

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

ENTROPİ KAVRAMININ İSTATİSTİKTEKİ BAZI UYGULAMALARI

GÖKHAN DİNÇER

**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANABİLİM DALI
İSTATİSTİK PROGRAMI**

**DANIŞMAN
DOÇ. DR. ATIF AHMET EVREN**

İSTANBUL, 2015

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ENTROPİ KAVRAMININ İSTATİSTİKTEKİ BAZI UYGULAMALARI

Gökhan DİNÇER tarafından hazırlanan tez çalışması 13.01.2015 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Atıf Ahmet Evren
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Atıf Ahmet Evren
Yıldız Teknik Üniversitesi

Prof. Dr. Ali Hakan BÜYÜKLÜ
Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Yavuz ILGAZ
İstanbul Üniversitesi

ÖNSÖZ

Entropi Kavramının İstatistik'teki Bazı Uygulamaları" başlıklı yüksek lisans tez çalışmamın oluşmasında bana destek olan ve her türlü sorunlarımla ilgilenip çözüm yolları arayan değerli tez danışmanım Doç. Dr. Atıf Ahmet Evren 'a teşekkür ederim.

Yüksek lisans derslerime ve tez çalışmama vakit ayırabileceğim bir çalışma ortamı sağlanmasında bana yardımcı olan ve karşılaştığım sorunlarla ilgili sürekli görüşlerinden yararlandığım İstatistik Bölüm Başkanı Prof. Dr. Ali Hakan BÜYÜKLÜ, teşekkürü bir borç bilirim.

Bu çalışmayı yaparken her zaman yanımda olan ve verileri ile sürekli beni destekleyen ALİZE Pazarlama San. Tic. Ltd Şirketi yetkililerine teşekkürlerimi sunarım.

Her zaman yanımda olan, maddi ve manevi desteklerini hiçbir zaman benden esirgemeyen bu günlere ulaşmamı sağlayan aileme de sabırları ve sonsuz güvenleri için ayrıca teşekkür ederim.

Ocak, 2015

Gökhan DİNÇER

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ.....	ix
ÇİZELGE LİSTESİ	x
ÖZET	xii
ABSTRACT	xiv
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti	1
1.2 Tezin Amacı	4
1.3 Hipotez	5
BÖLÜM 2	
ENTROPİ KAVRAMININ İNCELENMESİ	6
2.1 Bilgi Teorisi ve Entropi	6
2.2 Sosyal Entropi.....	8
2.2.1 İşletmelerde Sosyal Entropi	10
2.3 Entropi Türleri	12
2.3.1 Shannon Entropisi.....	12
2.3.1.1 Shannon Entropisinin Dağılımı	15
2.3.1.2 Shannon Entropi Birimi	16
2.3.1.3 Shannon Entropisinin Özellikleri	18
2.3.1.4 Shannon Entropi Fonksiyonunun Ekstremumu	19
2.3.1.5 Shannon Entropisi için Çok Değişkenli Genellemeler	20
2.3.1.6 Shannon Bileşik Entropi	21
2.3.1.7 Shannon Koşullu Entropi	22
2.3.2 Renyi Entropisi	26

2.3.2.1	Renyi Entropisinin Özellikleri	26
2.3.2.2	Renyi Sapması	26
2.3.3	Tsallis Entropisi	27
2.3.4	Bazı Olasılık Dağılımlarının Shannon Entropisi	27
BÖLÜM 3		
SHANNON ENTROPİSİ VE İSTATİSTİKSEL BAĞIMSIZLIK		29
3.1	Karşılıklı Bilgi Ölçüsü	29
3.2	İki Değişkenli Normal Dağılımda Karşılıklı Bilgi Ölçüsü	31
3.3	Karşılıklı Bilgi Miktarı için Varyasyonlar	31
3.4	Çok Değişkenli Hallere Genelleme	32
3.5	Karşılıklı Bilgi Uygulamaları	32
3.5.1	Nominal Veriye Uygulama	33
3.5.2	Metrik Veriye Uygulama	39
3.5.3	Nominal ve Metrik Verilerin Uygulama Sonuçları	46
BÖLÜM 4		
KULLBACK LEİBLER UZAKLIĞI		48
4.1	Metrik Tanımı ve Kullback Leibler Uzaklığının Metrik Fonksiyon Olup Olmadığının İncelenmesi	50
4.2	Görelî Entropi ve Kullback-Leibler Mesafesi Arasındaki İlişki	51
4.3	Kullback-Leibler Sapmasının Simetrik Bir Versiyonu: Jeffrey Sapması	52
4.4	Seçilmiş Bazı Dağılımlar için Kullback Leibler ve Jeffrey Sapmaları	52
BÖLÜM 5		
NİTEL DEĞİŞKENLİK ÖLÇÜLERİ		57
5.1	Nitel Değişkenlik Ölçülerine Giriş	57
5.2	Değişkenlik Oranı (DO)	57
5.3	Grup Sıklıklarının Varyansına Dayalı İndeks (GSVDI)	58
5.4	Nitel Değişkenlik İndeksi (NDI)	60
5.5	Benzemezlik Katsayısı	60
5.6	Simpson Çeşitlilik İndeksi (SCI)	61
5.6.1	Genişletilmiş Binom Dağılımı ile İlişkisi	62
5.7	Nitel Değişkenlik Ölçüleri ile İlgili Uygulama	62
5.7.1	Sonuçlar	65
BÖLÜM 6		
NİTEL DAĞILIMLARDA BİRLİKTELİK ÖLÇÜLERİ		66
6.1	Bazı Düzenlemeler	68
6.2	Tschuprow'un Kontenjans Katsayısı	68
6.3	Cramér'in Kontenjans Katsayısı	69
6.4	Uygulama	69
BÖLÜM 7		

SONUÇ VE ÖNERİLER	72
KAYNAKLAR.....	74
EK-A	
ENTROPİ DE R PROGRAMI	77
R PROGRAMINDA SHANNON ENTROPİSİNİ HESAPLAYAN PROGRAM KODLARI	77
EK-B	
ENTROPİ DE R PROGRAMI	79
R PROGRAMINDA KARŞILIKLI BİLGİ KODLARI VE ENTROPİDEKİ BAYES TAHMİNCİLERİNİ HESAPLAYAN PROGRAM KODLARI	79
ÖZGEÇMİŞ.....	82

SİMGE LİSTESİ

C_{XY}	Normalleştirme
D_{KL}	Kullback Leibler Uzaklığı
D_J	Jeffrey Sapması
df	Bağımsızlık derecesi
e_i	Beklenen frekans
E_n	Örnek Uzay
f_i	Gözlenen frekans
f_m	Maksimum frekans
F_m	Örnek Olay
H	Shannon Entropisi
$H(X,Y)$	Bileşke sistemin Shannon entropisi
$H(X Y)$	Koşullu Shannon entropisi
$H_\alpha(x)$	Tsallis Entropisi
$H_R(x)$	Renyi Entropisi
H'	Shannon indeksinin kestiricisi
$I(X)$	Bilgi içeriği
K	Kategori sayısı
k	Boltzman Sabiti
L	Lagrange Fonksiyonu
N	Toplam birim sayısı
p	Pearson istatistiği
$\hat{p}$	Pearson en çok olabilirlik tahmincisi
p^*	Sakoda istatistiği
S	Entropi
t	Test istatistiği
T	Tschuprow'un kontenjans katsayısı
u	Benzemezlik katsayısı
v	Cramer kontenjans katsayısı
ϕ^2	Fi-kare istatistiği
$\Psi(k)$	Digama fonksiyonu
χ^2	Ki-kare dağılımı
W	Termodinamik olasılık

KISALTMA LİSTESİ

AEP	Asimptotik Eş Bölmeleme Teoremi
DO	Değişkenlik Oranı
GSDVI	Grup Sıklıklarının Varyansına Dayalı İndeks
GYI	Gini Yoğunlaşma İndeksi
MSI	Moddan Sapma İndeksi
NDI	Nitel Değişkenlik İndeksi
SCI	Simpson Çeşitlilik İndeksi

ŞEKİL LİSTESİ

	Sayfa
Şekil 3. 1 Entropi ve Bilgi Kuramı.....	30
Şekil 3. 2 $H(x)$ için Renyi Entropilerinin farklı parametrelerdeki değişim grafiği	41
Şekil 3. 3 $H(x)$ için Tsallis Entropilerinin farklı parametrelerdeki değişim grafiği	41
Şekil 3. 4 $H(y)$ için Renyi Entropilerinin farklı parametrelerdeki değişim grafiği	43
Şekil 3. 5 $H(y)$ için Tsallis Entropilerinin farklı parametrelerdeki değişim grafiği.....	43
Şekil 3. 6 $I(X,Y)$ için Renyi Entropilerinin farklı parametrelerdeki değişim grafiği.....	45
Şekil 3. 7 $I(X,Y)$ için Tsallis Entropilerinin farklı parametrelerdeki değişim grafiği.....	45
Şekil 3. 8 $H(X/Y)$ için Renyi Entropilerinin farklı parametrelerdeki değişim grafiği.....	46
Şekil 3. 9 $H(X/Y)$ için Tsallis Entropilerinin farklı parametrelerdeki değişim grafiği....	46

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2. 1 Bazı olasılık dağılımlarının entropileri	28
Çizelge 3. 1 Nisan-Ekim ayları arası satılan ürün miktarı(kg).....	33
Çizelge 3. 2 Satılan Ürünlerin Farklı Türlerdeki Dağılımı	34
Çizelge 3. 3 Farklı α değerleri için Renyi Entropileri.....	35
Çizelge 3. 4 Farklı α değerleri için Koşullu Shannon Entropileri.....	36
Çizelge 3. 5 Farklı α değerleri için Koşullu Renyi Entropileri.....	36
Çizelge 3. 6 Farklı α değerleri için Tsallis Entropileri.....	37
Çizelge 3. 7 Farklı α değerleri için Koşullu Tsallis Entropileri	37
Çizelge 3. 8 Renyi ve Tsallis entropilerine dayalı istatistikler	38
Çizelge 3. 9 Renyi ve Tsallis entropilerine dayalı istatistikler	38
Çizelge 3. 10 X ve Y değişkenlerinin yüksek ve pozitif korelasyonlu olasılık dağılımı ...	39
Çizelge 3. 11 Shannon entropisi ile hesaplanan pozitif korelasyonlu bilgi ölçüleri.....	40
Çizelge 3. 12 Renyi ve Tsallis Entropilerinin farklı parametrelerdeki bilgi istatistikleri	40
Çizelge 3. 13 X ve Y metrik değişkenlerinin korelasyonsuz olasılık dağılımı	41
Çizelge 3. 14 X ve Y metrik değişkenlerinin Shannon entropisi ile hesaplanan korelasyonsuz bilgi ölçüleri	42
Çizelge 3. 15 X ve Y metrik değişkenlerinin Renyi ve Tsallis Entropilerinin farklı parametreler altında korelasyonsuz karşılıklı bilgi istatistikleri	42
Çizelge 3. 16 X ve Y değişkenlerinin yüksekve negatif korelasyonlu olasılık dağılımı ..	44
Çizelge 3. 17 X ve Y metrik değişkenlerinin Shannon entropisi ile hesaplanan yüksek fakat negatif korelasyonlu bilgi ölçüleri	44
Çizelge 3. 18 X ve Y metrik değişkenlerinin Renyi ve Tsallis Entropilerinin farklı parametreler altında korelasyonsuz karşılıklı bilgi istatistikleri	44
Çizelge 5. 1 X metrik değişkeninin maksimum entropili olasılık dağılımı.....	63
Çizelge 5. 2 X metrik değişkeninin Nitel Değişkenlik Ölçüleri	63
Çizelge 5. 3 X metrik değişkeninin Shannon, Renyi, Tsallis Entropileri	63
Çizelge 5. 4 X metrik değişkeninin minimum entropili olasılık dağılımı	63
Çizelge 5. 5 X metrik değişkeninin Nitel Değişkenlik Ölçüleri	64
Çizelge 5. 6 X metrik değişkeninin Shannon, Renyi, Tsallis Entropileri	64
Çizelge 5. 7 X metrik değişkeninin ortalama entropili olasılık dağılımı	64
Çizelge 5. 8 X metrik değişkeninin Nitel Değişkenlik Ölçüleri	64
Çizelge 5. 9 X metrik değişkeninin Shannon, Renyi, Tsallis Entropileri	65
Çizelge 6. 1 El örgü ipliği fabrikasının ürünlerin türlerinden sattığı miktarlar(ton)....	70
Çizelge 6. 2 Ki-Kareye Dayalı Nitel Birliklik İstatistikleri	70

Çizelge 6.3	Entropiye Dayalı Nitel Birliktelik İstatistikleri.....	71
-------------	--	----

ENTROPİNİN İSTATİSTİKTEKİ BAZI UYGULAMALARI

Gökhan DİNÇER

İstatistik Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Doç. Dr. Atıf Ahmet EVREN

İstatistiksel entropi, bir istatistiksel deneydeki belirsizliğin ölçüsüdür. Deneyin sonucu ne ölçüde öngörülebilir ise entropi o kadar düşük, deney sonucunun öngörülebilirliği ne ölçüde yüksekse entropi o kadar yüksek olmaktadır. Bunun dışında istatistiksel entropinin bazı ilginç uygulama alanları vardır.

İlk olarak, olasılık (ya da frekans) dağılımının nitel olması durumunda dağılma ölçüsü için istatistiksel entropiden yararlanılabilir. Bilindiği gibi nitel dağılımlarda merkezi toplanma ölçüsü olarak aritmetik ortalama kullanılamadığından, dağılma ölçüsü olarak varyans, standart sapma gibi istatistikler de hesaplanamamaktadır. Buna ek olarak değişkenler nitel olduğu zaman, klasik korelasyon ölçüleri de kullanılamamaktadır. Bu nedenle entropi ve karşılıklı bilgi ölçülerinin; değişkenler arasındaki bağımlılıkların analizinde kullanılması yerinde olacaktır.

İkinci olarak entropi kavramından türeyen görelî entropi, olabilirlik fonksiyonu ile ilgilidir. Genel olarak görelî entropi (Kullback-Leibler sapması) ve görelî entropiden türeyen bazı uyum iyiliğini ölçmeye yarayan istatistikler ki-kare dağılımları ile yakından ilişkilidir.

Tezde bu konular detaylıca tartışıldıktan sonra yapılacak bazı uygulamalarla entropi ve entropiden türeyen bazı istatistiklerin yukarıda sözü edilen temel istatistik problemlerine uygulanabilirliği tartışılmaktadır.

Anahtar Kelimeler: Shannon entropi, Rényi entropi, Tsallis entropi, Havrda-Charvat entropi, Kullback-Leibler mesafesi, Nitel Değişim, Karşılıklı Bilgi Ölçüsü

SOME APPLICATIONS IN STATISTICS ENTROPY

Gökhan DİNÇER

Department of Statistics

MSc. Thesis

Adviser: Assoc. Prof. Dr. Atif Ahmet EVREN

Statistical entropy is a measure of uncertainty that a statistical experiment possesses. The more predictable the outcome of an experiment, the lower the entropy; and the less predictable the outcome, the higher the entropy. In addition there are some interesting application areas for statistical entropy.

First of all, whenever the probability (or frequency) distribution is qualitative, entropy can be used as a measure of variation. In these cases, as known well, the statistics like variance and standard deviation can't be computed since arithmetic mean cannot be obtained as a measure of central tendency for qualitative distributions. Besides, whenever two variables are qualitative, classic correlation measures cannot be used. For this reason the application of measures of entropy and mutual information to the study of association between variables will be useful.

Entropy and relative entropy are closely related to likelihood function. In general, relative entropy (Kullback-Leibler divergence) is closely associated with some goodness of fit statistics like chi-square statistics.

In this thesis, after these issues are discussed in detail, by some applications, applicability of some of the measures based on entropy and relative entropy to basic problems of statistics mentioned previously is studied.

Keywords: Shannon entropy, Rényi entropy, Tsallis entropy, Havrda-Charvat entropy, Kullback-Leibler divergence, Qualitative variation, Mutual information

1.1 Literatür Özeti

Entropi kavramı bilgi kuramı ve belirsizlikle kuramı ile birlikte kullanılan bir kavramdır ve bir sistemdeki belirsizlik miktarını verir. Sistem düzensizliği arttıkça artan herhangi bir fonksiyon entropi fonksiyonu olabilir.

İstatistiksel entropi, bir istatistik deneyin belirsizliğinin ölçüsüdür. Deney sonucu ne ölçüde öngörülemez ise entropi o denli büyük olmaktadır. Örneğin hilesiz bir madeni paranın atılması deneyinde her iki sonucun ortaya çıkma olasılığı birbirine eşit olduğu için maksimum belirsizlik, maksimum entropi söz konusudur. Örneğin; tura gelme olasılığının yazı gelme olasılığından bariz bir şekilde büyük olduğu bir paranın atılmasından önceki (deneyin sonucunu tahmin etmeye yönelik) belirsizlik daha azdır. Çünkü büyük olasılıkla tura gelecektir. Bu anlamda bir kişinin böyle bir deneyi gerçekleştirmekle elde edeceği bilgi az olacaktır. Bu anlamda entropi ile (deneyden elde edilecek) bilgi miktarı birbirleri ile ilişkili iki kavram olmaktadır.

Entropiden türetilen bazı ölçüler nitel değişkenler arasındaki niteliksel varyasyonu ve ilişkiyi ölçmek için kullanılır. Ayrıca rölatif entropi, genel olarak uyum iyiliği testlerini ya da herhangi bir istatistiksel modelin yeterliliğini kontrol etmede kullanılır.

İstatistik mekaniğin ana sonuçlarından birisi, kapalı sistemlerin düzensizliğe eğilim gösterdiği ve entropinin de bu düzensizliğin bir ölçüsü olduğudur. Yani evrende kendisini dışarı kapatan, enerji ve madde kayıplarını dışarıdan yeni madde ve enerji girdileriyle telafi etmeyen bütün sistemler bu anlamda bir düzensizliğe, dağılmaya ve

ölüme doğru gitme eğilimindedir. Mesela günümüzde teknolojiyi kullanmayan bir fabrika buna bir örnektir.

Entropi, Fiziko-kimyanın önemli yasalarından olan Termodinamik yasalardan ikincisidir. Genel olarak termodinamik yasalar, bir sistemin enerjiyi yararlı işe dönüştürülmesi süreçleriyle ilgilidir. Bir sistem içerisinde yer alan, enerjileri farklı ya da benzer olabilen iyon, molekül ya da atom gibi taneciklerin etkileşimlerini (mekanik ve ısı olayları arasındaki bağlantıyı) inceleyen bir bilim dalıdır.

Entropi yasası, enerjinin bir halden başka bir hale dönüştürülmesinin belli bir maliyetinin olduğu ve bu maliyetin, "gelecekte aynı türden bir işin yürütülmesi için elde edilebilir enerji miktarındaki azalma" biçiminde ortaya çıktığını söylemektedir[1]. Yani ilave bir enerji olmaksızın aynı enerji kullanılarak aynı fayda elde edilemez. Enerjideki bu kayıplar, entropi kavramıyla genelleştirilmektedir. Dolayısıyla entropi, bir sistemde yeniden kullanılmayacak şekilde bir enerji kaybının ortaya çıkması durumunu ifade etmektedir. Diğer bir anlatımla entropi, "artık işe dönüştürülemeyen enerji miktarının ölçümüdür" [1].

Rudolf Clausius termodinamik biliminin başlıca kurucu kabul edilen Alman fizikçi ve matematikçi Carnot çevrimi olarak bilinen ilkesini ısı teorisi olarak yeniden şekillenmiştir. 1850 yılında yayımlanan en önemli makalesi Mekanik ısı teorisi üzerine termodinamiğin ikinci yasasındaki temel fikirlere açıklamalar getirmektedir. 1865 yılında bilimi entropi ile tanıştırdı. 1865 yılında termodinamiğin birinci ve ikinci yasasını aşağıdaki gibi özetlemiştir:

Evrenin enerjisi serbesttir.(Enerji yoktan var edilemez var olan enerjide yok edilemez. Sadece bir türden diğer bir türe dönüşebilir.)

Evrenin entropisi maksimuma ulaşma eğilimindedir.

Termodinamik'teki entropinin olasılığa dayalı anlamını ilk vurgulayan bilim adamlarından biri L. Boltzman'dır. Boltzman bir sistemin sahip olduğu belirsizlik ile entropisi arasındaki paralelliği vurgulaması bakımından İstatistiksel Bilgi (İnformasyon) Kuramının öncüleri arasında sayılabilir. Bir gazın sahip olduğu serbestlik derecesi gazın belirsizliği ile (dolayısıyla sistemi oluşturan parçacıkların belirsizliği ile) ilgilidir [2].

Bu bakış açısıyla Ludwig Boltzman İstatistiksel fizikte entropiyi, olasılık kavramını da gündeme getirerek doğal ve teknik süreçlerde geçerli olan termodinamik düzensizlik denklemini ortaya koymuş, ünlü Boltzmann bağıntısını $S=k \ln W$ olarak elde etmiştir. Burada $k=1.3806505 \times 10^{-23} \text{ J } \text{ K}^{-1}$ Boltzmann sabiti'dir. Sözü edilen mikroskopik durumların tanımı ve sayılması ise, sistemi oluşturan atomları tarif eden temel mekanik yasalar kullanılarak yapılır.

1948 yılı entropi alanındaki çalışmaların başlangıç yılı olarak kabul edilir. İnmasyon teorisinde Shannon'un, bir bilgi kaynağının belirsizliğinin ölçüm niceliğini keşfetmesinin ardından matematiksel formlarının ve fiziksel değerlerinin benzerliğinden dolayı bu niceliği de istatistiksel mekanik ve termodinamikteki gibi entropi olarak adlandırılmıştır. Shannon entropisi pratikte en sık kullanılan belirsizlik ölçüsüdür.

Alfred Renyi 1961 yılında Renyi entropisini tanımlamıştır. Rényi entropisi de biyolojik çeşitliliğin derecesinin ortaya konmasında kullanılmaktadır. Ayrıca Kuantum bilgi ölçüsü olarak kullanımı da önemlidir. Renyi entropisi Shannon entropisinin $\alpha \rightarrow 1$ yaklaşırken özel bir halidir. Yani Renyi entropileri özel bir durum olarak Shannon entropilerini içerir.

Fizikte Tsallis entropisi Boltzman ve Gibbs entropilerinin genelleştirilmiş halidir. Constantino Tsallis tarafından standart istatistiksel mekaniği genelleme için bir temel olarak 1988 yılında tanıtıldı. Tsallis entropisi genel bir olgu olarak Simpson çeşitlilik indeksinin özel konumunu doğrulamıştır.

Entropi, aynı zamanda bir sistemdeki "düzensizlik" in bir ölçüsüdür". Çünkü "sistemin düzensizliğini artıran her olay entropisini de artırmaktadır". Düzensizlik, bir sistem içerisinde öğeler arasındaki enerji akışını etkileyen bir faktördür. Bir sistemin öğeleri arasındaki etkileşim ya da bağlanma derecesiyle ters orantılıdır. Sistemi oluşturan öğeler arasındaki bağlanmalar zayıfladıkça entropi artmaktadır. Bu bakımdan "maksimum entropi en düzensiz durumdur"[1]. Bu anlamda düzensizlik, öğeler arasındaki muhtemel bağlanma derecelerini zayıflatarak bir sistem oluşmasını engelleyecek düzeyde muhtemel ilişki sayısındaki artışlardan kaynaklanan belirsizliktir. Bir sistemde entropinin artışıyla sistemdeki düzensizliğin artacağı (organizasyonun

kaybolacağı) ve tüm doğal ve fiziksel süreçlerde tersinmezlik yüzünden düzensizlik ve karmaşanın olasılığının hep en yüksek olduğu söylenebilir.

