T.C.
HARRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

SONLU MARKOV ZİNCİRLERİNİN ENTROPİLERİ İLE BİLGİ KAYNAKLARININ ENTROPİLERİNİN KARŞILAŞTIRILMASI

Selda HİÇDURMAZ

MATEMATİK ANABİLİM DALI

ŞANLIURFA
2009

T.C.
HARRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

SONLU MARKOV ZİNCİRLERİNİN ENTROPİLERİ İLE BİLGİ KAYNAKLARININ ENTROPİLERİNİN KARŞILAŞTIRILMASI

Selda HİÇDURMAZ

MATEMATİK ANABİLİM DALI

ŞANLIURFA
2009

Doç. Dr. Hasan AKIN danışmanlığında, Selda HİÇDURMAZ'ın hazırladığı "Sonlu Markov Zincirlerinin Entropileri ile Bilgi Kaynaklarının Entropilerinin Karşılaştırılması" konulu bu çalışma 10/07/2009 tarihinde aşağıdaki jüri tarafından Matematik Anabilim Dalı'nda yüksek lisans tezi olarak kabul edilmiştir.

Danışman : Doç. Dr. Hasan AKIN

Üye : Doç. Dr. Seyit TEMİR

Üye : Yrd. Doç. Dr. Necip ŞİMŞEK

Bu tezin Matematik Anabilim Dalında Yapıldığını ve Enstitümüz Kurallarına Göre Düzenlendiğin Onaylarım.

Prof. Dr. İbrahim BOLAT
Enstitü Müdürü

İÇİNDEKİLER

ÖZ

Yüksek Lisans Tezi

SONLU MARKOV ZİNCİRLERİNİN ENTROPİLERİ İLE BİLGİ KAYNAKLARININ ENTROPİLERİNİN KARŞILAŞTIRILMASI

Selda HİÇDURMAZ

Harran Üniversitesi
Fen Bilimleri Enstitüsü
Matematik Anabilim Dalı

Danışman: Doç.Dr. Hasan AKIN

Yıl: 2009, Sayfa: 53

Bu tezde sonlu Markov zincirlerinin ve bilgi kaynaklarının entropileri, Ekroot ve Cover (1993) ile Ash (1965)'in tanımlarından faydalanılarak incelenmiştir. Özellikle, klasik ergodik teori ve enformasyon teorideki entropilerin ifadesi üzerine yoğunlaşarak, bu fonksiyonların bazı özellikleri incelenmekte ve bu entropiler karşılaştırılmaktadır. Sonlu durum indirgenemez Markov zincirlerinin yörüngelerinin entropisi, Ekroot ve Cover (1993) tarafından $H(X) = -\sum_{i,j} \mu_i p_{ij} \log p_{ij}$ olarak formüle edilmiştir. Ash (1965), Markov kaynaklarının belli bir sınıfının belirsizliğini hesapladı ve $H(X)$ belirsizliği için alternatif bir ifade verdi. Ayrıca Ash (1965) gösterdi ki eğer bir kaynak M. mertebeden ise bu kaynağın belirsizliği, M+1 ardışık sembollerin belirsizliği ile M ardışık sembollerin belirsizliği arasındaki fark olarak ifade edilebilir. Bu sonuç yardımıyla, yeni örnekler elde ettik. Yaptığımız uygulamalarla, Ekroot and Cover (1993) tarafından verilen tanımla ilgili bir sorunun cevabının araştırılmasını inceledik. Ekroot and Cover (1993) tarafından verilen entropi ile Ash (1965) de bir kaynağın mertebesi yardımıyla hesaplanan entropi değerini karşılaştırdık. Bir örnekle, bu entropilerin aynı olduğunu gördük. Sonuç olarak bazı ilginç sonuçlar elde ettik. Ayrıca durdurulma zamanı tanımını vererek, Ekroot ve Cover (1991) tarafından incelenen rastgele durdurulmuş dizinin entropisi ile ilgili olan tanım ve teoremleri inceledik.

ANAHTAR KELİMELER: Entropi, Markov zinciri, Bilgi kaynağı, Mertebe, Durdurma zamanı

ABSTRACT

MSc Thesis

COMPARISON OF THE ENTROPIES OF FINITE MARKOV CHAINS WITH ENTROPIES OF INFORMATION SOURCES

Selda HİÇDURMAZ

Harran University
Graduate School of Natural and Applied Sciences
Department of Mathematics

Supervisor: Assoc. Prof. Dr. Hasan AKIN
Year: 2009, Page: 53

In this thesis, the entropies of Markov chains and information sources have been investigated by means of definitions of Ekroot and Cover (1993) and Ash (1965). In particular, focused on the notion of entropies in classical ergodic theory and information theory, some properties of these functions have been studied and these entropies have been compared. The entropy of trajectories of finite state irreducible Markov chains has been formulated as $H(X) = -\sum_{i,j} \mu_i p_{ij} \log p_{ij}$ by Ekroot and Cover (1993). Ash (1965) has computed the uncertainty of a certain class of Markov sources and he has given an alternate expression for the uncertainty $H(X)$. Also, Ash (1965) has shown that if a source is of order M, its uncertainty can be expressed as the difference of the uncertainty of M+1 successive symbols and that of M successive symbols. By means of this result, we have got new examples. By applications done by us, we have studied looking for an answer of a question regarding to the definition given by Ekroot and Cover (1993). We have compared the entropy given by Ekroot and Cover (1993) with the quantity of entropy computed via order of a source in Ash (1965). By an example, we have seen that these entropies coincide. Consequently we have obtained some interesting results. In addition, by given the stopping time, we have studied definitions and theorems regarding to the entropy of a randomly stopped sequence investigated by Ekroot and Cover (1991).

KEY WORDS: Entropy, Markov chain, Information source, Order, Stopping time

TEŞEKKÜR

Bu tezin hazırlanması aşamasında desteğini esirgemeyen, bilgi ve tecrübelerinden yararlandığım, saygı değer hocam Doç. Dr. Hasan AKIN'a sonsuz teşekkürlerimi sunarım.

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

İlk bölümde, daha sonraki bölümlerde kullanacağımız tanım ve teoremleri vereceğiz. Teoremlerin çoğu ilgili kaynaklarda ispatlandığından ispatlarını vermeyeceğiz.

1.1. Ölçüm

Tanım 1.1.1. X boş olmayan bir küme olsun. $\mathcal{A} \subset P(X)$ verilsin. ($\mathcal{A} \neq \emptyset$)

a) Aşağıdaki aksiyomları sağlayan $\mathcal{A}$ sınıfına X üzerinde bir cebir denir.

C1) $X \in \mathcal{A}$

C2) $A \in \mathcal{A}$ ise $X \setminus A \in \mathcal{A}$

C3) $(A_i)_{1 \le i \le n}$, $\mathcal{A}$'nın elemanlarından oluşan bir dizi ise $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$,

b) (C1) ve (C2) ile birlikte her i$\in$ I için $A_i \in \mathcal{A}$ olduğunda $\bigcup_{i \in I} A_i \in \mathcal{A}$ sağlanırsa $\mathcal{A}$ sınıfına σ – cebir denir.

Örnek 1.1.2. $X = \mathbb{N}$ olsun. $\mathcal{A} = \{\emptyset, \{1,3,5,\ldots,2n-1,\ldots\}, \{2,4,6,\ldots,2n,\ldots\}, \mathbb{N}\}$ alınırsa $\mathcal{A}$ kümesi, X üzerinde bir σ – cebirdir.

Tanım 1.1.3. X bir küme, $\mathcal{A}$'da X üzerinde bir σ– cebir olsun. (X, $\mathcal{A}$) ikilisine ölçülebilir uzay denir. A, $\mathcal{A}$ içindeki herhangi bir küme ise A'ya $\mathcal{A}$ ölçülebilir küme denir.

Tanım 1.1.4. (X, $\mathcal{A}$) ölçülebilir uzay olmak üzere, $\mathcal{A}$ üzerinde tanımlı genişletilmiş reel değerli μ fonksiyonu aşağıdaki özellikleri sağlıyorsa bu fonksiyona ölçüm fonksiyonu veya kısaca ölçüm denir.

a) $\mu(\emptyset)=0$

b) Her $A\in\mathcal{A}$ için $\mu(A)\geq 0$

c) Her ayrık (A_n) dizisi için $\mu\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\mu(A_n)$

Tanım 1.1.5. X boş olmayan herhangi bir küme ve $\mathcal{A}$, X kümesi üzerinde bir σ– cebir olsun. $\mu:\mathcal{A}\to[0,\infty)$ fonksiyonu tanımlansın. $\mu(X)=1$ ise $(X,\mathcal{A},\mu)$ ölçüm uzayına olasılık ölçüm uzayı denir.

Teorem 1.1.6. $(X,\mathcal{A})$ ölçülebilir uzay ve A, X' in bir alt kümesi olsun. Bu takdirde $f:\mathcal{A}\to(-\infty,\infty)$ fonksiyonu için aşağıdakiler denktir.

a) Her $t\in\mathbb{R}$ için $\{x\in A: f(x)<t\}\in\mathcal{A}$

b) Her $t\in\mathbb{R}$ için $\{x\in A: f(x)\leq t\}\in\mathcal{A}$

c) Her $t\in\mathbb{R}$ için $\{x\in A: f(x)\geq t\}\in\mathcal{A}$

d) Her $t\in\mathbb{R}$ için $\{x\in A: f(x)>t\}\in\mathcal{A}$ (Cohn, 1980).

Tanım 1.1.7. $(X,\mathcal{A})$ ölçülebilir uzay ve $f: X\to\mathbb{R}$ fonksiyonu verilsin. Eğer f fonksiyonu 1.1.6. teoreminin aksiyomlarından herhangi birini sağlarsa f 'ye ölçülebilir fonksiyon denir.

Tanım 1.1.8. $(X,\mathcal{A},\mu)$, $(Y,\mathcal{B},\nu)$ iki ölçüm uzayı ve $T: X\to Y$ ölçülebilir bir fonksiyon olsun. T ölçülebilir, 1–1, örten ve aynı zamanda T^{-1} de ölçülebilir ise T'ye terslenebilir dönüşüm denir. $A\in\mathcal{B}$ için $\mu\left(T^{-1}(A)\right)=\nu(A)$ oluyorsa buradaki T dönüşümüne ölçümü koruyan dönüşüm denir.

Tanım 1.1.9. $(X,\mathcal{A},\mu)$ ve $(Y,\mathcal{B},\nu)$ iki ölçüm uzayı olsun. $T: X\to Y$, T dönüşümü tersinir, 1–1 ve örten ise izomorfizmadır. X ve Y uzaylarına izomorf uzaylar denir ve kısaca $X\approx Y$ ile gösterilir. $T:X\to X$ dönüşümü, 1–1 ve örten ise otomorfizmadır.

Tanım 1.1.10. T birebir, örten ve ölçümü koruyan dönüşüm, T^{-1}' de ölçümü koruyan dönüşüm ise T' ye ölçümü koruyan tersinir dönüşüm denir. T sadece örten ise endomorfizmadır.

Tanım 1.1.11. ($X, \mathcal{A}$, μ) ölçüm uzayı ile T endomorfizması verilsin. Eğer her $A \in \mathcal{A}$ için $\mu\left(T^{-1}(A)\right) = \mu(A)$ eşitliğini sağlayan μ ölçümüne T– değişmez ölçüm denir.

1.2. Dinamik Sistemler

Tanım 1.2.1. ($X, \mathcal{A}$, μ) ölçüm uzayı, T, X' in bir otomorfizması ise ($X, \mathcal{A}$, μ, T) dörtlüsüne dinamik sistem denir. Kısaca (X, T) ile gösterilir.

Tanım 1.2.2. X_i boştan farklı herhangi bir küme olsun. X_i' lerin kartezyen çarpımını $\Omega = \prod_{i=1}^{\infty} X_i$ olarak gösterelim. $B_{i_1} \subset X_{i_1}, \ldots, B_{i_k} \subset X_{i_k}$ olmak üzere,

$$B = \left\{ X = \left(x_1, x_2, \ldots, x_n, \ldots\right) \in \Omega : x_{i_1} \in B_{i_1}, \ldots, x_{i_k} \in B_{i_k} \right\}$$

şeklinde tanımlanan $B \subset \Omega$ alt kümesine Ω'nın sonlu boyutlu silindir kümesi denir.

Uyarı 1.2.3. Herhangi iki tane sonlu boyutlu silindir kümenin kesişimi yine bir silindir kümedir (Denker ve ark. 1976).

Tanım 1.2.4. $\left\{ w \in \Omega : x_{i_1} = \alpha_1, x_{i_2} = \alpha_2, \ldots, x_{i_n} = \alpha_n \right\}$ kümesine n–boyutlu silindir kümesi denir.

Örneğin; E={0,1} alırsak tüm tek taraflı sonsuz dizilerin uzayı $\Omega = \prod_{i=0}^{\infty} E = E^{\mathbb{Z}^+}$ olur.

