu metodun ilk kullanımı Li (2005)'nin iki farklı makalesi, Ramirez Carta (2005) çalışmalarıyla gerçekleşmiştir. Bu üç çalışmada rüzgâr hızının dağılımının modellenmesinde MaxEnt metodu kullanılarak MaxEnt dağılımları elde edilmiş ve klasik istatistiksel dağılım modelleriyle karşılaştırılmıştır. Ayrıca bu makalelerde rüzgâr hızını modellemede MaxEnt dağılımlarının kullanılması önerilmiş ve rüzgâr gücü tahminlerinin belirlenen MaxEnt dağılımları kullanılarak, daha az hatayla tahmin edilebileceği önerilmiştir.

Son yıllarada ise özellikle Pandey (2000, 2001) MinxEnt metodunu kullanarak quantile fonksiyonunun doğrudan tahminini yapmıştır.

Ayrıca istatistikte önemli yeri olan parametre tahmini problemine de entropi optimizasyon metodlarıyla alternatif çözümler üretilebilmektedir (Kapur, Kesevan 1992 ve Lind 1997, Cheng 1984, Ranneby 1984). Özellikle Lind ve ark. (1989) MinxEnt metodunu parametre tahminleme metodu olarak geliştirmiştir. Ayrıca iki dağılım arasındaki bilgi teorisi dayalı ölçülerden olan Kullback-Leibler (Kullback ve ark 1959), Rengi (Rengi, 1961), Havrda-Rengi, Havrda-Chavrat ölçüleri (Havrda ve ark 1967) birçok çalışmada tanıtılmış ve uygulanmıştır. Bu ölçüler için mükemmel bir kaynak olarak (Kapur 1967; Kapur 1996; Cover 1991; Nailong 1977) verilebilir.

Bununla birlikte, olasılık dağılımlarının karakterizasyonu, kontenjans tablolarında bilgi analizi, rassal değişkenler arasında bağımsızlık ölçümü, normallik testleri (Park ve Park 2001) gibi istatistiğin önemli konuları da entropi optimizasyon metodlarıyla değerlendirilebilmektedir. Ayrıca iki dağılım arasındaki yönlendirilmiş K-L mesafesi (ölçüsü) yaygın olarak farklı alanlarda kullanılmaktadır. Bu mesafenin tahmini olarak türetilen Akaike-İnformasyon kriteri (AIC) (Akaike 1973, 1978), AIC'nin küçük örneklem (AIC_c, QAIC_c) ve Takeuchi bilgi kriteri (TIC) modelin belirlenmesinde kullanılan bilgi kriterleri kullanımlarından biridir. Bu kriterler model belirleme kriteri olarak kabul görmüş ve birçok alana uygulanmıştır.

Bu çalışmada, MaxEnt metodu yardımıyla moment fonksiyonlarına bağlı bir fonksiyonel tanımlanarak, bu fonksiyonele minimum değer veren ve

informasyonu maksimum yapan bir dağılım belirleme süreci geliştirilmiştir. Geliştirilen bu süreç MaxEnt metodunun bir genelleşmesi olarak düşünülebilir. Söz konusu süreç kesikli ve sürekli rassal değişkenler için geçerlidir. Bu sürecin MaxEnt metodundan farkı, MaxEnt ölçümü yoğunluk fonksiyonlarına bağlı iken, tanımlanan fonksiyonelin moment fonksiyonlarının bir fonksiyoneli olmasıdır. Bu süreçten elde edilen dağılıma MinMaxEnt dağılımı denilmiştir.

Uygulama olarak, geliştirilen süreç yardımıyla Eskişehir bölgesinde ölçülen rüzgâr hızının MinMaxEnt dağılımı bulunmuş ve klasik istatistiğin bir modeli olan ve rüzgâr alanında uygulanabilirliği kabul edilen Weibull dağılımıyla karşılaştırılmıştır (Celik 2004; Akpınar ve Akpınar 2005; Ulgen ve Hepbasli 2002).

MinMaxEnt dağılımının belirlenmesi sürecine benzer şekilde, Kullback-Leibler (K-L) ölçüsü yardımıyla moment fonksiyonlarına bağlı yeni bir fonksiyonel tanımlanmış ve bu fonksiyonele maksimum değer veren ve Kullback-Leibler ölçüsüne göre ana kütleinin gerçek dağılımına en yakın dağılım bulma süreci geliştirilmiştir. Söz konusu süreç kesikli ve sürekli rassal değişkenler için geçerlidir. Bu süreçte, MinxEnt ölçüsü yoğunluk fonksiyonlarına bağlı iken burada tanımlanan fonksiyonel moment fonksiyonlarına bağlıdır. Bu sebepten dolayı tanımlanan fonksiyonele moment fonksiyonlarına bağlı K-L fonksiyoneli denilmiştir. Bu şekilde ifade edilmesinin sebebi sürekli K-L ölçüsünün de fonksiyonel olmasıdır ve bu iki fonksiyonelin karıştırılmamasıdır. Geliştiren bu süreç rüzgâr dağılımını belirlemede kullanılmıştır.

Bu tezde, bundan sonraki bölümler şu şekilde düzenlenmiştir. İkinci bölümde Maksimum Entropi (MaxEnt) Metodu kesikli ve sürekli rassal değişken için verilmiş, metodun bazı özellikleri yine aynı bölümde ele alınmıştır. Amaç: MaxEnt metodu yardımıyla tanımlanacak olan fonksiyonelin var olması ve gereken özelliklere sahip olmasının kanıtı için temel oluşturmaktır.

Üçüncü bölümde Kullback-leibler ölçüsüne dayalı olarak geliştirilmiş olan minimum çapraz entropi (MinxEnt) metodu ele alınmıştır. Amaç: MinxEnt metodu yardımıyla tanımlanacak olan fonksiyonelin var olması ve gereken özelliklere sahip olmasının kanıtı için temel oluşturmaktır.

Dördüncü bölümde geliştirilen bu süreçlerden elde edilmiş dağılımları diğer dağılımları karşılaştırmak için literatürde bilinen bilgiye teorisyene dayalı model seçim kriterleri ve istatistiksel kriterler verilmiştir.

Beşinci bölümde MaxEnt metoduna bağı olarak yeni bir fonksiyonel tanımlanmış ve bu fonksiyonelin bazı özellikleri araştırılmıştır. Bu fonksiyonelin sahip olduğı özelliklerden yararlanarak bu fonksiyonele minimum değeri veren dağılımın maksimum bilgiyi içerdiiği gösterilmiştir. Ayrıca bu dağılıma özel olarak MinMaxEnt dağılımı denilmiştir. Yine aynı bölümde geliştirilen MinMaxEnt dağılımı kullanılarak Eskişehir bölgesinde ölçülen rüzgâr hızının dağılımı bulunmuş ve klasik istatistiğin bir modeli olan ve rüzgâr alanında uygulanabilirliğı kanıtlanmış Weibull dağılımıyla karşılaştırılmıştır.

Altıncı bölümde MinxEnt metodundan yararlanarak yeni bir fonksiyonel tanımlanmış ve bu fonksiyonele, moment fonksiyonlarına bağı Kullback-Leibler fonksiyoneli denilmiştir. Bu fonksiyonele maksimum değeri veren moment fonksiyonuna karşı gelen dağılımın Kullback-Leibler ölçüsüne göre ana kütlelin dağılımına (gerçek dağılıma) en yakın dağılım olduğı gösterilmiş, yeni bir dağılım belirleme süreci optimizasyon metodunun bir genelleşmesi gibi ifade edilmiştir. Belirlenen bu dağılıma MaxMinxEnt denilmiştir. Yine aynı bölümde geliştiren bu süreç rüzgâr dağılımını belirlemede kullanılmıştır.

Son bölümde ise tezde yapılanlar özetlenerek, gelecek araştırmalara ilişkin bazı düşünceler vurgulanmıştır.

Ayrıca bu tezde kullanılan rüzgâr hızı verisi Anadolu Meydan Meteoroloji Müdürlüğünden alınmıştır (Anadolu Üniversitesi, 2005).

2. MAKSİMUM ENTROPİ METODU

Rassal değişkenin moment değerleri verildiğinde, Shanon'un entropi ölçüsünü maksimize ederek sistemin dağılımını belirleme metoduna MaxEnt metodu, elde edilen dağılımda MaxEnt dağılımı denilmektedir. (Jaynes, 1948). Rassal örnekten elde edilen moment değerlerine sahip çok sayıda dağılım olmakla beraber entropiyi maksimum yapan dağılım bu dağılımlar arasından en yansız olanıdır (Jaynes, 1957; Kapur, 1992). MaxEnt metodunun istatistik disiplinlerine çok sayıda uygulamaları vardır. Daha önce belirtildiği gibi MaxEnt metodundan elde edilen en önemli sonuçlardan birisi, literatürde ele alınan teorik dağılımların bazı basit karakterize edici momentlerine göre maksimum entropi dağılımları gibi oluşturulabilmesidir.

Bu bölümde, MaxEnt metoduna ait bazı bilgiler verilmiştir. Amaç: MaxEnt metodu yardımıyla tanımlanacak olan fonksiyonelin var olması ve gereken özelliklere sahip olmasının kanıtı için temel oluşturmaktır.

2.1. Kesikli Rassal Değişkenler İçin MaxEnt Metodu

X rassal değişkeninin sonlu sayıda x_1, x_2, \dots, x_n durumlarında bulunma olasılıkları p_1, p_2, \dots, p_n olduğunda, sistemin entropisi

$$H(x) = -\sum_{i=1}^n p_i \ln p_i \quad (2.1)$$

şeklinde dir. Bu sistem hakkında aşağıda moment kısıtları şeklinde ifade edilen

$$\sum_{i=1}^n p_i g_r(x_i) = a_r, \quad r = 1, \dots, m \quad (2.2)$$

informasyonunun verildiğini varsayalım. Burada $g_r(x)$ 'e ($r = 1, \dots, m$) moment fonksiyonları, a_r 'ye ($r = 1, \dots, m$) moment değerleri ve (2.2) ye ise moment kısıtları denilmektedir.

Rassal değişkenin aldığı değerlerin olasılıkları toplamının da 1 olduğu, bilinen kısıt olduğuna göre (2.2)'ye ek olarak,

$$\sum_{i=1}^n p_i g_0(x_i) = 1, \quad g_0(x_1) = g_0(x_2) = \dots = g_0(x_n) = 1 \quad (2.3)$$

kısıtı her zaman kabul edilmelidir. Bu tezde moment fonksiyonlarının oluşturduğu

$$A = \begin{pmatrix} g_0(x_1) & g_0(x_2) & \cdots & g_0(x_m) & \cdots & g_0(x_n) \\ g_1(x_1) & g_1(x_2) & \cdots & g_1(x_m) & \cdots & g_1(x_n) \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ g_m(x_1) & g_m(x_2) & \cdots & g_m(x_m) & \cdots & g_m(x_n) \end{pmatrix} \quad (2.4)$$

matrisinin rankının $m+1$ olduğu kabul edilmiştir.

MaxEnt, (2.1) fonksiyonuna (2.2) ve (2.3) kısıtları altında maksimum değer veren p_1, p_2, \dots, p_n olasılıklarını belirleme metodudur. (2.1), (2.2), (2.3) problemi koşullu ekstremum problemidir ve bu problemin çözümlenmesinde Lagrange metodu uygulanmaktadır. Koşullu ekstremum probleminin çözüm metoduna göre verilmiş koşulları $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanları ile çarparak, yardımcı L fonksiyonu

$$L \equiv -\sum_{i=1}^n p_i \ln p_i - \sum_{r=1}^m \lambda_r \left(\sum_{i=1}^n g_r(x_i) p_i - a_r \right) - (\lambda_0 - 1) \left(\sum_{i=1}^n p_i - 1 \right) \quad (2.5)$$

şeklinde kurulur. Ekstremin varlık teoremine göre yardımcı (2.5) fonksiyonunun p_i lere göre kısmi türevlerinin sıfıra eşit olması gerekir. Böylelikle p_1, p_2, \dots, p_n olasılıkları $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanlarına bağlı olarak elde edilir. $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ çarpanları da (2.2), (2.3) eşitliği yardımıyla bulunur.

L yardımcı fonksiyonunun p_i lere göre kısmi türevleri;

$$\begin{aligned} \frac{\partial L}{\partial p_i} &= -\left(\ln p_i + \frac{1}{p_i} p_i \right) - \lambda_1 g_1(x_i) - \lambda_2 g_2(x_i) - \cdots - \lambda_m g_m(x_i) - (\lambda_0 - 1) \\ &= -(\ln p_i + 1) - \sum_{r=1}^m \lambda_r g_r(x_i) - (\lambda_0 - 1), \quad i = 1, 2, \dots, n \end{aligned} \quad (2.6)$$

elde edilir. (2.6) eşitliğinin sağ tarafı sıfıra eşitlenerek,

Buradan

$$p_i = \exp(-\lambda_0) \exp\left\{ -\sum_{r=1}^m \lambda_r g_r(x_i) \right\} \quad (2.7)$$

Bilinmeyen $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ değerleri ise (2.2) ve (2.3) eşitliklerinden bulunabilir. (2.7) eşitliği, (2.3)' de yerine yazıldığında

$$1 = \left(\exp\left\{-\lambda_0 - \sum_{r=1}^m \lambda_r g_r(x_1)\right\} + \dots + \exp\left\{-\lambda_0 - \sum_{r=1}^m \lambda_r g_r(x_n)\right\} \right)$$

ifadesi elde edilir. Bu ifade aşağıdaki gibi

$$\exp(-\lambda_0) \left(\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \right) = 1 \quad (2.8)$$

yazılır. (2.2) kısıtından

$$\sum_{i=1}^n g_r(x_i) \exp\left\{-\lambda_0 - \sum_{j=1}^m \lambda_j g_j(x_i)\right\} = a_r, \quad r = 1, \dots, m \quad (2.9)$$

şeklinde yazılır. (2.4) ve (Papuluos, 1991, sayfa:190).

(2.7)'den (2.10) ve

$$\exp(\lambda_0) = \left(\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \right) \quad (2.10)$$

ve (2.9)' dan

$$a_r \exp(\lambda_0) = \sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}, \quad r = 1, \dots, m \quad (2.11)$$

(2.11) olur. (2.10) ve (2.11) den

$$a_r = \frac{\sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}, \quad r = 1, 2, \dots, m \quad (2.12)$$

(2.12) olduğu elde edilir. (2.10) dan λ_0 çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ çarpanlarının bir fonksiyonu olduğu, (2.12) eşitliğinden a_1, a_2, \dots, a_m 'in $\lambda_1, \lambda_2, \dots, \lambda_m$ 'ların bir fonksiyonu olduğu görülür. (Kapur 1992)

Kesikli sistem MaxEnt metodu ile incelendiğinde dikkat edilmelidir ki burada $m+1$ sayıda kısıt, ancak n sayıda bilinmeyen olasılık vardır. Metodun uygulanabilmesi için gereken koşullardan birisi $m+1 < n$ olmasıdır. Bu koşul altında n sayıda p_1, \dots, p_n bilinmeyenlerin $n-m-1$ sayıda olanı serbest değişken olarak kalır ve bu ise bilinmeyenler için sonsuz sayıda değerlerin

bulunmasını gerektirir. MaxEnt metodu bu sonsuz sayıda değerler arasından entropi fonksiyonuna maksimum değeri veren p_1, \dots, p_n leri seçer.

Kesikli sistemler için MaxEnt dağılımı bulunması problemi çok değişkenli fonksiyonlar için koşullu bir ekstremum problemidir.

2.1.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı

$\lambda_1, \lambda_2, \dots, \lambda_m$ ' ların varlığını göstermek için (2.12) eşitliği

$$\begin{aligned} f_1(\lambda_1, \dots, \lambda_m) &= a_1 \\ &\vdots \\ f_m(\lambda_1, \dots, \lambda_m) &= a_m \end{aligned} \quad (2.13)$$

burada $f_r(\lambda_1, \dots, \lambda_m) = \frac{\sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}$, $f = (f_1, \dots, f_m)$ '

$\lambda = (\lambda_1, \dots, \lambda_m)$ ve (a_1, \dots, a_m) dir.

$f(\lambda) = a$ kapalı fonksiyon şeklinde yazılıp, $\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)}$ determinantının

sıfırdan farkı olduğunu göstermek gerekir.

Buna göre (2.13)'te λ_r 'ye göre türev alınırsa

$$\begin{aligned} \frac{\partial f_r}{\partial \lambda_r} &= \frac{-\sum_{i=1}^n g_r(x_i) g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2} \\ &\quad - \frac{-\sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2}, \end{aligned}$$

veya

$$\frac{\partial f_r}{\partial \lambda_r} = -E(g_r^2(X)) + E(g_r(X))E(g_r(X)) = -Var(g_r(X)); r = 1, \dots, m \quad (2.14)$$

$$\frac{\partial f_r}{\partial \lambda_s} = \frac{-\sum_{i=1}^n g_r(x_i)g_s(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2} - \frac{-\sum_{i=1}^n g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n g_s(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2},$$

veya

$$\frac{\partial f_r}{\partial \lambda_s} = -E(g_r(X)g_s(X)) + E(g_r(X))E(g_s(X)) = -Cov(g_r(X)g_s(X)); \quad (2.15)$$

$r \neq s, r = 1, \dots, m.$

(2.14) ve (2.15) den

$$\begin{pmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \dots & \frac{\partial f_1}{\partial \lambda_m} \\ \frac{\partial f_2}{\partial \lambda_1} & \frac{\partial f_2}{\partial \lambda_2} & \dots & \frac{\partial f_2}{\partial \lambda_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \dots & \frac{\partial f_m}{\partial \lambda_m} \end{pmatrix} = \begin{pmatrix} Var[g_1(X)] & Cov[g_1(X)g_2(X)] & \dots & Cov[g_1(X)g_m(X)] \\ Cov[g_2(X)g_1(X)] & Var[g_2(X)] & \dots & Cov[g_2(X)g_m(X)] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[g_m(X)g_1(X)] & Cov[g_m(X)g_2(X)] & \dots & Var[g_m(X)] \end{pmatrix} \quad (2.16)$$

Bulunan matris varyans-covaryans matrisinin tersidir. Varyans- kovaryans matrisi pozitif tanımlı simetrik olduğu için bu matrisin determinanı 0' dan farklıdır. Çünkü her simetrik A matrisi D köşegen matris olmak üzere, $Q'AQ = D$ olacak biçimde ortogonal bir Q vardır. Burada D matrisinin köşegen elemanları A 'nın özdeğerleridir. A pozitif tanımlı olduğu için özdeğerleride pozitiftir. Dolayısıyla $|Q'AQ| = |Q||A||Q| = |A| = |D| = \lambda_1 \dots \lambda_m$, burada $\lambda_1, \dots, \lambda_m$ lar A 'nın özdeğerleridir.

$$\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)} \neq 0. \quad (2.17)$$

Kapalı fonksiyonun varlık teoremine (Spivak, 1965) göre Jacobian sıfırdan farklı olduğu için $\lambda_1, \lambda_2, \dots, \lambda_m$ 'lar a_1, a_2, \dots, a_m ' lar yardımıyla ifade edilebilir. Böylece $f^{-1}(\mathbf{a}) = \boldsymbol{\lambda}$ vardır denilebilir.

2.1.2. Varyans-Covaryans matrisi

$X = (X_1, \dots, X_n)$ vektor rassal değişkeni, $m = (m_1, \dots, m_n)$ X rassal değişkeninin moment vektörü olsun. $Var(X_i) = E(X_i - m_i)^2$, $i, j = 1, \dots, n$;

$Cov(X_i, X_j) = E[(X_i - m_i)(X_j - m_j)]$, $i \neq j$ şeklinde tanımlanır. X rassal değişkeninin varyans-covaryans matrisi

$$\begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_1, X_2) & Var(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var(X_n) \end{pmatrix} \quad (2.18)$$

şeklinde tanımlanır. (2.4) ve (Papulou, syf 190).

$$E \left[\sum_i^n t_i (x_i - m_i) \right]^2 > 0 \quad (2.19)$$

olduğu açıktır. Bu beklenen değer aşağıdaki formda yazılabilir.

$$\begin{aligned} E \left[\sum_i^n t_i (x_i - m_i) \right]^2 &= \sum_{i=1}^n t_i^2 E(x_i - m_i)^2 + \sum_{i \neq j}^n t_i t_j E(x_i - m_i) E(x_j - m_j) \\ &= t' A t, \end{aligned} \quad (2.20)$$

burada $t = (t_1, \dots, t_n)$, A varyans-covaryans matrisidir. (2.20) ve (2.21) den A matrisinin pozitif tanımlı olduğu görülmektedir.