Bir termodinamik sistemin öğeleri arasında enerji düzeyleri bakımından farkın olmaması durumuna denge hali denmektedir. Denge hali ,“reaksiyonun(etkileşimin) hiç olmadığı halidir” [3]. Denge hali, bir sistemin yeniden aynı tür bir işi yapamayacak duruma gelmesi halidir. Çünkü enerji düzeyleri eşit olduğundan enerji akışı olmaz. Bu durumda sistemin yararlı bir iş üretebilmesi için ilave bir enerjiye gereksinimi ortaya çıkmaktadır. O halde ilave enerjinin olmadığı kapalı bir sistemde entropi, bir süre sonra maksimuma varmaktadır. Diğer bir ifadeyle, ”denge konumunda sistemi oluşturan öğelerin iç enerjisi azalırken entropisi artmaktadır”. Denge halindeki bir sistemde maksimum entropi vardır. Yani iki konum arasında enerji akışı yoktur ve dolayısıyla bir iş ortaya çıkmaz. Bu durum, ”elde edilemez veya bağlı enerji miktarının maksimuma ulaştığı anlamına gelmektedir. Yani kapalı sistemin iç enerjisi artmıştır.

1.2 Tezin Amacı

Entropi fizikten aldığımız bir kavramdır. Fiziko-kimyanın önemli yasalarından olan Termodinamik yasalardan ikincisidir. Yıllar içinde matematik biliminde kullanılarak entropi de gelişmeler elde edilmiştir. İstatistikte entropinin kullanımı belirsizliğin ölçüsü olarak Literatüre giriş yapmıştır. İstatistik biliminde entropi kullanılarak birçok çalışma yapılmıştır.

Bu çalışmada entropinin istatistikteki kullanım alanlarına değinilmiş ve bunlarla ilgili uygulamalarla bu kullanım alanları açıklanmıştır. Bu çalışmada genel olarak entropinin dört kullanım alanına değinilmiştir. Bunlar şunlardır:

- Nitel dağılımlar için değişkenliğin (variation) ölçülmesi
- Nitel değişkenler arasında birliktelik (association) dercesinin belirlenmesi
- Uyum iyiliğinin belirlenmesi
- Hipotez testleri.

1.3 Hipotez

Son yıllarda istatistik literatüründe entropi ve entropiye dayalı istatistikler üzerine pek çok çalışma yapıldığı görülmektedir. Özellikle bu istatistiklerin asimptotik (büyük örneklem) özelliklerinin çalışılmış olması istatistikçilere bazı problemlerin çözümünde katkı sağlamaktadır. Genel olarak entropi ve entropiye dayalı istatistikler nitel değişkenliğin ölçülmesinde, değişkenler arasındaki birliktelik (association) derecesinin belirlenmesinde, uyum iyiliğinin (goodness of fit) saptanmasında kullanılmaktadır. Bununla birlikte çoğu zaman entropi istatistiklerinin tartışılması soyut matematik bir düzlem ile sınırlı kalmaktadır. Halbuki entropi ve entropiden türeyen istatistikler, çıkarımda (statistical inference) alternatif olarak kullanılabilir. Bu anlamda hipotez testlerinde entropi istatistiklerinin alternatif olarak kullanılması konusunda daha somut çalışmalar yapılabilir. Bu çalışmada seçilmiş bazı problemler özelinde entropi istatistiklerinin birer çözümleme aracı olarak kullanılabileceğini, diğer geleneksel test istatistikleri (örneğin Ki-kare istatistikleri) ile benzer sonuçlar elde edilebileceğini göstermeye çalıştım.

ENTROPİ KAVRAMININ İNCELENMESİ

2.1 Bilgi Teorisi ve Entropi

İnformasyon Teorisinin (Bilgi Kuramının) amacı bilginin elde edilmesi (haber alınması), aktarılması, işlenmesi ve saklanmasına ilişkin nicel kanunları incelemektir. Bilgi kuramı 1940 'lı yıllarda telekomünikasyona bağlı problemlerin çözümü sırasında ortaya çıkmıştır. Bilgi kuramının amacı bilginin elde edilmesi, aktarılması, işlenmesi ve saklanmasına ilişkin kuralları incelemektir. Bilginin aktarılması sürecindeki rasgelelik olgusu bu süreçlerin incelenmesinde istatistik yöntemlerin kullanılmasını kaçınılmaz kılmaktadır. Entropi kavramının istatistiksel fizikte Boltzmann bağıntısı ile gelişiminin ardından Claude Shannon tarafından “İletişimin Matematiksel Teorisi”nde informasyon ile entropi kavramı arasında ilişki kurulmasıyla bu kavramın kullanımı hemen hemen tüm bilim dallarına yayılmıştır.

Bilgi teorisi üzerinde çalışan bir akademisyen olan Claude Shannon, bilgi ve entropinin aynı madalyonun farklı yüzleri olduğu fikrini ortaya atmıştır [4]. Bir sistemde entropi miktarı ne kadar düşükse, bilgi oranı o kadar fazladır.

Buhar makinesinin yanan kömürden elde ettiği enerjiyi dönüş hareketine çevirebilmesinin altında yatan neden, makinenin yüksek bilgi içeriği veya diğer bir ifade ile tasarımcısının ona aktarmış olduğu bilgidir. Aynı durum canlı organizmalar için de geçerlidir. Hayat DNA’da yazılı olan sayısal bilgidir. Canlı organizmalar DNA’daki bu bilgi sayesinde çevrelerinden enerji alarak kendilerini üretebilir ve entropiyi azaltarak bir düzen oluştururlar, yani yaşarlar. Bu örnekler bize bilginin, düzensizliği ve belirsizliği

azaltmada madde ve enerji ile birlikte yaşamsal bir önem taşıdığını göstermektedir. Bilgi, alıcı durumunda olan kişinin söz konusu sistem veya süreç hakkındaki kontrolünü artırır ve içinde bulunduğu belirsizliği azaltır. Diğer bir ifade ile yararlı bilgi, entropinin tersi bir işlev görerek sistemin entropisini azaltır ve “negatif entropi” olarak da tanımlanabilir [5].

Bilgi teorisi açısından “entropi”, bir durumun belirsizliğini ortadan kaldırmak için gerekli detaylı bilgi açığı olarak tanımlanmaktadır [5]. Evrendeki entropinin sürekli artması, kaçınılmaz olarak belirsizliklerin de artmasına neden olmakta ve bilgi açığını artırmaktadır. Bu durum, entropi artışıyla baş edebilmek ve negatif entropi üretebilmek için her geçen gün daha fazla yararlı bilgi üretilmesi ve etkin bir şekilde kullanılması gerektiğini göstermektedir.

Burada üzerinde durulması gereken diğer bir konu ise, bilgi düzeyini artırmak için enerji harcanması gerekliliğidir. Örneğin stoklardaki mal miktarını öğrenmek için saymak, bunun için de enerji harcamak gerekmektedir. Bu da yararlı bilgi üretebilmek için bir miktar kullanılabilir enerjinin kullanılamaz hale gelmesi, yani entropinin artması anlamını taşımaktadır. Görüldüğü gibi yararlı bilgi üreterek, bilginin negatif entropi gücünden yararlanmak ve belirsizlikleri azaltmak için de bir miktar entropi üretilmesi gerekmektedir [5]. Ancak bilgi kullanıldığı ve paylaşıldığı sürece entropiyi azaltacak, yani negatif entropi üreterek yaşamın devamlılığına katkıda bulunacaktır. Gizlenen, paylaşılmayan, kullanılmayan veya yanlış kullanılan bilginin ise insan yaşamına bir katkısı olmayacaktır. Bu durumda yaşamını devam ettirebilmek için sürekli olarak kullanılabilir enerjiyi tüketmek zorunda olan bir varlık olarak insan için yapılabilecek en faydalı işin, kendisi ve diğer canlılar için yararlı bilgi üretmek olduğu söylenebilir. Ayrıca bu bilginin diğer insanlarla paylaşılması ve yeni bilgilerin üretimine katkı sağlaması, yaşamın kalitesi ve insan neslinin devamlılığı açısından kaçınılmazdır.

Artan entropi ile birlikte detaylı bilgi ihtiyacının sürekli artması, bilgi üretimine katılacak kişi sayısının da sürekli olarak artmasını ve küresel işbirliğini kaçınılmaz kılmaktadır. Ancak günümüzde küçük bir azınlık bilgi üretimine katkıda bulunurken, özellikle az gelişmiş ülkelerdeki büyük bir çoğunluk sadece entropi üretimine katkıda bulunabilmektedir. Bu da üretilen bunca bilgiye rağmen dünyanın her geçen gün neden

daha düzensiz bir hal aldığını göstermektedir. Bu sebeple bilginin küresel ölçekte işbirliği ile üretilmesi ve paylaşılması, insan neslinin devamlılığı açısından kaçınılmazdır.

2.2 Sosyal Entropi

Sosyal entropi kuramı, bir sosyal sistem kuramıdır. Pynchon'nın da belirttiği gibi "sosyal entropi, insan ilişkileriyle ilgilidir". "Sosyal entropi teorisi, gerçek bir sosyal sistemin nasıl çalıştığını gösterir"[6]. Çünkü sosyal entropi, sistemi oluşturan öğeler arasındaki bütünleşememe ya da eklemleşememe (düzensizlik) nedeniyle toplumsal kaynakların, yararlı bir işe dönüştürülmesi sürecinde ortaya çıkan kayıpları ifade etmektedir. Sosyal entropi bağlamında enerji, bir toplumsal sistemin girdiden(ham madde ve yarı mamul madde), sermaye, işgücü, bilgi ve teknoloji kullanarak toplumların gereksinimlerini karşılayacak şekilde çıktılar(ürünler) elde edebilme kapasitesidir. Bailey'e göre sosyal entropinin nesnel ölçütü, "yaşama düzeyidir". Ona göre,"entropinin yani yaşama düzeyinin ölçümü, az gelişmiş toplumlarda tüketilen kalori miktarı; gelişmiş toplumlarda toplam refah düzeyi-dolar gibi" faktörlerdir [6]. Gerçekten de bir toplumun yaşama düzeyi(standartları), o toplumda artan verimlilikle birlikte sağlanan gelirlere karşılık gelmektedir. Bu bakımda minimum yaşama düzeyi, maksimum entropidir. Maksimum yaşama düzeyi ise minimum entropidir" [6]. Anlaşılacağı gibi entropi ve yaşama düzeyi ters orantılıdır. Sosyal entropi arttıkça yaşama düzeyi (refah düzeyi) düşmektedir.

Sosyal entropiyi etkileyen çeşitli faktörler bulunmaktadır. Sosyal entropi kuramı, bu değişkenleri iki ana grupta ele almaktadır. Makro(toplum) ve mikro (birey) olmak üzere iki ana grupta toplanan değişkenler şunlardır(Bailey):

A-Makro Değişkenler (Toplumsal alanda)

a-Nüfus

b-Bilgi

c-Yaşama standardı

d-Örgütlenme (organizasyon)

e-Teknoloji

B-Mikro Değişkenler (Bireysel alanda)

a-Gelir

b-Meslek

c-Yetenekler

d-Eğitim

e-İkametgah yeri

Bu faktörler birlikte toplumsal etkileşim düzeyi ve giderek ekonomik gelişme derecesi üzerinde etkili olmaktadır. Çünkü gerek sosyal, gerek biyolojik, gerekse termodinamik bir sistemin işleyebilmesi ya da varlığını sürdürebilmesi için bir enerji kullanması yani bir enerjiyi durumdan başka bir duruma dönüştürmesi gerekmektedir. Termodinamiğin enerjinin sakınımı yasası' na göre "enerji tüketilmez ancak bir formdan başka bir forma dönüştürülebilir". Biyolojik sistemler (organizma), dışarıdan aldıkları ilk maddeyi (girdiyi), bünyesi için yararlı ve gerekli olan son maddeye (çıktıya) dönüştürerek varlığını devam ettirir. Sosyal sistemler de, her bir alt sisteminin diğer alt sistemlerden sağladıkları girdiyi (ilk maddeyi), toplumsal gereksinimleri karşılayacak şekilde biçim, nicelik ya da nitelik dönüşümlerine uğratarak oluşturduğu 'çıktı' lardan yararlanarak varlıklarını ve gelişmelerini sürdürürler. Üç durum, toplumsal sistemin bu alandaki performansını etkilemektedir. Bunlar:

1- Sistemde yer alan ögeler arasında işlevsel farklılaşma derecesi. Öğeleri, arz ve talep düzeyleri (miktar-çeşit) bakımından farklılaşma derecesiyle doğru orantılı olarak sistem içindeki dolaşımın hızı artmaktadır. Dolayısıyla entropi minimum olmaktadır. Bunu sağlayan ise işlevsel farklılaşmadır. İşlevsel farklılaşma sistemin öğeleri arasında eşit olmayan bağılıklar meydana getirmektedir. Nitekim geleneksel köy topluluklarında ailelerin benzeri yaşama ve üretim biçimlerine sahip olması nedeniyle gerek köy halkı ve gerekse köyler arasındaki toplumsal enerjinin dolaşımı çok zayıftır. Çünkü "uzmanlaşma ve tabakalaşma düzeyleri düşük olan toplumlarda bütünleşme, dayanışma (consensus) ve homojenlik üzerine kuruludur". Yani denge halindeki topluluklardır.

2- İkinci koşul, toplumsal geçirgenlik derecesidir. Toplumsal geçirgenlik derecesi, aynı toplumun ögeleri olan parça ve bütünlerin birbirlerinden haberdar olma, derecesini ifade etmektedir. O bakımdan işlevsel farklılaşma gerekli, ancak yeterli değildir. Çünkü toplumsal sistemin çalışması, mekanik mekanizmalar gibi değildir. İşlevsel olarak farklılaşmış olsalar bile özel değerler etrafındaki gruplaşmalar karşılıklı bağlanma ve dayanışmayı azaltacağından sistem içindeki enerjinin dolaşım hızı düşecektir.

3- Üçüncü olarak da sistemin ögeleri arasında dolaşacak enerji kaynaklarının yenilenme derecesidir. Sistemin iş yapabilmesi için enerji tüketmek zorundadır. Sisteme giren her yeni enerji (teknolojik buluşlar, sermaye v.s.), entropiyi düşürmektedir. Ancak bir süre sonra sistem denge haline geldiğinden entropi yeniden yükselmektedir. Tüketilen enerji miktarında ilave artışların sağlanamaması sistemin entropisini artıracak yani denge haline getirecektir. Bu ekonomik gelişmenin yavaşlaması ve giderek toplumun refah düzeyinin düşmesi anlamına gelmektedir.

2.2.1 İşletmelerde Sosyal Entropi

Bir işletmenin ilişkiler alanı ya da sınırları hiçbir zaman kesin olarak çizilemez. Çünkü her işletme dış toplumla ilişki halindedir. Çünkü “her işletme dış toplumdan enerji alıp veren birimlerdir”[7]. Diğer bir deyişle işletmeler dış toplumdan girdi alıp çıktı sunan, böylece “enerjinin döngüsel biçimde tekrar enerjiye dönüşmesini sağlayan birimlerdir” [7].

Verimlilik, bir işletmede üretim performansının derecesini ölçerken sosyal entropi, performansın düşük ya da yüksek olma nedenlerini açıklamaya çalışmaktadır. Verimliliğin sosyal boyutu sosyal ilişkilerin, verimlilik üzerindeki etkilerinden ileri gelmektedir. Her türlü iş, işlevleri bakımından toplumsal olmakla birlikte işin gerçekleştirilmesi açısından bireyseldir. Ancak endüstri işletmelerinde iş, loncalarda olduğu gibi tek tek bireyler tarafından değil, gerek teknik işbölümü ve gerekse için küçük parçalara ayrılması nedeniyle birbirlerine eklemlenmiş iş gruplarının birlikte çalışması sonucunda gerçekleşmektedir. Bu durum entropiyi kaçınılmaz kılmaktadır. Nitekim Bailey de “bir olguyu oluşturan ögeler arasında sınırlar varsa orada entropi de vardır” demektedir [6]. Çünkü endüstri işletmelerinde ilk maddeden (ham madde ve yarı mamul madde), ürün elde edilmesi, gerek birbirlerinden işlevsel olarak ayrılmış

birçok alt birimin kendi içindeki ilişkileri ve gerekse bu birimler arasındaki eşgüdümü kapsayan üretim süreci sonucunda gerçekleşmektedir.

Girdi → emeğin işe dönüşme süreci → Çıktı

(İlk madde) (İş yapma performansı) (Ürün)

Üretim süreci (emeğin işe dönüşü süreci), aslında bir sistemin işleyiş sürecidir. Bu süreç aynı işletmede birden fazla çalışan arasındaki iş birliği, iş bölümü, işçi-makine ilişkisi, işçi-yönetici ilişkisi gibi alt sistemlerden oluşan alt grupların, gerek kendi aralarındaki gerekse dış çevreyle olan etkileşimleri kapsamaktadır. Alt gruplar, karşılıklılık ilişkileriyle birbirlerine bağlı olduklarından, birisindeki değişimler diğerlerinde de değişmelere neden olmaktadır.

Endüstri işletmelerinin verimliliği konusunda çoğunlukla iktisatçılar çalışmışlardır. O yüzden de verimlilik hesaplamalarında genellikle aşağıda belirtilen iktisadi parametreler kullanılmaktadır:

-İşyeri sayısı (adet)

-Girdi (Milyon TL.)

-Çıktı (Milyon TL.)

-Katma değer (Milyon TL.)

-Çalışan sayısı

-Çalışılan saatler

-Sabit sermayeye gayri safi ilaveler

-Ücret ödemeleri (milyon TL.)

-Stok değişimleri (Milyon TL.)

Bu ölçümlerde verimlilik, bağımlı bir değişkendir (açıklanan). Yani verimlilik oranı yukarıdaki değişkenlere göre hesaplanmakta ancak bu değişkenlerin hangi etkenlere göre değiştiği konusu ihmal edilmektedir. Oysa sosyal entropi verimliliğin azlık ya da çokluğunu ilişkiler sisteminin kuruluş ve işleyişine göre açıklayan bir bağımsız değişkendir. İşletmeler, dışarıdan aldıkları girdileri, başkalarının gereksinmelerini

karşılacak şekilde çıktıları dönüştürme yoluyla elde ettikleri gelirler vasıtasıyla varlıklarını ve gelişmelerini sürdüren örgütlerdir. İşletmeler bu işi gerçekleştirirken bir enerji harcama durumundadırlar. Buradaki kullanımıyla enerji, "bir sistemin iş yapabilme kapasitesidir". Sosyal entropi, söz konusu sosyal sistemin iş yapabilme performansının bir ölçütü olmaktadır. Başka bir anlatımla sosyal entropi, iş grupları arasındaki örgütlenme, eşgüdüm ve haberleşme alanlarındaki sorunlar nedeniyle iş yapabilme kapasitesindeki azalışları ifade etmektedir. Kapasite kullanımındaki azalışlar, kullanılan bilgi, teknoloji, işgücü, doğal kaynaklar gibi girdilerden çok azının faydalı işe dönüştürülmesi biçiminde ortaya çıkmaktadır. Minimum entropi, işletme girdilerinde bir artış olmadan çıktılarda en çok artışın olduğu durumdur. Maksimum entropi ise girdilerde bir artış sağlanmasına karşılık çıktıda bir artışın sağlanamaması durumudur. Görüldüğü gibi sosyal entropiyle verimlilik ters orantılıdır. Sosyal entropi düştükçe verimlilik artmaktadır.

2.3 Entropi Türleri

Uygulamada en sık kullanılan entropi Shannon entropisi olmakla beraber, Rényi entropisi de biyolojik çeşitliliğin derecesinin ortaya konmasında kullanılmaktadır. Yine Tsallis entropisi de son yıllarda istatistiksel çalışmalarda kendine ayrıcalıklı bir yer edinmiştir. Shannon, bir fiziksel sistemdeki durumların sayısı k iken ve sistemin bu durumlarda bulunma olasılıkları Kesikli bir örnek uzayı S içinde X rastlantı değişkeninin $x_1, x_2, \dots, x_n$ değerlerini alma olasılıkları sırası ile $p_1, p_2, \dots, p_n$ olsun. Bu durum için önerilen entropi ölçülerini aşağıda açıklayalım.

2.3.1 Shannon Entropisi

Kesikli bir dağılımda X rastlantı değişkeni $x_1, x_2, \dots, x_n$ değerlerini $p_1, p_2, \dots, p_n$ olasılıkları ile alsın. Bu durumda sisteminin Shannon entropisi

$$H_s = -\sum_{i=1}^n p_i \log p_i \quad (2.1)$$

olarak tanımlanır. Bu sistemi bir X rasgele değişkeni (rasgele değişkenin aldığı değerler sistemin bulunabileceği durumlar) olarak da ele alınabilir.

En heterojen durumda (X'in sözkonusu $x_1, x_2, \dots, x_n$ değerlerini alma olasılıklarının birbirine eşit olduğu ve bu olasılığın $1/K$ olduğu durumda) $H_s(p) = \log K$ olur. Buradan hareketle Shannon entropisinin alacağı üst değer grup sayısı K 'ya bağlı olduğu görülmektedir. Eğer $[0,1]$ aralığında bir entropi ölçüsü aranıyorsa Shannon entropisi " $\log K$ "ya bölünerek standardize edilebilir. Burada standardize etme deyimini indeksin maksimum heterojenlik halinde alacağı değere bölünmesi anlamındadır. Ancak bu durumda en homojen halde indeks sıfıra, en heterojen halde de bire eşit olabilecektir. Yoksa standardize edilmiş indeksin istatistikte kullanılan ortalaması sıfıra, varyansı bire eşit olan "standardize edilmiş" değişkenlerle bir ilişkisi bulunmamaktadır.

Sürekli haller için yukarıdaki formülasyonda toplama operatörü yerine integral operatörünü koymak yeterlidir. Genellikle logaritmanın tabanı önemsiz olmakla birlikte haberleşme mühendisliğinin pratik gereksinimleri çerçevesinde 2 olarak kabul edilmekte ve böylelikle entropi birimi bit olmaktadır. Baz e tabanı (Nauperian) olarak seçilecek olursa entropi birimi nat olmaktadır [8].

Kesikli hallerde örnek uzayı S içerisinde bütün sonuçların gerçekleşme şanslarının birbirine eşit olması halinde en büyük belirsizlik (en heterojen durum) gerçekleşecektir:

$$P(X=X_i) = \frac{1}{K} \quad i=1,2,\dots,K$$

ve bu duruma karşılık gelen maksimum entropi

$$H(X) = -\sum_{i=1}^K \frac{1}{K} \log\left(\frac{1}{K}\right) = \log K \quad (2.2)$$

değerine eşit olacaktır. Öbür uç durumda herhangi bir $\{X=X_k\}$ kesin olayının gerçekleşmesi halinde ise $H(X)=0$ olarak bulunur. Bu yüzden kesikli dağılımlar için entropinin 0 ve $\log K$ arasında olduğunu belirtmek gerekir. Sürekli hallerde ise Shannon entropi formülü

$$H = \int_{-\infty}^{\infty} f(x) \log(f(x)) dx \quad (2.3)$$

şeklinde dir. Sürekli hallerdeki entropi diferansiyel entropi olarak adlandırılmaktadır. Sürekli dağılımlardaki entropi hesaplamaları için yukarıdaki formüllerdeki toplama sembollerinin integral operatörleri ile yer değiştirmesi gerekmektedir. Yine de sürekli

dağılımlarda entropi hesabı kesikli hallerdeki kadar sorunsuz değildir. Bu sorunlar Reza tarafından özetlenmiştir. Sürekli hallerle ilgili sorunları Reza şöyle özetlemektedir:

“Sürekli durumlarda entropi negatif ya da sonsuz olabilir. Sürekli bir dağılımın kesikli bir dağılım ile yaklaşık olarak temsil edilmesinde özellikle grup sayısı arttırıldıkça bu durum daha net ortaya çıkabilir. Kesikli hallerin tersine sürekli hallerde değişken dönüşümüne gidildiğinde entropi sabit kalmayabilir. Başka bir ifade ile koordinat sisteminin dönüşümü halinde entropinin değişmeden kalacağının bir garantisi yoktur” [9].