Örnek 1.2.5. $A = \left\{ w \in \Omega : x_1 = 0, x_3 = 1, x_5 = 0 \right\}$ şeklindeki küme bir silindir kümesidir. Eğer A kümesinin x_i bileşenlerinin indisleri ardışık olarak geliyorsa bu silindir kümesine ince silindir kümesi denir. Kolaylık olması bakımından, $\left\{ x \in \Omega : x_n = i_1, x_{n+1} = i_2, \ldots, x_{n+s} = i_{s+1}, i_1, i_2, \ldots, i_{s+1} \in S \right\}$ ince silindir kümesini ${}_n\left[i_1, i_2, \ldots, i_{s+1}\right]_{n+s}$ şeklinde de gösterebiliriz.

Örnek 1.2.6. $A = \left\{ w \in \Omega : x_1 = 0, x_2 = 0, x_3 = 1 \right\}$ kümesi ince silindir kümedir ve ${}_1[001]_3$ şeklinde de gösterilebilir. Fakat $B = \left\{ w \in \Omega : x_1 = 0, x_2 = 1, x_4 = 0 \right\}$ kümesi ince silindir küme değildir. Çünkü x_3 elemanı sabitlenmemiştir.

Uyarı 1.2.7. İnce silindir olmayan kümeler, ince silindir kümelerin birleşimi olarak yazılabilir (Denker ve ark.1976).

Örnek 1.2.8. $\{w \in \Omega : x_1 = 0, x_3 = 1\}$ kümesi ince değildir. Ancak

$$\{w \in \Omega : x_1 = 0, x_3 = 1\} = \{w \in \Omega : x_1 = 0, x_2 = 1, x_3 = 1\}$$
$$\cup \{w \in \Omega : x_1 = 0, x_2 = 0, x_3 = 1\}$$

şeklindeki ince silindir kümelerin birleşimi olarak yazılabilir.

Tanım 1.2.9. $E = \{1, 2, \ldots, r\}$ olsun. $P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}_{rxr}$ matrisi ve $\forall i, j$ için $p_{ij} \geq 0$ ve $\sum_{j=1}^{r} p_{ij} = 1$ ise P' ye stokastik (olasılık) matris denir.

Tanım 1.2.10. P stokastik matris olsun. $\pi P = \pi$ eşitliğini sağlayan ve $\forall i = 1, 2, \ldots, k$ için $q_i \geq 0$ ve $\sum_{i=1}^{k} q_i = 1$ şartlarını sağlayan $\pi = (q_1, q_2, \ldots, q_k)$ vektörüne P'ye karşılık gelen olasılık vektörü denir.

Örnek 1.2.11. $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$ matrisini göz önüne alalım. Bu matrise karşılık gelen olasılık vektörü şöyle hesaplanır. $\pi = (q_1, q_2)$ olsun. $\pi P = \pi$ ise

$$(q_1, q_2) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} = (q_1, q_2) \text{ ve buradan } \left. \begin{aligned} \frac{q_1}{2} + \frac{q_2}{4} &= q_1 \\ \frac{q_2}{2} + \frac{3q_2}{4} &= q_2 \end{aligned} \right\} q_1 + q_2 = 1$$

olup $q_1 = \frac{1}{3}$ ve $q_2 = \frac{2}{3}$'tür. Böylece π olasılık vektörü $\pi = \left(\frac{1}{3}, \frac{2}{3}\right)$ olarak bulunur.

Bir stokastik matrisin herhangi bir kuvveti yine bir stokastik matristir.

Örneğin;

$$P^2 = P.P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} . \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{5}{16} & \frac{11}{16} \end{bmatrix}$$

$$P_{11}^{(2)} + P_{12}^{(2)} = 1 \text{ ve } P_{11}^{(2)} \geq 0, P_{12}^{(2)} \geq 0$$

$$P_{21}^{(2)} + P_{22}^{(2)} = 1 \text{ ve } P_{21}^{(2)} \geq 0, P_{22}^{(2)} \geq 0$$

olduğundan P^2 'de bir stokastik matristir.

Tanım 1.2.12. $A = \left(a_{ij}\right)_{1 \leq i,j \leq s}$, $s \times s$ tipinde bir matris olsun.

a) Eğer $\forall i, j = 1, 2, \ldots, s$ için $a_{ij} \geq 0$ ise A matrisine pozitif matris denir.

b) Her $1 \leq i, j \leq s$ için $A^n > 0$ yani $a_{ij}^{(n)} > 0$ ise A 'ya indirgenemez matris denir.

c) n_0 vardır öyle ki $\forall n > n_0$ için $P_{ij}^{(n)} > 0$ ise A 'ya aperiyodiktir denir.

1.3. Markov Ölçümü

Bu kesimde sonraki bölümlerde uygulamalarda kullanacağımız ve konunun iyi kavranabilmesi için bilinmesi gereken tanım ve teoremleri vereceğiz.

Tanım 1.3.1. $S = \{0, 1, 2, \ldots, r-1\}$ durum uzayı için $\Omega = \prod_{i=-\infty}^{\infty} S = S^{\mathbb{Z}}$ şeklinde tanımlansın ve $P(S)$ ise S'de tanımlanan bir σ-cebir olsun. $\mathcal{F}$ ince silindirik kümeler tarafından doğrulan minimal σ-cebir olsun. $\prod_{i=-\infty}^{\infty} P(S) = \mathcal{F}$; $w \in \Omega$ ve $w = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ için $C = \{w \in \Omega : x_1 = i_1, x_2 = i_2, \ldots, x_{i+n} = i_{i+n}\}$ şeklindeki küme için C$\in \mathcal{F}$ 'dir. $\pi = \{q_0, q_1, \ldots, q_{r-1}\}$ bir olasılık vektörü $P = (p_{ij})_{rxr}$ bir stokastik matris olmak üzere $\mu_{\pi p} : \mathcal{F} \to [0,1]$ küme fonksiyonu verilsin. C silindir kümesinin ölçümünü alırsak;

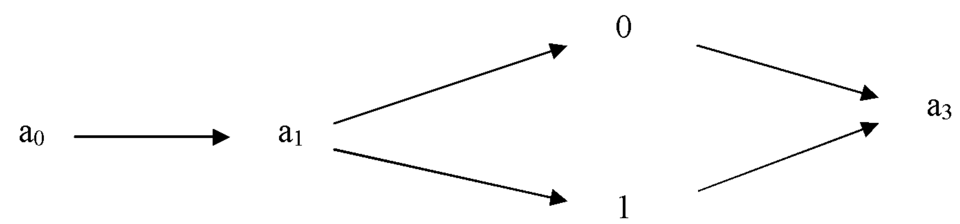
$$\mu_{\pi P}(C)=\mu_{\pi P}\left(\left\{w\in\Omega: x_1=i_1, x_2=i_2,\ldots,x_{i+n}=i_{i+n}\right\}\right)$$
$$=q_{i_1}p_{i_1 i_2}\cdots p_{i_n i_{n+1}}$$

şeklinde tanımlanan ölçüme Markov ölçümü denir.

Uyarı 1.3.2. Markov ölçümünde başlangıç durumu bellidir ve şartlı olasılık vardır.

Tanım 1.3.3. Tüm olası durumların oluşturduğu kümeye durum uzayı denir.

Örnek 1.3.4. $A=\left\{w\in\Omega: x_0=a_0, x_1=a_1,\ldots,x_3=a_3\right\}$ olarak verilsin. kümesinin Markov ölçümünü alalım. Burada durum uzayı $E=\{0,1\}$ olsun ve A kümesini ince silindir kümelerin birleşimi olarak yazalım.



$$A=\left\{w\in\Omega: x_0=a_0, x_1=a_1, x_3=a_3\right\}$$
$$=\left\{w\in\Omega: x_0=a_0, x_1=a_1, x_2=0, x_3=a_3\right\}$$
$$\cup\left\{w\in\Omega: x_0=a_0, x_1=a_1, x_2=1, x_3=a_3\right\}$$

olup A kümesinin Markov ölçümü-

$$\mu(A)=\mu_{\pi P}\left(\left\{w\in\Omega: x_0=a_0, x_1=a_1, x_2=0, x_3=a_3\right\}\right)$$
$$+\mu_{\pi p}\left(\left\{w\in\Omega: x_0=a_0, x_1=a_1, x_2=1, x_3=a_3\right\}\right)$$
$$=q_{a_0}p_{a_0a_1}p_{a_10}p_{0a_3}+q_{a_0}p_{a_0a_1}p_{a_11}p_{1a_3}$$
$$=q_{a_0}p_{a_0a_1}\left(p_{a_10}p_{0a_3}+p_{a_11}p_{1a_3}\right)$$

şeklindedir.

Örnek 1.3.5. 1.2.11 örneğindeki stokastik matris ve olasılık vektörünü kullanarak ${}_5[12121]_9$ kümesinin Markov ölçümünü bulalım. Burada ${}_5[12121]_9$ kümesi bir ince silindir kümedir ve kısa olması için bu şekilde yazılır.

$$\mu_{\pi p}\left({}_5[12121]_9\right)=\mu_{\pi p}\left(\left\{w\in\Omega: x_5=1, x_6=2, x_7=1, x_8=2, x_9=1\right\}\right)$$
$$=q_1p_{12}p_{21}p_{12}p_{21}$$

$$=\frac{1}{3}.\frac{1}{2}.\frac{1}{4}.\frac{1}{2}.\frac{1}{4}=\frac{1}{192} \text{ bulunur.}$$

Tanım 1.3.6. $E=\{0,1,2,\ldots,r-1\}$ kümesi için $\sigma: E^{\mathbb{Z}} \to E^{\mathbb{Z}}$ olsun. $(\sigma x)_i = x_{i+1}$ fonksiyonuna kaydırma dönüşümü denir.

Teorem 1.3.7. $M_\sigma\left(S^{\mathbb{Z}}\right)$ bütün σ-değişmez ölçümlerin kümesini göstermek üzere $\mu \in M_\sigma(S^{\mathbb{Z}})$ için aşağıdaki özellikler geçerlidir. Buradaki ölçüm Markov ölçümüdür.

1) $\sum_{a_0 \in S} \mu\left({}_0[a_0]\right)=1$ ve ${}_0[a_0]=\left\{w \in S^{\mathbb{Z}} : x_0 = a_0, a_0 \in S\right\}$ yani

$$\mu\left(\bigcup_{a_0 \in S}\left({}_0[a_0]\right)\right)=\mu\left(S^{\mathbb{Z}}\right)=1 \text{ şeklindedir.}$$

2) $(a_0,a_1,\ldots,a_k)$ bir blok $n \in \mathbb{Z}$ ise $\mu\left({}_n[a_0,a_1,\ldots,a_k]\right) \geq 0$

3) $\mu\left({}_m[a_0,a_1,\ldots,a_k]\right)=\sum_{i=1}^{n} \mu\left({}_m[a_0,a_1,\ldots,a_k,i]\right)$

4) $\mu\left({}_m[a_0,a_1,\ldots,a_k]\right)=\sum_{i=1}^{n} \mu\left({}_m[a_{-1},a_0,a_1,\ldots,a_k]\right)=\sum_{i=1}^{n} \mu\left({}_n[i,a_0,a_1,\ldots,a_k]\right)$

olacaktır (Denker ve ark. 1976).

1.4. Ayrışımlar

Tanım 1.4.1. $(X,\mathcal{A},\mu)$ olasılık ölçüm uzayı olsun. $\mathcal{A}$'nın ayrık alt kümelerinden oluşan ve birleşimi X'e eşit olan bir $\xi=\{A_1,A_2,\ldots,A_k\}$ ailesine X'in bir ölçülebilir ayrışımı denir. Burada $k \in I$ için I sonlu ise ayrışım sonludur.

Örnek 1.4.2. $\left([0,1), B(IR), \lambda\right)$ Lebesgue olasılık ölçüm uzayında

$$\xi=\left\{\left[\frac{i}{2^n},\frac{i+1}{2^n}\right);\ 0 \leq i \leq 2^n,\ n \in \mathbb{Z}\right\}$$

$[0,1)$ kümesinin bir ayrışımıdır. n=1 için $\left\{\left[0,\frac{1}{2}\right),\left[\frac{1}{2},1\right)\right\}$ olur.

Tanım 1.4.3. $(X, \mathcal{A}, \mu)$ uzayı üzerinde ξ ve η ayrışımları verilsin. Her $A \in \xi$ ve $B \in \eta$ için $\mu(A \cap B) = \mu(A).\mu(B)$ ise ξ ve η'ya bağımsız ayrışımlar denir.

Tanım 1.4.4. ξ ve η, $(X, \mathcal{A}, \mu)$ ' nın iki ayrışımı olsun. ξ nin her elemanı η nın elemanlarının birleşimi olarak yazılabiliyorsa η, ξ den daha incedir denir ve $\xi \leq \eta$ olarak gösterilir. Eğer $\xi \leq \eta$ ve $\eta \leq \xi$ ise $\xi = \eta$ olur.

Tanım 1.4.5. $(X, \mathcal{A}, \mu)$ uzayı için ξ ve η, X'in ayrışımları olsunlar. $\xi = \{A_1, A_2, \ldots, A_n\}, \eta = \{C_1, C_2, \ldots, C_k\}$ olursa bu iki ayrışımın katılımı yine bir ayrışımdır. $\xi \vee \eta = \{A_i \cap C_j : 1 \leq i \leq n, 1 \leq j \leq k\}$ şeklinde gösterilir.