2.1.3. Koşullu Ekstremumun Varlığı

Koşullu ekstremumunun varlığını göstermek için Taylor açılımının 2. teriminin negatif olduğunu göstermek yeterlidir. Bunun için (2.6)'in p_i ye ve p_j ye göre türevleri hesaplanır:

$$\frac{\partial^2 L}{\partial p_i^2} = -\frac{1}{p_i}; \quad \frac{\partial^2 L}{\partial p_i \partial p_j} = 0, \quad i \neq j. \quad (2.21)$$

Buradan,

$$p_i = \exp(-\lambda_0) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}, i = 1, \dots, n \text{ yi } p^0 = (p_1^0, \dots, p_n^0);$$

(2.5) deki L fonksiyonu $f(p)$ olarak göstererek ve p^0 noktasında Taylor serisine açılırsa

$$\begin{aligned} f(p) &= f(p^0) + df(p^0) + \frac{d^2 f(p^0)}{2} + o(\rho^2), \\ \Delta f &= f(p) - f(p^0) = df(p^0) + \frac{d^2 f(p^0)}{2} + o(\rho^2) \\ &= \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_n} dp_n\right) f(p^0) + \frac{1}{2} \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_n} dp_n\right)^2 f(p^0) + o(\rho^2) \\ &= 0 + \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_1} dp_n\right)^2 f(p^0) \\ &= \frac{1}{2} \left(\frac{\partial^2 f}{\partial p_1^2} dp_1 + \dots + \frac{\partial^2 f}{\partial p_n^2} dp_n\right) + o(\rho^2) \\ &+ \left(\frac{\partial^2 f}{\partial p_1 \partial p_1} dp_1 dp_2 + \dots + \frac{\partial^2 f}{\partial p_{n-1} \partial p_n} dp_{n-1} dp_n\right) \end{aligned} \quad (2.22)$$

$$\Delta f = -\frac{1}{2} \left(\frac{1}{p_1} dp_1^2 + \dots + \frac{1}{p_n} dp_n^2\right) + o(\rho^2),$$

$$\lim_{\rho \rightarrow \infty} \frac{o(\rho^2)}{\rho^2} = 0; \quad \rho = \sqrt{(dp_1)^2 + \dots + (dp_n)^2}$$

(2.2) ve (2.3) moment koşullarından

$$dp_1 + \dots + dp_n = 0; \quad \sum_{i=1}^n g_r(x_i) dp_i = 0, \quad r = 1, \dots, m \text{ elde edilir. Bu sonuçtan}$$

dp_1, \dots, dp_{m+1} 'ler dp_{m+2}, \dots, dp_n 'ler yardımıyla ifade edilebilir

$dp_i = \alpha_{m+2}^i dp_{m+2} + \dots + \alpha_n^i dp_n, i = 1, \dots, m+1$ olduğu dikkate alındığında Δf her zaman negatiftir. Böylece (2.7) değeri, (2.1) fonksiyonunun koşullu maksimumudur.

2.1.4. λ_0 Lagrange Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği

λ_0 Lagrange çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ çarpanlarına göre Konveksliği (2.10) eşitliğinden türev alınarak gösterilebilir. (2.10) eşitliğinden

$$\exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} = - \sum_{i=1}^n g_r(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\}, \quad (2.23)$$

(2.23)'den

$$\frac{\partial \lambda_0}{\partial \lambda_r} = - \frac{\sum_{i=1}^n g_r(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\}}{\exp(\lambda_0)} = -E[g_r(X)] = -a_r, \quad (2.24)$$

(2.24)'den

$$\frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_r} + \frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = \frac{\sum_{i=1}^n g_r^2(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\}}{\exp(\lambda_0)} = E[g_r^2(X)],$$

ve

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = E[g_r^2(X)] - E[g_r(X)]^2 = Var[g_r(x)] \quad (2.25)$$

bulunur. (2.23)'den λ_s e göre tekrar kısmi türev alınırsa

$$\exp(\lambda_0) \frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = \sum_{i=1}^n g_r(x_i) g_s(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\},$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = E[g_r(X) g_s(X)],$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} = E[g_r(X) g_s(X)] - E[g_r(X)] E[g_s(X)] = Cov[g_r(X) g_s(X)]. \quad (2.26)$$

λ_0 Lagrange çarpanı diğer $\lambda_1, \lambda_2, \dots, \lambda_m$ çarpanlarının fonksiyonudur. Bu nedenle Hessian matrisi (2.26) ve (2.25) den aşağıdaki şekilde ifade edilebilir.

$$\begin{pmatrix} \frac{\partial^2 \lambda_0}{\partial \lambda_1^2} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_m} \\ \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_m^2} \end{pmatrix} = \begin{pmatrix} \text{Var}[g_1(X)] & \text{Cov}[g_1(X)g_2(X)] & \dots & \text{Cov}[g_1(X)g_m(X)] \\ \text{Cov}[g_2(X)g_1(X)] & \text{Var}[g_2(X)] & \dots & \text{Cov}[g_2(X)g_m(X)] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[g_m(X)g_1(X)] & \text{Cov}[g_m(X)g_2(X)] & \dots & \text{Var}[g_m(X)] \end{pmatrix}. \quad (2.27)$$

(2.27) matrisi varyans-covaryans matrisidir. Bu matrisin pozitif tanımlı olduğu için λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'in konveks bir fonksiyonu olduğu gösterilmiş olur.

2.1.5. Entropi Fonksiyonunu Maksimum Değeri

(2.1) entropi fonksiyonunu maksimum yapan değer (2.7) eşitliği ile ifade edildiğine göre, entropinin maksimum değeri H_{\max} :

$$H_{\max} = -\sum_{i=1}^n p_i \ln p_i = -\sum_{i=1}^n p_i (-\lambda_0 - \lambda_1 g_1(x_i) - \dots - \lambda_m g_m(x_i)),$$

veya

$$H_{\max} = \lambda_0 + \lambda_1 a_1 + \dots + \lambda_m a_m \quad (2.28)$$

olur.

2.1.6. H_{\max} 'ın a_1, a_2, \dots, a_m 'e göre Konkavlığı

(2.28)' den H_{\max} 'ın $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ ve a_1, a_2, \dots, a_m 'in bir fonksiyonu olduğu görülmektedir. λ_0 parametresi $\lambda_1, \lambda_2, \dots, \lambda_m$ 'nin fonksiyonu olduğu için (2.12) den a_1, a_2, \dots, a_m 'lerinde fonksiyonudur. H_{\max} 'ın a_1, a_2, \dots, a_m lerin konkav bir fonksiyonu olduğunu göstermek için, H_{\max} 'ın a_1, a_2, \dots, a_m 'lere göre ikinci mertebeden kısmı türevlerinin oluşturduğu Hessian matrisi negatif tanımlı olduğu gösterilmelidir. Bu amaçla gereken türevler aşağıdaki şekilde bulunur.

$$\frac{\partial H_{\max}}{\partial a_r} = \sum_{s=1}^m \frac{\partial \lambda_0}{\partial \lambda_s} \frac{\partial \lambda_s}{\partial a_r} + \lambda_r + \sum_{s=1}^m a_s \frac{\partial \lambda_s}{\partial a_r}. \quad (2.29)$$

(2.24) den $\frac{\partial \lambda_0}{\partial \lambda_s} = -a_s$ olduğu bilindiğine göre

$$\frac{\partial H_{\max}}{\partial a_r} = \lambda_r \quad (2.30)$$

ve

$$\frac{\partial^2 H_{\max}}{\partial a_r^2} = \frac{\partial \lambda_r}{\partial a_r} \text{ ve } \frac{\partial^2 H_{\max}}{\partial a_r \partial a_s} = \frac{\partial \lambda_r}{\partial a_s}. \quad (2.31)$$

Sonuç olarak S_{\max} 'ın Hesian matrisi :

$$Z = \begin{pmatrix} \frac{\partial \lambda_1}{\partial a_1} & \frac{\partial \lambda_1}{\partial a_2} & \dots & \frac{\partial \lambda_1}{\partial a_m} \\ \frac{\partial \lambda_2}{\partial a_1} & \frac{\partial \lambda_2}{\partial a_2} & \dots & \frac{\partial \lambda_2}{\partial a_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \lambda_m}{\partial a_1} & \frac{\partial \lambda_m}{\partial a_2} & \dots & \frac{\partial \lambda_m}{\partial a_m} \end{pmatrix} \quad (2.32)$$

şekindedir. Z 'nin negatif tanımlı olduğunu göstermek için aşağıda tanımlanmış Q matrisini ile Z arasında $ZQ = I$ bağlantısı olduğunu göstermek gerekir.

Burada I birim matristir.

$$Q = \begin{pmatrix} \frac{\partial a_1}{\partial \lambda_1} & \frac{\partial a_1}{\partial \lambda_2} & \dots & \frac{\partial a_1}{\partial \lambda_m} \\ \frac{\partial a_2}{\partial \lambda_1} & \frac{\partial a_2}{\partial \lambda_2} & \dots & \frac{\partial a_2}{\partial \lambda_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial a_m}{\partial \lambda_1} & \frac{\partial a_m}{\partial \lambda_2} & \dots & \frac{\partial a_m}{\partial \lambda_m} \end{pmatrix}. \quad (2.33)$$

(2.33)'de

$$b_{kl} = \frac{\partial \lambda_k}{\partial \lambda_l} = \frac{\partial \lambda_k}{\partial a_1} \frac{\partial a_1}{\partial \lambda_l} + \frac{\partial \lambda_k}{\partial a_2} \frac{\partial a_2}{\partial \lambda_l} + \dots + \frac{\partial \lambda_k}{\partial a_m} \frac{\partial a_m}{\partial \lambda_l} \quad (2.34)$$

olduğu için, $ZQ = (b_{kl}) = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases}$ dir. Böylece $Q = Z^{-1}$ dir.

(2.12)'den λ_s 'e göre türev alınırsa

$$\frac{\partial a_r}{\partial \lambda_s} = - \frac{\left\{ \left[\sum_{i=1}^n \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right] \sum_{i=1}^n g_r(x_i) g_s(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} - \sum_{i=1}^n g_r(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \sum_{i=1}^n g_s(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right\}}{\left[\sum_{i=1}^n \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right]^2}$$

$$= -(E[g_r(X)g_s(X)] - E[g_r(X)]E[g_s(X)])$$

veya

$$\frac{\partial a_r}{\partial \lambda_s} = -Cov[g_r(X)g_s(X)] \quad (2.35)$$

$$\frac{\partial a_r}{\partial \lambda_r} = -Var[g_r(X)]. \quad (2.36)$$

(2.36) dan Q matrisinin varyans-covaryans matrisinin ters işaretlisi olduğu görülmektedir. Bu durumda Q matrisi negatif tanımlı, dolayısıyla onun tersi olan Z matrisi de negatif tanımlıdır. Z matrisi H_{\max} ın a_1, a_2, \dots, a_m lere göre ikinci mertebeden kısmi türevlerinin oluşturduğu matris olduğuna göre, H_{\max} a_1, a_2, \dots, a_m lerin konkav bir fonksiyonudur.

2.2. Sürekli Rassal Değişkenler İçin Maksimum Entropi Metodu

X sürekli rassal değişkeninin olasılık yoğunluk fonksiyonu $f(x)$ olduğunda, bu sistemin entropisi

$$H(x) = - \int_a^b f(x) \ln f(x) dx \quad (2.37)$$

şeklinde tanımlanır. X sistemi hakkında aşağıdaki moment kısıtları şeklinde ifade edilen informasyonun verildiği varsayalım:

$$\int_a^b f(x) g_r(x) dx = a_r \quad r = 1, \dots, m. \quad (2.38)$$

Burada kesikli sistemde olduğu gibi $g_r(x)$ moment fonksiyonları ve a_r moment değerleridir. Sürekli sistemler için de

$$\int_a^b f(x)dx = 1 \quad (2.39)$$

kısıtı geçerlidir.

(2.37) entropi fonksiyonelinin (2.39) kısıtı altında maksimum değeri $H_{\max} = \ln(b - a)$ şeklinde bulunur.

Sürekli sistem için MaxEnt, (2.37) fonksiyoneline (2.38) ve (2.39) kısıtları altında maksimum değer veren $f(x)$ fonksiyonunun bulunması metodudur. $f(x)$ 'in bulunmasında Euler-Lagrange metodu kullanılabilir. Öncelikle Lagrange yardımcı fonksiyoneli oluşturulur.

$$L = -\int_a^b f(x) \ln f(x) dx - (\lambda_0 - 1) \left[\int_a^b f(x) dx - 1 \right] - \sum_{r=1}^m \lambda_r \left[\int_a^b f(x) g_r(x) dx - a_r \right],$$

$$L = \int_a^b \underbrace{\left(-f(x) \ln f(x) - (\lambda_0 - 1)f(x) - \sum_{r=1}^m \lambda_r f(x) g_r(x) \right)}_F dx + \sum_{r=1}^m \lambda_r a_r + (\lambda_0 - 1),$$

$$L = \int_a^b F(x) dx, \quad (2.40)$$

L fonksiyonelin ekstremum değere sahip olması için gerek koşul bu fonksiyonelin birinci varyasyonunun 0' a eşit olmasıdır. Dolayısıyla $f(x)$ fonksiyonu L fonksiyoneline minimum değer verirse L Euler-Lagrange denklemini sağlamalıdır.

Euler-Lagrange eşitliğine göre L yi minimum veya maksimum yapan $f(x)$, aşağıdaki eşitliğini sağlayan fonksiyondur.

$$\frac{\partial F}{\partial f(x)} - \frac{d}{dx} \left(\frac{\partial F}{\partial f'(x)} \right) = 0 \quad (2.41)$$

(2.41) denkleminde $f(x)$ aşağıdaki şekilde

$$f(x) = \exp(-\lambda_0 - \lambda_1 g_1(x) - \dots - \lambda_m g_m(x)) \quad (2.42)$$

bulunur. $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ ler (2.38) ve (2.39) kısıtlarından bulunur. $f(x)$ fonksiyonunun L fonksiyoneline maksimum değer verdiği 2.2.2 bölümünde gösterilecektir. $f(x)$ MaxEnt fonksiyonu, moment kısıtları (2.38) ve (2.39) da yerine yazılırsa

$$a_r \exp(\lambda_0) = \int_a^b g_r(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx, \quad (2.43)$$

ve (2.44) ve (2.45) yardımıyla

$$\exp(\lambda_0) = \int_a^b \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx, \quad (2.44)$$

$$a_r = \frac{\int_a^b g_r(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx}{\exp(\lambda_0)} = \frac{\int_a^b g_r(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx}{\int_a^b \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx}, \quad (2.45)$$

$$r = 1, \dots, m$$

Kesikli sistemde olduğu gibi λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'in konveks bir fonksiyonu olduğu ve H_{\max} 'ın a_1, a_2, \dots, a_m 'lerin konkav bir fonksiyonu olduğu kolayca gösterilebilir.

2.2.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı

$\lambda_1, \lambda_2, \dots, \lambda_m$ 'in varlığı kesikli sistemde olduğu gibi (2.45) eşitliği

$$\begin{aligned} f_1(\lambda_1, \dots, \lambda_m) &= a_1 \\ &\vdots \\ f_m(\lambda_1, \dots, \lambda_m) &= a_m \end{aligned} \quad (2.46)$$

$$\text{burada } f_r(\lambda_1, \dots, \lambda_m) = \frac{\int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}, \quad f = (f_1, \dots, f_m)',$$

$\lambda = (\lambda_1, \dots, \lambda_m)$ ve (a_1, \dots, a_m) dir.

Sistemin $\lambda_1, \dots, \lambda_m$ lere göre çözüme sahip olması gösterilmelidir

$\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)}$ determinantının sıfırdan farkı olduğunu gösterilirse, $\lambda_1, \dots, \lambda_m$ varlığı

kanıtlanmış olur.

(2.46)' dan

$$\frac{\partial f_r}{\partial \lambda_r} = \frac{-\int_a^b g_r(x) g_r(x) \exp(-\sum_{j=1}^m \lambda_j g_j(x)) dx \int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2} - \frac{\int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\} \int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2}, \quad (2.47)$$

buradan

$$\frac{\partial f_r}{\partial \lambda_r} = -(E(g_r^2(X)) - E(g_r(X))E(g_r(X))) = -Var(g_r(X)); r = 1, \dots, m \quad (2.48)$$

$$\frac{\partial f_r}{\partial \lambda_s} = \frac{-\int_a^b g_r(x) g_s(x) \exp(-\sum_{j=1}^m \lambda_j g_j(x)) dx \int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2} - \frac{\int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\} \int_a^b g_s(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2},$$

buradan

$$\frac{\partial f_r}{\partial \lambda_s} = -(E(g_r(X)g_s(X)) - E(g_r(X))E(g_s(X))) = -Cov(g_r(X)g_s(X)); \quad (2.49)$$

$r \neq s, r = 1, \dots, m$

Bulunan matriks

$$\begin{pmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \dots & \frac{\partial f_1}{\partial \lambda_m} \\ \frac{\partial f_2}{\partial \lambda_1} & \frac{\partial f_2}{\partial \lambda_2} & \dots & \frac{\partial f_2}{\partial \lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \dots & \frac{\partial f_m}{\partial \lambda_m} \end{pmatrix} = \begin{pmatrix} Var[g_1(X)] & Cov[g_1(X)g_2(X)] & \dots & Cov[g_1(X)g_m(X)] \\ Cov[g_2(X)g_1(X)] & Var[g_2(X)] & \dots & Cov[g_2(X)g_m(X)] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[g_m(X)g_1(X)] & Cov[g_m(X)g_2(X)] & \dots & Var[g_m(X)] \end{pmatrix}$$

varyans-covaryans matrisinin tersidir. Varyans-kovaryans matrisi pozitif tanımlı olduğu için

$$\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)} = \neq 0. \quad (2.50)$$

dır. Dolayısıyla $f^{-1}(\mathbf{a}) = \boldsymbol{\lambda}$ vardır.

2.2.2. Koşullu Ekstremumun Varlığı

MaxEnt fonksiyonunun entropi fonksiyoneline maksimum değer verdiği, bilinen bir eşitsizlikten yola çıkılarak yapılacaktır. $\ln(x) \leq x-1$, her $x > 0$ için geçerli

olduğu bilinmektedir. Burada $k(x)$ yoğunluk fonksiyonu ve $t(x) \geq 0$, $\int_a^b t(x)dx \leq 1$

koşullarını sağlayan fonksiyon olsun. x yerine, $\frac{t(x)}{k(x)}$ ifadesi eşitsizlikte yerine

yazılırsa

$$\ln \frac{t(x)}{k(x)} \leq \frac{t(x)}{k(x)} - 1 \quad (2.51)$$

$$\ln t(x) - \ln k(x) \leq \frac{t(x)}{k(x)} - 1 \quad (2.52)$$

$$k(x) \ln t(x) - k(x) \ln k(x) \leq t(x) - k(x)$$

olur.

(2.52) eşitsizliğin her iki tarafından integral alındığında

$$\int_a^b k(x) \ln t(x) dx - \int_a^b k(x) \ln k(x) dx \leq \int_a^b t(x) dx - \int_a^b k(x) dx \leq 0$$

$$-\int_a^b k(x) \ln k(x) dx \leq -\int_a^b k(x) \ln t(x) dx \quad (2.53)$$

(2.53) eşitsizliği elde edilir.

(2.53) eşitsizliğinde $k(x)$ moment koşullarını sağlayan keyfi dağılım fonksiyonu, $t(x)$ 'de (2.42) deki MaxEnt fonksiyonu olsun. Bu durumda (2.53) eşitsizliği

$$\begin{aligned}
-\int_a^b k(x) \ln k(x) dx &\leq -\int_a^b k(x) \ln t(x) dx \\
-\int_a^b k(x) \ln k(x) dx &\leq -\int_a^b k(x) (-\lambda_0 - \lambda_1 g_1 - \dots - \lambda_m g_m) dx = H_{\max}
\end{aligned} \tag{2.54}$$

olarak şekillenir. Böylece (2.42) MaxEnt fonksiyonunun (2.37) entropi fonksiyoneline en büyük değer verdiği gösterilmiş olur.