Tanım 2.1 Asimptotik Eş Bölmeleme Teoremi (AEP) (Asymptotic Equipartition Property Theorem)

Bilgi kuramı olarak asimptotik bölmeleme özelliği (AEP) olarak adlandırılır. Bilgi kuramının asimptotik bölmeleme özelliği büyük sayılar yasasına benzerdir. Büyük sayılar kanunu bağımsız ve aynı dağıtılmış rasgele değişkenler için, $\frac{1}{n} \sum_{i=1}^n X_i$ n'nin büyük değerleri için beklenen değerin EX yakın olduğu belirtiyor. AEP bildiren $\frac{1}{n} \log \frac{1}{p(X_1, X_2, \dots, X_n)}$, $(X_1, X_2, \dots, X_n)$ bağımsız dağıtılmış rasgele değişkenleri için $p(X_1, X_2, \dots, X_n)$ gözlemlene olasılığına sahip H entropisi vardır. Bu durumda $p(X_1, X_2, \dots, X_n)$ gözlemlene olasılığı 2^{-nH} dizisine yakınsar. Bu durum diziler kümesini iki takım haline bölmemizi sağlar, entropiye yakın olan kısım tipik bir dizi, ve diğer kısmı da tipik olmayan bir dizidir. Entropi ile ilgili kısmı tipik dizilerdir. Tipik diziler için yüksek olasılıkla gerçek olacak büyük örneklerin ortalama davranışını belirlediği kanıtlanmıştır. Asimptotik eş bölmeleme teoremi aşağıdaki gibi tanımlanmıştır [10]:

Teorem 2.1 AEP

$X_1, X_2, \dots \sim p(x)$ bağımsız ve aynı dağılımlı rasgele değişkenler olmak üzere

$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \rightarrow H(X)$ olacak şekilde yakınsar.

İspat: Bağımsız rastgele değişkenlerin fonksiyonları da bağımsız rasgele değişkenlerdir. Böylece, X_i ayrık zaman bağımsız ve aynı dağılımlı rasgele değişkenler kaynakları için $\log(p_i)$ yakınsar. Bu nedenle büyük sayılar yasası ile,

$$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) = -\frac{1}{n} \sum_i \log p(X_i)$$

$$\rightarrow -E \log p(X)$$

$$= H(X)$$

2.3.1.1 Shannon Entropisinin Dağılımı

H' Shannon indeksinin kestiricisi H' olsun. H' P_i olasılıklarının, maksimum olabilirlik yöntemine göre bulunan P'_i değerlerinden hareketle hesaplanır [11].

Shannon entropisinin beklenen değeri ise aşağıdaki gibidir:

$$E(H') = -\sum p_i \ln p_i - \frac{S-1}{N} + \frac{1-\sum p_i^{-1}}{12N^2} + \frac{1-\sum(p_i^{-1}-p_i^{-2})}{12N^3} \quad (2.4)$$

Formülden de açıkça görülebileceği gibi maksimum olabilirlik yöntemine dayalı Shannon entropisinin kestiricisi yanlıdır. Bununla birlikte örnek büyüklüğü arttırıldığında yanlılık azalmaktadır.

H' varyansı aşağıdaki gibi hesaplanır:

$$\text{Var } H' = \frac{\sum p_i (\ln p_i)^2 - (\sum p_i \ln p_i)^2}{N} + \frac{S-1}{2N^2} \quad (2.5)$$

Shannon entropisinin asimptotik dağılım özellikleri literatürde incelenmiş bulunmaktadır[11]. Buna göre büyük örnekler için Shannon entropisinin dağılımı normal dağılım ile modellenmektedir.

İki Shannon entropisinin birbirine eşitliğinin test edilmesinde aşağıdaki test istatistiği kullanılır:

$$t = \frac{H'_1 - H'_2}{\sqrt{\frac{\text{Var } H'_1 + \text{Var } H'_2}{2}}} \quad (2.6)$$

Serbestlik derecesini de aşağıdaki formül ile hesaplanır:

$$df = \frac{(\text{Var } H'_1 + \text{Var } H'_2)^2}{\frac{(\text{Var } H'_1)^2}{N_1} + \frac{(\text{Var } H'_2)^2}{N_2}} \quad (2.7)$$

2.3.1.2 Shannon Entropi Birimi

(2.1) de tanımlanan Shannon entropisi bir istatistiksel deneyin belirsizliğinin bir ölçüsüdür. Logaritmanın tabanı 2 olduğunda entropinin birimi bit , “e” tabanı olduğunda ise “nat”tır. X rastlantı değişkeninin entropisi, X sonucunu ortaya koyabilmek için gerekli olan minimum ortalama bit sayısı olarak da görülebilir[8].

Entropi tanımındaki logaritmanın tabanı genellikle 2 olarak alınır. Bu durumda,

Durumlar	x_1	x_2
$p_i = P(X=x_i)$	$p_1 = \frac{1}{2}$	$p_2 = \frac{1}{2}$

gibi bir sistemin entropisi,

$$\begin{aligned} H(X) &= -\sum_{i=1}^2 p_i \log_2 p_i = -\left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{2} \log_2 \frac{1}{2}\right) \\ &= -\left(-\frac{1}{2} \log_2 2 - \frac{1}{2} \log_2 2\right) = 1 \end{aligned}$$

olur. Entropinin tanımında logaritmanın tabanı 2 olarak seçildiğinde, eşit olasılıklı iki durumlu bir sistemin entropisi, entropi ölçü birimidir. Birim entropiye “bit” denir.

Örneğin, düzgün bir paranın üç kez atılışındaki olasılıklara sahip, 8 durumlu bir sistemin entropisi,

YYY	YYT	YTY	TYT	TTY	TTT
1/8	1/8	1/8	1/8	1/8	1/8

$$H = -\left(\left(\frac{1}{8} \log_2 \frac{1}{8}\right) 8\right) = 3 \text{ (bit)}$$

dır. N tane eşit olasılıklı durum gösteren bir sistemin entropisi $\log_2 N$ bit’tir. Sadece bir durumda bulunan bir sistemin entropisi sıfırdır. Bundan sonraki kısımda bir X rasgele değişkeni sisteminin entropisi, (2.1) deki denklem olarak tanımlanacaktır.

Örnek 2.1 Bir X sistemi (rasgele değişkeni) ile ilgili durumlar ve bu durumlarda bulunma olasılıkları aşağıdaki gibi olsun.

$$X \equiv \begin{pmatrix} 1, 2, 3, 4 \\ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \end{pmatrix}$$

Bu sistemin entropisi,

$$H(X) = -\sum_{i=1}^4 p_i \log_2 p_i = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 2 \text{ (bit)}$$

dir. Entropi hesabında sadece durumların olasılıkları önemlidir.

$$X \equiv \begin{pmatrix} 1, 3, 3, 1 \\ \frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8} \end{pmatrix}$$

sisteminin entropisi

$$H(X) = -\sum_{i=1}^4 p_i \log_2 p_i = -\frac{1}{8} \log_2 \frac{1}{8} - \frac{3}{8} \log_2 \frac{3}{8} - \frac{3}{8} \log_2 \frac{3}{8} - \frac{1}{8} \log_2 \frac{1}{8} = 1.8 \text{ (bit) dir.}$$

Dört durumlu başka bir

$$X \equiv \begin{pmatrix} 1, 97, 1, 1 \\ \frac{1}{100}, \frac{97}{100}, \frac{1}{100}, \frac{1}{100} \end{pmatrix}$$

sisteminin entropisi,

$$H(X) = -\sum_{i=1}^4 p_i \log_2 p_i = -\frac{1}{100} \log_2 \frac{1}{100} - \frac{97}{100} \log_2 \frac{97}{100} - \frac{1}{100} \log_2 \frac{1}{100} - \frac{1}{100} \log_2 \frac{1}{100} \\ = 0.24 \text{ (bit)}$$

dir. Düzgün dağılım durumunda entropinin en fazla olduğuna dikkat edilmelidir.

Entropi tanımında iki tabanlı logaritma yerine e tabanlı logaritma da kullanılmaktadır.

$$-\sum_{i=1}^n p_i \ln p_i = -\sum_{i=1}^n p_i \ln(2^{\log_2 p_i}) = -\sum_{i=1}^n p_i \log_2 p_i \ln(2) = -\ln(2) \sum_{i=1}^n p_i \log_2 p_i$$

Entropi ölçü birimi bit olmak üzere, entropi tanımında iki tabanlı logaritma yerine e tabanlı logaritma kullanılırsa $\ln(2)$ katsayısını göz önünde tutmak gerekir. Doğal logaritma ve e tabanlı üslü ifadelerin türev ve integral hesaplamaları daha kolay

olmaktadır. Bundan sonraki kısımda bir $X \equiv \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix}$ sisteminin entropisi, (2. 1) denklemindeki gibi tanımlanacaktır.

2.3.1.3 Shannon Entropisinin Özellikleri

(i) Süreklilik: $H(p_1, p_2, \dots, p_n)$ bütün p_k değerlerini için süreklidir.

(ii) Simetri: $H(p_1, p_2, \dots, p_n)$ bütün p_k değerlerini için simetriktir.

(iii) Maksimum entropi: E_k olaylarının olasılıkları eşit olduğunda gerçekleşir.

(iv) Toplanabilirlik: $H(p_1, p_2, \dots, p_n)$ $E_1, E_2, \dots, E_n$ olayları ile ilgili entropi olsun. Bu olaylardan herhangi birisi ile ilgili daha ayrıntılı bir olasılık dağılımı şu şekilde elde edilir [9]:

$$E_n = F_1 \cup F_2 \cup \dots \cup F_m, p_n = \sum_{k=1}^m q_k P(F_k) = q_k \quad (2.8)$$

Şüphesiz burada $E_n, F_1, F_2, \dots, F_n$ olaylarından oluşan yeni bir örnek uzay olarak da düşünülebilir. Bu durumda bu olaylarla ilgili olasılıklar

$$\frac{q_i}{p_k} (i = 1, 2, \dots, m) \quad (2.9)$$

şeklinde hesaplanabilir. Dolayısıyla elde edilen yeni alt (örnek) uzay ile ilgili belirsizlik

ölçüsü $H_3\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, \dots, \frac{q_m}{p_n}\right)$ şeklindedir. Yine $H_1(p_1, p_2, \dots, p_n)$ ve

$H_2(p_1, p_2, \dots, p_{n-1}, q_1, q_2, \dots, q_m)$ olduğuna göre

$$H_2 = H_1 + p_n H_3 \quad (2.10)$$

Bu ölçü, entropinin doğrusal olma (ve dolayısıyla toplanabilme) özelliğinin yanısıra sezgilerimize de uygundur [9].

Yukarıdaki veya benzeri şartlara verilen özelliklere 1-4 ile uyumlu bir şekilde, arzu edilen bir belirsizlik işlev için bir fonksiyonel formu elde etmek için mümkün olmalıdır. Bu tür öneriler Feinstein, Khinchin, Shannon, Schützenberger'in ve diğerlerinin çalışmalarında ortaya çıkmıştır. Onların bulguları çok karışık değil, ancak ayrıntılı bir sunum için çok daha fazla mevcut çalışma alanı gereklidir. Literatürde aşağıdaki referanslar daha fazla okuma için tavsiye edilir.

1. Fadiev, çeşitli lemma özelliklerinden sonra 1, 2 ve 4 ile, kabul edilen H'ye önerilen biçiminde denk olması gerektiğini göstermektedir. (2. 1) denkleminde çarpımsal bir sabit dışında. (Feinstein [I] Bkz.)
2. Khinchin özellikleri 1, 3, ve 4 ve bir ekleme olması varsayar ve olaylar tam bir kümesi için boş olan entropi ile değiştirebilir olmamalıdır diye (2. 1) deki gibi pozitif sabit çarpanı şeklinde denklem türetir.
3. Schützenberger olayların tam kümesi ile ilişkili bilgilerin için daha genel bir aksiom amaçlamıştır. O denkleminin Shannon- Wiener entropisi dışındaki işlevlerini gösterir. (2. 1) deki kullanılabilir işlevlerini gösterir. (2. 1) denklemi mevcut zenginliği ve derinliği sayesinde teoride büyük oranda kolaylık sağlar.

2.3.1.4 Shannon Entropi Fonksiyonunun Ekstremumu

Teorem 2.2 Khinchin Teoremi: Sonlu sayıda durum gösteren bir sistemin entropisinin maksimum değeri durum sayısının logaritmasına eşittir ve durum olasılıklarının eşit olduğu halde ortaya çıkmaktadır.

İspat: Sistemin n tane durumu olsun.

$$X \equiv \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix} \quad i = 1, 2, \dots, n \text{ için } p_i > 0 \text{ ve } \sum_{i=1}^n p_i = 1$$

sisteminin,

$$H(X) = H(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n p_i \log p_i$$

entropisini maksimum değerini bulmak için,

$$\text{Amaç:} \quad \max H(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n p_i \ln p_i$$

$$\begin{aligned} \text{Kısıtlar:} \quad & \sum_{i=1}^n p_i = 1 \\ & p_i > 0, \quad i = 1, 2, \dots, n \end{aligned}$$

optimizasyon problemini ele alınır.

Problem Lagrange çarpanları yöntemi ile çözmeye çalışır. Lagrange fonksiyonu,

$$L = -\sum_{i=1}^n p_i \ln p_i + \lambda \left(\sum_{i=1}^n p_i - 1 \right)$$

ve

$$\frac{\partial L}{\partial p_i} = -\ln p_i - 1 + \lambda, \quad i = 1, 2, \dots, n$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^n p_i - 1$$

olmak üzere, türevlerin sıfıra eşitlenmesiyle elde edilen denklem sistemi,

$$\begin{cases} -\ln p_1 = 1 - \lambda \\ -\ln p_2 = 1 - \lambda \\ \vdots \\ -\ln p_n = 1 - \lambda \\ \sum_{i=1}^n p_i = 1 \end{cases}$$

olup, denklem sisteminin çözümü

$$p_1 = p_2 = \dots = p_n = \frac{1}{n}$$

$$\lambda = 1 + \ln \frac{1}{n}$$

ve

$$\max H(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n \frac{1}{n} \ln \frac{1}{n} = -\ln \frac{1}{n} = \ln n$$

dir. Sonlu sayıda duruma sahip sistemlerde maksimum entropi, durum olasılıklarının eşit olduğunda ortaya çıkmaktadır ve n durumlu bir sistemin maksimum entropisi $\ln n$ dir.

2.3.1.5 Shannon Entropisi için Çok Değişkenli Genellemeler

n değişkenli kesikli dağılımlarda birleşik dağılımın entropisi şu şekilde hesaplanmaktadır:

$$H(X_1, \dots, X_n) = -\sum_{x_1} \dots \sum_{x_n} \log(f(X_1, \dots, X_n)) f(X_1, \dots, X_n) \quad (2.11)$$

Sürekli hallerde ise

$$H(X_1, \dots, X_n) = -\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(X_1, \dots, X_n) \log(f(X_1, \dots, X_n)) dx_1 \dots dx_n \quad (2.12)$$

Çok değişkenli normal dağılımın entropisi de aşağıdaki gibi hesaplanmaktadır:

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x) \ln f(x) dx \\ &= \frac{1}{2} \ln | (2\pi e) \Sigma | \\ &= \frac{k}{2} (1 + (2\pi)) + \frac{1}{2} \ln | \Sigma | \end{aligned} \quad (2.13)$$

2.3.1.6 Shannon Bileşik Entropi

İki ya da daha çok sistemin bileşkesi olan sistemlerde entropi kavramı ele alınır.

$$X \equiv \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix}$$

ve

$$Y \equiv \begin{pmatrix} y_1, y_2, \dots, y_n \\ r_1, r_2, \dots, r_n \end{pmatrix}$$

gibi iki sistem söz konusu olsun. Bu iki sistemin bileşkesi dendiğinde,

$$\{(x_i, y_j) : i = 1, 2, \dots, n, j = 1, 2, \dots, m\}$$

kümesindeki durumlarda bulunabilen sistem anlaşılmaktadır. X ve Y gibi iki sistemin bileşkesi olan sistemi (X, Y) biçiminde gösterilir. Bu sistemin

$(x_i, y_j) : i = 1, 2, \dots, n, j = 1, 2, \dots, m$ durumlarında bulunması olasılıkları $P(X=x_i, Y=y_j), i = 1, 2, \dots, n, j = 1, 2, \dots, m$ olsun. Bileşke sistemin entropisi,

$$H(X, Y) = -\sum_{i=1}^n \sum_{j=1}^m P(X = x_i, Y = y_j) \ln P(X = x_i, Y = y_j) \quad (2.14)$$

olarak tanımlanmaktadır. İki sistem birbirinden bağımsız, yani

$$P(X=x_i, Y=y_j) = P(X=x_i)P(Y=y_j)$$

olduğunda,

$$\ln P(X=x_i, Y=y_j) = \ln P(X=x_i) + \ln P(Y=y_j)$$

olmak üzere,

$$H(X, Y) = H(X) + H(Y) \quad (2.15)$$

dır. Bağımsız iki sistemin bileşkesinin entropisi, sistemlerin entropi toplamına eşittir.

Bileşen sayısı $k \geq 2$ olmak üzere, k bileşenli bir $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}$ sistemi (k boyutlu rasgele

vektörü) için entropi,

$$H(\underline{X}) = -E \left[f_{\underline{X}}(X_1, X_2, \dots, X_k) \ln f_{\underline{X}}(X_1, X_2, \dots, X_k) \right] \quad (2.16)$$

olarak tanımlanır. Bileşenler bağımsız olduğunda,

$$H(\underline{X}) = H(X_1) + H(X_2) + \dots + H(X_k) \quad (2.17)$$

olur. Genelde,

$$H(\underline{X}) = H(X_1) + H(X_2 / X_1) + H(X_3 / X_1, X_2) + \dots + H(X_k / X_1, X_2, \dots, X_{k-1}) \quad (2.18)$$

dır. Sürekli durumda da yukarıdakilere benzer tanımlamalar yapılabilir.

2.3.1.7 Shannon Koşullu Entropi

(X, Y) gibi bileşik bir sistemde, X sisteminin x_i durumunda olduğu bilinsin. X sisteminin x_i durumunda olduğu bilindiğinde Y sisteminin koşullu entropisi

$$H(Y|X=x_i) = - \sum_{j=1}^m P(Y = y_j | X = x_i) \ln P(Y = y_j | X = x_i) \quad (2.19)$$

olarak tanımlanır.

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X = x)$$

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

$$= - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(y|x)$$

$$= -E_{p(x,y)} \log p(y|x)$$

olarak ifade edilir.

(X, Y) gibi bileşik bir sistemde, Y sisteminin X sistemine koşullu toplam entropisi

$$H(Y|X) = \sum_{i=1}^n p_i H(Y|X = x_i) \quad (2.20)$$

olarak tanımlanmaktadır.

(X, Y) bileşke sisteminde Y sistemi ile X sistemi bağımsız ise

$$\begin{aligned}
H(Y|X) &= \sum_{i=1}^n p_i H(Y|X = x_i) = - \sum_{i=1}^n p_i \sum_{j=1}^m P(Y = y_j | X = x_i) \ln P(Y = y_j | X = x_i) \\
&= - \sum_{i=1}^n p_i \sum_{j=1}^m P(Y = y_j) \ln P(Y = y_j) = - \sum_{j=1}^m P(Y = y_j) \ln P(Y = y_j) \\
&= H(Y)
\end{aligned}$$

olur.

(X, Y) bileşke sisteminin entropisi için

$$H(X, Y) = H(Y) + H(X|Y) = H(X) + H(Y|X) \quad (2.21)$$

yazılabilir. $(P(X=x_i, Y=y_j) = P(X=x_i)P(Y=y_j | X=x_i)$ zincir kuralından kolayca elde edilir.)

Ayrıca,

$$H(X, Y) \leq H(X) + H(Y) \quad (2.22)$$

olur.

Dikkat edilirse iki bileşenli bir sistem, iki değişkenli bir rasgele vektör olarak ele alınabilir. Bundan sonra, (X, Y) bileşke sistemi ile (X, Y) rasgele vektörü kavramlarını birbirinin yerine kullanılacaktır. (X, Y) nin olasılık fonksiyonu

$$f_{X,Y}(x_i, y_j) = p_{ij} = P(X=x_i, Y=y_j), \quad i=1,2,\dots,n, \quad j=1,2,\dots,m$$

olmak üzere (X, Y) sistemi, entropi açısından $p_{ij}, i=1,2,\dots,n, j=1,2,\dots,m$ olasılıkları ile anlatılabilir, yani $(X, Y) \equiv (p_{ij})_{n \times m}$ denebilir. Buna göre; X, Y rasgele değişkenlerinin ortak dağılımının entropisi (X ve Y sistemlerinin bileşkesinin entropisi)

$$H(X, Y) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \ln p_{ij} \quad (2.23)$$

olur.

Örnek 2.2 a) (X, Y) gibi bileşik bir sistem (iki boyutlu rasgele vektör) ile ilgili durumlar ve bu durumlarda bulunması olasılıkları aşağıdaki gibi olsun.

Dağılım 1		
Y/X	x_1	x_2
y_1	$\frac{1}{4}$	$\frac{1}{4}$

y_2	$\frac{1}{4}$	$\frac{1}{4}$
-------	---------------	---------------

Bu bileşke sistemin entropisi,

$$H(X,Y) = -\frac{1}{4}\ln\frac{1}{4} - \frac{1}{4}\ln\frac{1}{4} - \frac{1}{4}\ln\frac{1}{4} - \frac{1}{4}\ln\frac{1}{4} = \ln 4 = 2\ln 2 = 2(\text{bit})$$

olup, X sisteminin (X rasgele değişkeninin) entropisi,

$$H(X) = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2} = \ln 2 = 1(\text{bit})$$

Y sisteminin (Y rasgele değişkeninin) entropisi

$$H(Y) = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2} = \ln 2 = 1(\text{bit})$$

olmak üzere,

$$H(X,Y) = H(X) + H(Y)$$

dır.

b) (X,Y) gibi bileşik bir sistem (iki boyutlu rasgele vektör) ile ilgili durumlar ve bu durumlarda bulunması olasılıkları aşağıdaki gibi olsun.

Dağılım 2		
Y/X	x_1	x_2
y_1	$1/8$	$3/8$
y_2	$3/8$	$1/8$

Bu bileşke sistemin entropisi,

$$H(X,Y) = -\frac{1}{8}\ln\frac{1}{8} - \frac{3}{8}\ln\frac{3}{8} - \frac{3}{8}\ln\frac{3}{8} - \frac{1}{8}\ln\frac{1}{8} = 1.2555 = \frac{1.2555}{\ln 2} (\text{bit}) = 1.8113 (\text{bit})$$

olup, X sisteminin (X rasgele değişkeninin) entropisi,

$$H(X) = -\frac{1}{2} \ln \frac{1}{2} - \frac{1}{2} \ln \frac{1}{2} = \ln 2 = 1(\text{bit})$$

Y sisteminin (Y rasgele değişkeninin) entropisi

$$H(Y) = -\frac{1}{2} \ln \frac{1}{2} - \frac{1}{2} \ln \frac{1}{2} = \ln 2 = 1(\text{bit})$$

olmak üzere,

$$H(X, Y) \leq H(X) + H(Y) \text{ olur.}$$

X sisteminin (rasgele değişkeninin) x_1 durumunda olduğu bilindiğinde Y sisteminin (rasgele değişkeninin) koşullu dağılımı

Dağılım 3		
y_j	y_1	y_2
$P(Y = y_j X = x_1)$	$\frac{1}{4}$	$\frac{3}{4}$

ve koşullu entropisi,

$$H(Y / X = x_1) = -\sum_{j=1}^2 P\left(Y = y_j / X = x_1\right) \ln P\left(Y = y_j / X = x_1\right) = -\frac{1}{4} \ln \frac{1}{4} - \frac{3}{4} \ln \frac{3}{4} = 0.562 = 0.811 (\text{bit})$$

dır. Benzer şekilde X sisteminin (rasgele değişkeninin) x_2 durumunda olduğu bilindiğinde Y sisteminin (rasgele değişkeninin) koşullu entropisi,

$$H(Y / X = x_2) = -\sum_{j=1}^2 P\left(Y = y_j / X = x_2\right) \ln P\left(Y = y_j / X = x_2\right) = -\frac{3}{4} \ln \frac{3}{4} - \frac{1}{4} \ln \frac{1}{4} = 0.562 = 0.811 (\text{bit})$$

değerini alır.