Örnek 1.4.6. $\xi = \left\{\left[0, \frac{1}{3}\right), \left[\frac{1}{3}, \frac{2}{3}\right), \left[\frac{2}{3}, 1\right)\right\}$ ve

$\eta = \left\{\left[0, \frac{1}{4}\right), \left[\frac{1}{4}, \frac{2}{4}\right), \left[\frac{2}{4}, \frac{3}{4}\right), \left[\frac{3}{4}, 1\right)\right\}$, $[0,1)$ aralığının iki ayrışımı olsunlar. Bu ayrışımların katılımı

$$\xi \vee \eta = \left\{\left[0, \frac{1}{4}\right), \left[\frac{1}{4}, \frac{1}{3}\right), \left[\frac{1}{3}, \frac{1}{2}\right), \left[\frac{1}{2}, \frac{2}{3}\right), \left[\frac{2}{3}, \frac{3}{4}\right), \left[\frac{3}{4}, 1\right)\right\}$$

şeklindedir.

Tanım 1.4.7. Her $A_i \in \xi$ ve her $B_i \in \eta$ için $\mu(A_i \Delta B_i) = 0$ ise ξ ve η ayrışımlarına denk ayrışımlar denir. $\xi \approx \eta$ ile gösterilir. Eğer $\xi \subset \eta$ ve $\eta \subset \xi$ ise $\xi = \eta$'dır.

Şimdi bir ayrışımın ölçüm entropisi ile ilgili temel tanımları verelim.

Tanım 1.4.8. $(X, \mathcal{B}, m)$ bir olasılık ölçüm uzayı ve ξ, X' in bir sonlu ayrışımı olsun. ξ ayrışımının entropisi;

$$H(\xi) = -\sum_{i=1}^{k} m(A_i) \log m(A_i) \tag{1.1}$$

şeklindedir.

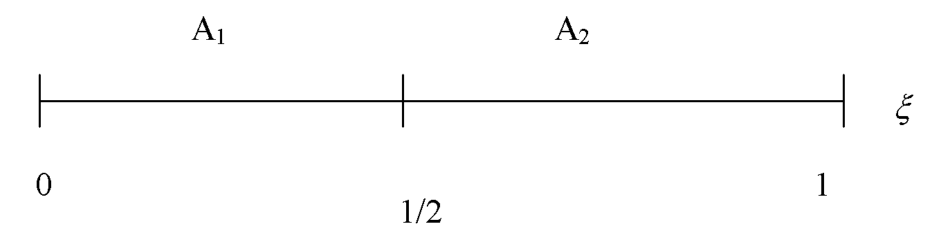
Tanım 1.4.9. $\xi=\{A_1,A_2,\ldots,A_k\}$, $\eta=\{C_1,C_2,\ldots,C_p\}$ X' in iki sonlu ayrışımı olsun. η verildiğinde ξ' nin şartlı ölçüm entropisi;

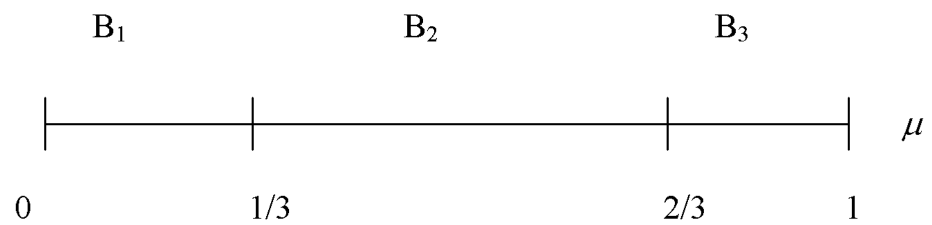
$$H(\xi\setminus\eta)=-\sum_{j=1}^{p}m\left(C_j\right)\sum_{i=1}^{k}\frac{m\left(A_i\cap C_j\right)}{m\left(C_j\right)}\log\frac{m\left(A_i\cap C_j\right)}{m\left(C_j\right)}$$

$$=-\sum_{i,j}m\left(A_i\cap C_j\right)\log\frac{m\left(A_i\cap C_j\right)}{m\left(C_j\right)} \tag{1.2}$$

olarak tanımlanır ve burada $m(C_j)\neq 0$'dır.

Örnek 1.4.10. X=[0,1), $B([0,1))$ Borel $\sigma-$cebir ve μ' de genel Lebesgue ölçümü olsun. $[0,1)$ aralığında $\xi=\left\{\left[0,\frac{1}{2}\right),\left[\frac{1}{2},1\right)\right\}$, $\eta=\left\{\left[0,\frac{1}{3}\right),\left[\frac{1}{3},\frac{2}{3}\right),\left[\frac{2}{3},1\right)\right\}$ gibi iki ayrışımı göz önüne alalım. Buna göre $H(\xi/\eta)$ şartlı ölçüm entropisini hesaplayalım.





$$\xi_{B_1}=\left\{\left[0,\frac{1}{3}\right),\phi\right\},\ \xi_{B_2}=\left\{\left[\frac{1}{3},\frac{1}{2}\right),\left[\frac{1}{2},\frac{2}{3}\right)\right\},\ \xi_{B_3}=\left\{\phi,\left[\frac{2}{3},1\right)\right\}$$

$$\left.\begin{array}{l}\mu_{B_1}(A_1)=\dfrac{\mu(A_1\cap B_1)}{\mu(B_1)}=\dfrac{\frac{1}{3}}{\frac{1}{3}}=1\\[2ex]\mu_{B_1}(A_2)=\dfrac{\mu(A_2\cap B_1)}{\mu(B_1)}=0\end{array}\right\}1+0=1$$

$$\left.\begin{array}{l} \mu_{B_2}(A_1)=\dfrac{\mu(A_1\cap B_2)}{\mu(B_2)}=\dfrac{\frac{1}{6}}{\frac{1}{3}}=\dfrac{1}{2} \\ \mu_{B_2}(A_2)=\dfrac{\mu(A_2\cap B_2)}{\mu(B_2)}=\dfrac{\frac{1}{6}}{\frac{1}{3}}=\dfrac{1}{2} \end{array}\right\}=\frac{1}{2}+\frac{1}{2}=1$$

$$\left.\begin{array}{l} \mu_{B_3}(A_1)=\dfrac{\mu(A_1\cap B_3)}{\mu(B_3)}=0 \\ \mu_{B_3}(A_2)=\dfrac{\mu(A_2\cap B_3)}{\mu(B_3)}=\dfrac{\frac{1}{3}}{\frac{1}{3}}=1 \end{array}\right\}=0+1=1$$

olduğundan olasılık ölçümleri elde edilmiş olur. Bulduğumuz sonuçları kullanarak şartlı ölçüm entropi değerini hesaplayalım.

$$\begin{aligned} H(\xi_{B_1}) &= -\sum_{i=1}^{2}\mu_{B_1}(A_i\cap B_1)\log\mu_{B_1}(A_i\cap B_1) \\ &= -\mu_{B_1}(A_1\cap B_1)\log\mu_{B_1}(A_1\cap B_1)-\mu_{B_1}(A_2\cap B_1)\log\mu_{B_1}(A_2\cap B_1) \\ &= 1\log 1 \\ &= 0 \end{aligned}$$

$$\begin{aligned} H(\xi_{B_2}) &= -\sum_{i=1}^{2}\mu_{B_2}(A_i\cap B_2)\log\mu_{B_2}(A_i\cap B_2) \\ &= -\mu_{B_2}(A_1\cap B_2)\log\mu_{B_2}(A_1\cap B_2)-\mu_{B_2}(A_2\cap B_2)\log\mu_{B_2}(A_2\cap B_2) \\ &= -\frac{1}{2}\log\frac{1}{2}-\frac{1}{2}\log\frac{1}{2} \\ &= 0,3 \end{aligned}$$

$$\begin{aligned} H(\xi_{B_3}) &= -\sum_{i=1}^{2}\mu_{B_3}(A_i\cap B_3)\log\mu_{B_3}(A_i\cap B_3) \\ &= -\mu_{B_3}(A_1\cap B_3)\log\mu_{B_3}(A_1\cap B_3)-\mu_{B_3}(A_2\cap B_3)\log\mu_{B_3}(A_2\cap B_3) \\ &= 1\log 1 \\ &= 0 \end{aligned}$$

ve böylece şartlı entropi değeri

$$H(\xi/\eta)=\mu(B_1)H(\xi_{B_1})+\mu(B_2)H(\xi_{B_2})+\mu(B_3)H(\xi_{B_3})$$

$$=\frac{1}{3}(0.3)=0.1$$

olarak elde edilir.

Aşağıdaki teorem hem ergodik teorideki şartlı ölçüm entropi hem de enformasyon teorideki şartlı entropinin özelliklerini vermektedir. Bu teorem daha sonraki bölümlerde bazı teoremlerin ispatında bize yol gösterecektir.

Teorem 1.4.11. $(X, \mathcal{B}, \mu)$ bir olasılık uzayı olsun. Eğer ξ, η, γ; X' in sonlu ayrışımları ise;

a) $H(\xi \vee \eta / \gamma) = H(\xi / \gamma) + H(\eta / \xi \vee \gamma)$

b) $H(\xi \vee \eta) = H(\xi) + H(\eta / \xi)$

c) $\xi \subseteq \eta \Rightarrow H(\xi / \gamma) \leq H(\eta / \gamma)$

d) $\xi \subseteq \eta \Rightarrow H(\xi) \leq H(\eta)$

e) $\eta \subseteq \gamma \Rightarrow H(\xi / \eta) \geq H(\xi / \gamma)$

f) $H(\xi) \geq H(\xi / \gamma)$

g) $H(\xi \vee \eta / \gamma) \leq H(\xi / \gamma) + H(\eta / \gamma)$

h) $H(\xi \vee \eta) \leq H(\xi) + H(\eta)$

i) Eğer T ölçümü koruyan dönüşüm ise $H\left(T^{-1}\xi / T^{-1}\eta\right) = H\left(\xi / \eta\right)$

j) $H\left(T^{-1}\xi\right) = H\left(\xi\right)$ (Walters, 1982).

Uyarı 1.4.12. Enformasyon teoride verilen $H(X, Y)$ kavramı ergodik teoride $H(\xi \vee \eta)$ şeklinde yorumlanabilir.

Lemma 1.4.13. $p_1, p_2, \ldots, p_M$ ve $q_1, q_2, \ldots, q_M$ keyfi pozitif sayılar ve $\sum_{i=1}^{M} p_i = \sum_{i=1}^{M} q_i = 1$ olsun. O zaman $-\sum_{i=1}^{M} p_i \log p_i \leq -\sum_{i=1}^{M} p_i \log q_i$ 'dir. Bütün i 'ler için $p_i = q_i$ ise eşitsizlik, eşitlik haline gelir (Ash, 1965).

Tanım 1.4.14. Olasılıksal değişkenlerin oluşturdukları $\{X_t : t \in T\}$ kümesine olasılıksal süreç denir. Zaman sürecin parametresidir.

Tanım 1.4.15. $T = \{0, \mp 1, \mp 2, \ldots\}$ ya da $T = \{0, 1, 2, \ldots\}$ olduğunda sürece kesikli değiştirgeli süreç, $T = \{t : -\infty < t < \infty\}$ ya da $T = \{t : t \geq 0\}$ olduğunda sürekli değiştirgeli süreç denir.

Tanım 1.4.16. Değeri bir deney sonucunda belirtilen bir değişkene rastgele değişken denir.

Teorem 1.4.17. $H(X,Y) \leq H(X)+H(Y)$'dir. Ancak ve ancak X ve Y bağımsız rastgele değişkenler olduklarında eşitsizlik, eşitlik haline gelir (Ash, 1965).

İspat: $p(x_i)=\sum_{j=1}^{L} p(x_i,y_j)$ ve $p(y_j)=\sum_{i=1}^{M} p(x_i,y_j)$ olduğundan

$$H(X)=-\sum_{i=1}^{M} p(x_i)\log p(x_i)=-\sum_{i=1}^{M}\sum_{j=1}^{L} p(x_i,y_j)\log p(x_i) \quad \text{ve}$$

$$H(Y)=-\sum_{j=1}^{L} p(y_j)\log p(y_j)=-\sum_{i=1}^{M}\sum_{j=1}^{L} p(x_i,y_j)\log p(y_j)$$ yazılabilir. Buradan

$$H(X)+H(Y)=-\sum_{i=1}^{M}\sum_{j=1}^{L} p(x_i,y_j)\left[\log p(x_i)+\log p(y_j)\right]$$

$$=-\sum_{i=1}^{M}\sum_{j=1}^{L} p(x_i,y_j)\log p(x_i)p(y_j)=-\sum_{i=1}^{M}\sum_{j=1}^{L} p_{ij}\log q_{ij},$$

burada $q_{ij}=p(x_i).p(y_j)$ dir.