2.2.3. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği

λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ çarpanlarının konveks bir fonksiyonu olduğunu göstermek için (2.44) eşitliğinden türev alınırsa,

$$\exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} = -\int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}, \tag{2.55}$$

$$\frac{\partial \lambda_0}{\partial \lambda_r} = -\frac{\int_a^b g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\exp(\lambda_0)} = E[g_r(X)] = -a_r,$$

$$\frac{\partial \lambda_0}{\partial \lambda_r} = -a_r, \tag{2.56}$$

(2.55)'den tekrar türev alınırsa

$$\begin{aligned}
\exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_r} + \exp(\lambda_0) \frac{\partial^2 \lambda_0}{\partial \lambda_r^2} \\
= \int_a^b g_r^2(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\} = E[g_r^2(X)],
\end{aligned} \tag{2.57}$$

veya

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = E[g_r^2(X)] - E[g_r(X)]^2 = \text{Var}[g_r(x)]. \tag{2.58}$$

(2.55)'den λ_s e göre tekrar türev alınırsa

$$\begin{aligned}
\exp(\lambda_0) \frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} \\
= -\int_a^b g_r(x) g_s(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\},
\end{aligned} \tag{2.59}$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\lambda_0}{\partial \lambda_s} = E[g_r(X)g_s(X)]. \quad (2.60)$$

Buradan,

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} = E[g_r(X)g_s(X)] - E[g_r(X)]E[g_s(X)] = Cov[g_r(X)g_s(X)].$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} = Cov[g_r(X)g_s(X)] \quad (2.61)$$

(2.44)'e göre Hessian matrisi aşağıdaki gibi bulunur.

$$\begin{pmatrix} \frac{\partial^2 \lambda_0}{\partial \lambda_1^2} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_m} \\ \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_m^2} \end{pmatrix} = \begin{pmatrix} Var[g_1(X)] & Cov[g_1(X)g_2(X)] & \dots & Cov[g_1(X)g_m(X)] \\ Cov[g_2(X)g_1(X)] & Var[g_2(X)] & \dots & Cov[g_2(X)g_m(X)] \\ \vdots & \vdots & & \vdots \\ Cov[g_m(X)g_1(X)] & Cov[g_m(X)g_2(X)] & \dots & Var[g_m(X)] \end{pmatrix}. \quad (2.62)$$

(2.62) matrisi varyans-kovaryans matrisi olarak pozitif tanımlı matristir. Budan dolayı Hessian matrisinin determinanı pozitiftir. Bu sonuç λ_0 çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ çarpanlarının konveks fonksiyonu olduğunu gösterir.

3. MİNİMUM ÇAPRAZ ENTROPİ METODU

Bu bölümde kesikli ve sürekli sistemler için MinxEnt metodu incelenerek bu metodun belirlediği minimum çapraz entropi (MinxEnt) dağılımının bazı özellikleri araştırılmıştır. Burada amaç: Kullback-Leibler fonksiyoneli ve moment fonksiyonları yardımıyla kurulacak olan yeni fonksiyonel için temel oluşturmaktır.

3.1. Kesikli Rassal Değişkenler için MinxEnt Metodu

$$D(p : q) = \sum_{i=1}^n p_i \ln \frac{p_i}{q_i} \quad (3.1)$$

$$\sum_{i=1}^n p_i = 1 \quad (3.2)$$

$$\sum_{i=1}^n p_i g_r(x_i) = a_r, \quad (3.3)$$

MinxEnt metodu, (3.1) Kullback-Leibler ölçümününü, (3.2), (3.3) kısıtlarına bağlı olarak minimize eden dağılımı bulma metodudur. Burada q_i dağılımı önsel dağılımdır. Söz konusu dağılım Lagrange çarpanları metoduyla bulunabilir. Bu amaçla

$$L = \sum_{i=1}^n p_i \ln \frac{p_i}{q_i} + (\lambda_0 - 1) \left(\sum_{i=1}^n p_i - 1 \right) + \sum_{r=1}^m \lambda_r \left(\sum_{i=1}^n p_i g_r(x_i) - a_r \right) \quad (3.4)$$

yardımcı fonksiyon oluşturulur. Ekstremumun varlık teoremine göre yardımcı (3.4) fonksiyonunun p_i 'lere göre kısmi türevlerinin sıfıra eşit olması gerekir. Böylelikle p_1, p_2, \dots, p_n olasılıkları $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanlarına bağlı elde edilir. $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ çarpanları ise (3.2), (3.3) koşulları yardımıyla bulunur.

L yardımcı fonksiyonunun p_i 'lere göre kısmi türevleri;

$$\frac{\partial L}{\partial p_i} = \left(\ln \frac{p_i}{q_i} + \frac{1}{p_i} \right) + \lambda_1 g_1(x_i) + \lambda_2 g_2(x_i) + \dots + \lambda_m g_m(x_i) + (\lambda_0 - 1) \quad (3.5)$$

$$\frac{\partial L}{\partial p_i} = \ln \frac{p_i}{q_i} + \sum_{r=1}^m \lambda_r g_r(x_i) + \lambda_0, \quad (3.6)$$

elde edilir. Eşitlik (3.6)'nın sağ tarafı sıfıra eşitlenerek,

$$p_i = q_i \exp(-\lambda_0 - \lambda_1 g_1(x_i) - \dots - \lambda_m g_m(x_i)), i = 1, \dots, n \quad (3.7)$$

sonucu elde edilir. Belirlenen bu dağılıma MinxEnt dağılımı denilmektedir (Kapur ve Kesevan, 1992).

Burada, λ_0 çarpanı (3.2) koşulundan, a_1, a_2, \dots, a_m 'in yanı sıra q_1, q_2, \dots, q_n 'e de bağlı olacaktır.

$$\exp(\lambda_0) = \left(\sum_{i=1}^n q_i \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right) \quad (3.8)$$

ve (3.3) koşulundan

$$a_r = \frac{\sum_{i=1}^n q_i g_r(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\}}{\sum_{i=1}^n q_i \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\}}, \quad r = 1, 2, \dots, m \quad (3.9)$$

bulunur. (3.8) eşitliğinden λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'in bir fonksiyonu olduğu, eşitlik (3.9)'den a_1, a_2, \dots, a_m 'in $\lambda_1, \lambda_2, \dots, \lambda_m$ 'lerin bir fonksiyonu olduğu görülür.

MinxEnt'de verilen bir dağılımdan bir olasılıksal uzaklığı minimize ederken MaxEnt'de belirsizliği maksimize ederiz.

Bununla beraber, önsel olasılık dağılımı en belirsiz dağılım olarak gerçekleştiği özel durumda iki prensip arasında çok yakın bir ilişki vardır. Bazı kısıtlara göre en belirsiz dağılıma yakın olmak aynı kısıtlara göre maksimal olarak belirsiz olma anlamına gelir.

3.1.1. Koşullu Ekstremumun Varlığı

(3.5) den.

$$\frac{\partial^2 L}{\partial p_i^2} = \frac{1}{p_i}; \quad \frac{\partial^2 L}{\partial p_i \partial p_j} = 0, \quad i \neq j \quad (3.10)$$

elde edilir. (3.7) de ifade edilen MinxEnt dağılımı $p^0 = (p_1^0, \dots, p_n^0)$ olsun. (3.4)

ile ifade edilen fonksiyonun p^0 'ın küçük komşuluğunda Taylor açılımı yazılsın.

$$L(p) = L(p^0) + dL(p^0) + \frac{d^2L(p^0)}{2} + o(\rho^2),$$

burada,

$$\rho = \sqrt{dp_1^2 + \dots + dp_n^2} \text{ ve } \lim_{\rho \rightarrow 0} \frac{o(\rho^2)}{\rho^2} = 0.$$

$$\Delta L = L(p) - L(p^0) = dL(p^0) + \frac{d^2L(p^0)}{2} + o(\rho^2), \quad (3.11)$$

$$\begin{aligned} \Delta L &= \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_n} dp_n \right) L(p^0) + \frac{1}{2} \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_n} dp_n \right)^2 L(p^0) + o(\rho^2) \\ &= 0 + \left(\frac{\partial}{\partial p_1} dp_1 + \dots + \frac{\partial}{\partial p_n} dp_n \right)^2 L(p^0) = \frac{1}{2} \left(\frac{\partial^2 L}{\partial p_1^2} dp_1 + \dots + \frac{\partial^2 L}{\partial p_n^2} dp_n \right) \\ &+ \left(\frac{\partial^2 L}{\partial p_1 \partial p_1} dp_1 dp_1 + \dots + \frac{\partial^2 L}{\partial p_{n-1} \partial p_n} dp_{n-1} dp_n \right) + o(\rho^2) \end{aligned}$$

$$\Delta L = \frac{1}{2} \left(\frac{1}{p_1} dp_1^2 + \dots + \frac{1}{p_n} dp_n^2 \right) + o(\rho^2), \quad (3.12)$$

(3.2) ve (3.3) koşullarından

$$dp_1 + \dots + dp_n = 0; \sum_{i=1}^n q_i g_r(x_i) dp_i = 0, \quad r = 1, \dots, m \text{ elde edilir. Bu sonu} \ddot{c} \text{tan}$$

dp_1, \dots, dp_m ' ler dp_{m+1}, \dots, dp_n ' ler yardımıyla ifade edilebilir.

$dp_i = \alpha_{m+2}^i dp_{m+2} + \dots + \alpha_n^i dp_n, \quad i = 1, \dots, m+1$ dikkate alındığında ΔL her zaman pozitifdir. Böylece (3.7) değeri, (3.5) fonksiyonelinin koşullu minimumudur.

3.1.2. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı

Bu bölümde moment koşullarını sağlayacak olan $\lambda_1, \lambda_2, \dots, \lambda_m$ ların var olduğu kanıtlanmıştır. $\lambda_1, \lambda_2, \dots, \lambda_m$ ' ların varlığını göstermek için (3.9) eşitliği

$$\begin{aligned} f_1(\lambda_1, \dots, \lambda_m) &= a_1 \\ &\vdots \\ f_m(\lambda_1, \dots, \lambda_m) &= a_m \end{aligned} \quad (3.13)$$

$$\text{burada } f_r(\lambda_1, \dots, \lambda_m) = \frac{\sum_{i=1}^n q(x_i) g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}, \quad f = (f_1, \dots, f_m)'$$

$\lambda = (\lambda_1, \dots, \lambda_m)$ ve (a_1, \dots, a_m) dir.

$f(\lambda) = a$ kapalı fonksiyon şeklinde yazılıp, $\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)}$ determinanın

sıfırdan farkı olduğunu göstermek gerekir.

$$\begin{aligned} \frac{\partial f_r}{\partial \lambda_r} &= \frac{-\sum_{i=1}^n g_r(x_i) g_r(x_i) q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2} \\ &\quad - \frac{\sum_{i=1}^n q_i g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n q_i g_r(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2}. \end{aligned} \quad (3.14)$$

Varyans-Kovaryans tanımlarından yararlanarak

$$\frac{\partial f_r}{\partial \lambda_r} = -\left(E(g_r^2(X)) - E(g_r(X))E(g_r(X))\right) = -\text{Var}(g_r(X)); \quad r = 1, \dots, m \quad (3.15)$$

şeklinde ifade edilebilir.

$$\begin{aligned} \frac{\partial f_r}{\partial \lambda_s} &= \frac{-\sum_{i=1}^n g_r(x_i) g_s(x_i) q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2} \\ &\quad - \frac{\sum_{i=1}^n g_r(x_i) q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\} \sum_{i=1}^n q_i g_s(x_i) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}}{\left\{\sum_{i=1}^n q_i \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x_i)\right\}\right\}^2}. \end{aligned} \quad (3.16)$$

Varyans-Kovaryans kavramları uygulanarak (3.16) (Cramer, 1966)

$$\begin{aligned} \frac{\partial f_r}{\partial \lambda_s} &= -\left(E(g_r(X) g_s(X)) - E(g_r(X))E(g_s(X))\right) \\ &= -\text{Cov}(g_r(X) g_s(X)); \quad r \neq s, \quad r = 1, \dots, m \end{aligned} \quad (3.17)$$

şeklinde ifade edilir.

$$\begin{pmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \dots & \frac{\partial f_1}{\partial \lambda_m} \\ \frac{\partial f_2}{\partial \lambda_1} & \frac{\partial f_2}{\partial \lambda_2} & \dots & \frac{\partial f_2}{\partial \lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \dots & \frac{\partial f_m}{\partial \lambda_m} \end{pmatrix}$$

matrisi varyans-covaryans matrisinin ters işaretlisidir. Varyans- kovaryans matrisi pozitif tanımlı olduğu için bu matrisin determinanı sıfırdan farklıdır.

3.1.3. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği

(3.8)'den diferansiyel almakla;

$$\exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} = - \left(\sum_{i=1}^n q_i g_r(x) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right), \quad (3.18)$$

elde edilir.

Buradan

$$\frac{\partial \lambda_0}{\partial \lambda_r} = - \frac{\left(\sum_{i=1}^n q_i g_r(x_i) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x_i) \right\} \right)}{\exp(\lambda_0)}$$

veya

$$\frac{\partial \lambda_0}{\partial \lambda_r} = -E[g_r(X)] = -a_r, \quad r = 1, 2, \dots, m \quad (3.19)$$

sonucuna varılır.

(3.18) eşitliğiden tekrar λ_r 'ye göre türev alınırsa,

$$\exp(\lambda_0) \frac{\partial^2 \lambda_0}{\partial \lambda_r^2} + \exp(\lambda_0) \left(\frac{\partial \lambda_0}{\partial \lambda_r} \right)^2 = \sum_{i=1}^n g_r^2(x_i) q_i \exp \left(- \sum_{j=1}^m \lambda_j g_j(x_i) \right), \quad (3.20)$$

veya

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} + \left(\frac{\partial \lambda_0}{\partial \lambda_r} \right)^2 = \sum_{i=1}^n p_i g_r^2(x_i) = \sum_{i=1}^n p_i g_r^2(x_i) = E[g_r^2(X)], \quad (3.21)$$

elde edilir. (3.21) de (3.19) dikkate alındığında

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = E[g_r^2(X)] - E[g_r(X)]^2 \quad (3.22)$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = Var[g_r(X)] \quad (3.23)$$

eşitliğine varılır. Yine (3.18)'de λ_s 'ye göre türev alındığında,

$$\exp(\lambda_0) \frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = \sum_{i=1}^n g_r(x_i) g_s(x_i) q_i \exp\left(-\sum_{j=1}^m \lambda_j g_{ji}\right), \quad (3.24)$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = \sum_{i=1}^n p_i g_r(x_i) g_s(x_i) = \sum_{i=1}^n p_i g_r(x_i) g_s(x_i) \quad (3.25)$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = E[g_r(X) g_s(X)] \quad (3.26)$$

olarak bulunur ve (3.19), (3.26) da yerine yazıldığında

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} = E[g_r(x) g_s(x)] - E[g_r(x)] E[g_s(x)] = Cov(g_r(x), g_s(x)) \quad (3.27)$$

eşitliği elde edilir.

Böylece özetle (3.23) ve (3.26) eşitliklerinden λ_0 parametresinin $\lambda_1, \lambda_2, \dots, \lambda_m$ lere bağlı bir fonksiyon olduğu görülmektedir. λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ lere göre Hessian matrisi aşağıdaki gibidir.

$$\begin{pmatrix} \frac{\partial^2 \lambda_0}{\partial \lambda_1^2} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_m} \\ \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_2^2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_m^2} \end{pmatrix} \quad (3.28)$$

gibi olur. Ve (3.28)'in yardımıyla da tekrar aşağıdaki gibi yazılabilir.

$$\begin{pmatrix} Var[g_1(X)] & Cov[g_1(X), g_2(X)] & \dots & Cov[g_1(X), g_m(X)] \\ Cov[g_2(X), g_1(X)] & Var[g_2(X)] & \dots & Cov[g_2(X), g_m(X)] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[g_m(X), g_1(X)] & Cov[g_m(X), g_2(X)] & \dots & Var[g_m(X)] \end{pmatrix} \quad (3.29)$$

Varyans-kovaryans matrisi daima pozitif tanımlı olduğundan, aynı şekilde λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'ye bağlı ikinci derece kısmi türevlerinin Hessian matrisi de pozitif

tanımlı olacaktır. Böylece λ_0 'ın, $\lambda_1, \lambda_2, \dots, \lambda_m$ 'nin konveks bir fonksiyonu olduğu kanıtlanmış olur.

3.1.4. D_{\min} 'in a_1, \dots, a_m 'e göre Konveksliği

$$D(p : q) = \sum_{i=1}^n p_i \ln \frac{p_i}{q_i} \quad (3.30)$$

Kullback-Leibler ölçümünde (3.7) dağılımı dikkate alınırsa,

$$D_{\min} = \sum_{i=1}^n p_i (-\lambda_0 - \lambda_1 g_1(x_i) - \lambda_2 g_2(x_i) - \dots - \lambda_m g_m(x_i)) \quad (3.31)$$

elde edilir, burada da (3.3) kısıtı göz önüne alınırsa,

$$D_{\min} = -\lambda_0 - \lambda_1 a_1 - \lambda_2 a_2 - \dots - \lambda_m a_m \quad (3.32)$$

eşitliğine ulaşılır.

Bu eşitlikte a_r 'ye göre türev alındığında,

$$\frac{\partial D_{\min}}{\partial a_r} = -\sum_{s=1}^m \frac{\partial \lambda_0}{\partial \lambda_s} \frac{\partial \lambda_s}{\partial a_r} - \lambda_r - \sum_{s=1}^m a_s \frac{\partial \lambda_s}{\partial a_r} = -\lambda_r \quad (3.33)$$

eşitliği bulunur ve bu eşitlik üzerinden de a_r 'ye ve a_s 'ye göre türevler ayrı ayrı alındığında

$$\frac{\partial^2 D_{\min}}{\partial a_r^2} = -\frac{\partial \lambda_r}{\partial a_r} \quad \text{ve} \quad \frac{\partial^2 D_{\min}}{\partial a_r \partial a_s} = -\frac{\partial \lambda_r}{\partial a_s} \quad (3.34)$$

elde edilir. Böylece a_1, a_2, \dots, a_m 'ye bağlı olarak, D_{\min} 'in a_1, a_2, \dots, a_m göre

Hessian matrisi

$$H = \begin{pmatrix} -\partial \lambda_1 / \partial a_1 & -\partial \lambda_1 / \partial a_2 & \dots & -\partial \lambda_1 / \partial a_m \\ -\partial \lambda_2 / \partial a_1 & -\partial \lambda_2 / \partial a_2 & \dots & -\partial \lambda_2 / \partial a_m \\ \vdots & \vdots & \ddots & \vdots \\ -\partial \lambda_m / \partial a_1 & -\partial \lambda_m / \partial a_2 & \dots & -\partial \lambda_m / \partial a_m \end{pmatrix} \quad (3.35)$$

şeklindedir.

D_{\min} 'in a_1, \dots, a_m 'lere göre Hessian matrisinin pozitif tanımlı olduğunu gösterebilmek için

$$H^{-1} = \begin{pmatrix} -\partial a_1 / \partial \lambda_1 & -\partial a_1 / \partial \lambda_2 & \cdots & -\partial a_1 / \partial \lambda_m \\ -\partial a_2 / \partial \lambda_1 & -\partial a_2 / \partial \lambda_2 & \cdots & -\partial a_2 / \partial \lambda_m \\ \vdots & \vdots & \ddots & \vdots \\ -\partial a_m / \partial \lambda_1 & -\partial a_m / \partial \lambda_2 & \cdots & -\partial a_m / \partial \lambda_m \end{pmatrix} \quad (3.36)$$

matrisi ele alınır. Çünkü $HH^{-1} = I$ matrisin çarpımı birim matristir.

$$\frac{\partial \lambda_k}{\partial \lambda_l} = \begin{cases} 0, & k \neq l \\ 1, & k = l \end{cases} \text{ olduğundan,}$$

$$HH^{-1} = (b_{ij}) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (3.37)$$

dir.

(3.36) matrisinin elemanları $\frac{\partial a_r}{\partial \lambda_s}$ (3.9) dan bulunursa

$$\begin{aligned} \frac{\partial a_r}{\partial \lambda_s} &= -\sum_{i=1}^n p_i \sum_{j=1}^n p_j g_r(x_i) g_s(x_i) + \sum_{i=1}^n p_i g_r(x_i) \sum_{j=1}^n p_j g_s(x_i) \\ &= -\{E[g_r(X)g_s(X)] - E[g_r(X)]E[g_s(X)]\} \end{aligned} \quad (3.38)$$

veya

$$\frac{\partial a_r}{\partial \lambda_s} = -Cov[g_r(X), g_s(X)] \quad (3.39)$$

$$\frac{\partial a_r}{\partial \lambda_r} = -Var[g_r(X)] \quad (3.40)$$

elde edilir. (3.39) ve (3.40) değerleri (3.37) deki H^{-1} matrisinde yerine yazılırsa elde edilen matris varyans-covaryans matrisi olur ki bu H^{-1} matrisinin pozitif tanımlı olduğu gösterir. Böylece pozitif tanımlı bir matrisin tersi de pozitif tanımlı olacağından, H matrisi pozitif tanımlanır ve de dolayısıyla D_{\min} 'in a_1, a_2, \dots, a_m 'nin konveks bir fonksiyonu olduğu ortaya çıkar.