Y sisteminin X sistemine koşullu toplam entropisi,

$$H(Y / X) = \sum_{i=1}^2 p_i H(Y / X = x_i) = \frac{1}{2} \times 0.811 + \frac{1}{2} \times 0.811 = 0.811(\text{bit})$$

olur.

2.3.2 Renyi Entropisi

Renyi entropisi adına Alfred Renyi'den almaktadır. Rényi entropisi de biyolojik çeşitliliğin derecesinin ortaya konmasında kullanılmaktadır. Ayrıca Kuantum bilgi ölçüsü olarak kullanımı da önemlidir. Renyi entropisi Shannon entropisinin $\alpha \rightarrow 1$ yaklaşırken özel bir halidir. Yani Renyi entropileri özel bir durum olarak Shannon entropilerini içerir. Renyi entropisi şu şekilde tanımlanmaktadır [2]:

$$H_R(p) = \frac{1}{1-\alpha} \log(\sum_{i=1}^n P_i^\alpha), \alpha > 0, \alpha \neq 1 \quad (2.24)$$

entropisinin alacağı değer, Shannon entropisinde olduğu gibi, $\log K$ 'ya eşit olacaktır. Rényi entropisinin değeri için limiti Shannon entropisidir. Genel olarak tüm ayrık rasgele değişkenlerin Renyi entropi için α artan olmayan bir fonksiyondur[16].

2.3.2.1 Renyi Entropisinin Özellikleri

Shannon entropisi ile benzer özelliklere sahiptir [13].

- 1) $H_R(X)$ Renyi entropisi negatif değerler almaz: $H_R(X) \geq 0$
- 2) $H_R(X)$ Renyi entropisi sınırlı, sürekli ve α değerleri için artan olmayan bir fonksiyondur.
- 3) Renyi entropileri farklı α değerleri için korelasyonludur.
- 4) $\alpha < 1$ değerleri için Renyi entropisi konkav, $\alpha > 1$ değerleri için Renyi entropileri konkav yada konveks değildir.
- 5) Renyi entropisi toplanabildir.
- 6) Renyi entropisinin maksimumu $P_i = 1/n$ değeri için $\ln n$ dir.

2.3.2.2 Renyi Sapması

Renyi sapması Kulback-Leibler uzaklığının özelleştirilmiş bir halidir.

Q ve P bir dağılım olmak üzere $\alpha > 0$ için Renyi sapması şu şekilde tanımlanır:

$$D_\alpha(P||Q) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^n \frac{p_i^\alpha}{q_i^{\alpha-1}}\right) = \frac{1}{1-\alpha} \log(\sum_{i=1}^n p_i^\alpha q_i^{\alpha-1}) \quad (2.25)$$

$\alpha > 0$ için Renyi sapması Kulback-Leibler uzaklığı gibi negatif değerler almaz. Bu durumda Renyi sapması, α sapması olarakta adlandırılır [14].

Bazı özel durumlar:

$$D_0(P||Q) = -\log Q(\{i : p_i > 0\})$$

$$D_{1/2}(P||Q) = -2 \log \sum_{i=1}^n \sqrt{p_i q_i}$$

$$D_1(P||Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} : \text{Kullback-Leibler uzaklığı}$$

$$D_2(P||Q) = \log \langle \frac{p_i}{q_i} \rangle : \text{Olasılıkların beklenen oranı}$$

$$D_\infty(P||Q) = \log \sup_i \frac{p_i}{q_i} : \text{Olasılıkların maksimum oranı}$$

2.3.3 Tsallis Entropisi

Fizikte Tsallis entropisi Boltzman ve Gibbs entropilerinin genelleştirilmiş halidir. Constantino Tsallis tarafından standart istatistiksel mekaniği genelleme için bir temel olarak 1988 yılında tanıtılmıştır. Tsallis entropisi genel bir olgu olarak Simpson çeşitlilik indeksinin özel konumunu doğrulamıştır. Tsallis entropisi şu şekilde tanımlanmıştır [15]:

$$H_\alpha(p) = \frac{1 - \sum p_i^\alpha}{\alpha - 1}, \alpha > 0, \alpha \neq 1 \quad (2.26)$$

Tsallis ve Renyi entropilerinin birbirleri ile ilişkilidirler. Bu 2 entropinin aralarındaki ilişki için aşağıdaki eşitlik geçerlidir:

$$H_R(\alpha) = \frac{\log(1 - (\alpha - 1) H_T(\alpha))}{1 - \alpha} \quad (2.27)$$

Diğer bir deyişle Renyi entropisi, Tsallis entropisinin monoton artan bir fonksiyonudur.

2.3.4 Bazı Olasılık Dağılımlarının Shannon Entropisi

Bazı seçilmiş ve çok kullanılan olasılık dağılımlarının entropileri aşağıda ele alınmıştır.

Çizelge 2. 1 Bazı olasılık dağılımların entropileri

Dağılımın adı	Olasılık fonksiyonu	Entropi
Üstel dağılım	$f(x) = \lambda e^{-\lambda x}, x, \lambda > 0$	$1 - \ln \lambda$
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{2} \ln(2\pi e \sigma^2)$
Düzgün	$f(x) = \frac{1}{\beta - \alpha}; \alpha \leq x \leq \beta$	$\ln(\beta - \alpha)$
Bernoulli	$f(k; p) = p^k(1 - p)^{1-k}; k = 0,1$	$-q \ln(q) - p \ln(p)$
Binom	$P(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}$	$\frac{1}{2} \ln(2\pi n e p(1 - p)) + O(\frac{1}{n})$
Geometrik	$P(K = k) = (1 - p)^{k-1} p$	$\frac{-(1 - p) \log_2(1 - p) - p \log_2(p)}{p}$
Gama	$f(x) = \frac{x^{k-1} \exp(-x/\theta)}{\Gamma(k) \theta^k}$	$k + \ln \theta + \ln \Gamma(k) + (1 - k) \Psi(k)$
Beta	$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1 - x)^{\beta-1}$	$\ln B(\alpha, \beta) - (\alpha - 1) \Psi(\alpha - (-1) \Psi(\beta) + (\alpha + \beta - 2) \Psi(\alpha + \beta))$

Yukarıdaki çizelgede Gama dağılımı [16] ile ilgili satırdaki $\Psi(k)$ terimi digama [17] fonksiyonunu ifade etmektedir ve gama fonksiyonunun logaritmasının k parametresine göre birinci mertebeden türevine eşittir [18]:

$$\Psi(k) = \frac{d \ln(\Gamma(k))}{dk} = \frac{\Gamma'(k)}{\Gamma(k)} \quad (2.28)$$

$$B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad (2.29)$$

olacak şekilde Beta fonksiyonudur.

BÖLÜM 3

SHANNON ENTROPİSİ VE İSTATİSTİKSEL BAĞIMSIZLIK

X ve Y bağımsız olduğunda koşullu dağılımların entropileri marjinal dağılımların entropilerine eşit olur. Dolayısıyla iki bağımsız rastlantı değişkeni için;

$$H(X|Y) = H(X) \quad (3.1)$$

$$H(Y|X) = H(Y) \quad (3.2)$$

Farklı türden dağılımları entropiler arasındaki ilişkileri de şu şekildedir:

$$H(X,Y) = H(Y) + H(X|Y) \quad (3.3)$$

$$H(X,Y) = H(X) + H(Y|X) \quad (3.4)$$

Yine (3. 3) ve (3. 4) no'lu denklemlerden yola çıkarak birleşik dağılımın entropisinin marjinal bir dağılımın entropisi ile koşullu bir dağılımın entropisinin toplamına eşit olduğunu belirtmek gerekir [19].

3.1 Karşılıklı Bilgi Ölçüsü

Buradaki bilgi kavramı gözlemcinin bilincinden bağımsız bir anlamda kullanılmaktadır. Dolayısıyla belirsizlik kavramı da daha çok deneyin sonucuna yönelik belirsizlik anlamında ya da nesnel bir bağlamdadır [2].

Değişkenlerin birbirleri hakkında verdiği karşılıklı bilgi miktarı C.E. Shannon tarafından

$$I(X,Y) = \sum_{j=1}^m \sum_{i=1}^n P(X = x_i, Y = y_j) \log \frac{P(X=x_i, Y=y_j)}{P(X=x_i)P(Y=y_j)} \quad (3.5)$$

olarak tanımlanmıştır. Sürekli hallerde yukarıdaki formüldeki toplama sembolleri integral sembolleri ile yer değiştirmektedir. X ve Y'nin bağımsız olması halinde $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$ veya tüm x_i ve y_j ler için $I(X,Y)=0$ olur. $\{X = x_i\}$ ve $\{Y = y_j\}$ çiftleri için $I(X,Y) \geq 0$ olduğu durumda X ve Y değişkenleri bağımlı olur ve karşılıklı bilgi miktarının negatif olabileceğine dikkat edilmelidir. Bazı işlemlerden sonra

$$I(X,Y)=H(X)+H(Y)-H(X,Y) \quad (3.6)$$

$$I(X,Y)=H(X)-H(X|Y) \quad (3.7)$$

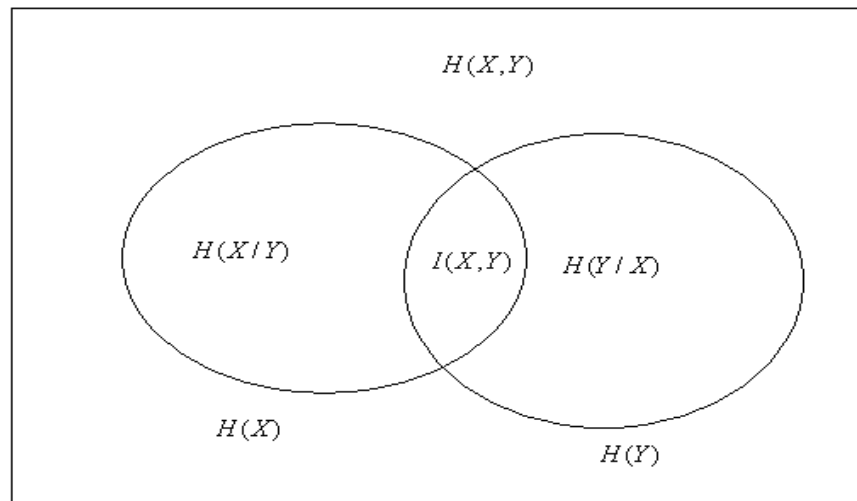
$$I(X,Y)=H(Y)-H(Y|X) \quad (3.8)$$

Burada entropi ve karşılıklı bilgi miktarının birbirleri ile yakından ilgili iki kavram olduğunun altı çizilmelidir [20].

$$I(X,Y)=0 \Leftrightarrow X \text{ ve } Y \text{ bağımsız değişkenlerdir} \quad (3.9)$$

$$I(X,Y)= I(Y,X) \quad (3.10)$$

$$I(X,X)=H(X) \quad (3.11)$$



Şekil 3. 1 Entropi ve bilgi kuramı

Buradan görebiliriz ki X ve Y bağımsız değişkenler ise $I(X,Y)=0$ eğer Y ve X bağımsız değişkenler ise tersi durum için bir şey söyleyemeyiz.

3.2 İki Değişkenli Normal Dağılımda Karşılıklı Bilgi Ölçüsü

(X,Y) değişkenleri $\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho$ parametrelili iki değişkenli normal dağılıma uysunlar. Bu durumda birleşik entropi

$$H(X,Y) = -\ln(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}) \quad (3.12)$$

X ve Y'nin marjinal entropileri de

$$H(X) = -\ln\sqrt{2\pi\sigma_x^2} \quad (3.13)$$

$$H(Y) = -\ln\sqrt{2\pi\sigma_y^2} \quad (3.14)$$

olarak bulunur. Hesaplanan bütün entropiler nat cinsindendir ve karşılıklı bilgi ölçüsü de

$$I(X,Y) = -\frac{1}{2}\ln(1-\rho^2) \quad (3.15)$$

olur. $\rho=0$ için $I(X,Y)=0$ olacağına dikkat edilmelidir [21].

3.3 Karşılıklı Bilgi Miktarı için Varyasyonlar

Karşılıklı bilgi miktarı için ($I(X,Y) \geq 0$) alt sınır vardır, fakat karşılıklı bilgi miktarının kesin bir üst sınırı yoktur. Normalleştirme istatistiğinin 0 ve 1 arasında kalmasını sağlamak için uygun sabite bölerek istatistiği dönüştüren bir işlemdir. Karşılıklı bilgi hakkında çeşitli varyasyonlar önerilmiştir. Bunlardan önemli olanları normalize varyantlar ve çok değişkenli genellemelerdir.

Karşılıklı bilgi miktarının ölçülebilmesi için önerilen istatistikler arasında “normalleştirilmiş” ve çok değişkenli hallere genellenmiş olanlar önemlidir. Bu bağlamda [22] ile [23] e bakılabilir. Bunların içerisinde

$$C_{XY} = \frac{I(X;Y)}{H(Y)} \quad (3.16)$$

$$C_{YX} = \frac{I(X;Y)}{H(X)} \quad (3.17)$$

Söz edilmeye değerdir. Bu iki katsayı birbirine eşit değildir. Simetrik bir versiyon ise

$$R = \frac{I(X;Y)}{H(X) + H(Y)} \quad (3.18)$$

şeklindedir. Bu katsayı bolluk (redundancy) katsayısıdır ve bağımsızlık durumunda 0, tam bağımlılık halinde ise maksimumu olan $\frac{1}{2}$ değerini almaktadır. Sözedilmesi gereken daha başka bağımlılık ölçüleri de şöyledir [24],[25]:

$$\frac{I(X;Y)}{\min\{H(X), H(Y)\}} \quad , \quad \frac{I(X;Y)}{H(X,Y)} \quad , \quad \frac{I(X;Y)}{\sqrt{H(X)H(Y)}} \quad (3.19)$$

3.4 Çok Değişkenli Hallere Genelleme

$X_1, X_2, \dots, X_n$, olasılık fonksiyonu $f(X_1, X_2, \dots, X_n)$ olacak şekilde birlikte dağılsın. Bu durumda birleşik dağılımın entropisi koşullu entropiler cinsinden ifade edilebilir [26]:

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i / X_{i-1}, \dots, X_1) \quad (3.20)$$

Sözelimi iki değişkenli halde

$$H(X_1, X_2) = H(X_1) + H(X_2 / X_1) \quad (3.21)$$

Ya da örneğin üç değişkenli halde

$$H(X_1, X_2, X_3) = H(X_1) + H(X_2, X_3 / X_1) \quad (3.22)$$

Benzer biçimde Z verildiğinde X ve Y arasındaki koşullu karşılıklı bilgi miktarı da

$$I(X;Y/Z) = H(X/Z) - H(X/Y,Z) \quad (3.23)$$

Karşılıklı bilgi miktarları için de

$$I(X_1, X_2, \dots, X_n; Y) = H(X_1, X_2, \dots, X_n) - H(X_1, X_2, \dots, X_n / Y) \quad (3.24)$$

$$= \sum_{i=1}^n H(X_i / X_{i-1}, \dots, X_1) - \sum_{i=1}^n H(X_i / X_{i-1}, \dots, X_1, Y) \quad (3.25)$$

$$I(X_1, X_2, \dots, X_n; Y) = \sum_{i=1}^n H(X_i; Y / X_{i-1}, X_{i-2}, \dots, X_1) \quad (3.26)$$

özdeşlikleri formüle edilebilir [27].

3.5 Karşılıklı Bilgi Uygulamaları

Birçok uygulama da koşullu entropiyi düşürmek için karşılıklı bilgi düzeyini en üst düzeye çıkartmak istenmektedir. Telekomünikasyon olarak, kanal kapasitesi karşılıklı

bilgi miktarına eşittir. Bunun yanı sıra, aynı zamanda makine öğrenme özellik seçimi için bir kriter olarak kullanılmaktadır. Ayrıca, bir veri seti, iki farklı kümelenebilir belirlenmesinde, benzerlik ve Bayesian ağlarda kullanılır. Entropi ve istatistiksel teori uygulamalarında, göreceli entropi ve karşılıklı bilgi miktarı hakkındaki genel bilgiler için Kullback [28]'ten faydalanılmalıdır.

3.5.1 Nominal Veriye Uygulama

Aşağıdaki frekans tablosu İstanbul'da faaliyet gösteren bir el örgü şirketinin verilerinden oluşturulmuştur. İki farklı türde müşterilere sunulan yedi farklı ürün grubu vardır. Nisan ve Ekim aylarında satılan ürünler gruplar şeklinde ayrılmıştır. Bu ürünlerin satış içindeki yüzdeleri bulunarak aşağıdaki tablo elde edilmiştir.

Çizelge 3. 1 Nisan-Ekim ayları arası satılan ürün miktarı(kg)

ÜRÜN	GENEL TOPLAM KG	DÜZ	DESENLİ
ANGORAGRUBU	235.334,50	206.912,00	28.422,50
%25YÜNLÜGRUBU	100.454,00	93.182,00	7.272,00
AKRİLİKGRUBU	35.099,00	7.413,50	27.685,50
POLYAMİDGRUBU	78.134,50	34.370,50	43.764,00
PAMUKLUGRUBU	22.674,00	19.490,00	3.184,00
%50YÜNLÜGRUBU	20.728,00	16.121,00	4.607,00
FANTEZİGRUBU	53.367,60	34.708,00	18.659,60
TOPLAM	545.791,60		

Ürün grubu düz ve desenli olarak 2 nitel değişkene ayrılmış ve kontenjans tablosu aşağıdaki gibi düzenlenmiştir.

Çizelge 3. 2 Satılan ürünlerin farklı türlerdeki dağılımı

ÜRÜN / TÜR	DÜZ	DESENLİ	TOPLAM
ANGORAGRUBU	0,379104	0,052076	0,431180
%25YÜNLÜGRUBU	0,170728	0,013324	0,184052
AKRİLİKGRUBU	0,013583	0,050725	0,064308
POLYAMİDGRUBU	0,062974	0,080184	0,143158
PAMUKLUGRUBU	0,035710	0,005834	0,041543
%50YÜNLÜGRUBU	0,029537	0,008441	0,037978
FANTEZİGRUBU	0,063592	0,034188	0,09778
TOPLAM	0,755228	0,244772	1

Bu 2 nitel değişkenin (ürün grubu ,ürün türü) anlamlı bir ilişkinin olup olmadığını belirlemek mümkün değildir. Burada X ve Y sırasıyla ürün grubu ve ürün türünü temsil etmektedir. Bu kesikli rassal dağılımın Shannon entropisi aşağıdaki gibidir. X, Y, (X,Y) entropilerinin bazı hesaplamaları aşağıda verilmiştir:

$$H(X)= 2,326$$

$$H(Y)= 0,80028$$

$$H(X,Y)= 2,9589 \text{ olarak bulunur.}$$

Buradan Shannon entropilerinden faydalanarak değişkenlerin birbirleri hakkındaki bilgi miktarları hesaplanır.

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \text{ olduğundan}$$

$$I(X;Y)= 2,326+0,8028-2,9589 \text{ (uyum iyiliği)}$$

$I(X;Y)=0,1699$ olarak bulunur.

Koşullu Shannon Entropileri aşağıdaki formüller ile hesaplanır.

$$C_{XY} = \frac{I(X;Y)}{H(Y)} \quad \text{ve} \quad C_{YX} = \frac{I(X;Y)}{H(X)}$$

$$C_{XY} = \frac{0,1699}{2,326} = 0,073 \quad \text{ve} \quad C_{YX} = \frac{0,1699}{0,8028} = 0,211 \quad \text{olarak bulunur.}$$

Şimdi Renyi ve Tsallis entropilerini α vereceğimiz farklı 4 değer için hesaplayabiliriz.

$$H_R(p) = \frac{\log \sum p_i^\alpha}{1-\alpha}, \alpha > 0, \alpha \neq 1 \quad \text{formülüyle hesapladığımızdan}$$

$\log \left[\frac{\sum p_i^{0,33}}{1-0,33} \right]$ olacak şekilde hesaplanarak aşağıdaki tablo elde edilir.

Çizelge 3. 3 Farklı α değerleri için renyi entropileri

α	$H_\alpha^R(X)$	$H_\alpha^R(Y)$
0,33	2,6443	0,9297
0,5	2,56	0,8952
2	1,959	0,6659
3	1,7317	0,5833

olarak bulunur. Koşullu Renyi entropilerini hesaplamak için gerekli koşullu Shannon entropilerini hesaplırsak.

$$\alpha = 0,33 \text{ için } H(X/Y) = 2,571$$

$$\alpha = 0,5 \text{ için } H(X/Y) = 2,455$$

Çizelge 3. 4 Farklı α değerleri için koşullu shannon entropileri

α	$H(X/Y)$
0,33	2,571
0,5	2,455
2	1,772
3	1,580

olarak bulunur. Buradan koşulla Renyi entropilerine geçerse:

$$C_{XY} = \frac{I(X,Y)}{H(Y)} = \frac{H(Y) - H(Y|X)}{H(Y)} = 1 - \frac{H(X|Y)}{H(Y)}$$

Çizelge 3. 5 Farklı α değerleri için koşullu renyi entropileri

α	C_{XY}	$H_{\alpha}^R(X/Y)$
0,33	$C_{XY} = 1 - \frac{2,571}{2,644}$	0,0276
0,5	$C_{XY} = 1 - \frac{2,455}{2,56}$	0,0411
2	$C_{XY} = 1 - \frac{1,772}{1,959}$	0,0953
3	$C_{XY} = 1 - \frac{1,58}{1,7317}$	0,0873

olarak bulunur.

Tsallis entropilerini α vereceğimiz 0,33, 0,5, 2 ve 3 değerleri için hesaplırsak:

$$H_{\alpha}(p) = \frac{1 - \sum p_i^{\alpha}}{\alpha - 1}, \alpha > 0, \alpha \neq 1$$

Çizelge 3. 6 Farklı α değerleri için tsallis entropileri

α	$H_{\alpha}(X)$	$H_{\alpha}(Y)$
0,33	3,603	0,805
0,5	2,858	0,727
2	0,742	0,369
3	0,454	0,277

olarak bulunur.

Buradan koşulla Tsallis entropilerine geçerse:

$$C_{XY} = \frac{I(X,Y)}{H(Y)} = \frac{H(Y) - H(Y|X)}{H(Y)} = 1 - \frac{H(X|Y)}{H(X)}$$

Çizelge 3. 7 Farklı α değerleri için koşullu tsallis entropileri

α	C_{XY}	$H_{\alpha}(X/Y)$
0,33	$C_{XY} = 1 - \frac{3,434}{3,6039}$	0,0469
0,5	$C_{XY} = 1 - \frac{2,685}{2,858}$	0,0600
2	$C_{XY} = 1 - \frac{0,7036}{0,7428}$	0,0528
3	$C_{XY} = 1 - \frac{0,44}{0,454}$	0,0313

olarak bulunur.