$H(X,Y)=-\sum_{i=1}^{M}\sum_{j=1}^{L} p_{ij}\log p_{ij}$ olduğu biliniyor. Lemma 1.4.13'den

$-\sum_{i=1}^{M}\sum_{j=1}^{L} p_{ij}\log p_{ij} \leq -\sum_{i=1}^{M}\sum_{j=1}^{L} p_{ij}\log q_{ij}$ elde edilir. Eğer bütün i,j'ler için $p_{ij}=q_{ij}$ ise buradan,

$$\sum_{i=1}^{M}\sum_{j=1}^{L} q_{ij}=\sum_{i=1}^{M} p(x_i)\sum_{j=1}^{L} q(y_j)=1.1=1\text{'dir.}$$

Böylece

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

eşitsizliği elde edilir ve bütün i,j' ler için $p(x_i,y_j)=p(x_i)p(y_j)$ olması durumunda yani X ve Y bağımsız değişkenler olduğunda bu eşitsizlik eşitlenir.

Sonuç 1.4.18. $H(X_1,X_2,\cdots,X_N) \leq H(X_1)+H(X_2)+\cdots+H(X_N)$'dir. Eğer $X_1,X_2,\cdots,X_N$'ler bağımsız değişkenler iseler eşitsizlik, eşitlik haline gelir (Ash, 1965).

Sonuç 1.4.19. $H(X_1,\ldots,X_n,Y_1,\ldots,Y_m) \le H(X_1,\ldots,X_n) + H(Y_1,\ldots,Y_m)$'dir. Eğer $X=(X_1,\ldots,X_n)$ ve $Y=(Y_1,\ldots,Y_m)$ serbest (rastgele) vektörleri bağımsız iseler eşitsizlik eşitlenir yani bütün $\alpha_1,\alpha_2,\ldots,\alpha_n$, $\beta_1,\beta_2,\ldots,\beta_m$'ler için

$$P\{X_1=\alpha_1,\ldots,X_n=\alpha_n,Y_1=\beta_1,\ldots,Y_m=\beta_m\}$$
$$=P\{X_1=\alpha_1,\ldots,X_n=\alpha_n\}P\{Y_1=\beta_1,\ldots,Y_m=\beta_m\}$$

dir (Ash, 1965).

Tanım 1.4.20. $X=x_i$ verildiğinde Y'nin şartlı belirsizliği (entropisi);

$$H(Y \mid X=x_i) = -\sum_{j=1}^{L} p(y_j \mid x_i)\log p(y_j \mid x_i) \tag{1.3}$$

ile verilir. $H(Y \mid X=x_i)$ belirsizliğinin bir ağırlıklı ortalamasına göre verilen X, Y'nin şartlı belirsizliğini tanımlar yani;

$$H(Y \mid X) = p(x_1)H(Y \mid X=x_1)+\ldots+p(x_M)H(Y \mid X=x_M)$$
$$=-\sum_{i=1}^{M} p(x_i)\sum_{j=1}^{L} p(y_j \mid x_i)\log p(y_j \mid x_i)$$

olur. Eğer $p(x_i,y_j)=p(x_i)p(y_j \mid x_i)$ ise o zaman,

$$H(Y \mid X) = -\sum_{i=1}^{M}\sum_{j=1}^{L} p(x_i,y_j)\log p(y_j \mid x_i) \tag{1.4}$$

elde edilir (Ash, 1965). Yine ergodik teoride olduğu gibi aşağıdaki tanımlar elde edilir.

Tanım 1.4.21. İki rastgele değişkenden daha fazlası için şartlı belirsizlik şöyle tanımlanır.

$$H(Y,Z \mid X) = -\sum_{i,j,k} p(x_i,y_j,z_k)\log p(y_j,z_k \mid x_i) \tag{1.5}$$

değerine, X verildiğinde Y ve Z'nin belirsizliği denir.

$$H(Z \mid X,Y) = -\sum_{i,j,k} p(x_i,y_j,z_k)\log p(z_k \mid x_i,y_j) \tag{1.6}$$

değerine de X ve Y verildiğinde Z'nin belirsizliği denir.

Tanım 1.4.22. 1.4.20 ve 1.4.21 tanımlarını genelleştirirsek;

$$H\left(Y_1,\ldots,Y_m \mid X_1,\ldots,X_n\right) = -\sum_{x_1,\ldots,x_n,y_1,\ldots,y_m} p\left(x_1,\ldots,x_n,y_1,\ldots,y_m\right)\log p\left(y_1,\ldots,y_m \mid x_1,\ldots,x_n\right)$$

entropi değeri $X_1,\ldots,X_n$ verildiğinde $Y_1,\ldots,Y_m$ 'lerin belirsizliğidir.

Teorem 1.4.23. $H\left(X,Y\right) = H\left(X\right) + H\left(Y \mid X\right) = H\left(Y\right) + H\left(X \mid Y\right)$ 'dir. (Ash, 1965).

İspat: $H\left(X,Y\right) = -\sum_{i=1}^{M}\sum_{j=1}^{L} p\left(x_i,y_j\right)\log p\left(x_i,y_j\right)$

$$= -\sum_{i=1}^{M}\sum_{j=1}^{L} p\left(x_i,y_j\right)\log p\left(x_i\right)p\left(y_j \mid x_i\right)$$

$$= -\sum_{i=1}^{M}\sum_{j=1}^{L} p\left(x_i,y_j\right)\log p\left(x_i\right) - \sum_{i=1}^{M}\sum_{j=1}^{L} p\left(x_i,y_j\right)\log p\left(y_j \mid x_i\right)$$

$$= -\sum_{i=1}^{M} p\left(x_i\right)\log p\left(x_i\right) + H\left(Y \mid X\right)$$

$= H\left(X\right) + H\left(Y \mid X\right)$ 'dir.

$H\left(X,Y\right) = H\left(Y\right) + H\left(X \mid Y\right)$ benzer şekilde ispatlanır.

Teorem 1.4.24. $H\left(X,Y,Z\right) = H\left(X\right) + H(Y \mid X) + H(Z \mid X,Y)$

$$= H\left(X,Y\right) + H\left(Z \mid X,Y\right)$$

$$= H\left(X\right) + H\left(Y,Z \mid X\right)$$

olur (Ash, 1965).

Teorem 1.4.25. $H\left(Y \mid X\right) \leq H\left(Y\right)$ 'dir. X ve Y bağımsız değişken olduklarında eşitsizlik eşitlenir (Ash, 1965).

2. ÖNCEKİ ÇALIŞMALAR

2.1. Sonlu Markov Zincirlerinin Entropisi

Bir kaynağın entropisini hesaplayabilmek için bazı tanım ve teoremlere ihtiyacımız vardır. Bu yüzden bu kesimde önce aşağıdaki tanım ve teoremleri verelim.

Tanım 2.1.1. Bir sonlu Π kümesi içinde değerler alan rastgele değişkenlerin bir durağan dizisine kaynak denir.

Enformasyon teoride ise, kaynak birim başına verilen bilgi miktarıdır.

Tanım 2.1.2. Dizin kümesindeki n sayıda zaman noktasının herhangi bir $t_1 < t_2 < \ldots < t_n$ kümesi için, X_{t_n}'in $X_{t_1}, X_{t_2}, \ldots, X_{t_{n-1}}$'in verilen değerlerine göre koşullu dağılımı yalnızca $X_{t_{n-1}}$'in değerine bağlı ise $\{X_t, t \in T\}$ sürecine Markov süreci denir. Buna göre herhangi bir gerçel $X_1, X_2, \ldots, X_n$ sayıları için;

$$P\left(X_{t_n} = x_n \mid X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n-1}\right) = P\left(X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}\right)$$

olur. Bu özelliğe Markov özelliği denir.

Tanım 2.1.3. Durum uzayında sonlu sayıda ya da sayılabilir sonsuzlukta durum varsa buna kesikli durum uzayı denir.

Tanım 2.1.4. Durum uzayı kesikli olan bir Markov sürecine Markov zinciri denir. Durum uzayı sonlu ise zincire sonlu Markov zinciri, sonsuz ise sonsuz Markov zinciri denir.

Tanım 2.1.5. Bir kesikli rastgele değişkenin olasılık dağılımı veya olasılık fonksiyonu rastgele değişkenin alabileceği değerlerin karşılık gelen olasılıklarla birlikte belirtilmesidir.

Tanım 2.1.6. *X*, sonlu sayıdaki $x_1, x_2, \ldots, x_N$ değerlerini $f(x_i) = P(X = x_i)$, $i = 1, 2, \ldots, N$ olasılıkları ile alabilen kesikli rastgele değişken olsun. Bu durumda aşağıdaki koşulları sağlayan $f(x)$ fonksiyonuna *X*'in olasılık fonksiyonu denir.

1. $f(x) \geq 0$, tüm *x*'ler için
2. $\sum_{i=1}^{N} f(x_i) = 1$.

Tanım 2.1.7. Bir *X* rastgele değişkeninin dağılım fonksiyonu *F*(*X*) ile gösterilir ve *X*'in *x*'e eşit ya da daha küçük olma olasılığıdır.

$$F(X) = P(x \leq X) = \sum_{x_i \leq X} f(x_i) \tag{2.1}$$

olup burada *f* bir olasılık fonksiyonudur.

Lioyd ve Pagels (1988) tarafından öne sürülen termodinamik derinlik fikri Markov yörüngelerinin entropisinin hesaplanmasını gerektirir. Bu anlamda indirgenemez sonlu Markov zinciri alınır. Bu Markov zincirinde *P* geçiş matrisi ve ilgili entropi değeri

$$H(X) = -\sum_{i,j} \mu_i p_{ij} \log p_{ij} \tag{2.2}$$

dir. Formüldeki μ_i değerleri $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ olmak üzere verilen durağan dağılım yani olasılık vektörüdür. Bu çalışmada *i* durumunda başlayan ve biten rastgele T_{ii} yörüngesinin entropisi

$$H(T_{ii}) = \frac{H(X)}{\mu_i} \tag{2.3}$$

ile verilir. $H(T_{ij})$ entropileri için genel bir kapalı çözüm formu

$$H = K - \tilde{K} + H_\Delta \tag{2.4}$$

ile verilmiştir. Burada *H* yörünge entropilerinin matrisidir. $H_{ij} = H(T_{ij})$,

$$K = (I - P + A)^{-1}(H^* - H_\Delta) \tag{2.5}$$

ve burada $\tilde{K}$, K'nın diyagonal $\tilde{K}_{ij}$ elemanına eşit olan *ij*. K_{ij} elemanlarının bir matrisidir. A ise $A_{ij} = \mu_j$ girişleriyle durağan ihtimaller matrisidir. H^*,

$$H^*_{ij} = H(P_i) = -\sum_k P_{ik} \log P_{ik} \tag{2.6}$$

girişleriyle tek basamaklı entropilerin matrisi, H_Δ ‘da

$$(H_\Delta)_{ii} = \frac{H(X)}{\mu_i} \tag{2.7}$$

girişleriyle bir diyagonal matristir. Burada birinci basamak entropilerinin matrisi;

$$H^* = \begin{bmatrix} H(P_1) & H(P_1) & \dots & H(P_1) \\ H(P_2) & H(P_2) & \dots & H(P_2) \\ \dots & \dots & \dots & \dots \\ H(P_m) & H(P_m) & \dots & H(P_m) \end{bmatrix}$$

ve H ile ilişkilendirilmiş diyagonal matriste;

$$H_\Delta = \begin{bmatrix} H_{11} & 0 & \dots & 0 \\ 0 & H_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & H_{mm} \end{bmatrix}$$

şeklindedir.

Bu bölümde sonlu indirgenemez Markov zincirlerinin yörüngelerinin entropisini inceleyeceğiz. P geçiş matrisi ve $X_1 = i$ başlangıç durumu olan bir sonlu indirgenemez Markov zincirini göz önüne alalım. Entropi oranı,

$$H(X) = \lim_{n\to\infty} \frac{H(x_1, x_2, \dots, x_n)}{n} \tag{2.8}$$

şeklinde ifade edilir (Ekroot ve Cover, 1993). Çünkü limit vardır ve bu limitin sembolik dinamik sistemdeki karşılığı

$$H(X) = -\sum_{i,j} \mu_i P_{ij} \log P_{ij}$$

şeklindedir. Burada μ tüm j değerleri için

$$\mu_j = -\sum_i \mu_i P_{ij} \tag{2.9}$$

denkleminin çözümüdür.

Tanım 2.1.8. Bir Markov zincirinin i durumundan j durumuna bir T_{ij} yörüngesi i başlangıç durumundan j son durumuna bir yoldur (Burada i’den j’ye olan tüm geçişlerde j durumu kesinlikle olmayacaktır).

Bir $t_{ij} = ix_2x_3 \dots x_k j$ yörüngesinin $p(t_{ij})$ olasılığı $p(t_{ij}) = p_{ix_2} p_{x_2x_3} \dots p_{x_k j}$ ile verilir. Bu $x_1 = i$ verildiğinde i’den j’ye t_{ij} yörüngesinin şartlı olasılığıdır. P’nin indirgenemezliği $\sum_{t_{ij}\in T_{ij}} P\left(t_{ij}\right)=1$ olduğunu gerçekler ve burada T_{ij} i'den j' ye bütün yörüngelerin kümesidir.