3.2. Sürekli Rassal Değişkenler İçin MinxEnt Metodu

Bu metot K-L ölçümünün

$$D(f : q) = \int_a^b f(x) \ln \frac{f(x)}{q(x)} dx \quad (3.41)$$

$$\int_a^b f(x)dx = 1 \quad (3.42)$$

$$\int_a^b g_r(x)f(x)dx = a_r, \quad r = 1, \dots, m, \quad (3.43)$$

(3.42) ve (3.43) kısıtları altında minimize eden dağılımı bulma metodudur. Sözkonusu dağılım Lagrange çarpanları metoduyla bulunabilir.

$$L = \int_a^b f(x) \ln \frac{f(x)}{q(x)} dx + (\lambda_0 - 1) \left(\int_a^b f(x)dx - 1 \right) + \sum_{r=1}^m \lambda_r \left(\int_a^b g_r(x)f(x)dx - a_r \right) \quad (3.44)$$

şeklinde kurulur.

$$L = \int_a^b \underbrace{\left(f(x) \ln \frac{f(x)}{q(x)} + (\lambda_0 - 1)f(x) + \sum_{r=1}^m \lambda_r g_r(x)f(x) \right)}_F dx - (\lambda_0 - 1) - \sum_{r=1}^m \lambda_r a_r \quad (3.45)$$

$$F = f(x) \ln \frac{f(x)}{q(x)} + (\lambda_0 - 1)f(x) + \sum_{r=1}^m \lambda_r g_r(x)f(x) \text{ olsun.}$$

Bu durumda (3.45)

$$L = \int_a^b F(x)dx \text{ olur. (3.45) } L \text{ fonksiyoneline minimum değer veren } f(x)$$

fonksiyonunun ekstremumunun varlığı için gerek koşula göre Euler-Lagrange koşulunu sağlamalıdır.

$$\frac{\partial F}{\partial f(x)} - \frac{d}{dx} \left(\frac{\partial F}{\partial f'(x)} \right) = 0. \quad (3.46)$$

(3.46) denkleminde

$$f(x) = q(x) \exp(-\lambda_0 - \lambda_1 g_1(x) - \dots - \lambda_m g_m(x)) \quad (3.47)$$

şeklinde bulunur. $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ ler (3.42) ve (3.43) koşullarını sağlamalıdır.

MinxEnt yoğunluk fonksiyonu moment koşullarında yerine yazılırsa aşağıdaki eşitlikler ortaya çıkar

$$\exp(\lambda_0) = \int_a^b q(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx \quad (3.48)$$

$$a_r \exp(\lambda_0) = \int_a^b g_r(x) q(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx, ,$$

$$a_r = \frac{\int_a^b g_r(x)q(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx}{\int_a^b \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx}, \quad r = 1, \dots, m \quad (3.49)$$

Kesikli sistemde olduğu gibi λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'in konkav bir fonksiyonu olduğu ve S_{\max} 'in a_1, a_2, \dots, a_m 'lerin konveks bir fonksiyonu olduğu kolayca gösterilebilir.

3.2.1. $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarının Varlığı

$\lambda_1, \lambda_2, \dots, \lambda_m$ 'in varlığı kesikli sistemde olduğu gibi (3.49) eşitliği

$$\begin{aligned} f_1(\lambda_1, \dots, \lambda_m) &= a_1 \\ &\vdots \\ f_m(\lambda_1, \dots, \lambda_m) &= a_m \end{aligned} \quad (3.50)$$

$$\text{burada } f_r(\lambda_1, \dots, \lambda_m) = \frac{\int_a^b q(x)g_r(x) \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}, \quad f = (f_1, \dots, f_m)'$$

$\lambda = (\lambda_1, \dots, \lambda_m)$ ve (a_1, \dots, a_m) dir.

Sistemin $\lambda_1, \dots, \lambda_m$ lere göre çözüme sahip olması gösterilmelidir.

$\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)}$ determinantının sıfırdan farklı olduğunu gösterilirse, $\lambda_1, \dots, \lambda_m$ varlığı

kanıtlanmış olur.

(3.50) den

$$\frac{\partial f_r}{\partial \lambda_r} = \frac{-\int_a^b g_r^2(x)q(x) \exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right) dx \int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2}$$

$$\frac{-\int_a^b g_r(x)q(x)\exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\} \int_a^b g_r(x)q(x)\exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2}, \quad (3.51)$$

veya

$$\frac{\partial f_r}{\partial \lambda_r} = -\left(E(g_r^2(X)) - E(g_r(X))E(g_r(X))\right) = -\text{Var}(g_r(X)); r = 1, \dots, m. \quad (3.52)$$

(3.50) de $r \neq s$ olduğunda

$$\frac{\partial f_r}{\partial \lambda_s} = \frac{-\int_a^b g_r(x)g_s(x)\exp\left(-\sum_{j=1}^m \lambda_j g_j(x)\right)dx \int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\sum_{i=1}^n \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2} \quad (3.53)$$

$$-\frac{\int_a^b g_r(x)\exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\} \int_a^b g_s(x)\exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}}{\left\{\int_a^b \exp\left\{-\sum_{r=1}^m \lambda_r g_r(x)\right\}\right\}^2};$$

$$\frac{\partial f_r}{\partial \lambda_s} = -\left(E(g_r(X)g_s(X)) - E(g_r(X))E(g_s(X))\right) = -\text{Cov}(g_r(X)g_s(X)); \quad (3.54)$$

$r \neq s, r = 1, \dots, m$

(3.52), (3.53), (3.54) yardımıyla

$$\begin{pmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \dots & \frac{\partial f_1}{\partial \lambda_m} \\ \frac{\partial f_2}{\partial \lambda_1} & \frac{\partial f_2}{\partial \lambda_2} & \dots & \frac{\partial f_2}{\partial \lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \dots & \frac{\partial f_m}{\partial \lambda_m} \end{pmatrix} = - \begin{pmatrix} \text{Var}[g_1(X)] & \text{Cov}[g_1(X)g_2(X)] & \dots & \text{Cov}[g_1(X)g_m(X)] \\ \text{Cov}[g_2(X)g_1(X)] & \text{Var}[g_2(X)] & \dots & \text{Cov}[g_2(X)g_m(X)] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[g_m(X)g_1(X)] & \text{Cov}[g_m(X)g_2(X)] & \dots & \text{Var}[g_m(X)] \end{pmatrix} \quad (3.55)$$

$\frac{D(f_1, \dots, f_m)}{D(\lambda_1, \dots, \lambda_m)} \neq 0$ dır. Çünkü (3.55) matrisi varyans-kovaryans matrisinin ters

işaretlisidir. (3.50) denklemini çözüme sahiptir. Dolayısıyla $f^{-1}(\mathbf{a}) = \boldsymbol{\lambda}$ vardır ve $\boldsymbol{\lambda}$ ' \mathbf{a} 'ya sürekli bağlıdır.

3.2.2. Koşullu Ekstreminin Varlığı

(3.47) MinxEnt fonksiyonunun Kulback-Leibler fonksiyoneline minimum değer verdiği bilinen bir eşitsizlikten yola çıkılarak yapılacaktır. $\ln(x) \leq x-1$, her $x > 0$ için geçerli olduğu bilinmektedir. Burada $k(x)$ yoğunluk fonksiyonu,

$t(x) \geq 0$ ve $\int_a^b t(x)dx \leq 1$ olsun. x yerine, $\frac{t(x)}{k(x)}$ ifadesi eşitsizlikte yerine yazılırsa

$$\ln \frac{t(x)}{k(x)} \leq \frac{t(x)}{k(x)} - 1 \quad (3.56)$$

$$\ln t(x) - \ln k(x) \leq \frac{t(x)}{k(x)} - 1 \quad (3.57)$$

$$k(x) \ln t(x) - k(x) \ln k(x) \leq t(x) - k(x).$$

Son eşitsizliğin her iki tarafından integral alındığında

$$\int_a^b k(x) \ln t(x) dx - \int_a^b k(x) \ln k(x) dx \leq \int_a^b t(x) dx - \int_a^b k(x) dx \quad (3.58)$$

(3.58)'in her iki tarafına $-\int_a^b k(x) \ln q(x) dx$ eklendiğinde

$$\begin{aligned} & \int_a^b k(x) \ln t(x) dx - \int_a^b k(x) \ln k(x) dx - \int_a^b k(x) \ln q(x) dx \\ & \leq \int_a^b t(x) dx - \int_a^b k(x) dx - \int_a^b k(x) \ln q(x) dx \\ & \int_a^b k(x) \ln \frac{t(x)}{q(x)} dx \leq \int_a^b k(x) \ln \frac{k(x)}{q(x)} dx \end{aligned} \quad (3.59)$$

elde edilir. (3.59) da $k(x) = f(x)$ yerine yazıldığında son eşitsizlik

$$\int_a^b k(x) \ln \frac{t(x)}{q(x)} dx \leq D(f : q) \quad (3.60)$$

şekline dönüşür. (3.60) eşitsizliği (3.41) Kullback-Leibler fonksiyonelinin

$\int_a^b k(x) \ln \frac{t(x)}{q(x)} dx$ ile sınırlı olduğunu göstermektedir. Burada $t(x)$ yerine

Lagrange çarpanları metodundan elde edilmiş dağılım yazılırsa

$$\int_a^b k(x) \ln \frac{q(x) \exp(-\lambda_0 - \lambda_1 g_1(x) - \dots - \lambda_m g_m(x))}{q(x)} \leq D(f : q) \quad (3.61)$$

veya

$$-\lambda_0 - \sum_{r=1}^m \lambda_r a_r \leq D(f : q) \quad (3.62)$$

sonucuna varılır. Bu sonuçtan $D(f : q)$ fonksiyonelinin alttan sınırlı olduğu belirlenmiş olmaktadır. Bu sonuç keyfi f yoğunluk fonksiyonu için geçerli olduğundan $f(x) = q(x) \exp(-\lambda_0 - \lambda_1 g_1(x) - \dots - \lambda_m g_m(x))$ fonksiyonu Kullback-Leibler fonksiyoneline minimum değer veren dağılımdır sonucuna ulaşılır.

3.2.3. λ_0 Çarpanının $\lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange Çarpanlarına göre Konveksliği

λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ konveks bir fonksiyonu olduğunu göstermek için (3.49) eşitliğinden λ_r 'ye göre türev alınırsa,

$$\exp(\lambda_0) \frac{\partial \lambda_0}{\partial \lambda_r} = - \int_a^b q(x) g_r(x) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x) \right\}, \quad (3.63)$$

$$\frac{\partial \lambda_0}{\partial \lambda_r} = - \frac{\int_a^b g_r(x) q(x) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x) \right\}}{\exp(\lambda_0)} = -E[g_r(X)] = -a_r. \quad (3.64)$$

(3.64)'den λ_r 'ye göre tekrar diferansiyel alınırsa

$$\frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_r} + \frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = \frac{\int_a^b g_r^2(x) q(x) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x) \right\}}{\exp(\lambda_0)} = E[g_r^2(X)], \quad (3.64)$$

böylece

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r^2} = E[g_r^2(X)] - E[g_r(X)]^2 = \text{Var}[g_r(x)]. \quad (3.65)$$

(3.64)'den λ_s 'ye ($r \neq s$) göre tekrar türev alınırsa

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\partial \lambda_0}{\partial \lambda_s} = \frac{- \int_a^b g_r(x) g_s(x) \exp \left\{ - \sum_{r=1}^m \lambda_r g_r(x) \right\}}{\exp(\lambda_0)}. \quad (3.66)$$

(3.64) den tekrar λ_s 'e göre tekrar türev alınırsa

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} + \frac{\partial \lambda_0}{\partial \lambda_r} \frac{\lambda_0}{\partial \lambda_s} = E[g_r(X)g_s(X)], \quad (3.67)$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_r \partial \lambda_s} = E[g_r(X)g_s(X)] - E[g_r(X)]E[g_s(X)] = Cov[g_r(X)g_s(X)] \quad (3.68)$$

sonucuna ulaşılır. λ_0 'ın $\lambda_1, \lambda_2, \dots, \lambda_m$ 'nın fonksiyonu olduğuna göre (3.65) ve

(3.68) yardımıyla oluşturulan Hessian matrisi:

$$\begin{pmatrix} \frac{\partial^2 \lambda_0}{\partial \lambda_1^2} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_m} \\ \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_1 \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_2 \partial \lambda_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_1} & \frac{\partial^2 \lambda_0}{\partial \lambda_m \partial \lambda_2} & \dots & \frac{\partial^2 \lambda_0}{\partial \lambda_m^2} \end{pmatrix} = \begin{pmatrix} Var[g_1(X)] & Cov[g_1(X)g_2(X)] & \dots & Cov[g_1(X)g_m(X)] \\ Cov[g_2(X)g_1(X)] & Var[g_2(X)] & \dots & Cov[g_2(X)g_m(X)] \\ \vdots & \vdots & & \vdots \\ Cov[g_m(X)g_1(X)] & Cov[g_m(X)g_2(X)] & \dots & Var[g_m(X)] \end{pmatrix}. \quad (3.69)$$

olduğu sonucuna ulaşılır. (3.69) matrisi varyans-kovaryans matrisi olduğu için λ_0 parametresi $\lambda_1, \lambda_2, \dots, \lambda_m$ 'nın konveks fonksiyonudur.

4. İNFORMASYON TEORİSİNE DAYALI MODEL BELİRLEME KRİTERLERİ

Bu bölümde, istatistiksel modellerin belirlenmesinde kullanılan bilgi teoriye dayalı bazı kriterler ve hataya dayalı kriterler verilerek, tezde belirlenen dağılım modellerini karşılaştırmak için kullanılmıştır.

Bu kriterlerden Akaike-İnformasyon kriteri (AIC) (Akaike, 1973, 1978), AIC'nin küçük örneklem (AIC_c, QAIC_c) ve Takeuchi bilgi teori kriteri (TIC) iki dağılım arasındaki yönlendirilmiş Kullback-Leibler mesafesinin tahmini olarak türetilir. Bayes bilgi teori kriteri (BIC), Minimum tarif uzunluğu (Minimum describe Length (MDL)) , Tutarlılık ve Fisher bilgi teori (Consistent and Fisher information) (QF) kriterleri ise boyut tutarlılığını belirleyen kriterlerdir. Ayrıca istatistikte hataya dayalı kriterler kök ortalama kare hata (Root mean square error, (RMSE)), Ki-Kare (χ^2) ve modelin belirlilik katsayısı olarak bilinen R^2 de model seçim kriteri olarak kullanılmaktadır.

4.1. Akaike İnfomasyon kriteri

AIC kriteri bilinmeyen gerçek dağılım ile önerilen model arasındaki K-L mesafesinin beklenen değerinin tahmini ile oluşturulur. AIC veriye uygun aynı zamanda en az parametreye sahip modeli önermektedir. Kriterin farklı kanıtları literatürde mevcut olmakla beraber (Burnham, 2002) kaynağında sade bir hali bulunmaktadır. Bu kriter, aday modeller arasından en küçük AIC değere sahip olan dağılımı seçer. Eğer model kümesi doğru seçilmemiş ise bu durumda kriter iyi bir model önermeyecektir (Kenneth,2002).

İki model arasındaki Kullback-Leibler bilgi teori ölçüsü

$$D(f(x); g(x, \theta)) = \int f(x) \ln \frac{f(x)}{g(x, \theta)} dx \quad (4.1)$$

şeklinindedir. Burada $f(x)$ gerçek dağılım, $g(x, \theta)$ önerilen model dağılımdır.

AIC:

$$AIC = -2 \log(L(\hat{\theta}|x)) + 2K \quad (4.2)$$

şeklinde. Burada L loglikelihood fonksiyonu, K modeldeki parametre sayısıdır. AIC örneklem sayısı (N) küçük olduğunda kötü sonuçlar verdiği için (Sigaura 1978; Sakamoto ve ark 1986; Hurvich ve Tsai,1989) tarafından küçük örneklem için modifiye edilmiştir. Küçük örneklem için AIC_c kriteri:

$$AIC_c = AIC + 2K + \frac{2K(K+1)}{N-K-1}, \quad (4.3)$$

(4.3) dikkat edildiği gibi, N , değeri K ya göre yeterince büyük ise son terim ihmal edilebilir.

Takeuchi 'nin kriteri olan TIC'nin ifadesi aşağıdaki gibidir.

$$TIC = -2 \log(L(\hat{\theta}|x)) + 2tr\{K^{-1}(\theta)J(\theta)\}. \quad (4.4)$$

$$\text{Burada } K = \text{Var}_f\left(\frac{\partial}{\partial \theta} \log(g(x, \theta))\right) \text{ ve } J = -E_f\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(g(x, \theta))\right) \quad (4.5)$$

$tr\{K^{-1}(\theta)J(\theta)\}$ ise çarpım matrisinin izidir. Burada Varyans ve beklenen değer $g(x, \theta)$ göre alınsaydı $2tr\{K^{-1}(\theta)J(\theta)\}=K$ parametre sayısına eşit olup $TIC=AIC$ olacağı görünmektedir.

4.2. Bayes İnfomasyon Kriteri

Schward “Boyut tutarlılığı” olarak bilinen BIC’i geliştirdi. Bu kriter tamamıyla Bayes teorisinden türetilmiş olup pek çok alana uygulanmaktadır.

Bayes infomasyon kriteri aşağıdaki gibi tanımlanır:

$$BIC = -2 \log(L(\hat{\theta}|x)) + K \log(N). \quad (4.6)$$

Yine burada, K modeldeki parametre sayısı, N ise örneklem sayısıdır.

4.3. İstatistiksel kriterler

İstatistikte daha çok hataya dayalı kriterler kullanılmaktadır. Bunlardan en önemlileri aşağıdakilerdir .

RMSE:

$$RMSE = \frac{\sum_{i=1}^N (y_i - x_i)^2}{N}; \quad (4.7)$$

Modelin tahmin gücünün bir ölçüsü olan Belirlilik katsayısının formülü

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - x_i)^2}{\sum_{i=1}^N (y_i - z_i)^2}; \quad (4.8)$$

Ki-Kare:

$$\chi^2 = \frac{\sum_{i=1}^N (y_i - x_i)^2}{N - K}. \quad (4.9)$$

Burada y_i , i . gerçek veri, x_i , i . dağılımından tahmin edilen veri, N gözlem sayısı, K parametre sayısı veya kısıt sayısıdır. Buraya kadar verilen kriterlerden R^2 dışındakiler, en küçük değerlerine göre en iyi dağılımı belirlerler. Ancak R^2 ise, bir modelin tahmin gücünün bir ölçüsü olarak 0 ve 1 arasında değişir ve 1'e yaklaşması modelin tahmin gücünün arttığının göstergesidir.

Ayrıca bu tezde kullanılmamakla birlikte bilgi teorisi dayalı kriterler mevcuttur. Daha çok Kodlama teorisinde kullanılan Minimum tanımlama uzunluğu (Minimum description Length (MDL)) en az karmaşıklığa sahip modeli seçer, regresyonda daha çok uygulanan Çarpıklık bilgi ölçüsü (Deviance information criterion, DIC), Autoregressive zaman serisi modellerini belirlemek için Hannan ve Qinn kriteri mevcuttur.

5. GENELLEŞTİRİLMİŞ MAKSİMUM ENTROPİ METODU VE GELİŞTİRİLMİŞ DAĞILIMLAR

Bu bölümde, MaxEnt fonksiyoneliinden faydalanarak, moment fonksiyonlarına bağlı olarak yeni bir fonksiyonel tanımlanmış ve onun bazı özellikleri araştırılmıştır. Ayrıca fonksiyonelin bazı özelliklerinden faydalanılarak yeni bir dağılım belirleme süreci geliştirilmiştir. Tanımlanan bu yeni fonksiyona MinMaxEnt dağılımı denilecektir. Bu süreç kesikli ve sürekli rassal değişkenler için ayrı ayrı incelenecektir.

5.1. Kesikli Rassal Değişkenler İçin Maksimum Entropi Fonksiyonu ve Moment Fonksiyonları Yardımıyla Tanımlanmış Bir Fonksiyonel

x_1, x_2, \dots, x_n rassal örnek (RÖ) ve bu değerlere karşı gelen olasılıklar p_1, p_2, \dots, p_n olsun.

$$H(x) = -\sum_{i=1}^n p_i \ln p_i \quad (5.1)$$

entropi fonksiyonunun

$$\sum_{i=1}^n p_i = 1, \quad (5.2)$$

$$\sum_{i=1}^n p_i g_r(x_i) = \mu_r, \quad r = 1, \dots, m \quad (5.3)$$

$g = (g_1, \dots, g_m)$ moment vektör fonksiyonu ve $\mu = (\mu_1, \dots, \mu_m)$ moment vektör değeri, $k = 1, \dots, m$; $m + 1 < n$ kısıtları altında maksimum değere sahiptir. Bu maksimum değeri bulmak için Lagrange çarpanları metodu uygulanarak entropinin maksimum değeri H_{\max} :

$$H_{\max} = \lambda_0 + \lambda_1 \mu_1 + \dots + \lambda_m \mu_m. \quad (5.4)$$

olduğu (2.1.5) in (2.28) formülüyle gösterilmiştir. Bu formülde $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanları, μ_1, \dots, μ_m ler $g_1(x), g_2(x), \dots, g_m(x)$ 'e uygun moment değerleridir. (5.4) eşitliğiyle tanımlanmış H_{\max} 'a (g_1, \dots, g_m) ' moment vektörlerine bağlı fonksiyonel gibi bakılabilir.

Tanım 5.1. $g = (g_1, \dots, g_m)$ moment vektör fonksiyonlarına bağlı H_{\max} değerine sahip

$$S(g) = H_{\max} \quad (5.5)$$

fonksiyoneline, moment fonksiyonlarına bağlı MaxEnt fonksiyoneli denilecektir. MaxEnt fonksiyoneli aşağıdaki teoremler şeklinde verilecek olan özelliklere sahiptir.