Bu bulduğumuz değerleri aşağıdaki tablolarda birleştirerek özetleyebiliriz.

Shannon entropisi için

$$H(X) = 2,326$$

$$H(X/Y) = 2,156, C_{XY} = 0,073$$

Çizelge 3. 8 Rényi ve tsallis entropilerine dayalı istatistikler

Rényi Entropileri	$\alpha=0,33$	$\alpha=0,5$	$\alpha=2$	$\alpha=3$
X'in Entropisi	2,6443	2,56	1,959	1,7317
Y verildiğinde X'in koşullu entropisi	2,571	2,455	1,772	1,58
C(X,Y)	0,02761	0,0411	0,0953	0,0873
Tsallis Entropileri	$\alpha=0,33$	$\alpha=0,5$	$\alpha=2$	$\alpha=3$
X'in Entropisi	3,603	2,858	0,7428	0,454
Y verildiğinde X'in koşullu entropisi	3,434	2,685	0,7036	0,44
C(X,Y)	0,0469	0,06	0,0528	0,0313

Shannon entropisi için

$H(Y)=0,8028$ için

$H(Y|X)= 0,6334$

$C_{XY}=0,211$

Çizelge 3. 9 Rényi ve tsallis entropilerine dayalı istatistikler

Rényi Entropileri	$\alpha=0,33$	$\alpha=0,5$	$\alpha=2$	$\alpha=3$
Y'nin Entropisi	0,929	0,8952	0,665	0,5833
X verildiğinde Y'nin koşullu entropisi	0,847	0,78	0,489	0,434
C(Y,X)	0,088	0,128	0,264	0,255
Tsallis Entropileri	$\alpha=0,33$	$\alpha=0,5$	$\alpha=2$	$\alpha=3$

Y'nin Entropisi	0,8059	0,727	0,369	0,277
X verildiğinde Y'nin koşullu entropisi	0,721	0,623	0,275	0,206
C(Y,X)	0,1	0,142	0,253	0,253

Shannon, Rényi ve Tsallis Entropileri için hesaplanan bilgiler istatistik sonuçlarını analiz ettiğimizde, değişken tarafından oluşturulan bilgiler düşük olduğu için entropi yüksek çıkmaktadır. Bu yüzden ürün grubu ve satılan ürün türü arasındaki ilişki oldukça düşük olduğu sonucuna varabilir.

3.5.2 Metrik Veriye Uygulama

Klasik doğrusal korelasyon katsayıları ve korelasyon oranları arasında bilgi istatistiklerini anlamlı karşılaştırmalar yapmak için üç farklı hipotetik metrik veri ile çalışmalarımızı yapmayı amaçladık.

Büyük olasılıklar aşağıda verilen kontenjans tablosunun ana diyagonal köşegeni üzerinde aşağıda verilen tabloda hızlı yerleştirilmiştir, X ve Y değişkenleri arasındaki yüksek pozitif korelasyon yansıtacaktır:

Çizelge 3. 10 X ve Y metrik değişkenlerinin yüksek ve pozitif korelasyonlu olasılık dağılımı

Dağılım 1			
X/Y	0	1	2
0	0,19	0,005	0,005
1	0,005	0,39	0,005
2	0,005	0,005	0,39

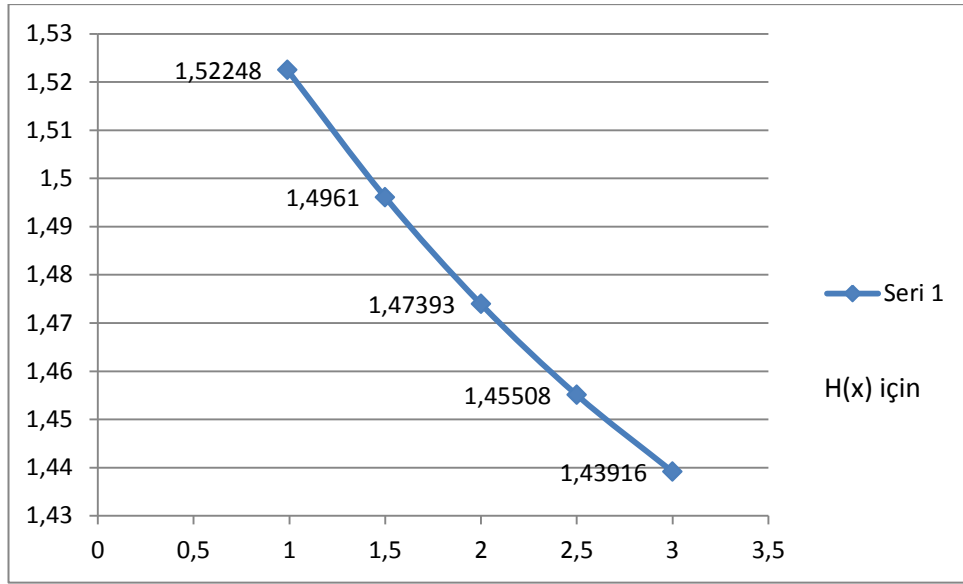
Çizelge 3. 11 Shannon entropisi ile hesaplanan pozitif korelasyonlu bilgi ölçüleri

Cor(X,Y)	Eta1(Y/X)	Eta2(X/Y)	Shannon	H(X,Y)	H(X)	H(Y)	I(X,Y)	U(X,Y)	U(Y,X)
0,94643	0,9465	0,9465	1,52193	1,7441	1,522	1,522	1,2997	0,854	0,854

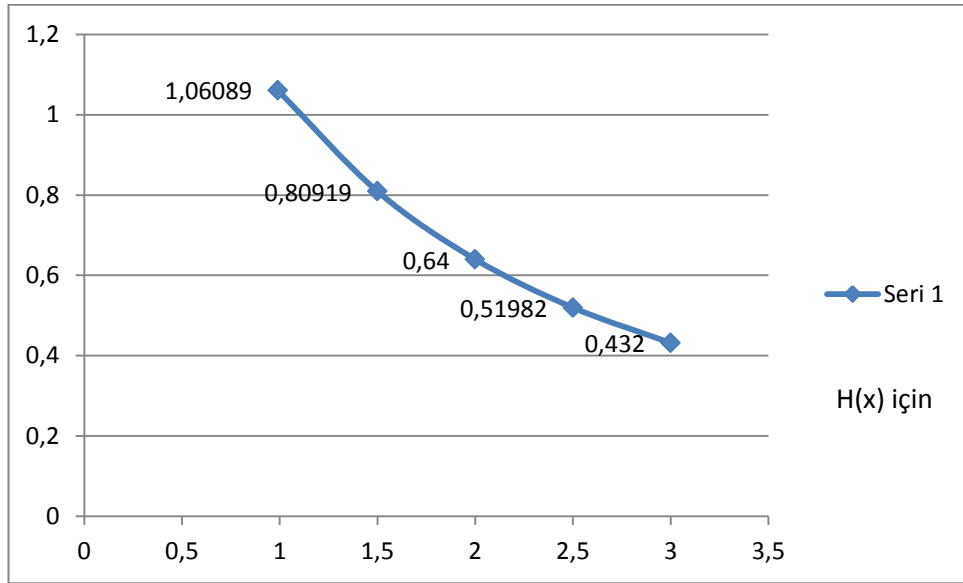
İşte bu iki doğrusal ilişki Eta1 (Y / X) ve Eta2 (X / Y) boyutlarında olduğunu belirtmek gerekir. İkinci bağımsız değişken olarak X'i alırsak "Eta bağımsız değişken olarak X görmektedir. Buradan klasik korelasyon ölçüsü ve bilgi ölçüleri belli bir dereceye kadar kabul edilebilir sonucuna varabiliriz.

Çizelge 3. 12 Renyi ve tsallis entropilerinin farklı parametreler altında karşılıklı bilgi istatistikleri

Alpha	Entropy	H(X,Y)	H(X)	H(Y)	H(X/Y)	H(Y/X)	I(X,Y)	U(X,Y)	U(Y,X)
0,99	Rényi	1,74849	1,52248	1,5225	0,068	0,068	1,2965	0,852	0,8516
0,99	Tsallis	1,21933	1,06089	1,0609	0,157	0,157	0,9025	0,851	0,8507
1,5	Rényi	1,61157	1,4961	1,4961	0,036	0,036	1,3806	0,923	0,9228
1,5	Tsallis	0,8559	0,80919	0,8092	0,082	0,082	0,7625	0,942	0,9423
2	Rényi	1,55449	1,47393	1,4739	0,026	0,026	1,3934	0,945	0,9453
2	Tsallis	0,65955	0,64	0,64	0,059	0,059	0,6205	0,969	0,9695
2,5	Rényi	1,52084	1,45508	1,4551	0,022	0,022	1,433	0,955	0,9548
2,5	Tsallis	0,52952	0,51982	0,5198	0,049	0,049	0,4711	0,981	0,9813
3	Rényi	1,49713	1,43916	1,4392	0,02	0,02	1,3812	0,96	0,9597
3	Tsallis	0,43725	0,432	0,432	0,044	0,044	0,4267	0,988	0,9878



Şekil 3. 2 Renyi entropilerinin farklı parametrelerdeki değişim grafiği



Şekil 3. 3 Tsallis entropilerinin farklı parametrelerdeki değişim grafiği

Çizelge 3. 13 X ve Y metrik değişkenlerinin korelasyonsuz olasılık dağılımı

Dağılım 2			
X/Y	0	1	2
0	0,2	0,05	0,15
1	0,05	0,1	0,05

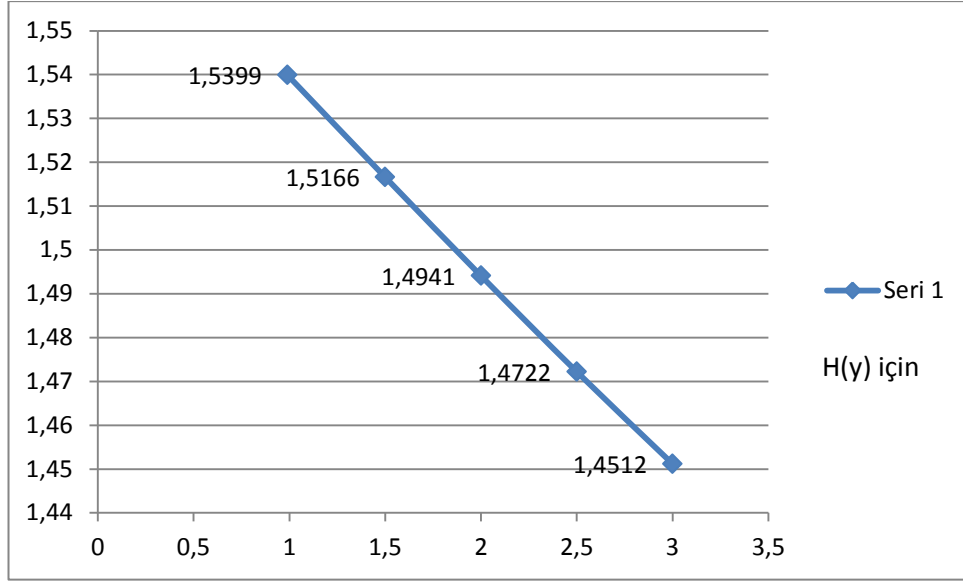
2	0,05	0,3	0,05
---	------	-----	------

Çizelge 3. 14 X ve Y metrik değişkenlerinin shannon entropisi ile hesaplanan korelasyonsuz bilgi ölçüleri

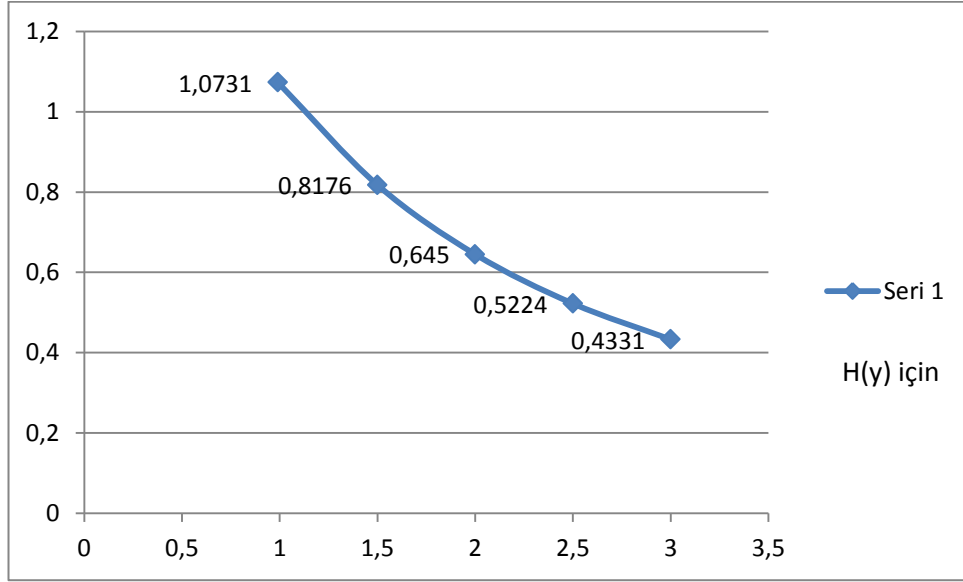
Cor(X,Y)	Eta1(Y/X)	Eta2(X/Y)	Shannon	H(X,Y)	H(X)	H(Y)	I(X,Y)	U(X,Y)	U(Y,X)
0,07555	0,08276	0,56335	1,52193	2,8087	1,522	1,539	0,2527	0,166	0,1642

Çizelge 3. 15 X ve Y metrik değişkenlerinin renyi ve tsallis entropilerinin farklı parametreler altında korelasyonsuz karşılıklı bilgi istatistikleri

Alpha	Entropy	H(X,Y)	H(X)	H(Y)	H(X/Y)	H(Y/X)	I(X,Y)	U(X,Y)	U(Y,X)
0,99	Rényi	2,81214	1,52248	1,5399	0,383	0,388	0,2503	0,164	0,1625
0,99	Tsallis	1,96835	1,06089	1,0731	0,886	0,898	0,1657	0,156	0,1544
1,5	Rényi	2,64834	1,4961	1,5166	0,343	0,355	0,3644	0,244	0,2403
1,5	Tsallis	1,20124	0,80919	0,8176	0,652	0,666	0,4256	0,526	0,5205
2	Rényi	2,51457	1,47393	1,4941	0,313	0,332	0,4535	0,308	0,3035
2	Tsallis	0,825	0,64	0,645	0,512	0,525	0,46	0,719	0,7132
2,5	Rényi	2,40719	1,45508	1,4722	0,289	0,316	0,5201	0,357	0,3533
2,5	Tsallis	0,6121	0,51982	0,5224	0,42	0,43	0,4301	0,827	0,8234
3	Rényi	2,32193	1,43916	1,4512	0,271	0,304	0,5684	0,395	0,3917
3	Tsallis	0,48	0,432	0,4331	0,356	0,363	0,3851	0,891	0,8892



Şekil 3. 4 Renyi entropilerinin farklı parametrelerdeki değişim grafiği



Şekil 3. 5 Tsallis entropilerinin farklı parametrelerdeki değişim grafiği

Bu durum içinde klasik korelasyon ölçüleri ve bilgi tabanlı entropi ölçülerinin benzer şekilde davrandığını söyleyebiliriz. Ayrıca biz değişkenler arasındaki bağımlılıkları analiz ederken, 2 tane parametrik temsil kullandığımızı dikkat etmeliyiz. Çünkü Renyi ve Tsallis yöntemlerden biriyle yapılan entropi tahminler seçilen alfa değerlerine bağlıdır. Renyi entropi kullanarak X ve Y değişkenleri arasında düşük değerli α değerleri için ($\alpha=0,99$ değeri için) orta derece bir ilişki vardır. Bu durum Tsallis entropisi için biraz daha gariptir: X ve Y değişkenlerinin arasındaki ilişki $\alpha=0,99$ değeri için çok az iken, tersine α değerleri arttıkça orta dereceli ilişkilere ulaşılacağı görülür. Nitekim $\alpha=3$

değeri için Tsallis entropisine göre X ve Y değişkenleri arasında orta dereceli bir ilişki olduğu sunucuna varılır. Bu, dikkate alınması gereken bir faktördür. Özellikle, Tsallis entropileri, Rényi ve Shannon entropilerine göre farklı parametrelere çok duyarlıdır.

Çizelge 3. 16 X ve Y metrik değişkenlerinin yüksek ve negatif korelasyonlu olasılık dağılımı

Dağılım 3			
X/Y	0	1	2
0	0,01	0,01	0,38
1	0,01	0,28	0,01
2	0,28	0,01	0,01

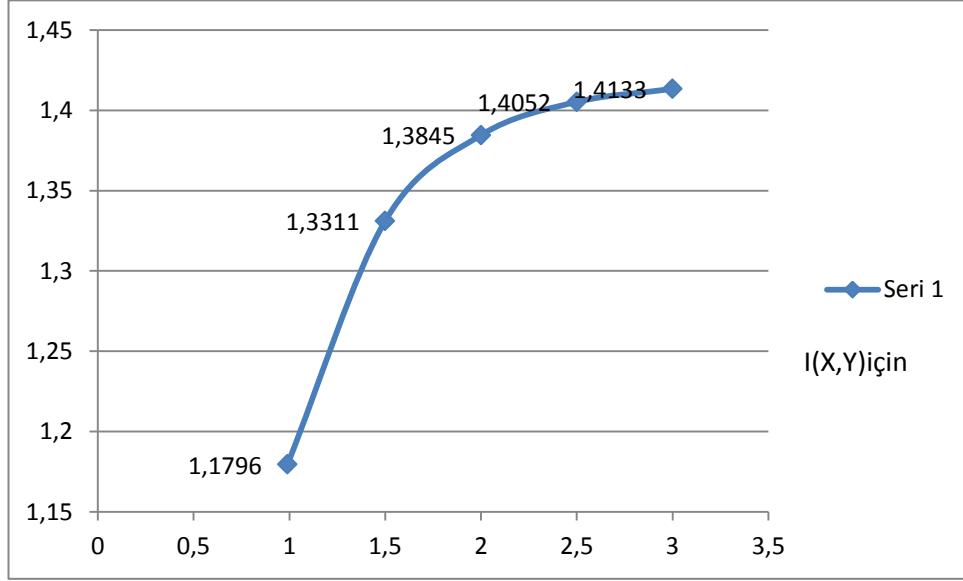
Çizelge 3. 17 X ve Y metrik değişkenlerinin shannon entropisi ile hesaplanan yüksek fakat negatif korelasyonlu bilgi ölçüleri

Cor(X,Y)	Eta1(Y/X)	Eta2(X/Y)	Shannon	H(X,Y)	H(X)	H(Y)	I(X,Y)	U(X,Y)	U(Y,X)
-0,913	0,91307	0,91307	1,57095	1,9575	1,571	1,571	1,1844	0,754	0,7539

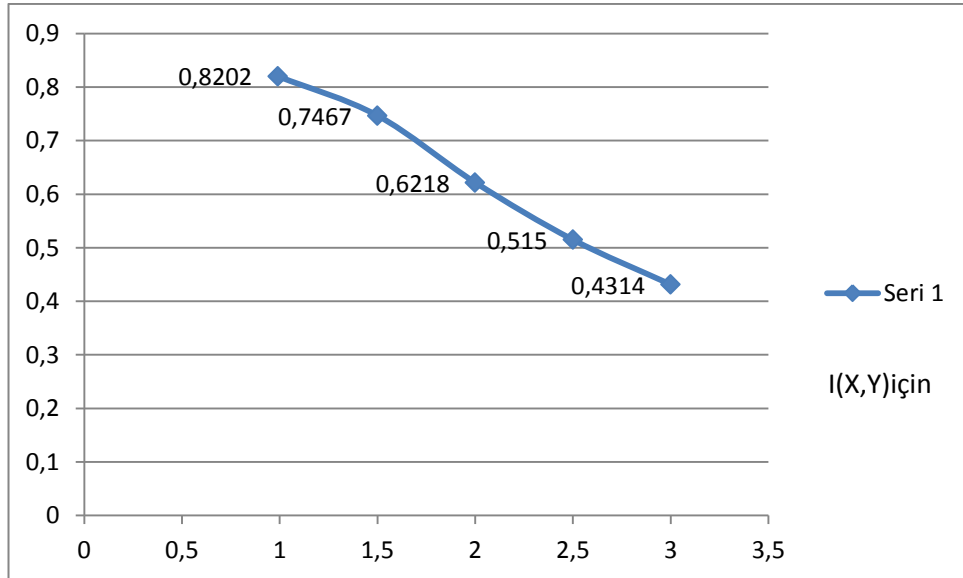
Çizelge 3. 18 X ve Y metrik değişkenlerinin renyi ve tsallis entropilerinin farklı parametreler altında korelasyonsuz karşılıklı bilgi istatistikleri

Alpha	Entropy	H(X,Y)	H(X)	H(Y)	H(X/Y)	H(Y/X)	I(X,Y)	U(X,Y)	U(Y,X)
0,99	Rényi	1,96258	1,57109	1,5711	0,118	0,118	1,1796	0,751	0,7508
0,99	Tsallis	1,36966	1,09495	1,095	0,272	0,272	0,8202	0,749	0,7491
1,5	Rényi	1,79631	1,56372	1,5637	0,071	0,071	1,3311	0,851	0,8513
1,5	Tsallis	0,92686	0,83677	0,8368	0,156	0,156	0,7467	0,892	0,8923
2	Rényi	1,72834	1,55639	1,5564	0,053	0,053	1,3845	0,89	0,8895
2	Tsallis	0,6982	0,66	0,66	0,115	0,115	0,6218	0,942	0,9421

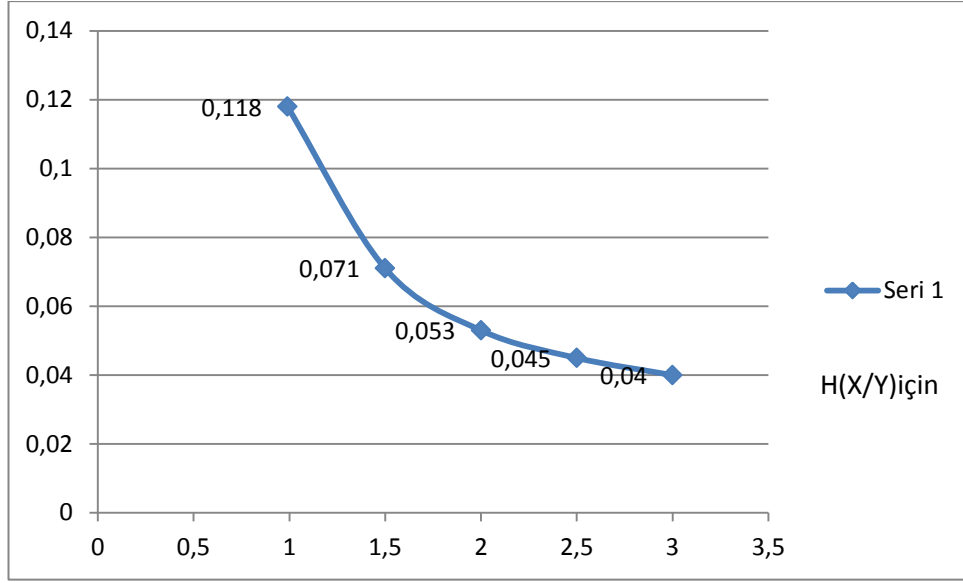
2,5	Rényi	1,69276	1,549	1,549	0,045	0,045	1,4052	0,907	0,9072
2,5	Tsallis	0,55197	0,53348	0,5335	0,095	0,095	0,515	0,965	0,9653
3	Rényi	1,6698	1,54157	1,5416	0,04	0,04	1,4133	0,917	0,9168
3	Tsallis	0,45061	0,441	0,441	0,085	0,085	0,4314	0,978	0,9782



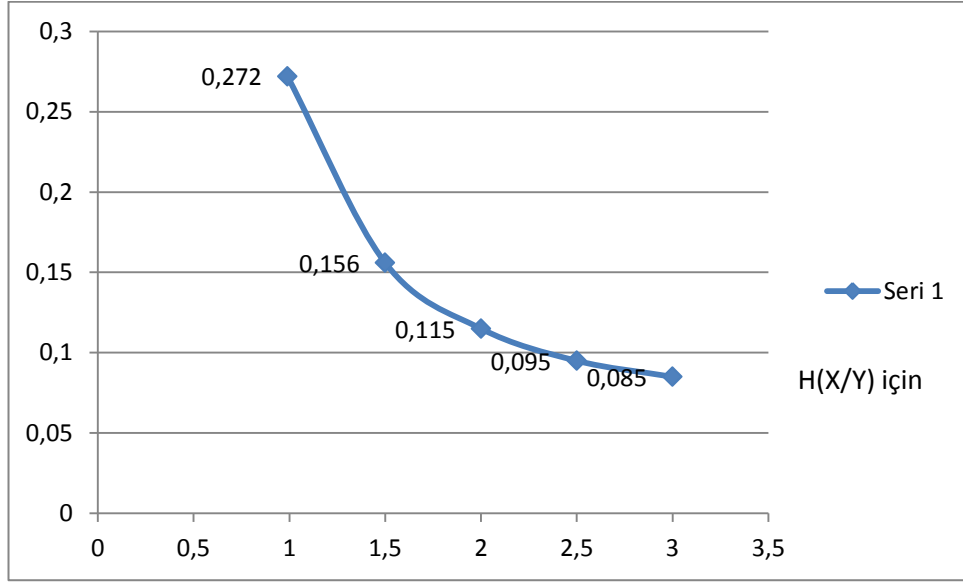
Şekil 3. 6 Rényi entropilerinin farklı parametrelerdeki değişim grafiği



Şekil 3. 7 Tsallis entropilerinin farklı parametrelerdeki değişim grafiği



Şekil 3. 8 Renyi entropilerinin farklı parametrelerdeki değişim grafiği



Şekil 3. 9 Tsallis entropilerinin farklı parametrelerdeki değişim grafiği

Burada klasik korelasyon ile karşılıklı bilgi ölçüsü arasındaki paralelliği izleyebiliriz.