Tanım 2.1.9. Yörüngenin i'den j'ye H_{ij} entropisi şöyle tanımlanır.

$$H_{ij}=H\left(T_{ij}\right)=-\sum_{t_{ij}\in T_{ij}} p\left(t_{ij}\right)\log p\left(t_{ij}\right) \tag{2.10}$$

Aşağıda vereceğimiz teoremler Ekroot ve Cover (1993) tarafından verilmiş olup sonraki bölümlerde yapacağımız uygulamalar için bize yol gösterecektir.

Teorem 2.1.10. Bir indirgenemez Markov zinciri için, i durumundan tekrar geri i durumuna rastgele yörüngenin H_{ii} entropisi $H_{ii}=\dfrac{H(X)}{\mu_i}$ biçimindedir. Burada μ_i, i durumu için durağan ihtimaldir ve

$$H(X)=-\sum_{i,j}\mu_i P_{ij}\log P_{ij}$$

şeklinde verilen entropi oranıdır (Ekroot ve Cover, 1993).

İspat: Matris

$$H=H^{*}+PH-PH_{\Delta} \tag{2.11}$$

şeklindedir. Bu denklemi soldan μ durağan dağılımı ile çarparsak, $\mu H=\mu H^{*}+\mu PH-\mu PH_{\Delta}$ olup $\mu P=\mu$ eşitliğini uygularsak, $\mu H=\mu H^{*}+\mu H-\mu H_{\Delta}$ elde edilir ve buradan $\mu H^{*}=\mu H_{\Delta}$ bulunur. μH^{*}, her biri $\sum_{k}\mu_k H\left(P_k\right)=H(X)$ ile verilen eşit bileşenli bir vektördür.

Burada $H(X)=-\sum_{k,j}\mu_k P_{kj}\log P_{kj}$ verilen Markov zincirinin entropi oranıdır ve $\mu H^{*}=\mu H_{\Delta}$ eşitliği aşağıdaki şekilde yazılabilir.

$$\left(H(X),H(X),\ldots,H(X)\right)=\left(\mu_1 H_{11},\mu_2 H_{22},\ldots,\mu_m H_{mm}\right)$$

bileşenlerin denkleştirilmesiyle, $H_{ii}=\dfrac{H(X)}{\mu_i}$ elde edilir.

Şimdi yineleme bağıntısının çözümü olan $H = H^* + PH - PH_\Delta$ ifadesini kullanarak önceden belirlenmiş diyagonal değerler $H_{ii} = \dfrac{H(X)}{\mu_i}$ ‘den elde edilen genel teoremi verelim.

Teorem 2.1.11. Eğer P, indirgenemez bir sonlu Markov zincirinin geçiş matrisi ise, o zaman yörünge entropilerinin H matrisi $H = K - \tilde{K} + H_\Delta$ biçiminde olur. Bu durumda

$$K = (I - P + A)^{-1}\left(H^* - H_\Delta\right)$$

$\tilde{K}_{ij} = K_{jj}$ tüm i,j’ ler için

$A_{ij} = \mu_j$ tüm i,j’ ler için

$H^*_{ij} = H\left(P_i\right)$ tüm i,j’ ler için ve

$$\left(H_\Delta\right)_{ij} = \begin{cases} \dfrac{H(X)}{\mu_i}, & i = j \\ 0, & i \neq j \end{cases}$$

denklemleri geçerlidir (Ekroot ve Cover, 1993).

İspat: İspat 3 kısımda incelenecektir. Öncelikle bir aperiyodik Markov zinciri için (2.11) yineleme bağıntısından bir çözüm elde edeceğiz. Sonra ispatı periyodik Markov zinciri için yapıp, son olarakta çözümün tekliğini ispatlayacağız. (2.11) eşitliğinin her iki tarafından H_Δ çıkarırsak;

$$H - H_\Delta = \left(H^* - H_\Delta\right) + P\left(H - H_\Delta\right)$$

veya

$$\hat{H} = \left(H^* - H_\Delta\right) + P\hat{H} \tag{2.12}$$

elde ederiz. $\hat{H} = H - H_\Delta$ olup burada $\hat{H}$, H ’ın yörünge entropilerinin matrisinin köşegen dışı elemanlarının matrisidir. (2.12) eşitliğinde $\hat{H}$ için yerine koyma tekrarlanırsa;

$$\hat{H} = P\hat{H} + \left(H^* - H_\Delta\right)$$

$$= P\left[\left(P\hat{H}\right)+\left(H^{*}-H_{\Delta}\right)\right]+\left(H^{*}-H_{\Delta}\right)$$

$$= P^{2}\hat{H}+P\left(H^{*}-H_{\Delta}\right)+\left(H^{*}-H_{\Delta}\right)$$

$$= P^{2}\left[P\hat{H}+\left(H^{*}-H_{\Delta}\right)\right]+P\left(H^{*}-H_{\Delta}\right)+\left(H^{*}-H_{\Delta}\right)$$

$$= P^{3}\hat{H}+P^{2}\left(H^{*}-H_{\Delta}\right)+P\left(H^{*}-H_{\Delta}\right)+\left(H^{*}-H_{\Delta}\right)$$

.

.

.

$$= P^{n}\hat{H}+P^{n-1}\left(H^{*}-H_{\Delta}\right)+\ldots+P\left(H^{*}-H_{\Delta}\right)+\left(H^{*}-H_{\Delta}\right)$$

$$= P^{n}\hat{H}+\left(P^{n-1}+P^{n-2}+\ldots+P+I\right)\left(H^{*}-H_{\Delta}\right)$$

$$= P^{n}\hat{H}+\left(I+P+\ldots+P^{n-1}\right)\left(H^{*}-H_{\Delta}\right) \tag{2.13}$$

elde edilir. $I+P+P^{2}+\ldots+P^{n-1}$ yakınsak olmadığından ve 2.1.10 teoreminden

$$\hat{H} = P^{n}\hat{H}+\left(I+P-A+P^{2}-A+\ldots+P^{n-1}-A\right)\left(H^{*}-H_{\Delta}\right) \tag{2.14}$$

bulunur. $(P-A)^{n}$ 'in bir iki terimli açılımında $AP = PA = A$ durumu uygulanırsa

$$P^{n}-A=(P-A)^{n} \tag{2.15}$$

olduğu görülür. (2.14) içinde (2.15)' i yazarsak;

$$\hat{H} = P^{n}\hat{H}+\sum_{k=0}^{n-1}(P-A)^{k}\left(H^{*}-H_{\Delta}\right) \tag{2.16}$$

elde ederiz. $P^{n}\to A$ iken P'nin aperiyodik olduğunu kabul edelim. (2.16) eşitliğinin limiti alındığında,

$$\hat{H} = A\hat{H}+\left(I-(P-A)\right)^{-1}\left(H^{*}-H_{\Delta}\right)$$

veya eşit bir biçimde

$$(I-A)\hat{H} = \left(I-(P-A)\right)^{-1}\left(H^{*}-H_{\Delta}\right) \tag{2.17}$$

sağlanır. $(I-A)$ nın tersinin olmadığı söylenebilir. Şimdi (2.17) eşitliğinin sağ tarafını ve $H_{ii}=0$ ifadesini kullanarak

$$(I-A)\hat{H}=K \tag{2.18}$$

ifadesini çözeceğiz. (2.18) eşitliğini ij. bileşenine göre

$$H_{ij}-\sum_{r}\mu_r \hat{H}_{rj}=K_{ij} \tag{2.19}$$

dir. $H_{jj}=0$ için jj. bileşen için ifade

$$-\sum_{r}\mu_r \hat{H}_{rj}=K_{jj} \tag{2.20}$$

şeklindedir. (2.19) ve (2.20)'den

$$\hat{H}_{ij}=K_{ij}-K_{jj} \tag{2.21}$$

elde edilir. (2.21)' i matris formunda yazarsak; $\hat{H}=K-\tilde{K}$ olur ve

$$\tilde{K}=\begin{bmatrix} K_{11} & K_{22} & \dots & K_{mm} \\ K_{11} & K_{22} & \dots & K_{mm} \\ \dots & \dots & \dots & \dots \\ K_{11} & K_{22} & \dots & K_{mm} \end{bmatrix}$$

dir. $H-H_{\Delta}=K-\tilde{K}$ olup aperiyodik indirgenemez Markov zincirleri için ispatlanmış olur.

Şimdi aperiyodiklik varsayımı çıkaralım. Eğer A'da P^n doğrudan yakınsak değilse fakat P^n, A'da Cesaro toplanabilir ise P periyodiktir. (2.13) eşitliğinin her iki tarafını n'e bölersek;

$$\hat{H}=\frac{1}{n}\sum_{k=1}^{n}P^k\hat{H}+\frac{1}{n}\sum_{k=1}^{n}\left(I+P+P^2+\dots+P^{k-1}\right)\left(H^*-H_{\Delta}\right) \tag{2.22}$$

dir. $A\left(H^*-H_{\Delta}\right)=0$ bağıntısını, (2.22) formu içinde yazarsak;

$$\hat{H}=\frac{1}{n}\sum_{k=1}^{n}P^k\hat{H}+\frac{1}{n}\sum_{k=1}^{n}\left(I+P-A+P^2-A+\dots+P^{k-1}-A\right)\left(H^*-H_{\Delta}\right)$$

dır. $P^n-A=(P-A)^n$ ifadesini tekrar uygularsak;

$$\hat{H}=\frac{1}{n}\sum_{k=1}^{n}P^k\hat{H}+\frac{1}{n}\sum_{k=1}^{n}\sum_{i=0}^{k-1}(P-A)^i\left(H^*-H_{\Delta}\right)$$

dir. P^n, A da Cesaro toplanabilir olduğundan; $\frac{1}{n}\sum_{k=1}^{n}P^k\to A$'dir. $I-P+A$'nın tersi vardır ve

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{l=0}^{k-1}(P-A)^{l} \rightarrow (I-P+A)^{-1}$$

ifadesine göre verilmiştir. $n \rightarrow \infty$ durumunda limit alınırsa,

$$\hat{H} = A\hat{H} + \left(I-(P-A)\right)^{-1}\left(H^{*}-H_{\Delta}\right)$$

elde edilir. Böylece indirgenemez periyodik ve aperiyodik Markov zincirleri için çözüm kanıtlanır. Şimdi (2.11)'de verilen H çözümünün tekliğini ispatlayıp ispatı sonlandıracağız. $H = H^{*} + PH - PH_{\Delta}$ bağıntısında iki çözüm H ve H' olsun;

O zaman ; $H = H^{*} + PH - PH_{\Delta}$ ve $H' = H^{*} + PH' - PH'_{\Delta}$ olur. Ancak 2.1.10 teoreminden $H_{\Delta} = H'_{\Delta}$ olduğundan

$$H - H' = P\left(H-H'\right) - P\left(H_{\Delta}-H'_{\Delta}\right) = P\left(H-H'\right)$$

sağlanır. Buradan,

$$\begin{aligned} H - H' &= P\left(H-H'\right) \\ &= P^{2}\left(H-H'\right) \\ &. \\ &. \\ &. \\ &= P^{n}\left(H-H'\right) \end{aligned}$$

'dir. İfadenin her iki tarafı n'e bölünürse

$H - H' = \frac{1}{n}\sum_{l=1}^{n} P^{l}\left(H-H'\right)$ bulunur. A' da P^{n} Cesaro-toplanabilirlik durumunu kullanırsak $H - H' = A\left(H-H'\right)$ bulunur ve $H - H' = 0$ olduğundan $H = H'$ olup çözüm tektir.

2.2. Bilgi Kaynakları

Bu kesimde enformasyon teori için çok önemli olan ve ileriki bölümlerde kullanacağımız kaynak kavramının tanımını vereceğiz.

Tanım 2.2.1. Enformasyon kaynağı mesaj üreten herhangi bir sistem olarak düşünülebilir. Üretilen bu mesaj sistemin sahip olduğu sonlu sayıdaki durumların birinin dışarı verilmesidir.

Tanım 2.2.2. $X_0, X_1, X_2, \ldots$ rastgele değişkenlerinin bir dizisi bir bilgi kaynağıdır. Öyle ki;

1. Bir sonlu Π kümesinde değerler alan her bir X_i kaynağın alfabesidir.

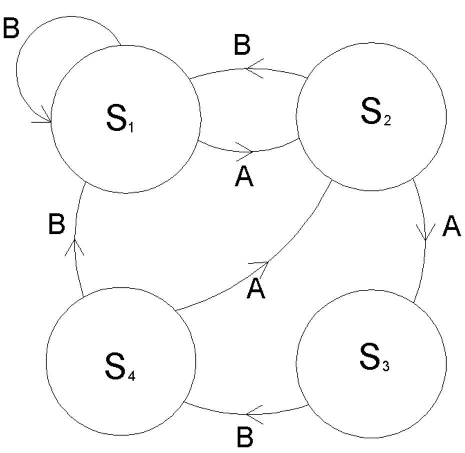
2. Bütün $\gamma_1, \gamma_2, \ldots, \gamma_k \in \Pi$ ve negatif olmayan $i_1, i_2, \ldots, i_k, h$ tamsayıları için

$P\left\{X_{i_1} = \gamma_1, \ldots, X_{i_k} = \gamma_k\right\} = P\left\{X_{i_1+h} = \gamma_1, \ldots, X_{i_k+h} = \gamma_k\right\}$ durağan dizisidir.