Teorem 5.1. (5.5) ile ifade edilen $S(g)$ fonksiyoneli sürekli vektör moment fonksiyonları kümesinde $(C[a,b])$ süreklidir.

Kanıt. Her bir $g(x)$ m -boyutlu moment vektör fonksiyonu için entropi fonksiyonu H (2.1.5) bölümünde kanıtlandığı gibi maksimum değere sahiptir. Böylece her moment vektör fonksiyonunu H 'ın maksimumu ile karşılaştıran fonksiyonel $S(g)$ (5.5) formülüyle tanımlansın. $g^{(1)}(x)$, $g^{(2)}(x)$ moment vektör fonksiyonları ve $\mu^{(1)}$, $\mu^{(2)}$ moment vektör değerleri olsun. Bu durumda

$$\mu^{(2)} - \mu^{(1)} = \sum_{i=1}^n p_i (g^{(2)}(x_i) - g^{(1)}(x_i)), \quad (5.6)$$

diğer bir deyişle

$$\begin{aligned} \|\mu^{(2)} - \mu^{(1)}\| &\leq \sum_{i=1}^n p_i \|g^{(2)}(x_i) - g^{(1)}(x_i)\| \leq \sum_{i=1}^n p_i \|g^{(2)} - g^{(1)}\|_C \\ &= \|g^{(2)} - g^{(1)}\|_C, \\ \|\mu^{(2)} - \mu^{(1)}\| &\leq \|g^{(2)} - g^{(1)}\|_C, \end{aligned} \quad (5.7)$$

burada $\|\cdot\|$ Euclid normu ve $\|g\|_C = \max_{x \in [a,b]} \|g(x)\|$ $C[a,b]$ uzayının normudur.

(5.7) den görülmektedir ki

$$\|g^{(2)} - g^{(1)}\|_C < \delta, \quad (5.8)$$

olduğunda

$$\|\mu^{(2)} - \mu^{(1)}\| < \varepsilon, \quad (\varepsilon = \delta). \quad (5.9)$$

eşitsizliği sağlanmaktadır. (5.9) da μ moment vektör değeri $g(x)$ 'e $C[a,b]$ normuna göre $g(x)$ 'e sürekli bağlıdır.

Burada λ_0 parametresi $\lambda_1, \dots, \lambda_m$ 'lere ve $\lambda_1, \dots, \lambda_m$ parametreleride $g(x)$ lere göre μ_1, \dots, μ_m 'lere sürekli bir şekilde bağlıdır. (bak. bölüm 2) (5.8) ve (5.9) dikkate alınırca $S(g)$ fonksiyoneli sürekli fonksiyonlar kümesinde sürekli bir fonksiyoneldir. Böylece Teorem 5.1. kanıtlanmış olur.

Teorem 5.2. (5.5) ile ifade edilen $S(g)$ fonksiyoneli, $(C[a,b])$ uzayının verilen kompakt kümesinde en büyük ve en küçük değerine ulaşır.

Bu teoremin kanıtı lineer uzayın kompakt kümesinde tanımlanmış sürekli fonksiyonellerin bir özelliğinden ortaya çıkmaktadır.

5.2. Kesikli Rassal Değişkenler için MinMaxEnt Dağılımı

(5.5) ile tanımlanmış $S(g) \subset C[a,b]$ fonksiyoneline, K kompakt moment vektör fonksiyonları kümesinde minimum değer veren moment vektör fonksiyonu $g^{(0)}$ olsun.

$$\min_{g \in K} S(g) = S(g^{(0)}). \quad (5.10)$$

(5.3) moment kısıtlarında $g = (g_1, \dots, g_m)$ yerine $g^{(0)}$ moment koşulu yazıldığında (5.1) fonksiyonuna (5.2), (5.3) koşulları altında maksimum değer veren $p = (p_1, \dots, p_n)$ dağılımına MinMaxEnt dağılımı denilecektir. Kısaca MinMaxEnt dağılımı S fonksiyoneline minimum değer veren $g^{(0)}$ moment vektör fonksiyonuna uygun MaxEnt dağılımıdır. Bu süreçte dikkat edilmesi gereken husus fonksiyonlar kümesinde tanımlanan MaxEnt fonksiyonundan farklı olarak, momentler kümesinde tanımlanan bir fonksiyonelin minimumunun bulunmasıdır. Sözkonusu MinMaxEnt dağılımı

$$p_i = \exp(-\lambda_0 - \sum_{j=1}^m \lambda_j g_j^{(0)}(x_i)), \quad i = 1, \dots, n \quad (5.11)$$

şeklinindedir. MinMaxEnt dağılımı, rassal örnekten elde edilen dağılımlar içerisinde en büyük bilgiyi içeren dağılımdır.

Teorem 5.3. $P^{(0)} = (p_1^{(0)}, \dots, p_n^{(0)})$ (5.11) eşitliğiyle tanımlanmış MinMaxEnt dağılımı olsun. Bu dağılım rassal örnekten elde edilen dağılımlar içerisinde en büyük informasyonu içeren dağılımdır.

Kanıt. n duruma sahip rassal değişkenin (5.2) koşulları altında maksimum entropisi $\ln n$ dir. (5.1) ve (5.2) koşulları altında entropi $H(p)$ olduğu için P dağılımının içerdiği informasyon $I = \ln n - H(p)$ şeklindedir.

MinMaxEnt dağılımı entropi fonksiyonuna en küçük değer veren dağılım olduğu için

$$H(p) \geq S(g^0) \quad (5.12)$$

koşulu sağlanır. Böylece Teorem 5.3. kanıtlanmış olur.

5.3. Sürekli Rassal Değişkenler İçin Maksimum Entropi Fonksiyonu ve Moment Fonksiyonları Yardımıyla Tanımlanmış Bir Fonksiyonel

$$H(x) = - \int_a^b f(x) \ln f(x) dx \quad (5.13)$$

aşğıdaki moment koşulları şeklinde ifade edilen

$$\int_a^b f(x) g_r(x) dx = \mu_r \quad r = 1, \dots, m \quad (5.14)$$

$$\int_a^b f(x) dx = 1 \quad (5.15)$$

(5.13) entropi fonksiyoneli (5.14) ve (5.15) kısıtları altında maksimuma sahiptir. Bu maksimum değeri bulmak için Lagrange çarpanları metodu uygulanarak entropi fonksiyonelinin maksimum değeri H_{\max} :

$$H_{\max} = \lambda_0 + \lambda_1 \mu_1 + \dots + \lambda_m \mu_m. \quad (5.16)$$

olduğı gösterilmiştir. Bu formülde $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_m$ Lagrange çarpanları, $\mu_1, \mu_2, \dots, \mu_m$ 'ler $g_1(x), g_2(x), \dots, g_m(x)$ 'e uygun moment değerleridir. (5.16) eşitliğiyle tanımlanmış H_{\max} 'a kesiklide olduğu gibi (g_1, \dots, g_m) ' moment vektörlerine bağlı fonksiyonel gibi bakılabilir.

Tanım 5.2. $g = (g_1, \dots, g_m)$ moment vektör fonksiyonlarına bağlı H_{\max} değerine sahip

$$S(g) = H_{\max} \quad (5.17)$$

fonksiyoneline moment fonksiyonlarına bağlı MaxEnt fonksiyoneli denilecektir.

MaxEnt fonksiyoneli 5.4. bölümde teoremler şeklinde verilecek olan özelliklere sahiptir.

5.4. Sürekli Rassal Değişkenler için MinMaxEnt Dağılımı

MinMaxEnt dağılımı (5.14) moment kısıtlarında $g = (g_1, \dots, g_m)$ yerine $g^{(0)}$ moment koşulu yazıldığında (5.13) fonksiyoneline (5.14), (5.15) koşulları altında maksimum değer veren dağılımına MinMaxEnt dağılımı denilecektir. Kısaca MinMaxEnt dağılımı S fonksiyoneline minimum değer veren $g^{(0)}$ moment fonksiyonuna karşı gelen MaxEnt dağılımıdır. (Shamilov ve ark. 2006)

Teorem 5.4. (5.17) ile ifade edilen $S(g)$ fonksiyoneli sürekli vektör moment fonksiyonları kümesinde $(C[a,b])$ süreklidir.

Kanıt. (5.14) eşitliğinden

$$\mu^{(2)} - \mu^{(1)} = \int_a^b (g^{(2)}(x) - g^{(1)}(x))f(x)dx, \quad (5.18)$$

veya

$$\begin{aligned} \|\mu^{(2)} - \mu^{(1)}\| &\leq \int_a^b \|g^{(2)}(x) - g^{(1)}(x)\| f(x)dx \leq \int_a^b \|g^{(2)} - g^{(1)}\|_C f(x)dx \\ &= \|g^{(2)} - g^{(1)}\|_C \int_a^b f(x)dx = \|g^{(2)} - g^{(1)}\|_C \end{aligned} \quad (5.19)$$

$$= \|g^{(2)} - g^{(1)}\|_C,$$

$$\|\mu^{(2)} - \mu^{(1)}\| \leq \|g^{(2)} - g^{(1)}\|_C. \quad (5.20)$$

(5.20) fonksiyonel olarak μ , $C[a,b]$ uzayında, $g(x)$ moment vektör fonksiyonuna sürekli bir şekilde bağlıdır. Diğer taraftan λ_0 parametresi $\lambda_1, \dots, \lambda_m$

lere ve $\lambda_1, \dots, \lambda_m$ parametreleri μ_1, \dots, μ_m 'lere sürekli bir şekilde bağlıdır (bkz. Bölüm 2). (5.20)'den $S(g)$ 'nin g 'nin sürekli bir fonksiyonu olduğu görülmektedir. Böylece Teorem 5.4 kanıtlanmış oldu.

Teorem 5.5. (5.17) ile ifade edilen $S(g)$ fonksiyoneli, $(C[a, b])$ sürekli moment fonksiyonlarının kompakt bir kümesinde en büyük ve en küçük değerini alır.

Bu teoremin kanıtı lineer uzayda kompakt kümede tanımlanan sürekli fonksiyonların bir özelliğinden ortalaya çıkmaktadır. (Kantoroviç ve Akilov, 1982).

Teorem 5.6. Sürekli sistemler için tanımlanmış MinMaxEnt dağılımı, rassal örnekten elde edilen dağılımlar içerisinde en büyük bilgiyi içeren dağılımdır.

Kanıt. (5.13) entropi fonksiyonelinin (5.15) koşulu altında maksimum değeri $H_{\max} = \ln(b - a)$ şeklinde olduğu Bölüm 2 de gösterilmişti. (5.13) ve (5.14) koşulları altında entropi $H(p)$ olduğu için P dağılımının içerdiği bilgi

$$I = \ln(b - a) - H(p) \quad (5.21)$$

şeklindedir.

MinMaxEnt dağılımı entropi fonksiyoneline en küçük değer veren dağılım olduğu için

$$H(p) \geq S(g^0) \quad (5.22)$$

koşulu sağlanır. Böylece Teorem 5.6 kanıtlanmış olur.

5.5. MinMaxEnt ve MaxMaxEnt dağılım dizisi

Kesikli sistem MaxEnt metodu ile incelendiğinde dikkat edilmelidir ki burada $m+1$ sayıda kısıt, ancak n sayıda bilinmeyen olasılık vardır. Metodun uygulanabilmesi için gereken koşullardan birisi $m+1 < n$ olmasıdır. Teorem 5 gösteriyor ki $H_{\max} = S(g)$ fonksiyoneli sonsuz sayıda moment fonksiyonlarının oluşturduğu kümede tanımlı olduğu için kendinin en büyük ve en küçük değerini

bu kümede alır. $S(g)$ fonksiyoneli \tilde{g} de minimum değerini, \tilde{g} de maksimum değerini alsın. Bu durumda $S(\tilde{g})$ S 'in minimum değeri, $S(\tilde{g})$ S 'in maksimum değeridir. Sonlu bir $\{S(\tilde{g}_k)\}$, $\{S(\tilde{g}_k)\}$ dizisi için, aşağıdaki eşitsizlik sağlanır.

$$S(\tilde{g}_1) \geq S(\tilde{g}_1) \geq S(\tilde{g}_2) \geq S(\tilde{g}_2) \geq \dots \geq S(\tilde{g}_{m+1}) \geq S(\tilde{g}_{m+1}).$$

\tilde{G}_k dağılımı, \tilde{g}_k moment fonksiyonuna karşılık gelen MaxEnt dağılımı, \tilde{G}_k dağılımı ise \tilde{g} momentine karşılık gelen MaxEnt dağılımı olsun. Diğer bir deyişle \tilde{G}_k MinMaxEnt, \tilde{G}_k MaxMaxEnt dağılımı denilsin. MinMaxEnt ve MaxMaxEnt dağılımları yardımıyla (5.21) eşitsizliği

$$S(\tilde{G}_1) \geq S(\tilde{G}_1) \geq S(\tilde{G}_2) \geq S(\tilde{G}_2) \geq \dots \geq S(\tilde{G}_{m+1}) \geq S(\tilde{G}_{m+1}) \geq H(P),$$

şeklinde yazılabilir. Burada P anakütlenin gerçek dağılımıdır.

Aynı eşitsizlikler sürekli durum içinde geçerlidir. Sadece sürekli durumda $\{S(\tilde{g}_k)\}$, $\{S(\tilde{g}_k)\}$, dizisi sonsuz dizidir.

5.6. Rüzgâr Hızınının MinMaxEnt dağılımı: Weibull Dağılımına Alternatif

Son yıllarda petrol ve gaz gibi tükenir enerji kaynaklarındaki azalış dikkate alınarak yenilenebilir enerji kaynaklarına olan ilgi oldukça artmıştır. Rüzgâr yenilenebilir enerji kaynaklarından en önemlilerinden biridir. Dolayısıyla onun dağılımı araştırması uygulama açısından düşünüldüğünde oldukça değerlidir. Bu anlamda burada sunulan dağılım belirleme süreci rüzgâr gücü ve enerjisi tahminlerinde kullanılabilir. Rüzgâr hızının dağılım fonksiyonu ayrıca rüzgâr enerjisini değerlendirmede önemlidir. Rüzgâr gücü, enerjisi ve rüzgârın pek çok parametresi onun dağılımı yardımıyla belirlenir. Rüzgâr verisine en uygun dağılım ne kadar iyi belirlenirse rüzgâr gücü ve enerjisi tahminleri o kadar az hataya sahip olur. (Celik, 2004,2005, Akpınar, 2004).

Bu bölümde rüzgâr hızı verisini modellemede çok yaygın bir şekilde kullanılan Weibull (Celik, 2003, 2005, Akpınar, 2004, Al-Nassar ve dk, 2005, Ulgen ark, 2002) dağılımına alternatif olarak MinMaxEnt bulunarak karşılaştırmalar

yapılmıştır. Bu karşılaştırmalar yapılırken RMSE, χ^2 , R^2 , AIC, BIC; K-L, kriterleri kullanılmıştır.

Entropi optimizasyon ilkelerinin rüzgâr alanında ilk uygulamaları Li ve Li (2005) de iki farklı yayınlara ve Ramirez ve Carta (2005) gerçekleştirmiştir. Li maksimum entropi prensibini dünyanın çeşitli bölgelerinden alınan rüzgâr verisine uygulamış ve Weibull dağılımı ile MaxEnt dağılımlarını karşılaştırmıştır. Li bu çalışmada moment koşulları olarak rüzgârın kütlesini, momentumunu, enerjisini almış ve bu koşullara göre MaxEnt dağılımını belirlemiştir. MaxEnt dağılımları günlük, aylık, sezonluk ve yıllık rüzgâr verisine uygulayarak karşılaştırmalar yapmıştır. Bu bölümde ise, Li (2005)'nin uyguladığı MaxEnt dağılımından farklı olarak Eskişehir bölgesinde ölçülen rüzgâr hızı verisi için MinMaxEnt dağılımı belirlenerek bu veriden bulunan Weibull dağılımıyla karşılaştırılmıştır.

Bu bölümde Eskişehir bölgesi rüzgâr hızı için Weibull, MinMaxEnt dağılımları belirlenerek karşılaştırılacaktır.

MinMaxEnt dağılımı belirlenmesi süreci aşağıdaki gibi verilir:

1. Moment kısıtları belirlenmesi
2. Her moment kısıtına uygun olarak MaxEnt dağılımını belirlenmesi
3. Belirlenen MaxEnt dağılımlarının entropisini hesaplanması
4. MaxEnt dağılımının içerdiği bilgiyi belirlemesi.
5. En büyük bilgiyi içeren moment kısıtının belirlenmesi.
6. Bu seçilen moment kısıtına karşı gelen dağılıma MinMaxEnt dağılımının bulunması.

1.de ifade edilen moment kısıtları x , x^2 , x^3 , $\ln(x)$, $\ln(x+1)$, $(\ln(x))^2$, $\ln(x^2+1)$ şeklinde belirlenmiştir. Bu kısıtların seçimi yapılırken, bilinen istatistiksel dağılımların karakterize edici momentleri düşünülmüştür. Örneğin x ve x^2 momentleri normal dağılımın, x ve $\ln x$ momentleri Gama dağılımının karakterize edici momentleridir.

1. deki tüm kısıtlar iki bileşkeye sahip moment vektör fonksiyonunu oluşturmaktadır. Çünkü tüm dağılımlar (5.15) le ifade edilen toplamsallık koşulunu sağlarlar.

1. deki tüm kısıtların ikişerli kombinasyonu yapılarak 2-6 adımlarını tekrarlanarak üç bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımı belirlenir. Burada seçilen ikili koşullar, üç bileşkeli moment vektörüne uygun koşulu belirler. Bu şekilde belirlenen iki bileşkeli moment vektör fonksiyonuna uygun MaxEnt, üç ve dört bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımları Çizelge 5.1, Çizelge 5.2 ve Çizelge 5.3 de listelenmiştir. Daha fazla bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımının entropisinin daha küçük olacağı açıktır.

Çizelge 5.1. Aralık ayında ölçülen rüzgâr hızı için iki bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımlarının entropileri

Kısıtlar	Entropi
(1, x)	3.1589
(1, x ²)	3.0445
(1, x ³)	3.0017
(1, ln x)	3.3497
(1, ln(1+x))	3.2924
(1, (ln x) ²)	3.1037

Çizelge 5.2. Aralık ayında ölçülen rüzgâr hızı için üç bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımlarının entropileri

Kısıtlar	Entropi	Kısıtlar	Entropi
(1, x, x ²)	2.9486	(1, x ³ , (ln x) ²)	2.9969
(1, x, x ³)	2.9667	(1, x ³ , ln(1+x ²))	2.9585
(1, x, ln x)	2.9253	(1, ln x, ln(x+1))	2.9702
(1, x, ln(x+1))	2.9284	(1, ln x, (ln x) ²)	2.9523
(1, x, (ln x) ²)	3.0548	(1, ln x, ln(1+x ²))	3.0036
(1, x, ln(1+x ²))	2.9428	(1, ln(x+1), (ln x) ²)	2.9486
(1, x ² , x ³)	2.9816	(1, ln(x+1), ln(1+x ²))	3.1428
(1, x ² , ln x)	2.9263	(1, (ln x) ² , ln(1+x ²))	2.9422
(1, x ² , ln(x+1))	2.9362		
(1, x ² , ln(x) ²)	3.0310		
(1, x ² , ln(1+x ²))	2.9443		
(1, x ³ , ln x)	2.9415		
(1, x ³ , ln(1+x))	2.9526		

Çizelge 5.3. Aralık ayında ölçülen rüzgâr hızı için dört bileşkeli moment vektör fonksiyonuna uygun MaxEnt dağılımlarının entropileri

Kısıtlar	Entropi
$(1, x, x^2, x^3)$	2.9421
$(1, x, x^2, \ln x)$	2.9217
$(1, x, x^2, \ln(x+1))$	2.9283
$(1, x, x^2, (\ln x)^2)$	2.9133
$(1, x, x^2, \ln(1+x^2))$	2.9422
$(1, x, x^3, \ln x)$	2.9194
$(1, x, x^3, \ln(x+1))$	2.9275
$(1, x, x^3, \ln(x)^2)$	2.9081
$(1, x, x^3, \ln(1+x^2))$	2.9423

Kısıtlar	Entropi
$(1, x, \ln x, \ln(x+1))$	2.925
$(1, x, \ln x, \ln(x)^2)$	2.9249
$(1, x, \ln x, \ln(1+x^2))$	2.9252
$(1, x, \ln(x+1), (\ln x)^2)$	2.9252
$(1, x, \ln(x+1), \ln(1+x^2))$	2.9223
$(1, x, (\ln x)^2, \ln(1+x^2))$	2.9278
$(1, x^2, x^3, \ln x)$	2.9257
$(1, x^2, x^3, \ln(1+x))$	2.9339
$(1, x^2, x^3, \ln(x)^2)$	2.9254
$(1, x^2, x^3, \ln(1+x^2))$	2.9421
$(1, x^2, \ln x, \ln(1+x))$	2.9179
$(1, x^2, \ln x, (\ln x)^2)$	2.9186
$(1, x^2, \ln x, \ln(1+x^2))$	2.9164
$(1, x^2, \ln(1+x), (\ln x)^2)$	2.9191
$(1, x^2, \ln(1+x), \ln(1+x^2))$	2.9124
$(1, x^2, (\ln x)^2, \ln(1+x^2))$	2.9212
$(1, x^3, \ln x, \ln(1+x))$	2.9124
$(1, x^3, \ln x, (\ln x)^2)$	2.9136
$(1, x^3, \ln x, \ln(1+x^2))$	2.9103
$(1, x^3, \ln(1+x), (\ln x)^2)$	2.9141
$(1, x^3, \ln(1+x), \ln(1+x^2))$	2.9071
$(1, x^3, \ln(x)^2, \ln(1+x^2))$	2.9162
$(1, \ln x, \ln(1+x), (\ln x)^2)$	2.9278
$(1, \ln x, \ln(1+x), \ln(1+x^2))$	2.9396
$(1, \ln x, (\ln x)^2, \ln(1+x^2))$	2.9332
$(1, \ln(1+x), (\ln x)^2, \ln(1+x^2))$	2.9346

Çizelge 5.1, Çizelge 5.2 ve Çizelge 5.3' den sonuç olarak iki bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımı ($f_{mx2}(x)$), $(1, x^3)$ moment kısıtına karşılık gelen dağılım, üç bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımı ($f_{mx3}(x)$), $(1, x, \ln x)$ moment koşuluna karşılık gelen dağılım ve dört bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımı ($f_{mx4}(x)$), $(1, x^3, \ln x, \ln(1+x^2))$ moment kısıtına karşılık gelen dağılımdır denilebilir.