3.5.3 Nominal ve Metrik Verilerin Uygulama Sonuçları

Bu çalışmada, biz nominal ve metrik verilere entropi örnekleri uygulayarak klasik korelasyon ölçüsü ile entropi tabanlı ilişki ölçüleri arasındaki paralellik gösterilmek istendi. Değişkenler nitel olduğu zaman, biz klasik korelasyon ölçülerini kullanamayız. Bu nedenle karşılıklı bilgi ölçüleri bağımlılıkların analizinde etkili olabilir. Genellikle (Shannon, Renyi veya Tsallis entropilerinin birini kullanarak) koşullu entropi ve karşılıklı

bilgi ölçüsü hesaplamaları Shannon ve Renyi entropilerinde benzer sonuçlar üretirken Tsallis entropisinde bu durum farklılık gösterir. Tsallis entropileri parametrelere çok duyarlı oldukları için nominal ilişkilendirme hususlarında yanıltıcı sonuçlar verebilir.

BÖLÜM 4

KULLBACK LEİBLER UZAKLIĞI

Şu ana kadar $X \equiv \begin{pmatrix} x_1, x_2, \dots, x_n \\ p_1, p_2, \dots, p_n \end{pmatrix}$ gibi kesikli bir sistem veya rasgele değişken için entropi ve bilgi içeriği kavramları ($x_1, x_2, \dots, x_n$ durumları üzerinde $p_1, p_2, \dots, p_n$ olasılık dağılımının veya kısaca $p_1, p_2, \dots, p_n$ olasılık dağılımının entropisi ve bilgi içeriği kavramları) ele alınmıştır.

$$H(X)=I(X)=-\sum_{i=1}^n p_i \log p_i = -\sum_{i=1}^n f_x(x_i) \ln f_x(x_i) = -E [\ln f_x(X)] \quad (4.1)$$

ve sürekli rasgele değişkenler (sistemler) için

$$H(X) = - \int_{-\infty}^{\infty} f_x(x) \ln f_x(x) dx \quad (4.2)$$

ve

$$I(X) = - \int_{-\infty}^{\infty} f(x) \ln f(x) dx = E [f_x(X)] \quad (4.3)$$

olarak tanımlanmıştır.

Yukarıdaki $H(X)$ ve $I(X)$ gösterimi yerine $H(f_x)$ ve $I(f_x)$ gösterimi de kullanılabilir,

yani

$$H(f_x) = I(f_x) = E [-\ln f_x(X)] \quad (4.4)$$

dır. Şimdi, yukarıdaki kesikli sistemin $x_1, x_2, \dots, x_n$ durumları üzerinde başka bir

$$g_x(x_i) = P(X=x_i)=q_i, \quad i=1,2, \dots, n$$

olasılık dağılımının da söz konusu olduğunu düşünülebilir. Bu dağılımın bilgi içeriği ve entropisi,

$$I(g_x) = H(g_x) = E [-\ln g_x(X)] = -\sum_{i=1}^n g_x(x_i) \ln g_x(x_i) = \sum_{i=1}^n q_i \log q_i \quad (4.5)$$

dır. Aynı $x_1, x_2, \dots, x_n$ durumları üzerinde söz konusu olan f_x ve g_x gibi iki dağılımı karşılaştırmak için tanımlanan,

$$D_{KL}(f_x \| g_x) = E_{f_x} \left[\ln \frac{f_x(X)}{g_x(X)} \right] = \sum_{i=1}^n f_x(x_i) \ln \frac{f_x(x_i)}{g_x(x_i)} = \sum_{i=1}^n p_i \ln \frac{p_i}{q_i} \quad (4.6)$$

değerine Kullback-Leibler uzaklığı (Kullback-Leibler bilgi kazancı, görelî entropi) denir[27].

Sürekli durumda Kullback-Leibler uzaklığı,

$$D_{KL}(f_x \| g_x) = E_{f_x} \left[\ln \frac{f_x(X)}{g_x(X)} \right] = \int_{-\infty}^{\infty} f_x(x) \ln \frac{f_x(x)}{g_x(x)} dx \quad (4.7)$$

dır. Genelde, $D_{KL}(f_x \| g_x) \neq D_{KL}(g_x \| f_x)$ olduğundan Kullback-Leibler uzaklığı bir metrik değildir.

f_x ve g_x bir X sistemi (rasgele değişkeni) ile ilgili iki hipotez ($H_0 : X \sim f_x, H_1 : X \sim g_x$) altındaki dağılımları gösterdiğinde Kullback-Leibler uzaklığı

$$\ln \frac{f_x(X)}{g_x(X)} \quad (4.8)$$

karşıtlık oranının logaritmasının sıfır hipotezi altında beklenen değeridir.

(X, Y) bileşik sistemi, başka bir ifade ile (X, Y) rasgele vektörü ile ilgili öne sürülen $f_{X,Y}$ ve $g_{X,Y}$ gibi iki dağılım için Kullback-Leibler uzaklığı [26],

$$D_{KL}(f_{X,Y} \| g_{X,Y}) = E_{f_{X,Y}} \left[\ln \frac{f_{X,Y}(X,Y)}{g_{X,Y}(X,Y)} \right] \quad (4.9)$$

olarak tanımlanır.

(X, Y) bileşik sisteminin $g_{X,Y}$ dağılımı altında bileşenleri bağımsız, yani $g_{X,Y} = g_X g_Y$ ise,

$$\begin{aligned} D_{KL}(f_{X,Y} \| g_{X,Y}) &= E_{f_{X,Y}} \left[\ln \frac{f_{X,Y}(X,Y)}{g_X(X)g_Y(Y)} \right] = \left[\ln \frac{f_{X,Y}(X,Y)}{g_X(X)} \right] - E_{f_{X,Y}} [\ln g_Y(Y)] \\ &= E_{f_{X,Y}} [\ln f_{X/Y}(Y/X)] + H(Y) = H(Y) - H(X/Y) = I(X, Y) \end{aligned} \quad (4.10)$$

olup, bu durumda $D_{KL}(f_{X,Y} \| g_{X,Y})$ Kullback-Leibler uzaklığı aynı zamanda $I(X, Y)$ karşılıklı bilgi değeridir [27].

N tane durumu olan kesikli (X, Y) bileşik sistemi için $g_{X,Y}$ dağılımı düzgün dağılım olduğunda,

$$\begin{aligned}
D_{KL}(f_{X,Y} || g_{X,Y}) &= E_{f_{X,Y}} \left[\ln \frac{f_{X,Y}(X,Y)}{g_X(X)g_Y(Y)} \right] \\
&= E_{f_{X,Y}} \left[\ln \frac{f_{X,Y}(X,Y)}{\frac{1}{N}} \right] = \ln N + E_{f_{X,Y}} [\ln f_{X,Y}(X,Y)] \\
&= \ln N - H(f_{X,Y})
\end{aligned} \tag{4.11}$$

olur.

4.1 Metrik Tanımı ve Kullback Leibler Uzaklığının Metrik Fonksiyon Olup Olmadığının İncelenmesi

X boş olmayan bir küme olsun. Bu küme üzerinde reel değerli, negatif olmayan bir

$d: X \times X \rightarrow R^+$ fonksiyonu aşağıdaki aksiyomları sağlasın:

- i. Her $x, y \in X$ için $d(x, y) \geq 0$.
- ii. Her $x, y \in X$ için ancak ve ancak $x = y$ ise $d(x, y) = 0$
- iii. Her $x, y, \in X$ için $d(x, y) = d(y, x)$ (simetri)
- iv. Her $x, y, z \in X$ için $d(x, y) \leq d(x, z) + d(z, y)$ (üçgen eşitsizliği)

Böyle bir $d(x, y)$ fonksiyonuna X kümesi üzerinde bir metrik adını verecek ve X kümesinin bundan böyle nokta adını vereceğimiz x ve y gibi elemanları arasındaki uzaklık olarak yorumlayacağız.

İki kesikli dağılım arasındaki Kullback-Leibler mesafesi (ya da görelî entropi) (4.6) denklemindeki gibi tanımlandığından bu fonksiyonun metrik olup olmadığını inceleyebiliriz.

Kullback leibler uzaklığı

Her $x, y, \in X$ için $d(x, y) = d(y, x)$ (simetri) özelliğini yani;

$D(p || q) = D(q || p)$ eşitliğini sağlamadığı için, yani İki dağılım özdeş olduğunda bu mesafe sıfıra eşit olur. Bu ölçü genellikle asimetriktir. Başka bir ifade ile

$$D(p || q) \neq D(q || p) \tag{4.12}$$

olduğundan dolayı metrik fonksiyon değildir.

4.2 Göreli Entropi ve Kullback-Leibler Mesafesi Arasındaki İlişki

$D(p \parallel q)$ ya da göreli entropi gerçek dağılım “p” iken, dağılımın “q” sayılmasından kaynaklanan yetersizliğin bir ölçüsüdür. Aynı zamanda “p” ve “q” gibi iki dağılım arasındaki mesafenin bir ölçüsü olarak da değerlendirilebilir. Göreli entropi olasılık için bir geometri tanımlamak için kullanılabilir. Göreli entropi büyük sonuçlar hakkında birçok yorum yapmamızı sağlar[29]. İki kesikli dağılım arasındaki Kullback-Leibler mesafesi (ya da göreli entropi) (4.6) denkleminde tanımlanmıştır. İki dağılım özdeş olduğunda bu mesafe sıfıra eşit olur. Bu ölçü genellikle asimetriktir. Başka bir ifade ile (4.12) deki eşitsizliği sağlar.

Bu istatistik, belirli düzenlilik koşullarında asimptotik olarak ki-kare dağılmaktadır. Bu konu ile ilgili olarak [28]’e başvurulabilir. Son olarak karşılıklı bilgi ölçüsü ile göreli entropi arasında bir paralellik kurulabilir. Karşılıklı bilgi miktarı, birleşik olasılık fonksiyonu ile iki marjinal olasılık fonksiyonunun çarpımı arasındaki Kullback-Leibler mesafesi ya da göreli entropiye eşit olmaktadır. Başka bir ifade ile

$$\begin{aligned} I(X;Y) &= \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= D(p(x,y) \parallel p(x)p(y)) \\ &= E_{p(x,y)} \log \frac{p(x,y)}{p(x)p(y)} \end{aligned} \quad (4.13)$$

Örnek 4.1 $X=(0,1)$ aralığında tanımlı bir küme ve p ile q bu kümenin 2 dağılımı olsun. $p(0)=1-r$, $p(1)=r$ ve $q(0)=1-s, q(1)=s$ şeklinde tanımlansın. Bu 2 kesikli dağılımın arasındaki karşılıklı bilgi miktarını inceleyelim.

Çözüm

$$D(p \parallel q) = (1-r) \log \frac{1-r}{1-s} + r \log \frac{r}{s}$$

$$D(q \parallel p) = (1-s) \log \frac{1-s}{1-r} + s \log \frac{s}{r} \quad \text{olduğundan}$$

Eğer $r=s$ ise $D(p \parallel q)=D(q \parallel p)=0$ olur. Eğer $r=1/2$ ve $s=1/4$ alıp Kullback Leibler mesafelerini hesaplarsak:

$$D(p \parallel q) = \left(\frac{1}{2}\right) \log \frac{\frac{1}{3}}{\frac{1}{4}} + \frac{1}{2} \log \frac{\frac{1}{2}}{\frac{1}{4}} = 1 - \frac{1}{2} \log 3 = 0,2075 \text{ bit}$$

$$D(q \parallel p) = \left(\frac{3}{4}\right) \log \frac{\frac{3}{4}}{\frac{1}{2}} + \frac{1}{4} \log \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{3}{4} \log 2 - \frac{1}{4} = 0,1817 \text{ bit çıkar.}$$

Buradan da $D(p \parallel q) \neq D(q \parallel p)$ olduğundan dolayı Kullback Leibler uzaklığı metrik fonksiyon değildir.

4.3 Kullback-Leibler Sapmasının Simetrik Bir Versiyonu: Jeffrey Sapması

Jeffrey sapması

$$D_J = D_{KL}(p \parallel q) + D_{KL}(q \parallel p) \quad (4.14)$$

şeklindedir [30]. İki kesikli dağılım p ve q için Jeffrey sapması

$$D_J = \sum (p(x) - q(x)) \log \left(\frac{p(x)}{q(x)} \right) \quad (4.15)$$

olur [2]. Sürekli haller için de benzer bir biçimde Jeffrey sapması tanımlanabilir:

$$D_J = \int_{-\infty}^{\infty} (f(x) - g(x)) \log \left(\frac{f(x)}{g(x)} \right) dx \quad (4.16)$$

Jeffrey sapması da bir metrik olmanın üçüncü koşulunu (üçgen eşitsizliği) yerine getirmediği için mesafe olarak değil de “sapma” (divergence) olarak adlandırılmaktadır [28].

4.4 Seçilmiş Bazı Dağılımlar için Kullback Leibler ve Jeffrey Sapmaları

1) Bernoulli: $X \sim (\text{Bernoulli}(p_1))$ ve $Y \sim (\text{Bernoulli}(p_2))$ için

$$D_{KL} = (1 - p_1) \log \frac{(1-p_1)}{(1-p_2)} + p_1 \log \frac{p_1}{p_2} \quad (4.17)$$

$$D_J = (p_1 - p_2) \left[\log \frac{(1-p_2)}{p_2} - \log \frac{1-p_1}{p_1} \right] \quad (4.18)$$

2) Binom : $X \sim (\text{Binom}(p_1; n))$ ve $Y \sim (\text{Binom}(p_2; n))$ için

$$D_{KL} = n \left[(1 - p_1) \log \frac{(1-p_1)}{(1-p_2)} + p_1 \log \frac{p_1}{p_2} \right] \quad (4.19)$$

$$D_J = n(p_1 - p_2) \left[\log \frac{(1-p_2)}{p_2} - \log \frac{1-p_1}{p_1} \right] \quad (4.20)$$

3)Poisson: $X \sim \text{Poisson}(\lambda_1)$ ve $Y \sim \text{Poisson}(\lambda_2)$ için

$$D_{KL} = (\lambda_2 - \lambda_1) + \lambda_1 \ln \frac{\lambda_1}{\lambda_2} \quad (4.21)$$

$$D_J = (\lambda_1 - \lambda_2) \ln \frac{\lambda_1}{\lambda_2} \quad (4.22)$$

4) Normal : $X \sim (\text{Normal}(\mu_X, \sigma_X^2))$ $Y \sim (\text{Normal}(\mu_Y, \sigma_Y^2))$ için

$$D_{KL} = \frac{1}{2} \left[2 \ln \frac{\sigma_Y}{\sigma_X} + \frac{\sigma_X^2}{\sigma_Y^2} + \frac{(\mu_X - \mu_Y)^2}{\sigma_Y^2} - 1 \right] \quad (4.23)$$

$$D_J = \frac{(\sigma_X^2 + \sigma_Y^2)}{2\sigma_X^2\sigma_Y^2} [(\sigma_X^2 - \sigma_Y^2) + (\mu_X - \mu_Y)^2] \quad (4.24)$$

İki dağılımın varyansları eşit kabul edilerek $\sigma^2 = \sigma_1^2 = \sigma_2^2$ deneyecek olursa

$$D_J = \frac{(\mu_1 - \mu_2)^2}{\sigma^2} \quad (4.25)$$

şeklinde daha yalın bir formülasyona ulaşılmış olur.

5)Üstel: $X \sim \text{Üstel}(\lambda_1)$ $X \sim \text{Üstel}(\lambda_2)$ için

$$f_1(x) = \lambda_1 e^{-\lambda_1 x} \quad f_2(x) = \lambda_2 e^{-\lambda_2 x} \quad x > 0 \text{ için}$$

$$D_{KL} = \frac{(\lambda_2 - \lambda_1)}{\lambda_1} + \ln \frac{\lambda_1}{\lambda_2} \quad (4.26)$$

$$D_J = \frac{(\lambda_2 - \lambda_1)^2}{\lambda_1 \lambda_2} \quad (4.27)$$

6)Gama : $X \sim \text{Gama}(\theta_0, \beta_0)$ $X \sim \text{Gama}(\theta, \beta)$ olsun. Başka bir deyişle

$$f_1(x) = \frac{\beta_0^{\theta_0} x^{\theta_0-1} \exp(-\beta_0 x)}{\Gamma(\theta_0)} \text{ ve } f_2(x) = \frac{\beta^{\theta} x^{\theta-1} \exp(-\beta x)}{\Gamma(\theta)} \quad \beta, \beta_0, \theta, \theta_0 > 0, X \geq 0 \text{ olsun.}$$

f_1 sıfır hipotezinde geçerli kabul edilen dağılım, f_2 de bu dağılıma uygunluğu test edilen dağılım olduğuna göre dağılımlar arasındaki Kullback-Leibler sapması [26]

$$D_{KL} = \ln \left[\frac{\beta_0^{\theta_0} \Gamma(\theta)}{\beta^{\theta} \Gamma(\theta_0)} \right] + (\theta_0 - \theta) \Psi(\theta_0) - \ln(\beta_0) + \theta_0 \frac{\beta - \beta_0}{\beta_0} \quad (4.28)$$

şeklindedir.

7. Beta: $X \sim \text{Beta}(\theta_0, \beta_0)$ $X \sim \text{Beta}(\theta, \beta)$ $0 \leq X \leq 1$ olsun.

$$D_{KL} = \ln \left[\frac{B(\alpha_0, \beta_0)}{B(\alpha, \beta)} \right] - (\alpha_0 - \alpha) \Psi(\alpha) - (\beta_0 - \beta) \Psi(\beta) + (\alpha_0 - \alpha + \beta_0 - \beta) \Psi(\alpha + \beta) \quad (4.29)$$

(α_0, β_0) parametrelili Beta dağılımı H_0 hipotezinde geçerli olduğu düşünülen dağılımı temsil etmektedir.

Örnek 4.2 Bir para 50 kere atılıyor. 15 kere tura gelme olasılığını Ki-Kare testi ile belirleyip bunları Kullback Leibler sapması ile hesaplayıp karşılaştıralım.

Çözüm:

$$H_0: \text{Para hilesiz} = P(T) = \frac{1}{2}$$

$$H_A: \text{Para hileli} \quad P(T) \neq \frac{1}{2}$$

DURUM	TURA	YAZI
O_i	15	35
e_i	25	25

O_i : Gözlenen frekanslar

e_i : Beklenen frekanslar

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - e_i)^2}{e_i}$$

$$\chi^2 = \left(\frac{(15-25)^2}{25} \right) + \left(\frac{(35-25)^2}{25} \right)$$

$$= 4+4$$

$$= 8$$

İki Bernoulli dağılımını Kullback Leibler sapmasıyla hesaplırsak

$$D_{KL} = \frac{1}{2} \cdot \log_2 \frac{\frac{1}{2}}{\frac{15}{50}} + \frac{1}{2} \cdot \log_2 \frac{\frac{1}{2}}{\frac{35}{50}}$$

$$= \frac{1}{2} \cdot \log_2 \frac{5}{3} + \frac{1}{2} \cdot \log_2 \frac{5}{7}$$

$$= \frac{1}{2} \cdot 0,74 + \frac{1}{2} \cdot (-0,48) = 0,13$$

$$\chi^2 = 0,13 \times 50 = 6,5$$

İki Bernoulli dağılımını Jeffrey sapmasıyla hesaplırsak:

$$D_J = \left(\frac{1}{2} - \frac{35}{50} \right) \left[\log \frac{\left(\frac{15}{50} \right)}{\frac{35}{50}} - \log \frac{\frac{1}{2}}{\frac{1}{2}} \right]$$

$$= (-0,2)(-1,22239)$$

$$= 0,244418 \cong 0,25$$

$$\chi^2 = 0,25 \times 50 = 12,5$$

İki dağılımın farklı olduğu ortaya çıkmaktadır. Çünkü elde edilen χ^2 değeri oldukça büyüktür.

Örnek 4.3 Bir fabrikada üretilen ürünlerin %2 si defoludur. Seçilen 100 üründen 3 tanesinin kusurlu olma olasılığını Poisson dağılımı ile belirleyip Kullback Leibler sapması ile karşılaştıralım.

Çözüm: $n=100$ ve $p=0,02$ olduğu için $\lambda = n.p = 100.0,02 = 2$ bulunur. Böylece seçilen 100 üründen 3 tanesinin kusurlu üretilme olasılığı Poisson Dağılımından

$$P(3) = \frac{2^3 \cdot e^{-2}}{3!} = 0,18 \text{ bulunur.}$$

Kullback Leibler sapması ile hesaplarsak:

$$H_0 = 2$$

$$H_A \neq 2$$

olduğundan $\lambda_1 = 2$ ve $\widehat{\lambda}_2 = 3$ olarak alırsak

$$D_{KL} = (3-2) + 2 \cdot \ln \frac{2}{3}$$

$$= 0,18907 \text{ olarak bulunur.}$$

$$D_J = (2-3) \ln \frac{2}{3} = 0,40547 \text{ olarak bulunur.}$$

Buradan da görüldüğü gibi Kullback Leibler sapması ve Jeffry sapmalarını kullanarak H_0 red edilir.

Örnek 4.4 Bir fabrikadaki bir üretilen bir ürünün üretim ortalaması 0,448 , varyansı 0,863 tür. Bu ürün üretimde standart normal dağıldığına göre üretimi Kullback Leibler sapmasıyla inceleyelim.