Tanım 2.2.3. $H\{X_n, n = 0,1,\ldots\}$ veya kısalık için $H\{X\}$ ile ifade edilen kaynağın belirsizliği, bir bilgi kaynağı $X_0, X_1, \ldots$ ile verilir ve $\lim_{n\to\infty} H\left(X_n \mid X_0, X_1, \ldots, X_{n-1}\right)$ şeklinde tanımlıdır.

Tanım 2.2.4. Π alfabesi, $S = \{S_1, S_2, \ldots, S_r\}$ durumlarının kümesi olsun. $f : S \to \Pi$ fonksiyonu ve $w = [w_1, w_2, \ldots, w_r]$ durağan dağılımı ile bir Markov bilgi kaynağını göz önüne alalım. Her bir S_k durumu için S_k'dan bir adımda erişilebilecek durumlar $S_{k_1}, S_{k_2}, \ldots, S_{k_{n_k}}$ olsun. Her bir S_k durumu için $f\left(S_{k_1}\right), \ldots, f\left(S_{k_{n_k}}\right)$ ayrık iseler kaynağın unifilar olduğu söylenebilir.

Örnek 2.2.5.



Şekil 2.1. Bilgi kaynağı

a) Unifilar Markov kaynağı

$$f(S_2) = f(S_3) = A$$

$$f(S_1) = f(S_4) = B$$

b) Unifilar olmayan Markov kaynağı

$$f(S_3) = A$$

$$f(S_1) = f(S_2) = f(S_4) = B$$

Bu örnekte verilen bilgi kaynağı karşılaştırma yapabilmemiz ve bazı sonuçlar elde edebilmemiz için sonraki bölümlerde de kullanılacaktır.

3. MATERYAL ve YÖNTEM

Bu tezde uygulanan temel yöntem şöyledir: İnternet üzerinden ve farklı kütüphanelerden elde edilen makaleler ve kitaplar temin edilerek bu kaynaklar iyice incelenip değerlendirildi. Daha önce yapılan çalışmalar karşılaştırılarak yeni sonuçlar ilave edilmeye çalışıldı. Temel olarak Ekroot ve Cover (1993) ile Ash'ın (1965) çalışmalarından yararlanıldı. Ekroot ve Cover (1993) sonlu Markov zinciri ve entropi kavramlarını incelediler. Ash (1965) ise bilgi kaynağını ve bu bilgi kaynağının entropisini inceledi. Bu makale ve kitaptaki bir takım sorular üzerinde bazı uyarlamalar yapılarak bu sorular yeniden çözüldü.

Ash'ın 1965'te çıkarmış olduğu "Information Theory" adlı kitabından ve Ekroot ile Cover'ın 1993'de yayınlamış oldukları "The entropy of Markov trajectories" adlı makalesinden faydalanılarak bir Markov yörüngesinin entropisi ile bir bilgi kaynağının entropisi karşılaştırıldı. Yapılan bu karşılaştırmalar bu tezin önemli bir kısmını teşkil etmektedir.

Bir bilgi kaynağının mertebesi, Ash'ın (1965) kitabında ki açıklamalardan ve örneklerden yararlanılarak elde edildi. Ayrıca eğer bir kaynağın mertebesi M ise bu kaynağın orijinal entropi değerinin ard arda gelen $M+1$ sembolleri ile ard arda gelen M sembollerinin belirsizliğinin farkı olduğu uygulamalarla gösterildi.

Son olarak Ekroot ve Cover (1991) tarafından yayınlanmış olan "The entropy of a randomly stopped sequence" adlı makaleden faydalanılarak bağımsız özdeşçe dağılmış bir dizinin durdurma zamanı kavramı ve bununla ilgili teoremler incelendi.

4. ARAŞTIRMA BULGULARI ve TARTIŞMA

Çalışmamızın asıl kısmını içeren bu bölümde Ekroot ve Cover (1993) ile Ash (1965) tarafından bağımsız olarak incelenen entropi tanımlarını ve uygulamalarını detaylı olarak inceleyeceğiz. Ekroot ve Cover (1993) entropiyi,

$$H(X) = -\sum_{i,j} \mu_i P_{ij} \log P_{ij} \tag{4.1}$$

şeklinde tanımlamış ve

$$H = K - \tilde{K} + H_{\Delta} \tag{4.2}$$

olacak şekilde bir kapalı çözüm formu geliştirmiştir. Ash (1965) ise bilgi kaynaklarının entropi değerini hesaplarken mertebe kavramını da kullanarak verilen bilgi kaynağının entropisini hesaplamıştır. Çalışmamızın bu kısmında Ekroot ve Cover (1993) ile Ash (1965)'ın entropi tanımlarına göre uygulamalar yaparak, bu iki entropi tanımının birbirini karşıladığını ve mertebe hesaplama işlemine katılınca entropi hesabının nasıl yapılacağını araştıracağız.

4.1. Sonlu Markov Zincirlerinin Entropisi ve Uygulamaları

Tanım 4.1.1. Enformasyon teoride bir deneyin belirsizlik derecesinin ölçüsüne veya bilgi miktarına entropi denir.

Uyarı 4.1.2. Eğer entropi tanımı bu şekilde göz önüne alınırsa, bir sonraki deneyin bir önceki deneye göre daha belirsiz olduğu açıktır.

Tanım 4.1.3. P geçiş matrisi ve μ durağan dağılım olmak üzere, entropi değeri $H(X) = -\sum_{i,j} \mu_i P_{ij} \log P_{ij}$ şeklinde tanımlıdır.

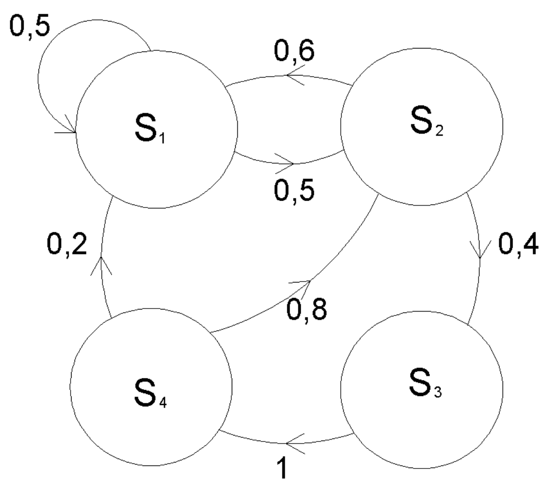
Tanım 4.1.4. i durumunda başlayan ve i durumunda biten rastgele T_{ii} yörüngesinin entropisi

$$H\left(T_{ii}\right)=\frac{H(X)}{\mu_i} \tag{4.3}$$

şeklinde tanımlıdır (Ekroot ve Cover,1993).

Şimdi alacağımız örneği hem Ekroot ve Cover (1993) makalesindeki çözüm yollarını kullanarak hem de Ash (1965) kitabındaki çözüm yollarını kullanarak inceleyeceğiz. Ancak Ash (1965) kitabındaki çözüm yollarını kullanabilmemiz için aldığımız kaynağın unifilar olup olmadığına bakmalıyız. Kaynağın unifilar olduğu Örnek 2.2.5'te verilmiştir.

Örnek 4.1.5.



Şekil 4.1. Durum geçiş diyagramı

Bu diyagrama karşılık gelen stokastik matris; $P=\begin{bmatrix}0.5 & 0.5 & 0 & 0\\ 0.6 & 0 & 0.4 & 0\\ 0 & 0 & 0 & 1\\ 0.2 & 0.8 & 0 & 0\end{bmatrix}$ şeklindedir.

Bu zincir tek bir durağan dağılıma sahiptir ve bu durağan dağılım $\mu P=\mu$ formülü ile hesaplanır. Durağan dağılım, $\mu=(0.42,0.32,0.13,0.13)$'dür. X bir durağan Markov zinciri olmak üzere entropi oranını hesaplayalım.

$$H(X)=-\sum_{i,j=1}^{4}\mu_i P_{ij}\log_2 P_{ij}$$

$$=-\sum_{i=1}^{4}\mu_i\sum_{j=1}^{4}P_{ij}\log_2 p_{ij}$$

$$
\begin{aligned}
&= -\Big[\mu_1 \sum_{j=1}^{4} P_{1j}\log_2 P_{1j} + \mu_2 \sum_{j=1}^{4} P_{2j}\log_2 P_{2j} \\
&\quad + \mu_3 \sum_{j=1}^{4} P_{3j}\log_2 P_{3j} + \mu_4 \sum_{j=1}^{4} P_{4j}\log_2 P_{4j}\Big] \\
&= -\mu_1\left[P_{11}\log_2 P_{11} + P_{12}\log_2 P_{12} + P_{13}\log_2 P_{13} + P_{14}\log_2 P_{14}\right] \\
&\quad -\mu_2\left[P_{21}\log_2 P_{21} + P_{22}\log_2 P_{22} + P_{23}\log_2 P_{23} + P_{24}\log_2 P_{24}\right] \\
&\quad -\mu_3\left[P_{31}\log_2 P_{31} + P_{32}\log_2 P_{32} + P_{33}\log_2 P_{33} + P_{34}\log_2 P_{34}\right] \\
&\quad -\mu_4\left[P_{41}\log_2 P_{41} + P_{42}\log_2 P_{42} + P_{43}\log_2 P_{43} + P_{44}\log_2 P_{44}\right] \\
&= -(0.42)\big[(0.5)\log_2(0.5) + (0.5)\log_2(0.5)\big] \\
&\quad -(0.32)\big[(0.6)\log_2(0.6) + (0.4)\log_2(0.4)\big] \\
&\quad -(0.13)\big[1\log_2 1\big] \\
&\quad -(0.13)\big[(0.2)\log_2(0.2) + (0.8)\log_2(0.8)\big] \\
&= 0.83 \text{ parçadır.}
\end{aligned}
$$

1. adım entropilerinin matrisini yazabilmek için aşağıdaki ifadeleri hesaplayalım.

$$
\begin{aligned}
H(P_1) &= -\sum_{j=1}^{4} P_{1j}\log_2 P_{1j} \\
&= -\left[P_{11}\log_2 P_{11} + P_{12}\log_2 P_{12} + P_{13}\log_2 P_{13} + P_{14}\log_2 P_{14}\right] \\
&= -\big[(0.5)\log_2(0.5) + (0.5)\log_2(0.5)\big] = 1
\end{aligned}
$$

$$
\begin{aligned}
H(P_2) &= -\sum_{j=1}^{4} P_{2j}\log_2 P_{2j} \\
&= -\left[P_{21}\log_2 P_{21} + P_{22}\log_2 P_{22} + P_{23}\log_2 P_{23} + P_{24}\log_2 P_{24}\right] \\
&= -\big[(0.6)\log_2(0.6) + (0.4)\log_2(0.4)\big] = 0.97
\end{aligned}
$$

$$
\begin{aligned}
H(P_3) &= -\sum_{j=1}^{4} P_{3j}\log_2 P_{3j} \\
&= -\left[P_{31}\log_2 P_{31} + P_{32}\log_2 P_{32} + P_{33}\log_2 P_{33} + P_{34}\log_2 P_{34}\right] \\
&= -\big[1\log_2 1\big] = 0
\end{aligned}
$$

$$
H(P_4) = -\sum_{j=1}^{4} P_{4j}\log_2 P_{4j}
$$

$$= -\left[P_{41}\log_2 P_{41} + P_{42}\log_2 P_{42} + P_{43}\log_2 P_{43} + P_{44}\log_2 P_{44}\right]$$

$$= -\left[(0.2)\log_2(0.2) + (0.8)\log_2(0.8)\right] = 0.73$$

olarak elde edilir. Böylece 1.adım entropilerinin matrisi;

$$H^* = \begin{bmatrix} H(P_1) & H(P_1) & H(P_1) & H(P_1) \\ H(P_2) & H(P_2) & H(P_2) & H(P_2) \\ H(P_3) & H(P_3) & H(P_3) & H(P_3) \\ H(P_4) & H(P_4) & H(P_4) & H(P_4) \end{bmatrix} \text{ yani}$$

$$H^* = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0.97 & 0.97 & 0.97 & 0.97 \\ 0 & 0 & 0 & 0 \\ 0.73 & 0.73 & 0.73 & 0.73 \end{bmatrix}$$

olarak bulunur. *H* ile ilişkilendirilmiş diyagonal matrisi de yazabilmek için

$H_{ii} = \dfrac{H(X)}{\mu_i}$ ifadelerini hesaplarsak;

$$H_{11} = \frac{H(X)}{\mu_1} = \frac{0.83}{0.42} = 1.98$$

$$H_{22} = \frac{H(X)}{\mu_2} = \frac{0.83}{0.32} = 2.59$$

$$H_{33} = \frac{H(X)}{\mu_3} = \frac{0.83}{0.13} = 6.39$$

$$H_{44} = \frac{H(X)}{\mu_4} = \frac{0.83}{0.13} = 6.39$$

bulunur ve

$$H_\Delta = \begin{bmatrix} H_{11} & 0 & 0 & 0 \\ 0 & H_{22} & 0 & 0 \\ 0 & 0 & H_{33} & 0 \\ 0 & 0 & 0 & H_{44} \end{bmatrix} \text{ olup diyagonal matris,}$$

$$H_\Delta = \begin{bmatrix} 1.98 & 0 & 0 & 0 \\ 0 & 2.59 & 0 & 0 \\ 0 & 0 & 6.39 & 0 \\ 0 & 0 & 0 & 6.39 \end{bmatrix} \text{ şeklindedir.}$$

Buradan görülür ki;

1 durumundan 1 durumuna rastgele yörüngenin entropisi 1.98,

2 durumundan 2 durumuna rastgele yörüngenin entropisi 2.59,

3 durumundan 3 durumuna rastgele yörüngenin entropisi 6.39,

4 durumundan 4 durumuna rastgele yörüngenin entropisi 6.39'dur.