Çizelge 5.4' de örnek olarak 2005 yılı Aralık ayına ait birimi m/s olan hız (v) verileri kullanılarak elde edilen frekans dağılımları, iki bileşekli moment vektör fonksiyonuna uygun MinMaxEnt ($f_{mx2}(x)$), üç bileşekli moment vektör fonksiyonuna uygun MinMaxEnt ($f_{mx3}(x)$), dört bileşekli moment vektör fonksiyonuna uygun MinMaxEnt ($f_{mx4}(x)$), ve Weibull ($f_w(x)$) dağılımlarından hesaplanan teorik frekanslar verilmiştir. Diğer aylar içinde benzer Çizelgeler oluşturulmuştur. Bu Çizelgede dağılımların veriyi modelleme derecesini istatistiksel olarak belirleyen kriterler verilmiştir. Ayrıca tüm 2005 yılı ayları için bu kriterler hesaplanarak Çizelge 5.5 şeklinde verilmiştir.

Bu şekilde 2005 yılındaki on iki ay için MinMaxEnt dağılımı belirlenmiş, örnek olarak aralık ayı için frekans dağılımları verilerek Çizelge 5.5 deki gibi düzenlenmiştir. Bununla birlikte her ay için bu Çizelgeler oluşturulmuş, hesaplanan MinMaxEnt dağılımları ve Weibull dağılımı Bölüm 4 te verilen RMSE, χ^2 , R^2 , AIC, BIC, K-L kriterlere göre karşılaştırılmıştır. Bu kriterlerden, R^2 dışındakiler en küçük değerlerine göre en iyi dağılımı belirlerken, R^2 de ise en büyük değerine göre seçim yapılır. Çizelge 5.4' de en büyük R^2 f_{mx3} ve f_{mx4} de gerçekleşmiştir. Aynı şekilde χ^2 , AIC, BIC, K-L kriterleri en küçük f_{mx3} ve f_{mx4} dağılımlarında elde edilmiştir.

Çizelge 5.4. Aralık ayında ölçülen rüzgâr hızı için MinMaxEnt dağılımları ve Weibull dağılımından elde edilen teorik frekanslar

v	frekans	f_i	f_w	f_{mx2}	f_{mx3}	f_{mx4}
0.5	86	0.1230	0.1011	0.1555	0.1224	0.1250
1.5	92	0.1316	0.1699	0.1542	0.1306	0.1302
2.5	101	0.1445	0.1785	0.1494	0.1422	0.1404
3.5	107	0.1531	0.1588	0.1393	0.1504	0.1488
4.5	91	0.1302	0.1272	0.1229	0.1474	0.1469
5.5	92	0.1316	0.0942	0.1012	0.1276	0.1282
6.5	73	0.1044	0.0654	0.0765	0.0930	0.0939
7.5	36	0.0515	0.0429	0.0523	0.0544	0.0548
8.5	15	0.0215	0.0268	0.0318	0.0243	0.0241
9.5	6	0.0086	0.0160	0.0169	0.0079	0.0076
RMSE			0.0249	0.0192	0.0068	0.0067
χ^2			0.0008	0.0004	0.0000	0.0000
R ²			0.752	0.852	0.981	0.981
AIC			55.24	54.07	15.75	8
BIC			55.85	54.68	16.66	9.21
K-L			0.0718	0.0281	0.0030	0.0029

Çizelge 5.5 de bir yıllık rüzgâr hızı verisi için MinMaxEnt dağılımlarının, Weibull dağılımıyla çeşitli kriterlere göre karşılaştırması verilmiştir. Çizelge 5.5’den RMSE, χ^2 , R², AIC, BIC, K-L kriterlerine göre tüm aylar için dört bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımının rüzgarı modellemede iyi sonuçlar verdiği görülmüştür. Bu tabloda iki, üç ve dört bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımları ayrı ayrı Weibull dağılımıyla karşılaştırıldığında, f_{mx2} MinMaxEnt dağılımının Weibull’ a göre veriye uyumunun iyi olmadığı, kriterlerin sonuçlarına göre söylenebilir. Bununla birlikte f_{mx3} MinMaxEnt dağılımının Weibull’ a göre rüzgar verisine uyumunun tüm aylar gözlemlendiğinde oldukça iyi olduğu yine kriterlerin bir sonucu olarak görülebilir. Ayrıca üç bileşkeli moment vektör fonksiyonuna uygun

MinMaxEnt dağılımlarının Weibull'a göre çok daha iyi sonuçlar vermesi, kısıt bileşen sayısının da üç olmasının yeterli olduğu kanısına da götürmektedir.

Çizelge 5.5. 2005 yılı tüm aylarında ölçülen rüzgâr hızı için MinMaxEnt dağılımlarının ve Weibull dağılımının karşılaştırılması

Ocak						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0190	0.0004	0.934	71.92	72.72	0.023
f_{mx2}	0.0400	0.0019	0.710	69.58	70.37	0.098
f_{mx3}	0.0146	0.0002	0.961	19.24	20.43	0.021
f_{mx4}	0.0052	0.0000	0.995	8	9.59	0.003
Şubat						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.014	0.0002	0.925	63.78	64.58	0.0435
f_{mx2}	0.0196	0.0004	0.853	62.20	62.99	0.0321
f_{mx3}	0.0118	0.0001	0.946	16.40	17.59	0.0194
f_{mx4}	0.0116	0.0002	0.948	8	9.59	0.0115
Mart						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0178	0.0003	0.937	80.22	81.19	0.0252
f_{mx2}	0.0384	0.0017	0.709	78.72	79.69	0.1072
f_{mx3}	0.0149	0.0002	0.956	18.21	19.67	0.0176
f_{mx4}	0.0079	0.0000	0.987	8	9.93	0.0069
Nisan						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0225	0.0006	0.8908	66.11	66.91	0.0424
f_{mx2}	0.0423	0.0021	0.6158	63.96	64.76	0.1186
f_{mx3}	0.0136	0.0002	0.9602	15.81	17.01	0.0151
f_{mx4}	0.0108	0.0001	0.9749	8	9.59	0.0070

Mayıs						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0180	0.0003	0.952	86.48	87.45	0.0340
f_{mx2}	0.0417	0.0020	0.742	86.63	87.60	0.1236
f_{mx3}	0.0109	0.0001	0.982	18.33	19.78	0.0159
f_{mx4}	0.0080	0.0000	0.990	8	9.93	0.0138

Haziran						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0227	0.0006	0.9152	60.64	61.25	0.0238
f_{mx2}	0.0570	0.0040	0.4694	56.09	56.70	0.1841
f_{mx3}	0.0195	0.0005	0.9373	16.11	17.02	0.0151
f_{mx4}	0.0185	0.0005	0.9436	8	9.21	0.0131

Temmuz						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0185	0.0004	0.9236	67.92	68.71	0.0305
f_{mx2}	0.0439	0.0023	0.5693	64.65	65.44	0.1463
f_{mx3}	0.0185	0.0005	0.9233	18.74	19.94	0.0277
f_{mx4}	0.0165	0.0005	0.9387	8	9.59	0.0203

Ağustos						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0194	0.0005	0.925	81.48	82.45	0.0376
f_{mx2}	0.0522	0.0032	0.467	75.16	76.13	0.2121
f_{mx3}	0.0135	0.0002	0.964	22.20	23.66	0.0157
f_{mx4}	0.0125	0.0002	0.969	8	9.93	0.0145

Eylül						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0165	0.0004	0.9529	50.69	51.09	0.0246
f_{mx2}	0.0508	0.0033	0.5533	48.03	48.42	0.1286
f_{mx3}	0.0168	0.0004	0.9509	15.04	15.63	0.0178
f_{mx4}	0.0150	0.0004	0.9611	8	8.78	0.0159

Ekim						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0180	0.0004	0.9317	59.37	59.97	0.0298
f_{mx2}	0.0316	0.0012	0.7905	58.96	59.56	0.0572
f_{mx3}	0.0101	0.0001	0.9785	18.74	19.65	0.0077
f_{mx4}	0.0089	0.0001	0.9832	8	9.21	0.0066

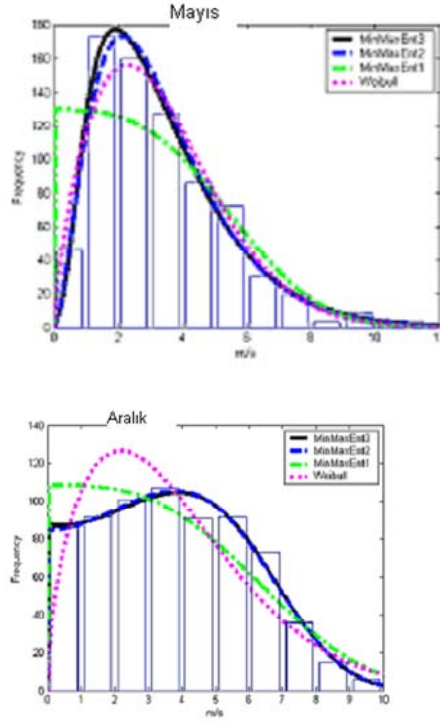
Kasım						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0230	0.0006	0.790	63.63	64.43	0.0808
f_{mx2}	0.0128	0.0002	0.935	62.72	63.51	0.0172
f_{mx3}	0.0109	0.0001	0.952	16.07	17.26	0.0114
f_{mx4}	0.0102	0.0001	0.958	8	9.59	0.0099

Aralık						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0249	0.0008	0.752	55.24	55.85	0.0718
f_{mx2}	0.0192	0.0004	0.852	54.07	54.68	0.0281
f_{mx3}	0.0068	0.0000	0.981	15.75	16.66	0.0030
f_{mx4}	0.0067	0.0000	0.981	8.00	9.21	0.0029

Örnek olarak Mayıs ve Aralık ayları rüzgâr verisi için elde edilen MinMaxEnt ve Weibull olasılık yoğunluk fonksiyonlarının grafiğini Şekil 5.1 de verilmiştir. Şekil 5.1 den de iki, üç ve dört bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımının veriye uyumu görülmektedir. Bununla birlikte

moment koşulları arttığında bu modellemenin daha iyi olacağı aşıkardır. Şekil 5.1’de üç ve dört bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımlarının veriye ne derece iyi uyum gösterdiği oldukça net sergilenmiştir.

Şekil 5.1. Mayıs ve Aralık ayında ölçülen rüzgâr hızının MinMaxEnt dağılımları ve Weibull dağılımı



Son uygulama olarak, incelenen (Celik 2004) makalesindeki rüzgar verilerinin MinMaxEnt dağılımı belirlenmiştir. Celik, (Celik 2004) makalesinde rüzgar dağılımı olarak çeşitli dağılımları test ederek Weibull dağılıma uygun olduğu sonucuna ulaşmıştır Son olarak yukarıda yapılan uygulama (Çelik, 2003) de verilen rüzgâr verisi için uygulanarak sonuçlar yorumlanmıştır. Çelik rüzgârı modellerken iki farklı dağılım kullanarak çeşitli hataya bağlı kriterlere göre Weibull dağılımının incelediği rüzgâr verisini daha iyi modellediğini görmüştür. Bu uygulamada aynı veriyi modellemek için MinMaxEnt dağılımları kullanılarak karşılaştırmalar yapılarak Çizelge 5.6 şeklinde verilmiştir.

Çizelge5.6 ‘den rüzgârı modellemek için kullanılan Weibull dağılımına MinMaxEnt dağılımlarının alternatif olduğu gösterilmiş olur.

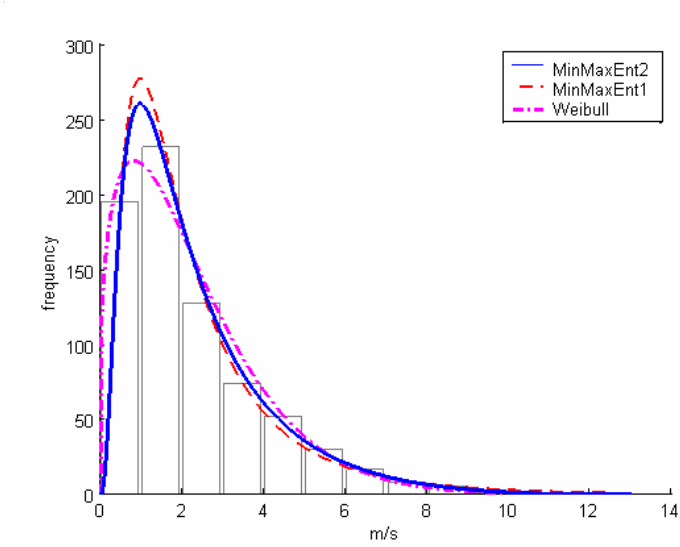
Şekil 5.2, rüzgâr verisi için elde edilen MinMaxEnt ve Weibull olasılık yoğunluk fonksiyonlarının grafiğini vermektedir.

Çizelge 5.6. (Celik 2004) alınan rüzgar hızı verisi için MinMaxEnt dağılımlarının Weibull dağılımları ile karşılaştırılması.

V	frekans	f_i	f_w	f_{mx2}	f_{mx3}
0-1	195	0.2624	0.2609	0.2491	0.2588
1-2	231	0.3109	0.2711	0.3246	0.2983
2-3	127	0.1709	0.1958	0.1843	0.1878
3-4	74	0.0996	0.1237	0.1000	0.1103
4-5	52	0.0700	0.0716	0.0560	0.0637
5-6	30	0.0404	0.0388	0.0326	0.0364
6-7	17	0.0229	0.0200	0.0198	0.0205
7-8	8	0.0108	0.0098	0.0124	0.0114
8-9	6	0.0081	0.0046	0.0080	0.0062
9-10	1	0.0013	0.0021	0.0053	0.0033
10-11	1	0.0013	0.0009	0.0036	0.0017
11-12	1	0.0013	0.0004	0.0025	0.0009
12-13	0	0	0.0002	0.0018	0.0004
RMSE			0.0147	0.0081	0.0070
χ^2			0.0003	0.0000	0.0000
R^2			0.9792	0.9937	0.9952
AIC			115.96	102.25	111.45
BIC			117.09	103.38	113.15
K-L			0.0147	0.0081	0.0070

Şekil 5.2' den iki ve üç bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımlarının veriye iyi uyum gösterdiği görülmektedir. Özellikle üç bileşekli moment vektör fonksiyonuna uygun MinMaxEnt dağılımının veriye oldukça iyi uyduğu Şekil 5.2'de gözlenmektedir.

Şekil 5.2. (Celik 2004) alınan rüzgar hızı verisi için MinMaxEnt dağılımları ve Weibull dağılımı



6. GENELLEŞTİRİLMİŞ MİNİMUM ÇAPRAZ ENTROPİ METODU VE GELİŞTİRİLMİŞ DAĞILIMLAR

Bu bölümde Kullback-Leibler informasyon ölçüsü kullanılarak yeni bir fonksiyonel tanımlanmış (moment fonksiyonlarına bağlı Kullback-Leibler fonksiyoneli) ve onun özellikleri araştırılmıştır. Kesikli ve sürekli rassal değişkenler için tanımlanan fonksiyonel yardımıyla genelleştirilmiş Kullback Leibler optimization ilkesi önerilmiştir. Önerilen bu optimizasyon ilkesi yardımıyla sistemin gerçek dağılımına Kullback-Leibler ölçüsüne göre en yakın dağılımı belirleme süreci gerçekleştirilmiştir. Ayrıca, bu süreç yardımıyla elde edilen dağılımın içerdiği informasyon için alt ve üst sınır elde edilmiştir.

6.1. Kesikli Rassal Değişkenler için Kullback-Leibler Fonksiyonu ve Moment Fonksiyonları Yardımıyla Tanımlanmış Bir Fonksiyonel

Kullbak'in minimum çapraz entropi metodunda, bazı moment kısıtlarına göre önsel dağılım arasındaki K-L mesafesi minimize edilerek bir fonksiyon belirlenir. Burada moment koşulları önceden verilir ve buna göre önsel dağılıma en yakın dağılım belirlenir. Bu metotta önsel dağılım uniform alınır, MaxEnt metodu, MinxEnt metodunun bir özel sonucu olarak ortaya çıkar.

Bu metodun amacı, Bölüm 3. de verildiği gibi Lagrange çarpanları metodu yardımıyla bir yardımcı fonksiyonel kurularak, bu fonksiyonun minimum değerini bulmaktır. Buradan farklı bir bakış açısıyla, K-L ölçümünün minimum değerlerine karşılık gelen moment fonksiyonlarına bağlı bir fonksiyonel tanımlanabilir. Bu fonksiyonel K-L ölçüsünden farklı olarak moment fonksiyonlarına bağlıdır. Bu fonksiyonel "moment fonksiyonlarına bağlı K-L fonksiyoneli" olarak adlandırılmıştır. Dikkat edilmelidir ki sürekli sistemler için K-L ölçüsü dağılım fonksiyonlarının bir fonksiyoneli. Ancak burada tanımlanan fonksiyonel moment fonksiyonları kümesinde tanımlanmıştır.

Tanımlanan fonksiyonelin moment fonksiyonları kümesinde sürekli olduğu ve moment fonksiyonlarının kompakt bir kümesinde en büyük ve en küçük değere ulaştığı kanıtlanacaktır. Ayrıca fonksiyonelin maksimum değer

aldığı moment fonksiyonun belirlediği dağılım K-L mesafesine göre gerçek dağılıma en yakın olacaktır. Bu dağılıma MaxMinxEnt dağılımı denilecektir.

Bu süreçte dikkate edilmelidir ki: k sayıda moment koşulu verildiğinde, $(m < k)$, $\binom{k}{m}$ sayıda farklı moment seçilerek bunlara karşı gelen dağılımlar belirlenebilir. Böylece bu şekilde kurulan moment fonksiyonları test edilerek K-L ölçüsüne göre gerçek dağılıma en yakın dağılım belirlenmiş olur. Kullback-Leibler fonksiyoneli aşağıdaki gibi kurulur.

$$D(p : q) = \sum_{i=1}^n p_i \ln \frac{p_i}{q_i} \quad (6.1)$$

$$\begin{aligned} \sum_{i=1}^n p_i &= 1 \\ \sum_{i=1}^n p_i g_r(x_i) &= \mu_r \end{aligned}, \quad r = 1, \dots, m \quad (6.2)$$

(6.1) K-L fonksiyonu, (6.2) moment kısıtları altında minimize edilir. Burada $\mathbf{g}(\mathbf{x}) = (g_1(x), \dots, g_m(x))'$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)'$ verilen moment fonksiyonları ve moment değeridir.

(6.1)'e (6.2) koşulları altında minimum değer veren $\mathbf{p} = (p_1, \dots, p_n)$ dağılımını bulmak için yardımcı fonksiyon

$$\sum_{i=1}^n p_i \ln \frac{p_i}{q_i} + (\lambda_0 - 1) \left(\sum_{i=1}^n p_i - 1 \right) + \left(\sum_{j=1}^m \lambda_j \sum_{i=1}^n g_j(x_i) p_i - \mu_j \right) \quad (6.3)$$

kurularak (6.3) den p_i lere göre kısmi türev alınarak 0'a eşitlendiğinde

$$p_i = q_i \exp\left(\lambda_0 + \sum_{j=1}^m \lambda_j g_j(x_i)\right), \quad i = 1, \dots, n \quad (6.4)$$

alınır.