Çözüm:

$\mu_1=0, \sigma_1^2=1, \mu_2=0,448, \sigma_2^2=0,863$ verildiğine göre

$$D_{KL} = \frac{1}{2} \left[2 \ln \sqrt{\frac{1}{0,863}} + \frac{1}{0,863} + \frac{(1-0,448)^2}{\sqrt{0,863}} - 1 \right]$$

$$= \frac{1}{2} [0,294 + 1,158 + 0,327 - 1]$$

$$= 0,39 \text{ bulunur.}$$

Örnek 4. 5 Bir fabrikada üretilen ürünün kullanım süresi yaklaşık olarak 12,5 yıl olan bir üstel dağılımla temsil edildiğine göre bu üründen bir tane satın alan müşterinin 25 yıl kullanma olasılığını bulunuz ve Kullback Leibler ve Jeffrey sapmaları ile de karşılaştırınız.

Çözüm:

Üstel dağılımın beklenen değer tanımından

$$E_x = \frac{1}{\alpha} = \frac{1}{12,5} \text{ olarak bulunur.}$$

$f(x)$ olasılık yoğunluk fonksiyonu ise;

$$f(x, \alpha) = \frac{1}{12,5} e^{-\frac{x}{12,5}}$$

$$Pr\{x \geq 25\} = \int_0^{25} \frac{1}{12,5} e^{-\frac{x}{12,5}} dx = \frac{1}{12,5} (-12,5 e^{-\frac{x}{12,5}}) \Big|_0^{25}$$

$$= 1 - e^{-2} = 0,14$$

$$D_{KL} = \frac{(\lambda_2 - \lambda_1)}{\lambda_1} + \ln \frac{\lambda_1}{\lambda_2}$$

$$D_{KL} = \frac{(25 - 12,5)}{12,5} + \ln \frac{12,5}{25}$$

$$= 0,30$$

$$D_J = \frac{(\lambda_2 - \lambda_1)^2}{\lambda_1 \lambda_2} = \frac{(25 - 12,5)^2}{12,5 \cdot 25} = 0,50 \text{ olarak bulunur.}$$

NİTEL DEĞİŞKENLİK ÖLÇÜLERİ

5.1 Nitel Değişkenlik Ölçülerine Giriş

Kader ve Perry temel istatistik kitaplarının nitel değişkenlik ölçüleri konusuna fazla önem vermediğini vurgulamaktadır. Yine de bu trende uymayan bazı kaynaklar vardır [31]. Wilcox, bir nitel değişkenlik ölçüsünün sahip olması arzu edilen özelliklerini şu şekilde özetlemektedir: İndeksin sahip olacağı en küçük ve en büyük değerler nitel dağılımın sahip olacağı kategori sayısına ve kategorilerin büyüklüklerine bağlı olmamalıdır. Ayrıca bir indekxin standart bir değer aralığı bulunmalıdır. Bu koşul özellikle farklı indekslerin birbirleri ile kıyaslanabilmesi açısından önemlidir [32]. Wilcox konu ile ilgili kapsamlı çalışmasında bir indekxin anlamlı bir nitel değişkenlik ölçüsü olabilmesi için şu koşulların altını çizmektedir:

- i) Değişkenlik sıfır ile bir arasındadır.
- ii) En çok benzerlik ya da homojenlik sözkonusu olduğunda değişkenlik sıfıra eşittir.
- iii) En çok benzemezlik ya da heterojenlik sözkonusu olduğunda değişkenlik bire eşittir.

5.2 Değişkenlik Oranı (DO)

L. Freeman tarafından önerilen değişkenlik oranı:

$$DO = 1 - \frac{f_m}{N} \quad (5.1)$$

şeklinde. Burada K kategori sayısı, N toplam birim sayısı olsun. Bütün gözlemler aynı kategoriden geldiğinde DO sıfıra eşit olur. İndeks maksimum değerini

$f_1 = f_2 = \dots = f_K = \frac{N}{K}$ koşulu altında alarak $\frac{K-1}{K}$ değerine eşit olur. Dolayısıyla DO,

bu değere bölünecek olursa

$$DO \div \left(\frac{K-1}{K} \right) = \frac{NK - K f_m}{N(K-1)} \quad (5.2)$$

Wilcox tarafından önerilen bu indeks tanesi moddan sapmalar indeksidir (MSİ).

$$MSI = \frac{NK - K f_m}{N(K-1)} \quad (5.3)$$

5.3 Grup Sıklıklarının Varyansına Dayalı İndeks (GSVDI)

Wilcox tarafından önerilen bu istatistik şu şekilde hesaplanmaktadır [32]:

$$GSVDI = 1 - \frac{\sum_{i=1}^K \left(f_i - \frac{N}{K} \right)^2}{\frac{N^2(K-1)}{K}} \quad (5.4)$$

Bu formül klasik varyans formülü ile analogi kurularak anlaşılabilir. En çok heterojenlik halinde bütün kategorilerin sıklıkları N/K'ya eşit olacaktır. Bütün gözlemler aynı kategoriden geldiğinde ise

$$\sum_{i=1}^K \left(f_i - \frac{N}{K} \right)^2 = \frac{N^2(K-1)}{K} \quad (5.5)$$

eşitliği geçerlidir. Bu yüzden (5.4)'deki $\frac{N^2(K-1)}{K}$ çarpımı standardizasyonu sağlamak

içindir. Yine (5.4)'ün sağ tarafındaki 1'den çıkarma işlemi Wilcox'un sözünü ettiği (ii) ve (iii) no'lu koşulların yerine gelmesi içindir. (5.4) no'lu denklem üzerinde cebirsel olarak bir miktar manipülasyona gidilerek aşağıdaki daha basit form elde edilebilir:

$$GSVDI = \frac{K \left(N^2 - \sum_{i=1}^K f_i^2 \right)}{N^2(K-1)} \quad (5.6)$$

Bir dağılımda K tane kategoriye düşen gözlem sayıları O_i , beklenen gözlem sayıları da e_i olsun. Başka bir deyişle $H_0 : e_i = \frac{N}{K}$, $i = 1, 2, \dots, K$ şeklinde sıfır hipotezi oluşturulsun.

Bu durumda $\sum_{i=1}^K \frac{(O_i - e_i)^2}{e_i}$ istatistiği K-1 serbestlik dereceli bir ki-kare değişkeni olacaktır. $O_i = f_i$ değerleri yerine konacak olursa

$$\sum_{i=1}^K \frac{(O_i - e_i)^2}{e_i} = \frac{\sum_{i=1}^K \left(f_i - \frac{N}{K} \right)^2}{\left(\frac{N}{K} \right)} \quad (5.7)$$

bulunur. Bu oran en çok homojenlik halinde (bütün gözlemlerin j. gruptan gelmesi durumunda bütün $i \neq j$ için $f_i = 0$ ve j.grup için $f_j = N$ olacağından)

$$\frac{(K-1) \left(\frac{N}{K} \right)^2 + \left(N - \frac{N}{K} \right)^2}{\left(\frac{N}{K} \right)} = N(K-1) \quad (5.8)$$

Bu χ^2 değeri standardizasyon için $N(K-1)$ 'e bölünecek olursa

$$\frac{\sum_{i=1}^K \left(f_i - \frac{N}{K} \right)^2}{\left(\frac{K-1}{K} \right) N^2} \text{ olur. Bu oran en çok homojenlik durumunda 1, en az homojenlik}$$

durumunda ise 0 değerine eşit olmaktadır. Dolayısıyla Wilcox'un sözünü ettiği (ii) ve (iii) no'lu koşulların yerine gelebilmesi için bu oranı birden çıkarmak gerekmektedir. Bu durumda ise elde edilen indeks (5.4) 'dekine özdeştir. Buradan hareketle

$$GSVDI = 1 - \frac{\chi_{K-1}^2}{N(K-1)} \quad (5.9)$$

5.4 Nitel Değişkenlik İndeksi (NDI)

$$NDI = \frac{K}{K-1} \left(1 - \sum_{i=1}^K p_i^2 \right) \quad (5.10)$$

şeklindedir. Bu denklemde $p_i = f_i / N$ denecek olursa

$$NDI = \frac{K}{K-1} \frac{\left(N^2 - \sum_{i=1}^K f_i^2 \right)}{N^2} \quad (5.11)$$

Fakat bu formül (5. 9)'daki gibidir. Başka bir deyişle GSVDI ile NDI birbirine eşit olmaktadır. Bazen bu indeksin standardize edilmemiş hali $\sum_{i=1}^K p_i(1-p_i)$ toplamı kullanılmaktadır. Bu forma sahip indeks aynı zamanda Gini Yoğunlaşma İndeksi (GYI) olarak da adlandırılmaktadır. Başka bir deyişle NDI, GYI'nin standartlaştırılmış hali olarak da değerlendirilebilir [33].

5.5 Benzemezlik Katsayısı

“Benzemezlik her bir gözlem değerinin bir diğerinden ne sıklıkla farklılaştığı anlamına gelir” [34]. Benzemezlik katsayısı (the coefficient of unalikeability) şu şekilde tanımlanmaktadır:

$$u = \frac{\sum_{i \neq j}^N C(x_i, x_j)}{N^2 - N} \quad (5.12)$$

Burada eğer x_i ve x_j farklı kategorilerden geliyorsa $C(x_i, x_j) = 1$ kabul edilir. Aksi takdirde $C(x_i, x_j) = 0$ olarak alınır.

Her biri sırası ile $f_1, f_2, \dots, f_K$ sıklıklarına sahip K tane ayrık kategori varsayalım. Bu durumda (5. 12) no'lu denklemdeki payda ifade edilen toplam (birlerin ve sıfırların

toplamı) $f_1(N - f_1) + f_2(N - f_2) + f_3(N - f_3) + \dots + f_K(N - f_K) = \sum_{i=1}^K f_i(N - f_i)$

$$\sum_{i=1}^K f_i(N - f_i) = N^2 - \sum_{i=1}^K f_i^2 \quad (5.13)$$

En çok heterojenlik halinde

$$N^2 - \sum_{i=1}^K f_i^2 = N^2 \left(\frac{K-1}{K} \right) \quad (5.14)$$

ve bu duruma karşılık gelen benzemezlik indeksi u_{MH} ile gösterilecek olursa

$$\{ 2 \binom{N}{2} = N(N-1) = N^2 - N \}$$

$$u_{MH} = \frac{\sum_{i \neq j}^N C(x_i, x_j)}{N^2 - N} = \frac{N(K-1)}{K(N-1)} \text{ bulunur. İndeksi standartlaştırmak için}$$

$$\frac{u}{u_{MH}} = \frac{K}{(K-1)} \left(\frac{N^2 - \sum_{i=1}^K f_i^2}{N^2} \right) \quad (5.15)$$

Ancak bu da (5.4) 'deki sonuçlarla aynıdır. Başka bir ifade ile benzemezlik katsayısının standartlaştırılmış hali nitel grup sıklıklarının varyanslarına dayalı indekse (GSVDI) eşit olmaktadır.

5.6 Simpson Çeşitlilik İndeksi (SCI)

Simpson çeşitlilik indeksi bir ekolojik sistem içerisindeki çeşitliliği ölçmek üzere geliştirilmiştir.

$$SCI = 1 - \sum_{j=1}^K \frac{f_j(f_j-1)}{N(N-1)} \quad (5.16)$$

Ya da bir miktar cebirsel işlemde sonra

$$SCI = \frac{N(1 - \sum_{j=1}^K p_j^2)}{(N-1)} \quad (5.17)$$

En çok heterojenlik ya da biyolojik çeşitlilik söz konusu olduğunda bu toplam

$$SCI_{MH} = \frac{N(K-1)}{K(N-1)} \quad (5.18)$$

$$\frac{SCI}{SCI_{MH}} = \frac{K}{(K-1)} \left(1 - \sum_{j=1}^K p_j^2 \right) \quad (5.19)$$

Bu da hiç şüphesiz NDI (ya da GSVDI) ile aynı indekstir.

5.6.1 Genişletilmiş Binom Dağılımı ile İlişkisi

Rastgele yapılan bir gözlemin belirli bir kategoriden gelme (ortalama) olasılığı $p=1/K$ ve gelmeme olasılığı $q=1-(1/K)$ olacaktır. Aynı Bernoulli deneyi K kere tekrarlanırsa ve rastlantı değişkeni belirli bir kategoriden gelen toplam gözlem sayısı olarak tanımlanırsa rastlantı değişkeninin varyansı

$$K \frac{1}{K} \left(1 - \frac{1}{K} \right) = \frac{(K-1)}{K} \text{ bulunacaktır.}$$

Şimdi de her denemede i . kategoriden gelme (başarı) olasılıklarının değişken olduğu bir durumu ele alalım. Her denemede başarı ve başarısızlık olasılıkları sırası ile p_i ve $q_i = 1 - p_i$ $i=1,2,...,K$ olsun. Her denemenin bağımsız olduğu varsayımı altında toplam başarı sayısı X 'in varyansı

$$Var(X) = \sum_{i=1}^K p_i q_i = \sum_{i=1}^K p_i (1 - p_i) = 1 - \sum_{i=1}^K p_i^2 \text{ olarak bulunacaktır. Dolayısıyla } 1 - \sum_{i=1}^K p_i^2$$

ifadesi genelleştirilmiş binom dağılımının varyansı , $\frac{(K-1)}{K}$ de $p=1/K$ ve $n=K$ parametrelili bir binom dağılımının varyansıdır. Bu bakımdan (5. 19)'i iki varyansın oranı olarak ele almak da mümkündür.

5.7 Nitel Değişkenlik Ölçüleri ile İlgili Uygulama

Nitel değişkenlik ölçüleri arasında bilgi istatistiklerini anlamlı karşılaştırmalar yapmak için üç farklı metrik veri ile çalışmalarımızı yapmayı amaçladık.

Oluşturduğumuz tablolar Maksimum Entropi, Minimum Entropi ve ortalama bir entropiyi içerecek şekilde aşağıda verilen tablolarda gösterilmiştir. X değişkeniyle 100 farklı sayıda gözlem gruplandırılarak Nitel değişkenlik değerleri hesaplanacak ve bu değerler Shannon, Tsallis ve Renyi entropileri ile karşılaştırılacaktır.

Çizelge 5. 1 X metrik değişkeninin maksimum entropili olasılık dağılımı

x	P(x)	Frekanslar
0	0,01	1
1	0,01	1
2	0,01	1
3	0,95	95
4	0,01	1
5	0,01	1
	Toplam	100
	Maks	95

Çizelge 5. 2 X metrik değişkeninin nitel değişkenlik ölçüleri

DEĞİŞKENLİK ÖLÇÜLERİ	DEĞERLER
DO	0,0500
MSI	0,0600
GSVDI	0,1160
NDI	0,1164
SCI	0,0980

Çizelge 5. 3 X metrik değişkeninin shannon, renyi, tsallis entropileri

ENTROPİ	DEĞERLER	DEĞERLER2	DEĞERLER3
SHANNON	0,278987139		
	$\alpha=2$	$\alpha=1,5$	$\alpha=0,5$
RENYİ	0,102032726	0,143109165	0,776881267
TSALLİS	0,097000000	0,138109074	0,949358869

Çizelge 5. 4 X metrik değişkeninin minimum entropili olasılık dağılımı

x	P(x)	Frekanslar
0	0,17	17
1	0,17	17
2	0,17	17
3	0,17	17
4	0,16	16
5	0,16	16

	Toplam	100
	Maks	17

Çizelge 5. 5 X metrik değişkeninin nitel değişkenlik ölçüleri

DEĞİŞKENLİK ÖLÇÜLERİ	DEĞERLER
DO	0,8300
MSI	0,9960
GSVDI	0,9998
NDI	0,9998
SCI	0,8416

Çizelge 5. 6 X metrik değişkeninin shannon, renyi, tsallis entropileri

ENTROPİ	DEĞERLER	DEĞERLER2	DEĞERLER3
SHANNON	1,791356721		
	$\alpha=2$	$\alpha=1,5$	$\alpha=0,5$
RENYİ	1,790959789	1,791157514	1,791557382
TSALLİS	0,833200000	1,183257635	2,898484500

Çizelge 5. 7 X metrik değişkeninin ortalama entropili olasılık dağılımı

x	P(x)	Frekanslar
0	0,05	5
1	0,2	20
2	0,3	30
3	0,1	10
4	0,15	15
5	0,2	20
	Toplam	100
	Maks	30

Çizelge 5. 8 X metrik değişkeninin nitel değişkenlik ölçüleri

DEĞİŞKENLİK ÖLÇÜLERİ	DEĞERLER
DO	0,700
MSI	0,840
GSVDI	0,954

NDI	0,954
SCI	0,803

Çizelge 5. 9 X metrik değişkeninin shannon, renyi, tsallis entropileri

ENTROPİ	DEĞERLER	DEĞERLER2	DEĞERLER3
SHANNON	1,669580127		
	$\alpha=2$	$\alpha=1,5$	$\alpha=0,5$
RENYİ	1,5847453	1,623410709	1,725174457
TSALLİS	0,7950000	1,111799856	2,738565294

5.7.1 Sonuçlar

1. Nitel değişkenlik ölçülerinin bazıları yukarıdaki tablolarda da görüldüğü gibi birbirleriyle yakından ilişkilidir.
2. Bu tür örneklem dağılımları gerçekleştirmek için ve daha güvenilir sonuçlar elde etmek için nitel değişkenlik ölçülerini kullanmak kaçınılmazdır. Bu gerçeğe rağmen analitik çözümler elde etmek kolay değildir. Bu nedenle Simülasyon çalışmaları ile bu niteliksel önlemlerin “örneklem dağılımları” elde edilebilir.
3. Maksimum entropinin olduğu birinci örnekte GSDVI ve SCI değişkenlik ölçülerinin Tsallis entropisinin $\alpha=2$ parametrelili değeri ile benzer sonuçlar verdiği görülür.
4. Minimum entropinin olduğu ikinci örnekte GSDVI ve MSI benzer sonuçlar vermektedir. Ayrıca DO ve SCI değişkenlik ölçülerinin Tsallis entropisinin $\alpha=2$ parametrelili değeri ile benzer sonuçlar verdiği görülür.
5. Ortalama entropili üçüncü örnekte ise tüm değişkenlik ölçüleri birbirine yakın değerler vermektedir. SCI değişkenlik ölçüsünün Tsallis entropisinin $\alpha=2$ parametrelili değeri ile benzer sonuçlar verdiği görülür.
6. Değişkenler nitel olduğu zaman, biz klasik korelasyon ölçülerini kullanamayız. Bu nedenle karşılıklı bilgi ölçüleri bağımlılıkların analizinde etkili olabilir. Genellikle (Shannon, Renyi veya Tsallis entropilerinin birini kullanarak) karşılıklı bilgi ölçüsü hesaplamaları Shannon ve Renyi entropilerinde benzer sonuçlar üretirken Tsallis entropisinde bu durum farklılık gösterir. Tsallis entropileri parametrelere çok duyarlı oldukları için nominal ilişkilendirme hususlarında yanıltıcı sonuçlar verebilir.

BÖLÜM 6

NİTEL DAĞILIMLARDA BİRLİKTELİK ÖLÇÜLERİ

İki nitel değişkenin birlikte sıklık dağılımı bir kontenjans tablosu şeklinde düzenlensin. Birinci değişken, X_i $i=1,2,...,n$ değerlerini ikincisi de Y_j $j=1,2,...,m$ değerlerini alsın. $X_{ij} \{X = X_i, Y = Y_j\}$ birleşik durumunu ifade etsin. f_{ij} ve e_{ij} X_{ij} 'nin gözlenen ve (bir hipotez çerçevesinde) beklenen sıklıkları [35]

$$\sum_{i=1}^n \sum_{j=1}^m f_{ij} = \sum_{i=1}^n \sum_{j=1}^m e_{ij} = N \quad (6.1)$$

$$P(X = X_i) = \pi_i = \frac{1}{N} \sum_{j=1}^m f_{ij} \quad (6.2)$$

$$P(Y = Y_j) = \pi_j = \frac{1}{N} \sum_{i=1}^n f_{ij} \quad (6.3)$$

$$P(X = X_i, Y = Y_j) = P(X_{ij}) = \pi_{ij} = \frac{f_{ij}}{N} \quad (6.4)$$

olsun. Değişkenlerin bağımsız olabilmeleri için bütün $i=1,2,...,n$ ve $j=1,2,...,m$ için $\pi_{ij} = \pi_i \pi_j$ olması gerekmektedir. Birliktelik derecesinin ölçülmesinde fi-kare istatistiği

$$\phi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{\pi_{ij}^2}{\pi_i \pi_j} - 1 \quad (6.5)$$

şeklinde tanımlanmaktadır. Bu değer değişkenler bağımsız ise 0 değerini almaktadır.

Alabileceği maksimum değer $q-1$ olur. Burada $q = \min\{n, m\}$ olur. Dolayısıyla $\frac{\phi^2}{(q-1)}$

istatistiği bir nitel birliktelik indeksi olarak değerlendirilebilir. Bu istatistik değişkenler arasında birliktelik yokken 0, bağımlılık varken 1 değerini alacaktır. (6. 5)'de π_{ij}, π_i ve π_j söz konusu olasılıkların maksimum olabilirlik tahmincileri ve seçilen N birimlik örnekten hareketle olabilirlik fonksiyonunu [36]

$$L = \prod_{i,j} (\pi_i \pi_j)^{f_{ij}} \quad (6.6)$$

olsun.

Burada $\sum_{i=1}^n \pi_i = 1$ ve $\sum_{j=1}^m \pi_j = 1$ ve $\text{Log} L - \lambda \sum_{i=1}^n \pi_i - \mu \sum_{j=1}^m \pi_j$ maksimize edildiğinde

$$\pi_i = \frac{\sum_{j=1}^m f_{ij}}{N} \quad (6.7)$$

$$\pi_j = \frac{\sum_{i=1}^n f_{ij}}{N} \quad (6.8)$$

En çok olabilirlik tahmincilerini kullanarak $e_{ij} = N\pi_i\pi_j$ bulunur.

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \quad (6.9)$$

istatistiği yaklaşık (n-1)(m-1) serbestlik derecesinde bir ki-kare dağılımına uymaktadır.

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(f_{ij} - N\pi_i\pi_j)^2}{N\pi_i\pi_j} \quad (6.10)$$

eşitliğin sağ tarafı açılacak ve bir miktar cebirsel işlem yapılacak olursa

$$\chi^2 = N \left(\sum_{i=1}^n \sum_{j=1}^m \frac{(\pi_{ij})^2}{\pi_i\pi_j} - 1 \right) \quad (6.11)$$

$$\phi^2 = \frac{\chi^2}{N} \quad (6.12)$$

bulunur. Değişkenler bağımsız olduğunda bu istatistik 0'a eşit olmaktadır. Bu istatistiğin büyüklüğünün ki-kare tablosundaki hücre sayısına bağlı olması yorumlama gücünü getirmektedir. Dolayısıyla istatistiğin yeniden düzenlenmesi gerekmektedir [37].

6.1 Bazı Düzenlemeler

Bu gücünü ortadan kaldırmak için Pearson

$$p = \left(\frac{\phi^2}{1 + \phi^2} \right)^{1/2} \quad (6.13)$$

istatistiğini önermiştir P de 0 ile 1 arasında değer almaktadır. Bu istatistik, değişkenler birbirleriyle tam ilişkili görünse bile 1 değerine ulaşamamaktadır. Multinomial örnekleme planı çerçevesinde p'nin en çok olabilirlik tahmincisi $\hat{p}$ ile gösterilecek olursa

$$\hat{p} = \left(\frac{\chi^2 / N}{1 + \chi^2 / N} \right)^{1/2} = \sqrt{\frac{\chi^2}{\chi^2 + N}} \quad (6.14)$$

Kontenjans tablosunun satır sayısı sütun sayısına eşit olduğunda $\hat{p}$ nin alacağı maksimum değer $\sqrt{(q-1)/q}$, aksi hallerde ise 1'den küçük olabilmektedir. Bu bakımdan formülde düzenlemeler önerilmiştir. Örneğin Sakoda şu istatistiği geliştirmiştir:

$$p^* = \frac{P}{P_{Maks}} = \left(\frac{q\phi^2}{(q-1)(1+\phi^2)} \right)^{1/2} \quad (6.15)$$

p^* , değişkenlerin bağımsızlığı halinde 0, tam birlikteliği halinde 1 değerini almaktadır [38].