Şimdi yörünge entropilerinin matrisini hesaplayalım ve genel bir kapalı çözüme ulaşmaya çalışalım:

Sırasıyla; $P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 1 \\ 0.2 & 0.8 & 0 & 0 \end{bmatrix}$, $H^* = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0.97 & 0.97 & 0.97 & 0.97 \\ 0 & 0 & 0 & 0 \\ 0.73 & 0.73 & 0.73 & 0.73 \end{bmatrix}$,

$$H_\Delta = \begin{bmatrix} 1.96 & 0 & 0 & 0 \\ 0 & 2.59 & 0 & 0 \\ 0 & 0 & 6.39 & 0 \\ 0 & 0 & 0 & 6.39 \end{bmatrix}, \quad A = \begin{bmatrix} 0.42 & 0.32 & 0.13 & 0.13 \\ 0.42 & 0.32 & 0.13 & 0.13 \\ 0.42 & 0.32 & 0.13 & 0.13 \\ 0.42 & 0.32 & 0.13 & 0.13 \end{bmatrix}$$

matrisleri verildiğinde $K = (I - P + A)^{-1}\left(H^* - H_\Delta\right)$ ifadesinden,

$$K = \begin{bmatrix} -0.74 & 1.11 & 5.04 & 2.27 \\ 0.91 & -0.96 & 1.52 & -0.93 \\ 0.28 & -0.38 & -5.76 & 0.01 \\ 0.95 & 0.39 & -4.65 & -5.53 \end{bmatrix}$$

ve $\tilde{K}_{ij} = K_{jj}$'den

$$\tilde{K} = \begin{bmatrix} -0.74 & -0.96 & -5.76 & -5.53 \\ -0.74 & -0.96 & -5.76 & -5.53 \\ -0.74 & -0.96 & -5.76 & -5.53 \\ -0.74 & -0.96 & -5.76 & -5.53 \end{bmatrix}$$

olur. Sonuç olarak 2.1.11 teoreminden yörünge entropilerinin matrisi,

$$H = \begin{bmatrix} 1.98 & 2.07 & 10.8 & 7.8 \\ 1.65 & 2.59 & 7.28 & 4.6 \\ 1.02 & 0.58 & 6.39 & 5.54 \\ 1.69 & 1.35 & 1.11 & 6.39 \end{bmatrix}$$

şeklindedir. Örnek olarak; 2 durumundan 4 durumuna yörüngenin entropisi 4.6,
1 durumundan 3 durumuna yörüngenin entropisi 10.8
şeklindedir.

4.2. Bilgi Kaynakları ve Kaynağın Belirsizliği

Enformasyon teori iletişim (komünikasyon) konularını inceler. Bu teorinin başlangıcı Shannon ‘un "A Mathematical Theory of Communication" (1948, The Bell System Technical Journal vol.27, pp 379–423) adlı çalışması olarak kabul edilir. Enformasyon teorinin ilgilendiği konular, kaynağın ürettiği mesajın enformasyon miktarı, kanalın iletebileceği maksimum enformasyon miktarı, iletim sırasında oluşan hatanın düzeltilmesi ve daha verimli bir iletim için kodlama yapılması vb. şeklindedir. Bu teoriye göre kaynaktaki p olasılıklı bir durumun ortaya çıkması $\log_2\left(\frac{1}{p}\right)$’lik bir enformasyon oluşumudur.

Tanım 4.2.1. Kaynağın entropisi, sahip olduğu durumların taşıdığı enformasyon miktarlarının ortalamasıdır. Entropi kavramı ile belirsizlik kavramı bu teoride aynı anlama gelmektedir yani entropi yerine belirsizlik ifadesi kullanılabilir.

Tanım 4.2.2. Bir bilgi kaynağı Markov zincirinden türetilebilir. Bir sonlu Markov zinciri $Z_0, Z_1, \ldots$ ve bir f fonksiyonunun tanım kümesi zincirin durumlarının kümesi, değer kümesi de bir sonlu Π kümesi olsun. Kabul edelim ki $w = [w_1, \ldots, w_r]$durağan dağılımı ile uyum içinde seçilen ilk durum Z_0’dır yani tüm S_j durumları için $P\{Z_0 = S_j\} = w_j$’ dir. O zaman $X_n = f(Z_n)$, $n = 0,1,\ldots$ durağan dizisi bir Markov bilgi kaynağıdır.

Tanım 4.2.3. Eğer verilen zincir sabit durum olasılıklarına sahipse, bilgi kaynağının, bir regüler Markov bilgi kaynağı olduğunu söyleyebiliriz. Eğer zincir indirgenemezse, bilgi kaynağının indirgenemez bir Markov kaynağı olduğunu söyleyebiliriz.

Tanım 4.2.4. Π=S ve f birim fonksiyon olsun. $k = 0,1,\ldots,$ $i, j = 1,2,\ldots,r$ için $P_{ij} = w_j = P\{X_k = S_j\}$ alırsak, düzenli (regüler) bir Markov bilgi kaynağıdır ve bir sonlu $S=\{s_1,\dots,s_r\}$ kümesi içinde değerler alan aynı dağılımlı $X_0,X_1,\dots$ serbest değişkenleri, bir dizinin bağımsızlığını belirtir.

Teorem 4.2.5. Eğer $X_0,X_1,\dots$ bir bilgi kaynağı ise,

$$H\{X\}=\lim_{n\to\infty}\frac{H(X_0,X_1,\dots,X_n)}{n+1} \tag{4.4}$$

dir (Ash, 1965).

İspat. $k_n=H(X_n\mid X_0,X_1,\dots,X_{n-1}),n=1,2,\dots;\ k_0=H(X_0)$ ise o zaman 1.4.24 teoremine göre;

$$\frac{H(X_0,X_1,\dots,X_n)}{n+1}=\frac{H(X_0)+H(X_1\mid X_0)+\dots+H(X_n\mid X_0,X_1,\dots,X_{n-1})}{n+1}$$

$$=\frac{k_0+k_1+\dots+k_n}{n+1}\text{ 'dir.}$$

Fakat $H(X)$'in tanımına göre $k_n\to H(X)$ ve buradan

$$\left\{\frac{H(X_0,X_1,\dots,X_n)}{n+1},n=0,1,2,\dots\right\}$$

bir yakınsak dizi olan $H(X)$' in aritmetik ortalamasıdır.

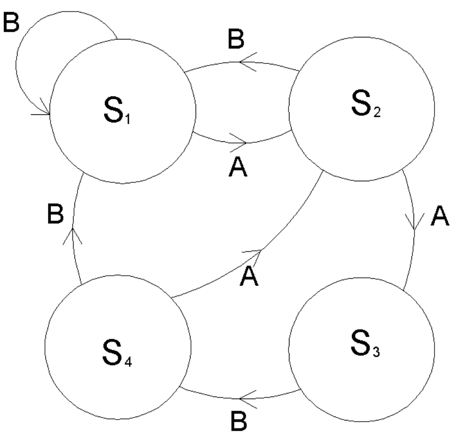
$\frac{h_{n+1}}{n+1}=\frac{H(X_0,X_1,\dots,X_n)}{n+1}$ artmayan dizidir. $h_{n+1}=\sum_{i=0}^{n}k_i$ olduğundan

$$\frac{h_{n+1}}{n+1}-\frac{h_n}{n}=\frac{\sum_{i=0}^{n}k_i}{n+1}-\frac{\sum_{i=0}^{n-1}k_i}{n}=\frac{nk_n-\sum_{i=0}^{n-1}k_i}{n(n+1)}$$

dir. $\{k_n\}$ artmayan dizi olduğundan $i=0,1,\dots,n-1$ için $k_i\geq k_n$' dir. Buradan

$$nk_n-\sum_{i=0}^{n-1}k_i\leq nk_n-nk_n=0$$

olup ispat tamamlandı. Şimdi amacımız bir kaynağın mertebesini de kullanarak bir bilgi kaynağının entropisini hesaplamaktır. Bu aşamada sonlu mertebeli bir unifilar Markov kaynağından, orijinal kaynağa ulaşmaya çalışacağız.



Şekil 4.2. 3. mertebeden bir unifilar Markov bilgi kaynağı

İlk olarak şekil 4.2.1'deki bilgi kaynağını analiz edelim. Diyagramda eğer $f\left(S_j\right)=\gamma$ ise, o zaman S_j durumuna giren bütün oklar γ ile belirlenir. Diyagramda geçiş olasılıkları belirtilmemiş olup, eğer $P_{ij}>0$ ise S_i ve S_j durumları birbirlerine doğrudan bağlıdırlar. Diyagrama göre verilmiş ilk durum için uygun son durumlarla bütün mümkün dizileri sıralayıp bir tablo oluşturacağız. Örnek olarak eğer $Z_{t-3}=S_2, X_{t-2}=A, X_{t-1}=B, X_t=B$ ise $Z_t=S_1$'dir. İmkânsız geçişlerde benzer girişler boştur, örneğin $Z_{t-2}=S_3$ ise $X_{t-1}=A,\quad X_t=A$ imkânsızdır. Eğer $X_{t-2}=B, X_{t-1}=A, X_t=A$ ise o zaman $Z_t=S_3$'tür çünkü bütün girişler BAA kolonu içindedir. Burada önceki iki girdiye göre mevcut durum belli değildir, çünkü eğer $X_{t-1}=A, X_t=B$ ise o zaman Z_t, S_1 veya S_4'ten birinde olabilir. Bu da kaynağın mertebesini ifade eder. O yüzden şekil 4.2.1' de verilen kaynak 3. mertebedendir.

Şimdi Şekil 4.2.1 için kaynağın geçiş durumlarının tablosunu oluşturalım.

	X_{t-1} X_t			
Z_{t-2}	AA	AB	BB	BA
S_1	S_3	S_1	S_1	S_2
S_2	-	S_4	S_1	S_2
S_3	-	-	S_1	S_2
S_4	S_3	S_1	S_1	S_2

	$X_{t-2}\ X_{t-1}\ X_t$							
Z_{t-3}	AAA	AAB	ABA	ABB	BAA	BAB	BBA	BBB
S_1	-	S_4	S_2	S_1	S_3	S_1	S_2	S_1
S_2	-	-	S_2	S_1	S_3	S_1	S_2	S_1
S_3	-	-	-	-	S_3	S_1	S_2	S_1
S_4	-	S_4	S_2	S_1	S_3	S_1	S_2	S_1

Çizelge 4.1. Şekil 4.2. nin kaynağı için geçiş durumları

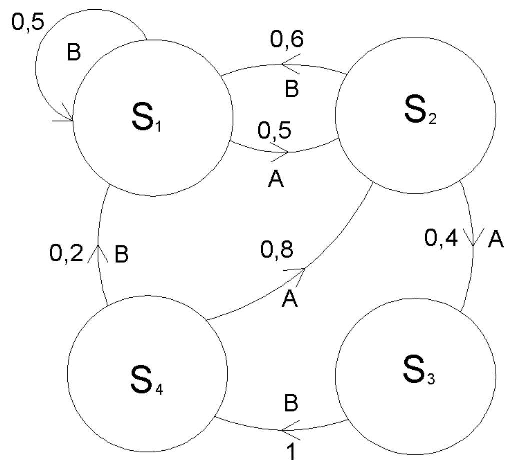
Tanım 4.2.6. Z_t durumu eğer M mertebeli ise $i < M-1$ olduğunda $X_t, X_{t-1}, \ldots, X_{t-M+1}$ sembollerine bağlı olan ancak $X_t, X_{t-1}, \ldots, X_{t-i}$ sembollerine bağlı olmayan bir unifilar Markov kaynağı tanımlanır. Bu nedenle şekil 4.2.1'deki kaynağın mertebesi 3'tür.

Uyarı 4.2.7. Eğer bir kaynağın mertebesi M ise, onun belirsizliği ard arda gelen M sembollerinin ve ard arda gelen $M+1$ sembollerinin belirsizliğinin farkıdır (Ash, 1965).