Böylece

$$D_{\min} = \sum_{j=1}^m \lambda_j \mu_j + \lambda_0 \quad (6.5)$$

$$D_{\min} = \lambda_0 + \sum_{j=1}^m \lambda_j \mu_j = U(\mathbf{g}) \quad (6.6)$$

(6.4)' ün K-L ölçüsüne minimum değer verdiği gösterilebilir.

Herhangi bir g moment fonksiyonuna karşın D_{\min} belli bir değere sahiptir. Bu nedenle de $U(g)$ moment fonksiyonlar kümesinde tanımlanmış bir fonksiyoneldir.

(6.4) den p dağılımının entropisi,

$$H(p) = -\sum_{i=1}^n p_i \ln p_i = -\sum_{i=1}^n p_i \left(\ln q_i + \lambda_0 + \sum_{j=1}^m \lambda_j g_j(x_i) \right) \quad (6.7)$$

$$= -\sum_{i=1}^n p_i \ln q_i - \left(\lambda_0 + \sum_{j=1}^m \lambda_j g_j(x_i) \right)$$

$$H(p) = -\sum_{i=1}^n p_i \ln q_i - U(g) \quad (6.8)$$

İnformasyon entropi değişimi olduğuna göre $\ln n$ ile (6.8) farkı informasyondur.

$$I = \ln n - H = \ln n + \sum_{i=1}^n p_i \ln q_i + U(g). \quad (6.9)$$

Tanım 6.1.

$$D_{\min} = U(g) \quad (6.10)$$

Tanımlanan $U(g)$ fonksiyonelinin özellikleri aşağıda teoremler şeklinde verilmiştir.

Teorem 6.1. (6.10) $U(g)$ fonksiyoneli moment vektör fonksiyonlarının $(C[a,b])$ 'e ait kümesi üzerinde süreklidir.

Kanıt. Her bir $g(x)$ m -boyutlu moment fonksiyonu için K-L fonksiyonu minimum değere sahiptir. Böylece her moment fonksiyonunu U 'un minimumu ile karşılaştıran fonksiyonel $U(g)$ ile tanımlansın. $g^{(1)}(x)$, $g^{(2)}(x)$ moment vektör fonksiyonları ve $\mu^{(1)}$, $\mu^{(2)}$ uygun moment vektör değerleri olsun. (6.2) den

$$\mu^{(2)} - \mu^{(1)} = \sum_{i=1}^n p_i (g^{(2)}(x_i) - g^{(1)}(x_i)), \quad (6.11)$$

diğer bir deyişle

$$\begin{aligned} \|\mu^{(2)} - \mu^{(1)}\| &\leq \sum_{i=1}^n p_i \|g^{(2)}(x_i) - g^{(1)}(x_i)\| \leq \sum_{i=1}^n p_i \|g^{(2)} - g^{(1)}\|_C \\ &= \|g^{(2)} - g^{(1)}\|_C, \\ \|\mu^{(2)} - \mu^{(1)}\| &\leq \|g^{(2)} - g^{(1)}\|_C, \end{aligned} \quad (6.12)$$

burada $\|\cdot\|$ Euclid normu ve $\|g\|_C = \max_{x \in [a,b]} \|g(x)\|$ $C[a,b]$ uzayının normudur.

(6.12)'den görüldüğü gibi, $\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0$ vardır ki

$$\|g^{(2)} - g^{(1)}\|_C < \delta, \quad (6.13)$$

$$\|\mu^{(2)} - \mu^{(1)}\| < \varepsilon, \quad (\varepsilon = \delta). \quad (6.14)$$

(6.14)'den μ moment vektör fonksiyonu $g(x)$ 'e $C[a,b]$ normuna göre bağlıdır.

(6.2) koşullarına göre D 'yi minimumunu elde etmek için Lagrange metodu uygulanırsa, aşağıdaki MaxMinxEnt dağılımı elde edilir.

$$p_i = q_i \exp(-\lambda_0 - \sum_{j=1}^m \lambda_j g_j(x_i)), \quad i = 1, \dots, n. \quad (6.15)$$

(6.2) ve (6.1) koşullarına göre D 'yi minimum yaparak elde edilen $U(g)$ fonksiyoneli aşağıdaki formdadır.

$$U(g) = \lambda_0 + \sum_{j=1}^m \lambda_j \mu_j. \quad (6.16)$$

Burada λ_0 parametresi $\lambda_1, \dots, \lambda_m$ 'lere ve $\lambda_1, \dots, \lambda_m$ parametreleri μ_1, \dots, μ_m 'lere sürekli bir şekilde bağlıdır. Böylece $U(g)$ fonksiyoneli sürekli fonksiyonlar kümesi üzerinde sürekli bir fonksiyoneldir. Böylece Teorem 6.1. kanıtlanmış olur.

Teorem 6.2 (6.10) ile verilen $U(g)$ fonksiyoneli verilen kompakt kümede en büyük ve en küçük değerine ulaşır.

Bu teoremin kanıtı lineer uzayda kompakt kümede tanımlanan sürekli fonksiyonellerin bir özelliğinden ortaya çıkmaktadır.

Teorem 6.3. $\min_i q_i = \alpha$ ve $\max_i q_i = \beta$ olsun. Bu durumda

$$\ln n\alpha + U(g) \leq I \leq \ln n\beta + U(g) \quad (6.17)$$

eşitsizliği geçerlidir.

Kamt. (6.4) MinxEnt dağılımının entropisi

$$H(p) = -\sum_{i=1}^n p_i \ln p_i = -\sum_{i=1}^n p_i \ln q_i - U(g) \quad (6.18)$$

şeklindedir.

İnformasyon $I = \ln n - H(p)$ olduğuna göre

MinxEnt dağılımının informasyonu

$$I = \ln n - H(p) = \ln n - \left\{ -\sum_{i=1}^n p_i \ln q_i - U(g) \right\} \quad (6.19)$$

$$I = \ln n + \sum_{i=1}^n p_i \ln q_i + U(g) \quad (6.20)$$

Son eşitlikte $\min_i q_i = \alpha$ ve $\max_i q_i = \beta$ dikkate alınır

$$\ln n\alpha + U(g) \leq I \leq \ln n\beta + U(g) \quad (6.21)$$

şeklinde bulunur.

6.2. Kesikli Rassal Değişkenler için MaxMinxEnt Dağılımı

(6.10) eşitliğiyle tanımlanan $U(g)$ fonksiyoneline $K \subset C[a, b]$ kompakt moment vektör fonksiyonlar kümesinde maksimum değer veren moment vektör fonksiyonu $g^{(0)}$ olsun:

$$\max_{g \in K} U(g) = U(g^{(0)}). \quad (6.22)$$

Bu $g^{(0)}$ vektör moment fonksiyonuna karşılık gelen dağılım, MaxMinxEnt dağılımı yada en küçük MaxMinxEnt dağılımı olarak adlandırılacaktır. Kısaca $g^{(0)}$ moment fonksiyonuna uygun olan ve (6.15) formülüyle belirlenen

$$p^{(0)} = (p_1^{(0)}, \dots, p_n^{(0)})$$

$$p_i = q_i \exp\left(-\lambda_0 - \sum_{j=1}^m \lambda_j g_j^{(0)}(x_i)\right), \quad i = 1, \dots, n \quad (6.23)$$

MaxMinxEnt dağılımıdır. MaxMinxEnt dağılımının bulunması sürecinde dikkat edilmesi gereken fonksiyonlar kümesi üzerinde tanımlanan MinxEnt

fonksiyonundan farklı olarak, momentler kümesinde tanımlanan bir fonksiyonelin minimumunun bu moment fonksiyonuna göre bulunabilmesidir.

MaxMinxEnt dağılımı, rassal örnekten elde edilen dağılımlar içerisinde anakütlenin gerçek dağılımına K-L ölçüsüne göre en yakın dağılımdır.

6.3. Sürekli Rassal Değişkenler için Kullback-Leibler Fonksiyonu ve Moment Fonksiyonları Yardımıyla Tanımlanmış Bir Fonksiyonel

Dağılım fonksiyonları üzerinde tanımlanmış Kullback-Leibler fonksiyoneli

$$D(f(x); q(x)) = \int_a^b f(x) \ln \frac{f(x)}{q(x)} dx, \quad (6.24)$$

gibi ifade edilir.

$$\int_a^b g(x) f(x) dx = \mu, \quad (6.25)$$

$$\int_a^b f(x) dx = 1$$

moment kısıtları altında (6.24)'e minimum değer veren fonksiyon (3.7) formülüyle verilmişti. Burada $\mathbf{g}(x) = (g_1(x), \dots, g_m(x))'$ moment fonksiyonları,

$\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)$ moment değeridir. Bu durumda D_{\min}

$$D_{\min} = \int_a^b f(x) \ln \frac{f(x)}{q(x)} dx = \int_a^b f(x) \left(\lambda_0 + \sum_{j=1}^m \lambda_j g_j(x) \right) dx \quad (6.26)$$

$$D_{\min} = \lambda_0 + \sum_{j=1}^m \lambda_j \mu_j. \quad (6.27)$$

Tanım 6.2.

$$U(g) = D_{\min} \quad (6.28)$$

(6.28) eşitliğiyle tanımlanan $U(g)$ fonksiyoneline moment vektör fonksiyonları na bağlı K-L fonksiyoneli adlandırılınsın. Moment fonksiyonlarına bağlı K-L fonksiyoneli aşağıdaki teoremlerle ifade edilen özelliklere sahiptir.

Teorem 6.4 (6.28) $U(g)$ fonksiyoneli moment vektör fonksiyonlarının kompakt kümesinde ($K \subset C[a, b]$) süreklidir.

Kanıt. (6.25) kısıtına göre

$$\mu^{(2)} - \mu^{(1)} = \int_a^b (g^{(2)}(x) - g^{(1)}(x))f(x)dx \quad (6.29)$$

buradan

$$\begin{aligned} \|\mu^{(2)} - \mu^{(1)}\| &\leq \int_a^b \|g^{(2)}(x) - g^{(1)}(x)\| f(x)dx \leq \int_a^b \|g^{(2)} - g^{(1)}\|_c f(x)dx \\ &= \|g^{(2)} - g^{(1)}\|_c \int_a^b f(x)dx = \|g^{(2)} - g^{(1)}\|_c \\ &= \|g^{(2)} - g^{(1)}\|_c, \end{aligned} \quad (6.30)$$

veya

$$\|\mu^{(2)} - \mu^{(1)}\| \leq \|g^{(2)} - g^{(1)}\|_c. \quad (6.31)$$

(6.31) den moment vektörü μ 'nün $g(x)$ 'e sürekli bir şekilde bağlı olduğu görülmektedir. Böylece $U(g)$ fonksiyoneli moment fonksiyonları kümesinde süreklidir.

Teorem 6.4 kanıtlanmış olur.

Teorem 6.5 (6.28) $U(g)$ fonksiyoneli verilen kompakt kümede en büyük ve en küçük değerine ulaşır.

6.4. Sürekli Rassal Değişkenler için MaxMinxEnt Dağılımı

(6.28) ile tanımlanmış $U(g)$ fonksiyoneline K kompakt moment vektör fonksiyonlar kümesinde maksimum değer veren moment vektör fonksiyonu $g^{(0)}$ olsun:

$$\max_{g \in K} U(g) = U(g^{(0)}). \quad (6.32)$$

Bu $g^{(0)}$ vektör moment fonksiyonuna karşılık gelen dağılım, MaxMinxEnt dağılımı yada en büyük MinxEnt dağılımı olarak adlandırılacaktır. Kısaca

MaxMinxEnt dağılımı $U(g)$ fonksiyoneline maksimum değer veren $g^{(0)}$ moment fonksiyonuna karşı gelen MinxEnt dağılımıdır. Bu süreçte dikkat edilmesi gereken fonksiyonlar kümesi üzerinde tanımlanan Kullbak-Leibler fonksiyonundan farklı olarak, momentler kümesinde tanımlanan bir fonksiyonelin minimumunun bu moment fonksiyonuna göre bulunabilmesidir.

$$f(x) = q(x) \exp\left(\lambda_0 + \sum_{j=1}^m \lambda_j g_j^{(0)}(x)\right) \quad (6.33)$$

şeklindedir. MaxMinxEnt dağılımı, rassal örnekten elde edilen dağılımlar içerisinde K-L ölçümüne göre gerçek dağılıma en yakın dağılımdır.

Bu teoremlerden sonuç olarak aşağıdaki sonuç verilebilir.

$$H(f) = -\int_a^b f(x) \ln f(x) dx \text{ entropi fonksiyonelinin}$$

$$\int_a^b f(x) dx = 1 \quad (6.34)$$

koşulu altında $H_{\max} = \ln(b - a)$ ve entropideki değişim informasiyon olduğuna göre $U(g)$ ye minimum değer veren dağılımın entropisi yani (6.4) dağılımının entropisi

$$H(f) = -\int_a^b f(x) \ln f(x) dx = -\int_a^b f(x) \left(\ln q(x) + \lambda_0 + \sum_{j=1}^m \lambda_j g_j(x) \right) dx \quad (6.35)$$

$$H = -\int_a^b f(x) \ln q(x) dx - U(g) \quad (6.36)$$

olarak elde edilir.

İnformasyon ise,

$$I = \ln(b - a) + \int_a^b f(x) \ln q(x) dx + U(g), \quad (6.37)$$

şeklinde ifade edilir.

Sonuc. Eğer q dağılımı için $0 \leq \alpha \leq q(x) \leq \beta$ dolayısıyla $\ln \alpha \leq \ln q(x) \leq \ln \beta$ ise

$$\ln(b - a)\alpha + U(g) \leq I \leq \ln(b - a)\beta + U(g) \quad (6.38)$$

geçerlidir.

Kanıt. (6.37) informasyon formülünde, (6.38) eşitsizliği dikkate alınır,

$$\ln(b-a) + \int_a^b f(x) \ln \alpha dx + U(g) \leq I \leq \ln(b-a) + \int_a^b f(x) \ln \beta dx + U(g) \quad (6.39)$$

eşitsizliği gerçekleşir. Buradan

$$\ln(b-a)\alpha + U(g) \leq I \leq \ln(b-a)\beta + U(g). \quad (6.40)$$

Bu sonuç, $\max U(g) = U(g^{(0)})$ eşitliğini sağlayan $g^{(0)}$ karşı gelen MaxMinxEnt dağılımı için alt ve üst sınırdır.

6.5. Rüzgâr Hızınının MaxMinxEnt dağılımı: Weibull dağılımına Alternatif

Bu bölümde, geliştirilen MaxMinxEnt dağılımı kullanılarak Eskişehir bölgesinde, 10m yükseklikte ölçülen rüzgâr hızının dağılımı bulunmuş ve klasik istatistiğin bir dağılım modeli olan ve rüzgâr alanında uygulanabilirliği kabul edilmiş Weibull dağılımıyla karşılaştırılmıştır. Bu uygulamada Bölüm 5.den farklı olarak önsel bir dağılımın bilinmesi de sözkonusudur. Yani aranan dağılımın öncesinde elimizde bir önsel dağılım mevcut olmalıdır. Burada ilk olarak önsel dağılımının seçimi için incelenen ayın bir önceki ayı düşünülmüş ancak mevsimselliğin olduğu düşüncesiyle bunun doğru olmayacağı kanısına varılmıştır. Çünkü aylarla birlikte değişen mevsimden kaynaklanan verilerin dağılımında oldukça farklılık görülmektedir. Bu durumda optimal seçim incelenen ayın dağılımını ararken önsel olarak bir önceki yılın aynı ayının dağılımını düşünülebilir fikri ağır basmaktadır. Bununla birlikte elimizde bunu sağlayacak 2 yıllık data mevcut değildir. Başka bir fikir rüzgâr çalışmalarında yaygın olarak kullanılan Weibull dağılımı olabileceğidir. Bu tezde ise önsel dağılımın MinMaxEnt dağılımı olacağı fikri önerilmiştir. Bu şekilde önerilen önsel dağılıma bağlı olarak MaxMinxEnt dağılımları belirlenmiştir. Bulunan sonuçlar Çizelge 6 şeklinde özetlenmiştir.

MaxMinxEnt dağılımı belirlenmesi süreci aşağıdaki gibi verilir:

1. Önsel dağılımı belirlenmesi (Burada MinMaxEnt dağılımı alınır)
2. Moment kısıtlarını belirlenmesi
3. Her moment kısıtına uygun olarak MinxEnt dağılımını belirlenmesi
4. Belirlenen MinxEnt dağılımlarının K-L ölçüsünü hesaplanması

5. En büyük K-L ölçüsünü veren moment koşulunu belirlenmesi
6. Bu seçilen moment kısıtına karşı gelen dağılıma MaxMinxEnt dağılımı denmektedir.

1. deki moment kısıtları $(x, x^2, \ln(x), \ln(x+1), (\ln(x))^2, \ln(x^2+1))$

şeklinde belirlenmiştir. 1. deki tüm moment kısıtlarına ek olarak her zaman olasılıkların toplamının 1 olduğu kısıtı vardır. Bir deki tüm fonksiyonlar, iki bileşkeli moment vektörüne uygun kısıtı belirler. 1. deki tüm moment kısıtlarının ikişerli kombinasyonları ele alınarak 2-5 adımları tekrarlanır ve böylece üç bileşkeli moment vektörüne uygun kısıt belirlenir. Tüm bu moment vektörlerine uygun MaxMinxEnt dağılımı belirlenir.

Bu şekilde belirlenen iki ve üç bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımların entropisi Çizelge 6.1 ve Çizelge 6.2 de listelenmiştir.

Çizelge 6.1. Aralık ayında ölçülen rüzgâr hızı için, iki bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımlarından hesaplanan K-L ölçümü

Moment Kısıtlar	K-L ölçüsü
(1, x)	0.0000
(1, x ²)	0.0000
(1, ln x)	3.3398e-6
(1, ln(1+x))	2.224e-6
(1, (ln x) ²)	0.0000
(1, ln(1+x ²))	2.587e-6

Çizelge 6.2. Aralık ayında ölçülen rüzgâr hızı için üç bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımlarından hesaplanan K-L ölçümü

Moment Kısıtlar	K-L ölçüsü
(1, x, x ²)	1.2467e-5
(1, x, ln x)	1.2609e-5
(1, x, ln(x+1))	1.8328e-5
(1, x, (ln x) ²)	0.0000
(1, x, ln(1+x ²))	2.55e-5
(1, x ² , ln x)	1.3339e-5
(1, x ² , ln(x+1))	1.5886e-5
(1, x ² , ln(x) ²)	1.323e-5

Moment Kısıtlar	K-L ölçüsü
$(1, x^2, \ln(1+x^2))$	1.9254e-004
$(1, \ln x, \ln(x+1))$	6.9281e-004
$(1, \ln x, (\ln x)^2)$	7.0650e-004
$(1, \ln x, \ln(1+x^2))$	6.1391e-004
$(1, \ln(1+x), (\ln x)^2)$	7.0707e-004
$(1, \ln(1+x) \& \ln(1+x^2))$	8.0013e-005
$(1, (\ln x)^2, \ln(1+x^2))$	7.2263e-004

Çizelge 6.1 ve Çizelge 6.2' den sonuç olarak iki bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımı $(1, \ln x)$ moment koşuluna uygun dağılımdır ve üç bileşkeli moment vektör fonksiyonuna uygun MinMaxEnt dağılımı $(1, (\ln x)^2, \ln(1+x^2))$ moment koşuluna karşı gelen dağılımdır denilebilir.

Bu şekilde 2005 yılındaki on iki ay için MaxMinxEnt dağılımı belirlenmiş, örnek olarak aralık ayı için frekans dağılımları belirlenerek Çizelge 6.3' deki gibi düzenlenmiştir. Bununla birlikte her ay için bu Çizelgeler oluşturulmuş, hesaplanan MaxMinxEnt dağılımları ve Weibull dağılımı Bölüm 4' te verilen kriterlere göre karşılaştırılmıştır.

Çizelge 6.3'e örnek olarak 2005 yılı Aralık ayına ait hız verileri kullanılarak elde edilen frekans dağılımları, iki bileşkeye sahip moment kısıtı altında MaxMinxEnt $(f_{kx2}(x))$ üç bileşkeye sahip moment kısıtı altında MaxMinxEnt $(f_{kx3}(x))$ ve Weibull $(f_w(x))$ dağılımlarından hesaplanan teorik frekanslar verilmiştir. Diğer aylar içinde benzer Çizelgeler oluşturulmuştur. Ayrıca Çizelge 6.3' de dağılımların veriyi modelleme derecesini istatistiksel olarak belirleyen kriterler verilmiştir. Çizelge 6.3 dikkate alındığında, tüm kriterlerin MaxMinxEnt dağılımları lehinde olduğu görülmektedir. Ayrıca iki ve üç bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımları, kriterlere göre karşılaştırıldığında sonuçların birbirine çok yakın olduğu gözlenmektedir.