6.2 Tschuprow'un Kontenjans Katsayısı

Bir diğer versiyon Tschuprow'un T kontenjans katsayısıdır:

$$T = \left(\frac{\phi^2}{\sqrt{(m-1)(n-1)}} \right)^{1/2} \quad (6.16)$$

T de 0 ile (bağımsızlık halinde) 1 arasında değerler almaktadır. Bununla birlikte maksimum değerine yalnızca kare formülü kontenjans katsayılarında ulaştığını belirtmek gerekmektedir.

6.3 Cramér'in Kontenjans Katsayısı

Cramér, p ve T katsayılarına karşılık olarak

$$v = \left(\frac{\phi^2}{q-1} \right)^{1/2} = \left(\frac{\chi^2}{N(q-1)} \right)^{1/2} \quad (6.17)$$

istatistiğini önerdi. v tam birliktelik halinde 1, bağımsızlık halinde ise 0 değerini alması ve satır sayısının sütun sayısına eşit olmadığı kontenjans tablolarında bile alabileceği maksimum değere ulaşabilmesi nedeniyle bahsedilmeye değer bir istatistiktir [37].

p , p^* , T ve v 'nin tahmincilerinin tam olasılık dağılımlarının belirlenmesi oldukça zordur ama yaklaşık büyük örneklem dağılımları belirlenebilir. Bağımsızlık varsayımı altında ($\phi^2 = 0$) kontenjans katsayılarının dağılımlarının çeşitli yüzde noktaları aşağıdaki gibi kolaylıkla hesaplanabilir:

$$P(T \geq t_0) = P(T^2 \geq t_0^2) = P\left(\left[\frac{\chi^2}{N\sqrt{(n-1)(m-1)}} \right] \geq t_0^2 \right) \quad (6.18)$$

$$P(T \geq t_0) = P\left[\chi^2 \geq N t_0^2 \sqrt{(n-1)(m-1)} \right] \quad (6.19)$$

Benzer olasılık hesaplamaları p , p^* ve v için de gerçekleştirilebilir. $\phi^2 \neq 0$ koşulunda ise bu tahmincilerin standart hataları için Liebetrau'ya bakılabilir [37]. Biz bu formüller çok karmaşık olduğu gerekçesi ile buraya almadık.

6.4 Uygulama

Aşağıdaki frekans tablosu İstanbul'da faaliyet gösteren bir el örgü şirketinin verilerinden oluşturulmuştur. İki farklı türde müşterilere sunulan yedi farklı ürün grubu vardır. Nisan ve Ekim aylarında satılan ürünler gruplar şeklinde ayrılmıştır. Ürün grubu

2 nitel deęiřkene ayrılmıř ve kontenjans tablosu ařaęıdaki gibi dzenlenmiřtir. Bu rnlerin satıř kilogramları ton olarak hazırlanarak ařaęıdaki tablo elde edilmiřtir.

izelge 6. 1 El rg iplięi fabrikasının rnlerin trlerinden sattıęı miktarlar(ton)

RN	DZ	DESENLİ	GENEL TOPLAM (TON)
ANGORAGRUBU	207	28	235
AKRİLİKGRUBU	7	28	35
%25YNLGRUBU	93	7	100
POLYAMİDGRUBU	34	44	78
PAMUKLUGRUBU	20	3	23
%50YNLGRUBU	16	5	21
FANTEZİGRUBU	35	18	53
TOPLAM	412	133	545

Bu fabrikanın 2013 yılı 3. eyreęindeki rn grubunun ton olarak satıř kilogramları ıkarılarak kontenjans tablosu oluřturulmuř ve rn grubu ile tr arasındaki iliřkisinin birliktelik ls bulunmaya alıřılmıřtır. İlgili daęılım ve bu daęılımla ilgili birliktelik istatistikleri izelge 6. 2 ve izelge 6. 3 de zetlenmektedir.

izelge 6. 2 Ki-Kareye dayalı nitel birliktelik istatistikleri

Ki-Kare	108,5024553
Fi-kare	0,199087074
Pearson p	0,656478187
Sakoda	0,690974293
Tschuprow	0,182157017
Cramér	0,446191746

Yukarıdaki çizelgede de görüldüğü gibi (6. 12) eşitsizliği 0,19 çıktığından dolayı nitel değişkenler arasındaki bağımlılık zayıftır. Pearson eşitsizliğine göre nitel değişkenler arasındaki ilişkinin 0,65 çıkması nitel değişkenler arasındaki bağımlılığı arttırmıştır. Tschuprow kontenjans katsayısı (6.12) eşitsizliğine yakın bir değer verdiği yukarıdaki çizelgede verilir. 0,18 lik Tschuprow kontenjans katsayısı ile nitel değişkenler arasındaki ilişkinin zayıf olduğu sonucuna varılır. Cramer kontenjans katsayısına göre ise Tschuprow ve fi kare, Ki kare istatistiklerine oranla daha yüksek bir bağımlılık çıkmasına karşı, Pearson katsayısına oranla nitel değişkenler arasında daha düşük bir ilişki vardır. Cramer kontenjans katsayısına göre nitel değişkenler arasındaki birliktelik ölçüsü zayıftır. Sakoda kontenjans katsayısına göre nitel değişkenler arasındaki ilişki ise en yüksek çıktığı görülür.

Çizelge 6. 3 Entropiye dayalı nitel birliktelik istatistikleri

H(X)	2,326
H(Y)	0,8
H(X,Y)	2,958
I(X,Y)	0,169
C(X,Y)	0,073
C(Y,X)	0,211
Bolluk(Redundancy)	0,054
Kullback-Leibler	0,707
Jeffreys	1,414

Çizelgelerden de izlenebileceği gibi iki değişken arasında anlamlı bir birliktelik söz konusu değildir. Bu yüzden ürün grubu ve satılan ürün türü arasındaki ilişki oldukça düşük olduğu sonucuna varabilir.

SONUÇ VE ÖNERİLER

Entropinin istatistikte oldukça yaygın uygulama alanları bulunmaktadır. Bunları şu şekilde sıralayabiliriz:

- Nitel dağılımlarda, değişkenliğin ölçülmesi
- Nitel değişkenler arasındaki birlikteliğin (association) derecesinin belirlenmesi
- Özellikle Bayesçi istatistikte dağılımlar arasında entropiye dayalı sapma ölçülerinin kullanılması ile örneklem bilgisinin değerlendirilmesi
- Uyum iyiliğinin belirlenmesi
- Hipotez testleri

şeklinde sıralayabiliriz.

Entropi, belirsizlik demektir. Belirsizlik ve bilgi öyle yakın iki kavramdır ki bir deney hakkında edinilen bilgi yok edilen belirsizlik miktarına eşittir. Gerçekleştirilen çalışmayla ilgili olan “olasılıksal belirsizlik durumu” ele alındığında entropi “bir sisteme ait belirsizlik düzeyinin ölçümü” olarak ifade edilebilir. Olasılık teorisi sadece gelecekteki olayın beklentisi ile ilişkili olan belirsizlik ile ilgilenir. Bir başka deyişle herhangi bir olayın baş gösterme olasılığı, belli düzeyde bu olayın gerçekleşip gerçekleşmeyeceği hakkında belirsizliğin göstergesidir. Shannon, “bir olasılık uzayının belirsizliğinin ölçüsü” olarak tanımladığı entropi için, bir deneyin sonuçlarının içerdiği ortalama bilgi miktarını önermiştir.

Değişkenlerin birbirleri hakkında verdiği karşılıklı bilgi miktarı Shannon tarafından tanımlanmıştır. Göreli entropi ve karşılıklı bilgi miktarının birbirleri ile yakından ilgili iki

kavramdır. Görelî entropi ya da Kullback Leibler mesafesi Shannon entropisine baęlı olarak hesaplanmaktadır. Belirsizlik ölçütü olarak Shannon entropisinin dışında önerilen bir dizi entropi ölçüsü içerisinde Renyi ve Tsallis entropilerini özellikle belirtmek gerekir. Literatürde bu entropi ölçülerine dayalı olarak hesaplanan karşılıklı bilgi ölçüleri de bulunmaktadır. Bu çalışmada, nominal ve metrik verilere bu üç entropi ölçüsü uygulayarak klasik korelasyon ölçüsü ile entropi tabanlı ilişki ölçüleri arasındaki paralellik gösterilmek istenmiştir.

Değişkenler nitel olduğu zaman, Pearson korelasyon katsayısı kullanılamaz. Bu nedenle nitel değişkenler arasındaki karşılıklı bağımlılıkların analizinde entropiye dayalı ölçüler yararlı olabilir. Genellikle farklı entropi türlerini kullanarak elde edilen karşılıklı bağımlılık ölçüleri paralel sonuçlar verseler de bulunan sonuçları yorumlarken ihtiyatlı olmak gerekir. Çünkü Renyi ve Tsallis entropileri parametrik türden entropilerdir. Başka bir deyişle bulunan sonuçlar parametre seçiminden etkilenebilmektedir.

Çalışmanın uygulama aşamasında bir el örgü iplięi fabrikasının ürettięi farklı çeşitlerde ürün gruplarının satışları incelenmiş ve elde edilen verilere Shannon, Renyi ve Tsallis entropileri uygulanmıştır. Elde edilen sonuçlarda üretilen ürün grupları ile ürün çeşitlerinin birliktelik göstermedięi saptanmıştır. Her ürün grubunda her çeşidin aynı oranda talep görmedięi çalışmada belirlenmiştir. Özellikle üretim planlamasında bu elde edilen bu sonuçlar yararlı olacaktır.

Bu çalışmada ikinci olarak, literatürden derlenen bazı nitel değişkenlik indekslerinin birbirleri ile ilişkili olduğu ele alınan örnekler çerçevesinde gösterilmiştir.

Üçüncü olarak entropiye dayalı birliktelik ölçülerinin seçilmiş bazı örnekler için verdikleri sonuçları, ki-kare gibi geleneksel birliktelik ölçüleri ile elde edilen sonuçlarla kıyasladım. Bu sonuçların paralellik gösterdiğini gözledim. Yine de bulunan sonuçlar arasında bazı farklılıklar olması kullanılan birliktelik ölçülerinin farklı metrik yöntemlere dayanmasından kaynaklanmaktadır. Bu ölçüleri birer analiz aracı olarak kullanacak olanların bu olguya dikkat etmesi gerekir.

Son olarak entropiye dayalı ölçülerin rahatlıkla çok değişkenli hallere uygulanabilmesi de bu birliktelik ölçülerinin içerisinde kendilerine ayrıcalıklı bir konum sağlamaktadır. Bu noktanın altı çizilmelidir.

KAYNAKLAR

- [1] Rifkin, J. ve Howard, T.,(1992). Entropi Dünyaya Yeni Bir Bakış, (Çev.H.Okay) Ağaç Yayıncılık, 43, İstanbul.
- [2] Renyi, A.,(2007). Probability Theory, Dover Publications,554-563, New York.
- [3] Alpaut, O.,(1971). Teorik Analitik Kimya, H.Ü. Yayınları, 156, Ankara.
- [4] Shannon, C.E.,(1948). “A Mathematical Theory of Communication”, The Bell System Technical Journal,27: 379–423, 623–656.
- [5] Brillouin L.,(1960). Science and Information Theory (2. Baskı), Academic Press, New York.
- [6] Bailey, D.K.,(1990). Social Entropy Theory, State University of New York Press 38, New York.
- [7] Katz, D. ve Kahn, R.L.,(1978). The Social Psychology of Organization, Willey, 26, New York.
- [8] Garcia, A.L.,(1994). Probability and Random Processes for Electrical Engineering, Addison Wesley Longman (Second Edition), 169, Boston.
- [9] Reza, F.M.,(1994). An Introduction to Information Theory, Dover Publications, 80, 289, New York.
- [10] Cover, T.M, ve Thomas, J.A.,(2006). Elements of Information Theory, Wiley Interscience (Second Edition), 226 New Jersey.
- [11] Magurran, A.E,(1988). Ecological Diversity and Its Measurement, Princeton University Press, 35, New Jersey.
- [12] Cachin, C.,(1997). Entropy Measures and Unconditional Security in Cryptography, Doktora Tezi, Swiss Federal Institute of Technology Zurich, 16, Zurich.
- [13] Principe J.C.,(2009). Information Theory, Florida University Press, 5, Florida
- [14] Wikipedia, http://en.wikipedia.org/wiki/Kullback-Leibler_distance, 20 Şubat 2014.

- [15] Tsallis, C., ve Mendes R., ve Plastino A.,(1988). "The Role of Constraints within Generalized Nonextensive Statistics", Physica 261A, 535-554.
- [16] Wikipedia, http://en.wikipedia.org/wiki/Gamma_distribution, 5 Ekim 2014.
- [17] Wikipedia, http://en.wikipedia.org/wiki/Digamma_distribution, 5 Ekim 2014.
- [18] Evren, A.,(2010). "Entropinin İstatistikteki Bazı Uygulamaları", II. Ulusal Konya Ereğli Kemal Akman Melek Yüksek Okulu Tebliğ Günleri, 13-14 Mayıs, Sayı 2, 2010, 414-428.
- [19] Klir, G.J.,(2006). Uncertainty and Information, John Wiley Interscience Publication, 80-83, New Jersey.
- [20] Wikipedia, http://en.wikipedia.org/wiki/Digamma_distribution, 10 Ocak 2014.
- [21] Brillinger, D.R.,(2004). "Some Data Analyses Using Mutual Information" Brazilian Journal of Probability and Statistics, 18: 163-182.
- [22] Coombs, C.H., ve Tversky, R.M.,(1970). Mathematical Psychology Elementary Introduction, Prentice Hall, New Jersey.
- [23] Press, W.H., ve Flannery, B.P.,(1988). Numerical Recipes in C, Cambridge University Press, 634, Cambridge.
- [24] Yao, Y.Y.,(2003). "Information theoretic Measures for Knowledge Discovery and Data Mining in Entropy Measures", Maximum Entropy Principle and Emerging Applications, 115-136.
- [25] Strehl, A., ve Ghosh, J.,(2002). "Cluster Ensembles Knowledge Reuse Framework for Combining Multiple Participations", Journal of Machine Learning Research, 3:583-617.
- [26] Evren, A., ve Genç, A., ve Tuna, E.,(2010). "Entropinin İstatistik ve Ekonometrideki Kullanım Alanları Üzerine", 12th International Symposium on Econometrics, Statistics and Operations Research, Mayıs 2011, Denizli, 11-21.
- [27] Gray, R.M.,(1990). Entropy and Information Theory, Springer Verlag Publications, 21-41, New York.
- [28] Kullback, S.,(1996). Information Theory and Statistics, Dover Publications, 6, New York.
- [29] Renyi, A.,(2007). Foundations of Probability, Dover Publications, 148, New York.
- [30] Everitt, B.S.,(2006). The Cambridge Dictionary of Statistics, Third Edition Cambridge University Press, 220, UK.
- [31] Evren, A., ve Tuna, E.,(2011). "On Some Measures of Qualitive Variation and Entropy", Selçuk Journal of Applied Mathematics, 61-71.
- [32] Wilcox, A.R.,(1967). Indices of Qualitive Variation, Oak Ridge National Laboratory, USA.
- [33] Wikipedia, http://en.wikipedia.org/wiki/Qualitative_variation, 15 Nisan 2014.

- [34] Kader, G.D., ve Perry M.,(2007). "Variability for Categorical Variables", Journal of Statistics Education, 15:2.
- [35] Evren, A.,(2011). "İstatistik'te Entropiye Dayalı Uyum Ölçülerinin Diğer Uyum Ölçüleri ile Kıyaslanması", 7. İstatistik Günleri Sempozyumu, Bildiri Tam Metinler Kitabı, Orta Doğu Teknik Üniversitesi, Mart 2011, Ankara, 58-67.
- [36] Evren, A., ve Tuna, E.,(2012)." A Comparison Between Entropy Based Association Measures and Other Qualitative Association Measures", Selçuk Journal of Applied Mathematics, 3-17.
- [37] Liebetrau, A.M.,(1983). Measures of Association, Series Quantitative Applications in the Social Sciences, Sage University Publications, 12-16, Kanada.
- [38] Evren, A., ve Tuna, E.,(2012)." On Some Applications of Goodness of Fit Measures Based on Statistical Entropy", International Journal of Research and Reviews in Applied Sciences, 13:192-205.

ENTROPİ DE R PROGRAMI

R PROGRAMINDA SHANNON ENTROPİSİNİ HESAPLAYAN PROGRAM KODLARI

Description

entropy estimates the Shannon entropy H of the random variable Y from the corresponding observed counts y .

freqs estimates bin frequencies from the counts y .

Usage

```
entropy(y, lambda.freqs, method=c("ML", "MM", "Jeffreys", "Laplace", "SG",  
"minimax", "CS", "NSB", "shrink"), unit=c("log", "log2", "log10"), verbose=TRUE, ...)
```

```
freqs(y, lambda.freqs, method=c("ML", "MM", "Jeffreys", "Laplace", "SG",  
"minimax", "CS", "NSB", "shrink"), verbose=TRUE)
```

Arguments

y vector of counts.

method the method employed to estimate entropy (see Details).

unit the unit in which entropy is measured. The default is "nats" (natural units). For computing entropy in "bits" set unit="log2".

lambda.freqs shrinkage intensity (for "shrink" option).

verbose verbose option (for "shrink" option).

... option passed on to entropy.NSB.

Details

The entropy function allows to estimate entropy from observed counts by a variety of methods:

- method="ML": maximum likelihood, see entropy.empirical

- method="MM":bias-corrected maximum likelihood, see entropy.MillerMadow
- method="Jeffreys":entropy.Dirichlet with $a=1/2$
- method="Laplace":entropy.Dirichlet with $a=1$
- method="SG":entropy.Dirichlet with $a=a=1/\text{length}(y)$
- method="minimax":entropy.Dirichlet with $a=\sqrt{\text{sum}(y)}/\text{length}(y)$
- method="CS":see entropy.ChaoShen
- method="NSB":see entropy.NSB
- method="shrink":see entropy.shrink

The freqs function estimates the underlying bin frequencies. Note that estimated frequencies are

not available for method="MM", method="CS" and method="NSB". In these instances a vector containing

NAs is returned.

Value

entropy returns an estimate of the Shannon entropy.

freqs returns a vector with estimated bin frequencies (if available).

Author(s)

Korbinian Strimmer (<http://strimmerlab.org>).

See Also

entropy-package, discretize.

Examples

```
# load entropy library
library("entropy")

# observed counts for each bin
y = c(4, 2, 3, 0, 2, 4, 0, 0, 2, 1, 1)

entropy(y, method="ML")
entropy(y, method="MM")
entropy(y, method="Jeffreys")
entropy(y, method="Laplace")
entropy(y, method="SG")
entropy(y, method="minimax")
entropy(y, method="CS")
#entropy(y, method="NSB")
entropy(y, method="shrink")
```

ENTROPİ DE R PROGRAMI

R PROGRAMINDA KARŞILIKLI BİLGİ KODLARI VE ENTROPİDEKİ BAYES TAHMİNCİLERİNİ HESAPLAYAN PROGRAM KODLARI

Description

freqs.Dirichlet computes the Bayesian estimates of the bin frequencies using the Dirichlet multinomial pseudocount model.

entropy.Dirichlet estimates the Shannon entropy H of the random variable Y from the corresponding observed counts y by plug-in of Bayesian estimates of the bin frequencies using the Dirichlet-multinomial pseudocount model.

KL.Dirichlet computes a Bayesian estimate of the Kullback-Leibler (KL) divergence from counts y_1 and y_2 .

chi2.Dirichlet computes a Bayesian version of the chi-squared statistic from counts y_1 and y_2 .

mi.Dirichlet computes a Bayesian estimate of mutual information of two random variables.

chi2indep.Dirichlet computes a Bayesian version of the chi-squared statistic of independence from a table of counts y_{2d} .

Usage

```
freqs.Dirichlet(y, a)
```

```
entropy.Dirichlet(y, a, unit=c("log", "log2", "log10"))
```

```
KL.Dirichlet(y1, y2, a1, a2, unit=c("log", "log2", "log10"))
```

```
chi2.Dirichlet(y1, y2, a1, a2, unit=c("log", "log2", "log10"))
```

```
mi.Dirichlet(y2d, a, unit=c("log", "log2", "log10"))
```

```
chi2indep.Dirichlet(y2d, a, unit=c("log", "log2", "log10"))
```

Arguments

y vector of counts.

y1 vector of counts.

y2 vector of counts.

y2d matrix of counts.

a pseudocount per bin.

a1 pseudocount per bin for first random variable.

a2 pseudocount per bin for second random variable.

unit the unit in which entropy is measured. The default is "nats" (natural units). For computing entropy in "bits" set unit="log2".

entropy.Dirichlet 9

Details

The Dirichlet-multinomial pseudocount entropy estimator is a Bayesian plug-in estimator: in the

definition of the Shannon entropy the bin probabilities are replaced by the respective Bayesian

estimates of the frequencies, using a model with a Dirichlet prior and a multinomial likelihood.

The parameter a is a parameter of the Dirichlet prior, and in effect specifies the pseudocount per

bin. Popular choices of a are:

- a=0: maximum likelihood estimator (see entropy.empirical)
- a=1/2: Jeffreys' prior; Krichevsky-Trochimov (1991) entropy estimator
- a=1: Laplace's prior
- a=1/length(y): Schurmann-Grassberger (1996) entropy estimator
- a=sqrt(sum(y))/length(y): minimax prior

The pseudocount a can also be a vector so that for each bin an individual pseudocount is added.

Value

freqs.Dirichlet returns the Bayesian estimates of the frequencies .

entropy.Dirichlet returns the Bayesian estimate of the Shannon entropy.

KL.Dirichlet returns the Bayesian estimate of the KL divergence.

chi2.Dirichlet returns the Bayesian version of the chi-squared statistic.

mi.Dirichlet returns the Bayesian estimate of the mutual information.

chi2indep.Dirichlet returns the Bayesian version of the chi-squared statistic of independence.

Author(s)

Korbinian Strimmer (<http://strimmerlab.org>).

References

Agresti, A., and D. B. Hitchcock. 2005. Bayesian inference for categorical data analysis. *Stat.*

Methods. Appl. 14:297–330.

Krichevsky, R. E., and V. K. Trofimov. 1981. The performance of universal encoding. *IEEE Trans.*

Inf. Theory 27: 199-207.

Schurmann, T., and P. Grassberger. 1996. Entropy estimation of symbol sequences. *Chaos* 6:41-427.

See Also

`entropy`, `entropy.shrink`, `entropy.empirical`, `entropy.plugin`, `mi.plugin`, `KL.plugin`, `discretize`

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Gökhan DİNÇER
Doğum Tarihi ve Yeri :01.01.1987-Keşan
Yabancı Dili :İngilizce
E-posta :gkhandnrcer87@hotmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Matematik	YILDIZ TEKNİK	2011
Lise	Fen Bilimleri	KEŞAN ANADOLU LİSESİ	2005

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2012	YÜNTEKS TEKSTİL	SATIŞ VE PAZARLAMA YÖNETİCİSİ

YAYINLARI

Bildiri

- 1.IGS 2014 **Applications of Some Entropy Measures on Multinomial Populations**

- 2.ICEOS 2012 **Uç Değerler Yöntemiyle İstanbul Menkul Kıymetler Borsası Üzerine Bir Uygulama**