Çizelge 6.3. Aralık ayında ölçülen rüzgâr hızı için MaxMinxEnt dağılımları ve Weibull dağılımından elde edilen teorik frekanslar

v	frekans	f_i	f_w	$f_{k \times 2}$	$f_{k \times 3}$
0.5	86	0.1230	0.1011	0.1231	0.1239
1.5	92	0.1316	0.1699	0.1309	0.1310
2.5	101	0.1445	0.1785	0.1422	0.1416
3.5	107	0.1531	0.1588	0.1503	0.1495
4.5	91	0.1302	0.1272	0.1471	0.1466
5.5	92	0.1316	0.0942	0.1273	0.1272
6.5	73	0.1044	0.0654	0.0927	0.0930
7.5	36	0.0515	0.0429	0.0542	0.0547
8.5	15	0.0215	0.0268	0.0242	0.0246
9.5	6	0.0086	0.0160	0.0079	0.0081
RMSE			0.0249	0.0068	0.0067
χ^2			0.0007	0.0000	0.0000
R ²			0.7523	0.9811	0.9816
AIC			55.24	4.01	5.94
BIC			55.85	4.62	6.84
K-L			0.0718	0.0030	0.0030

Ayrıca tüm 2005 yılına ait aylar için bu kriterler hesaplanarak Çizelge 6.4 şeklinde verilmiştir.

Çizelge 6.4. 2005 yılı tüm aylarında ölçülen rüzgâr hızı için MaxMinxEnt dağılımlarının ve Weibull dağılımının karşılaştırılması

	Ocak					
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0190	0.0004	0.9345	71.92	72.72	0.0231
$f_{k \times 2}$	0.0148	0.0002	0.9604	4.30	5.09	0.0215
$f_{k \times 3}$	0.0161	0.0003	0.9526	7.05	8.24	0.0203

Şubat						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0140	0.0002	0.9252	63.78	64.58	0.0435
$f_{k \times 2}$	0.0117	0.0001	0.9470	4.05	4.84	0.0193
$f_{k \times 3}$	0.0109	0.0001	0.9544	6.14	7.33	0.0185

Mart						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0178	0.0003	0.9372	80.22	81.19	0.0252
$f_{k \times 2}$	0.0149	0.0002	0.9552	4.1916	5.16	0.0176
$f_{k \times 3}$	0.0143	0.0002	0.9591	6.1487	7.60	0.0170

Nisan						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0137	0.0006	0.8908	66.11	66.91	0.0424
$f_{k \times 2}$	0.0225	0.0002	0.9593	4.23	5.03	0.0150
$f_{k \times 3}$	0.0144	0.0002	0.9550	6.51	7.70	0.0141

Mayıs						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0180	0.0003	0.9521	86.48	87.45	0.0344
$f_{k \times 2}$	0.0108	0.0001	0.9827	4.18	5.15	0.0159
$f_{k \times 3}$	0.0099	0.0001	0.9854	6.29	7.75	0.0150

Haziran						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0227	0.0006	0.9152	60.64	61.25	0.0238
$f_{k \times 2}$	0.0194	0.0004	0.9379	4.07	4.67	0.0151
$f_{k \times 3}$	0.0187	0.0005	0.9424	6.24	7.14	0.0142

Temmuz						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0185	0.0004	0.9236	67.92	68.71	0.0305
$f_{k \times 2}$	0.0185	0.0004	0.9234	4.01	4.80	0.0277
$f_{k \times 3}$	0.0181	0.0004	0.9263	6.09	7.28	0.0274

Ağustos						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0154	0.0002	0.9557	84.78	85.75	0.0245
$f_{k \times 2}$	0.0141	0.0002	0.9629	3.61	4.58	0.0165
$f_{k \times 3}$	0.0135	0.0002	0.9660	6.59	8.04	0.0159

Eylül						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0165	0.0003	0.9529	50.69	51.09	0.0246
$f_{k \times 2}$	0.0168	0.0003	0.9509	4.03	4.43	0.0178
$f_{k \times 3}$	0.0166	0.0004	0.9518	6.15	6.74	0.0175

Ekim						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0180	0.0004	0.9317	59.37	59.97	0.0298
$f_{k \times 2}$	0.0101	0.0001	0.9785	4.01	4.61	0.0077
$f_{k \times 3}$	0.0100	0.0001	0.9788	6.07	6.98	0.0076

Kasım						
	RMSE	X ²	R ²	AIC	BIC	K-L
f_w	0.0230	0.0006	0.7901	63.63	4.91	0.0808
$f_{k \times 2}$	0.0108	0.0001	0.9534	4.12	64.43	0.0113
$f_{k \times 3}$	0.0103	0.0001	0.9578	5.73	6.93	0.0110

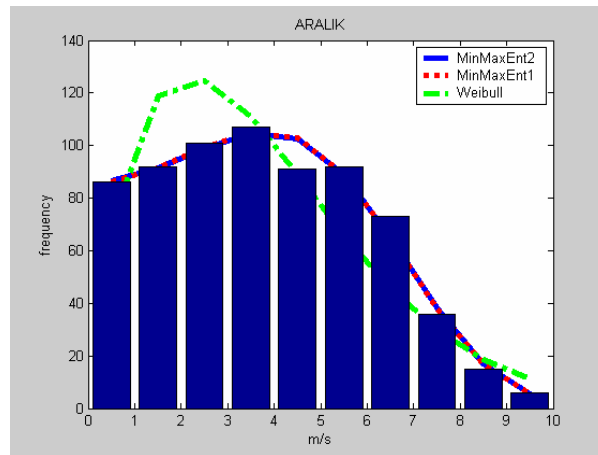
	Aralık					
	RMSE	X^2	R^2	AIC	BIC	K-L
f_w	0.0249	0.0007	0.7523	55.24	55.85	0.0718
$f_{k \times 2}$	0.0068	0.0000	0.9811	4.01	4.62	0.0030
$f_{k \times 3}$	0.0067	0.0000	0.9816	5.94	6.84	0.0030

MaxMinxEnt ve Weibull dağılımlarının istatistiksel olarak karşılaştırılması Çizelge 6.4 da verilmiştir. Bu Çizelgeye göre en büyük R^2 değeri 0.9854 ile Mayıs ayı üç bileşkeye sahip moment kısıtı altında MaxMinxEnt kullanılığında elde edilmiştir. Eylül ayı dışında, hemen hemen tüm kriterler MaxMinxEnt dağılımlarının veriye iyi uyumunun Weibull dağılımına göre daha iyi olduğunu göstermektedir. Bununla birlikte, iki ve üç bileşkeli moment vektör fonksiyonuna uygun MaxMinxEnt dağılımları kendi aralarında karşılaştırıldığında, ikisinin açıkça birbirlerine göre üstünlüğü vardır denilemez.

Kısacası tüm aylar dikkate alındığında Weibull dağılımı MaxMinxEnt dağılımları karşısında üstünlük sağlayamamıştır denilebilir. Bu Çizelgeden, dağılımların karşılaştırılması dışında, koşulların artırılması veriye uygunluğu artırır düşüncesinin doğru olmadığı da görülmektedir.

Ayrıca örnek olarak Aralık ayı rüzgâr versisine uygun MaxMinxEnt ve Weibull olasılık yoğunluk fonksiyonlarının grafiği verilmektedir.

Şekil 6.1. Aralık ayında ölçülen rüzgâr hızının MaxMinxEnt dağılımları ve Weibull dağılımı



Şekil 6.1 dikkate alındığında, MaxMinxEnt dağılımlarının veriye uyumunun oldukça iyi olduğu görülmektedir.

7. SONUÇLAR VE ÖNERİLER

Bu çalışmada elde edilen sonuçlar aşağıdaki gibi özetlenebilir.

1. Bu çalışmada, MaxEnt metodu yardımıyla moment fonksiyonlarına bağlı bir fonksiyonel tanımlanarak, bu fonksiyonele minimum değer veren ve informasyonu maksimum yapan bir dağılım belirleme süreci geliştirilmiştir. Belirlenen dağılıma MinMaxEnt dağılımı denilmiştir.

2. Söz konusu süreç kesikli ve sürekli rassal değişkenler için geçerlidir. Bu sürecin MaxEnt metodundan farkı, MaxEnt fonksiyoneli yoğunluk fonksiyonlarına bağlı iken, tanımlanan fonksiyonel moment fonksiyonlarına bağlıdır.

3. Moment fonksiyonlarına bağlı olan bu fonksiyonele minimum değer veren moment fonksiyonuna karşılık gelen dağılımın içerdiği informasyonun en büyük olduğu gösterilmiştir.

4. Kullback-Leibler (K-L) ölçüsü yardımıyla moment fonksiyonlarına bağlı yeni bir fonksiyonel tanımlanmış ve bu fonksiyonele maksimum değer veren ve K-L ölçüsüne göre ana kütlelin gerçek dağılımına en yakın dağılım bulma süreci geliştirilmiştir. Belirlenen dağılıma MaxMinxEnt dağılımı denilmiştir.

5. Söz konusu süreç kesikli ve sürekli rassal değişkenler için geçerlidir. Bu süreçte, MinxEnt fonksiyoneli yoğunluk fonksiyonları üzerinde tanımlıyken, tanımlanan fonksiyonel moment fonksiyonları üzerinde tanımlıdır

6. Uygulamada bir grup moment yardımıyla değişen bir fonksiyonel tanımlanmış ve bu momentlerden elde edilen bir dizi yardımıyla

i. maksimum informasyona sahip dağılımın belirlenmesi süreci sağlanmıştır

ii. Kullback-Leibler fonksiyoneline en büyük değer veren ve ana kütlelin gerçek dağılımına K-L ölçüsüne göre en yakın dağılım belirleme süreci sağlanmıştır.

7. Tanımlanan süreçler rüzgâr hızının dağılımını modellemek için kullanılmış, rüzgârın MinMaxEnt ve MaxMinxEnt dağılımları belirlenmiştir.

8. Önerilen bu dağılımlar, rüzgâr çalışmalarında uygulanabilirliği kabul edilmiş olan Weibull dağılımıyla karşılaştırılmıştır. Yapılan karşılaştırmalar

sonucunda rüzgar hızını modellemede MinMaxEnt ve MaxMinxEnt dağılımlarının Weibull' a göre veriye uyumunun çok daha iyi olduğu, çeşitli istatistiksel kriterler kullanılarak görülmüştür. Böylece önerilen dağılımların rüzgar hızı için Weibull'a alternatif olarak kullanılabilceğini göstermiştir. Bu anlamda burada sunulan MinMaxEnt ve MaxMinxEnt dağılımları, rüzgâr gücü ve enerjisi tahminlerinde kullanılabilir.

9. Burada geliştirilmiş süreç, genel bir süreçtir ve bu süreçler sonucu elde edilen MinMaxEnt ve MaxMinxEnt dağılımları, tüm istatistiksel dağılımlara alternatif olarak kullanılabilir.

KAYNAKLAR

- Akaike, H.(1973), "Information theory and extension of the maximum likelihood principle," *In 2nd International symposium on Information Theory* (Ed.B.Petrov and F.Csaki), 267-281, Akademiai Kiado, Budapest.
- Akaike. H.(1978), "A new look at Bayes procedure," *Biometrika*, **65**, 53-59.
- Akpınar E.K., Akpınar S. (2005), "An assessment on seasonal analysis of wind energy characteristics and wind turbine characteristics," *Energy conversion management*, **46**, 1848-1867.
- Al-Nassar W., Alhajraf S, Al-Enizi A., Al-Awadhi L. (2005), "Potential wind power generation in the State of Kuwait," *Renewable energy*, **30**, 2149-2161.
- Ali S.M. Silvey S.D. (1966), "A general class of coefficient of divergence of one distribution from another," *J.Roy .Statist. Soc.*, **28 (B)**, 131-142.
- Buchen P. W. ve Kelly M. (1996), "The maximum entropy distribution of an asset inferred from option prices," *The Financial and Quantitative Analysis*, **31 (1)**, 143-159.
- Cramer, H. *Mathematical Methods of Statistics*, Princeton University Pres, 1966
- Celik A.N. (2003), "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey," *Renewable energy*, **29**, 593-604
- Cover, T. M. *Elements of information theory*. John Wiley, New York. 1991
- Cozzolino J. M. ve Zahner M. J. (1973), "The maximum entropy distribution of the future market price of a stock," *Operations Research*, **21 (6)**, 1200-1211.
- Ebrahimi N., Pflughoeft K. ve Soofi E. S. (1994), "Two measure of sample entropy," *Statistics and Probability Letters*, **20**, 225-234.
- Ebrahimi N. (2000), "The maximum entropy method for lifetime distributions," *Sankhya: The Indian Journal of Statistics*, **62 (A)**, Pt. 2, 236-243.
- Ebrahimi N. (2001), "Families of distributions characterized by entropy", *IEEE Transactions on Information Theory*, **89 (10)**, 1368 - 1381
- Ebrahimi N., Maasoumi E. Ve Sofi (1999), "Ordering univariate distributions by entropy and variance," *Journal of Econometrics*, **90**, 317-336.

- Elsgolts, L. (1970), *Differential Equations and the calculus of variations*, Mir Publication, Moskov.
- Fisk C. ve Brown G. R. (1975), "A note on the entropy formulation of distribution models," *Operations Research Quarterly*, **26 (4)**, 755-758.
- Freund D. ve Saxena U. (1984), "An algorithm for a class of discrete maximum entropy problems," *Operations Research*, **32 (1)**, 210-215.
- Havrda H., Chavrat F. (1967), "Quantifications methods of classification Processes: Concepts of Structural α entropy," *Kybernetika*, **3**, 30-35.
- Hurvich C.M. and Tsai C.L. (1990), "The impact of model selection on inference in linear regression," *Am Statistician*, **44**, 214–217.
- Jaynes E.T.(1957), "Information theory and statistical mechanics," *Phys. Rev.*, **106**, 620-630.
- Jon L. (1998), "Constrained maximum-entropy sampling," *Operations Research*, **46 (5)**, 655-664.
- Kapur, J.N. and Kesevan, H.K.(1992), *Entropy Optimization Principle with Applications*, Academic Press.
- Kantoroviç, L. V., Akilov, G. P.(1982), *Functional Analysis*, Pergamon Press
- Kapur J.N.(1967), "Generalized entropy of order α and type β ," *The mathematical seminar*, **4**, 78-94.
- Kenneth B.(2002), *Model selection and Multi-Model Inference*, Springer-Verlag.
- Kolmogorov A. N. Ve Fomin S. V. (1972). *Elements of Theory of Functions and Functional Analysis*, 469-472, Moskow
- Kullback, S. (1959), *Information Theory and Statistics*. Wiley.
- Kullback, S. and Leibler, R. A. (1951), "On Information and Sufficiency", *Annals of Mathematical Statistics*, **22**, 79–86.
- Kumar U., Kumar V., Kapur J.N (1986), "Some normalized measures of directed divergence," *Int.J.Gen.Syst*, **13**, 5-16.
- Li M., Li X. (2005), "MEP-type distribution function: a beter alternative to Weibull function for wind speed distribution", *Renewable energy*, **30(8)**, 1221-1240.

- Li M., Li X. (2005). Investigation of wind characteristics and assessment of wind energy potential for Waterloo region, Canada, *Energy conversion management*, **46**, 3014-3033.
- Lind N.C. (1997). "Three information theoretical methods to estimates a random variable," *Journal of Environmental Management*, **49**, 43-51.
- Lind N.C., Hong H.P. ve Solona V. (1989) "A cross entropy method for Flood Frequency Analysis," *Stochastic Hydrology and Hydraulics*, **3**, 191-202.
- Nailong W. (1977), *The Maximum Entropy Method*, Springer-Verlag, Berlin.
- Niven R. K. (2003), "The constrained entropy and cross-entropy functions," **334 (A)**, 444-458.
- Pandey M.D. (2000), "Direct estimation of quantile functions using the maximum entropy principle," *Structural Safety*, **22 (1)**, 61-79.
- Pandey M.D. (2001), "Extreme quantile estimation using order statistics with minimum cross-entropy principle," *Probabilistic Engineering mechanics*, **16 (1)**, 31-42 .
- Pandey M.D. (2001), "Minimum cross-entropy method for extreme value estimation using peaks-over-threshold data," *Structural Safety* , **23 (4)** , 345-363.
- Park S. ve Park D. (2001), "Correcting moment for goodness of fit tests based on two entropy estimates," *Journal of statistical computations and simulations*, **73 (9)**, 682-694.
- Ramirez P. Ve Carta J.A. (2005), "The use of wind probability distributions derived from the maximum entropy principle in the analysis of wind energy. A case study," *Energy and Conversion Management,(basimda)*.
- Ranneby B.(1984), "The maximum spacing method: an estimated method related to the maximum likelihood method." *Scand.J.Statists*, **11** ,93-112.
- Rengi A. (1961), "On measures of entropy and information," *Proc. 4. Berkeley Syms.Math. Stat.Prob.*, **1**, 547-561.
- Papuluos A. (1991), *Probability, Random variables, and stochastic process*, McGraw-Hill Book.
- Sakamoto, Yosiyuki, Ishiguro M., and Kitagawa G. (1986), *Akaike Information Criterion Statistics*. Tokyo: KTK Scientific Publishers.

- Schwarz, G. (1978), "Estimating the Dimension of a Model", *Annals of Statistic*, **6**, 461-464.
- Shibata R. (1997), "Bootstrap estimates of Kullback-Leibler information for model selection," **7**, 375-394.
- Shamilov A. (2004), "On a new concept of generalized probability density function to systems of random variables," *The Scientific and Pedagogical News of Odlar Yurdu University*, **12**, 3-12.
- Shamilov A., Yuzer A.F., Agaoglu A., Mert Y. (2005), "Method of Obtaining Distributions of Transformed Random Variables by Using the Heaviside and the Dirac Generalized Functions," *Journal of Statistical Research (in printed)*
- Shamilov A., Kantar Mert Y. (2005), "Some Applications of Generalized Functions in the Theory of Distributions", *The Scientific and Pedagogical News of Odlar Yurdu University*, **14**, 68-75.
- Shamilov A. (2005), "Comparison of classical estimate methods based on maximum entropy" *Ordered Statistical Data: Approximations, Bounds and Characterizations*, Izmir University of Economics,
- Shamilov A. ve Kantar Mert Y. (2005), "On a Distribution Minimizing Maximum Entropy", *Ordered Statistical Data: Approximations, Bounds and Characterizations*, Izmir University of Economics,
- Shamilov A. (2006), "An Approach to Generalize Entropy Optimization Principles and Related Aspects of Newton's Method," *The 9th WSEAS International Conference on Applied mathematics (MATH 06)* (Kabul edildi, basımda)
- Shamilov A., Kantar Mert Y., Usta İ. (2006), "On a Functional defined by means of Kullback-Leibler Measure and Its Statistical Applications," *The 9th WSEAS International Conference on Applied mathematics (MATH 06)* (Kabul edildi, basımda)
- Shamilov A., Usta İ., Kantar Mert Y. (2006), "The Distribution of Minimizing Maximum Entropy: Alternative to Weibull distribution for wind speed," *The 9th WSEAS International Conference on Applied mathematics (MATH 06)* (Kabul edildi, basımda)

- Shamilov A., Yolacan S. (2005), "Various Binary Codes For Probability Distribution of Turkish Letters", *Ordered Statistical Data: Approximations, Bounds and Characterizations*, Izmir University of Economics
- Shannon C. E. (1948). "Mathematical theory of communication," **27**, 379-423, 623-659, *Bell System Tech. J.*
- Shen E. Z. Ve Perloff J. M. (2001). "Maximum entropy and Bayesian approaches to ratio problem," *Journal of Econometrics*, **104**, 289-313.
- Sugiura N.(1978). Further analysis of the data by Akaike's information criterion and the finite corrections," *Communications in Statistics*, **A7**, 13-16
- Spivak, M.(1965). *Calculus on manifolds : a modern approach to classical theorems of advanced calculus*, Anadolu Üniversitesi yayın
- Takeuchi K. (1976), "Distribution of informational statistics and a criterion of model fitting," *Suri-Kagaku (mathematical Science)* **153**, 12-18 [In japanese].
- Ulgen K, Hepbasli, A. (2002), "Determination of Weibull parameters for wind energy analysis of Izmir, Turkey," *Int J Energy Res*, **26**, 495-506
- Wagner, U.ve Geyer A. L. J.(1995), "A maximum entropy method for inverting Laplace transforms of probability density functions," *Biometrika*, **82 (4)**, 887-92.
- Ximing, W. (2003). "Calculation of maximum entropy densities with application to income distribution," *Journal of Econometrics*, **115**, 347-354.
- Zellner A. (1997). "The Bayesian method of moment (BMOM): theory and applications," *Advanges in Econometrics*, **12**, 85-105.
- Zellner A., Tobias J. (2001). "Further result on Bayesian methods of moments analysis of the multiple regression model," *International Economic Review* **42 (1)**, 121-139.
- Anadolu Üniversitesi (2005)-Anadolu Meydan Meteoroloji Müdürlüğü