Tanım 4.2.8. Genel olarak, bir kaynağa göre üretilen ard arda gelen n sembollerinin $H(X_1, \ldots, X_n)$ belirsizliği kaynağın "n- gram belirsizliği" diye anılabilir (Ash, 1965).

Şimdi 4.1.5. örneğindeki diyagramı kullanarak bir kaynağın mertebesini de hesaba katarak sonlu mertebeli kaynaktan genel bilgi kaynağına yaklaşımı inceleyeceğiz. Burada keyfi bir bilgi kaynağı alıp doğal bir yöntemle orijinal kaynağa ulaşmaya çalışacağız ve son olarakta 4.1.5 örneğindeki ve aşağıdaki örnekte bulduğumuz sonuçları karşılaştırıp Ekroot ve Cover ile Ash'ın entropi tanımlarından önemli sonuçlar elde edeceğiz.

Örnek 4.2.9. Önce aşağıdaki grafiği alalım.



Şekil 4.3. 3.mertebeden unifilar Markov kaynağı

$$f(S_2) = f(S_3) = A$$

$$f(S_1) = f(S_4) = B$$

X_{t-1} X_t

Z_{t-2}	AA	AB	BB	BA
S_1	S_3	S_1	S_1	S_2
S_2	-	S_4	S_1	S_2
S_3	-	-	S_1	S_2
S_4	S_3	S_1	S_1	S_2

X_{t-2} X_{t-1} X_t

Z_{t-3}	AAA	AAB	ABA	ABB	BAA	BAB	BBA	BBB
S_1	-	S_4	S_2	S_1	S_3	S_1	S_2	S_1
S_2	-	-	S_2	S_1	S_3	S_1	S_2	S_1
S_3	-	-	-	-	S_3	S_1	S_2	S_1
S_4	-	S_4	S_2	S_1	S_3	S_1	S_2	S_1

	X_{t-3} X_{t-2} X_{t-1} X_t							
Z_{t-4}	AAAA	ABAA	AABA	AAAB	ABBA	AABB	ABAB	ABBB
S_1	-	S_3	S_2	-	S_2	S_1	S_1	S_1
S_2	-	S_3	-	-	S_2	-	S_1	S_1
S_3	-	-	-	-	-	-	-	-
S_4	-	S_3	S_2	-	S_2	S_1	S_1	S_1

	X_{t-3} X_{t-2} X_{t-1} X_t							
Z_{t-4}	BBBB	BABB	BBAB	BBBA	BAAB	BBAA	BABA	BAAA
S_1	S_1	S_1	S_1	S_2	S_4	S_3	S_2	-
S_2	S_1	S_1	S_1	S_2	S_4	S_3	S_2	-
S_3	S_1	S_1	S_1	S_2	S_4	S_3	S_2	-
S_4	S_1	S_1	S_1	S_2	S_4	S_3	S_2	-

Çizelge 4.2. Şekil 4.3 ün kaynağı için geçiş durumları

Kararlı durum olasılıkları; $w_1 = 0.42, w_2 = 0.32, w_3 = 0.13, w_4 = 0.13$ olur.

$$P(A) = P\{X_t = A\} = \sum_{i=1}^{4} P\{Z_{t-1} = S_i\} P\{X_t = A \mid Z_{t-1} = S_i\}$$

$$= (0.42)(0.5) + (0.32)(0.4) + (0.13)(0.8)$$

$$= 0.21 + 0.128 + 0.104$$

$$= 0.442$$

$$P(B) = P\{X_t = B\} = \sum_{i=1}^{4} P\{Z_{t-1} = S_i\} P\{X_t = B \mid Z_{t-1} = S_i\}$$

$$= (0.42)(0.5) + (0.32)(0.6) + (0.13)1 + (0.13)(0.2)$$

$$= 0.21 + 0.192 + 0.13 + 0.026$$

$$= 0.558$$

$$P(AA) = P\{X_t = A, X_{t+1} = A\}$$

$$= \sum_{i=1}^{4} P\{X_t = A, X_{t+1} = A \mid Z_{t-1} = S_i\}$$

$$= (0.42)(0.5)(0.4) + (0.13)(0.8)(0.4)$$

$$= 0.084 + 0.0416$$
$$= 0.1256$$
$$P(AB) = P\{X_t = A, X_{t+1} = B\}$$
$$= \sum_{i=1}^{4} P\{X_t = A, X_{t+1} = B \mid Z_{t-1} = S_i\}$$
$$= (0.42)(0.5)(0.6) + (0.32)(0.4)1 + (0.13)(0.8)(0.6)$$
$$= 0.216 + 0.128 + 0.0624$$
$$= 0.3164$$
$$P(BB) = P\{X_t = B, X_{t+1} = B\}$$
$$= \sum_{i=1}^{4} P\{X_t = B, X_{t+1} = B \mid Z_{t-1} = S_i\}$$
$$= (0.42)(0.5)(0.5) + (0.32)(0.6)(0.5) + (0.13)1(0.2) + (0.13)(0.2)(0.5)$$

$$= 0.105 + 0.096 + 0.026 + 0.013$$
$$= 0.24$$
$$P(BA) = P\{X_t = B, X_{t+1} = A\}$$
$$= \sum_{i=1}^{4} P\{X_t = B, X_{t+1} = A \mid Z_{t-1} = S_i\}$$
$$= (0.42)(0.5)(0.5) + (0.32)(0.6)(0.5) + (0.13)1(0.8) + (0.13)(0.2)(0.5)$$
$$= 0.105 + 0.096 + 0.104 + 0.013$$
$$= 0.318$$
$$P(AAA) = P\{X_t = A, X_{t+1} = A, X_{t+2} = A\}$$
$$= \sum_{i=1}^{4} P\{X_t = A, X_{t+1} = A, X_{t+2} = A \mid Z_{t-1} = S_i\}$$
$$= 0$$

ve benzeri işlemler yapılırsa;

$$P(AAB) = 0.1256$$

$$P(ABA) = 0.1966$$

$$P(ABB) = 0.1198$$

$$P(BAA) = 0.1272$$

$P(BAB) = 0.1908$

$P(BBA) = 0.12$

$P(BBB) = 0.12$

bulunur.

$$
\begin{aligned}
P(AAAA) &= P\{X_t = A, X_{t+1} = A, X_{t+2} = A, X_{t+3} = A\} \\
&= \sum_{i=1}^{4} P\{X_t = A, X_{t+1} = A, X_{t+2} = A, X_{t+3} = A \mid Z_{t-1} = S_i\} \\
&= 0
\end{aligned}
$$

ve benzeri işlemler yapılırsa,

$P(ABAA) = 0.07864$

$P(AABA) = 0.1$

$P(AAAB) = 0$

$P(ABBA) = 0.0599$

$P(AABB) = 0.02512$

$P(ABAB) = 0.11796$

$P(ABBB) = 0.0599$

$P(BBBB) = 0.06$

$P(BABB) = 0.0954$

$P(BBAB) = 0.072$

$P(BBBA) = 0.06$

$P(BAAB) = 0.1272$

$P(BBAA) = 0.048$

$P(BABA) = 0.0954$

$P(BAAA) = 0$

değerlerini elde ederiz. Elde ettiğimiz tüm değerleri kullanarak J_i değerlerini hesaplayalım.

$$J_1 = -P(A)\log_2 P(A) - P(B)\log_2 P(B)$$
$$= -(0.442)\log_2(0.442) - (0.558)\log_2(0.558)$$
$$= 0.522 + 0.471$$
$$= 0.993$$

$$J_2 = -P(AA)\log_2 P(AA) - P(AB)\log_2 P(AB) \ - P(BB)\log_2 P(BB)$$
$$-P(BA)\log_2 P(BA)$$
$$= -(0.1256)\log_2(0.1256) - (0.3164)\log_2(0.3164) - (0.24)\log_2(0.24)$$
$$-(0.318)\log_2(0.318)$$
$$= 0.377+0.527+0.496+0.527$$
$$= 1.9272$$

$$J_3 = -P(AAA)\log_2 P(AAA) - P(AAB)\log_2 P(AAB) - P(ABA)\log_2 P(ABA)$$
$$-P(ABB)\log_2 P(ABB) - P(BAA)\log_2 P(BAA) - P(BAB)\log_2 P(BAB)$$
$$-P(BBA)\log_2 P(BBA) - P(BBB)\log_2 P(BBB)$$
$$= -(0.1256)\log_2(0.1256) - (0.1966)\log_2(0.1966) - (0.1198)\log_2(0.1198)$$
$$-(0.1272)\log_2(0.1272) - (0.1908)\log_2(0.1908) - (0.12)\log_2(0.12)$$
$$-(0.12)\log_2(0.12)$$
$$= 0.377 + 0.463 + 0.368 + 0.38 + 0.458 + 0.368 + 0.368$$
$$= 2.78$$

$$J_4 = -P(AAAA)\log_2 P(AAAA) - P(ABAA)\log_2 P(ABAA)$$
$$-P(AABA)\log_2 P(AABA) - P(AAAB)\log_2 P(AAAB)$$
$$-P(ABBA)\log_2 P(ABBA) - P(AABB)\log_2 P(AABB)$$
$$-P(ABAB)\log_2 P(ABAB) - P(ABBB)\log_2 P(ABBB)$$
$$-P(BBBB)\log_2 P(BBBB) - P(BABB)\log_2 P(BABB)$$
$$-P(BBAB)\log_2 P(BBAB) - P(BBBA)\log_2 P(BBBA)$$
$$-P(BAAB)\log_2 P(BAAB) - P(BBAA)\log_2 P(BBAA)$$
$$-P(BABA)\log_2 P(BABA) - P(BAAA)\log_2 P(BAAA)$$
$$= -(0.07864)\log_2(0.07864) - (0.1)\log_2(0.1) - (0.0599)\log_2(0.0599)$$
$$-(0.02512)\log_2(0.02512) - (0.11796)\log_2(0.11796)$$
$$-(0.0599)\log_2(0.0599) \ -(0.06)\log_2(0.06)$$

$$-(0.0954)\log_2(0.0954)-(0.072)\log_2(0.072)$$
$$-(0.06)\log_2(0.06)-(0.1272)\log_2(0.1272)$$
$$-(0.048)\log_2(0.048)\ -(0.0954)\log_2(0.0954)$$
$$=0.289+0.333+0.244+0.134+0.365+0.244+0.244+0.325+0.274$$
$$+0.244\ +0.379+0.211+0.325$$
$$=3.611$$

olarak elde edilir. Böylece 1. entropi tahmini;

$$H_1(X)=J_2-J_1=1.93-0.99\ =0.94$$

2. entropi tahmini;

$$H_2(X)=J_3-J_2\ =2.78-1.93\ =0.85$$

ve orijinal entropi değeri;

$$H(X)=J_4-J_3\ =3.611-2.78\ =0.83$$

bulunur.

Sonuç 4.2.10. Eğer bir kaynağın mertebesi M ise entropi değeri

$$H(X)=J_{M+1}-J_M \tag{4.5}$$

dir (Ash, 1965).

Sonuç 4.2.11. Eğer kaynak 1. mertebeden ise $H(X)=J_2-J_1$ orijinal entropi değerini verir.

Sonuç 4.2.12. Genel olarak; eğer kaynak n. mertebeden ise ard arda gelen n sembollerinin ve ard arda gelen $n+1$ sembollerinin belirsizliğinin farkı orijinal entropi değerini verir.

Sonuç 4.2.13. 1. entropi tahmini yani $H_1(X)$, orijinal entropi değerinden yani $H(X)$'den büyük eşittir. Yani $H_1(X)\geq H(X)$'dir.

Sonuç 4.2.14. Genel olarak $H_n(X)\geq H_{n-1}(X)\geq\ldots\geq H_1(X)\geq H(X)$'dir.

Sonuç 4.2.15. Ekroot ve Cover (1993) tarafından tanımlanan entropi kavramı ile Ash (1965) tarafından tanımlanan entropi kavramları birbirini karşılar. Çünkü 4.1.5 örneğinde ve 4.2.9 örneğinde aldığımız kaynağa göre her iki örnekte de entropi değerleri eşit çıkmıştır. Yani kaynağın mertebesini hesaba katmadan yaptığımız hesaplama ile kaynağın mertebesini hesaba katarak yaptığımız hesaplama aynı çıkmıştır.

4.3. Rastgele Durdurulmuş Dizinin Entropisi

Tanım 4.3.1.

$X = x$	$x_1, x_2, x_3 \ldots x_N$
$f(x) = \mathrm{P}(X = x)$	$f(x_1), f(x_2), \ldots, f(x_N)$

X yukarıdaki olasılık fonksiyonuna sahip kesikli bir değişken olsun. X'in $E(X)$ ile gösterilen beklenen değeri aşağıdaki gibi tanımlanır.

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_N f(x_N) = \int_{i=1}^{N} x_i f(x_i)\, dx \qquad (4.6)$$

X rastgele değişkeni sayılabilir sonsuzluktaki $x_1, x_2, \ldots, x_N, \ldots$ sonuçlarını alıyorsa;

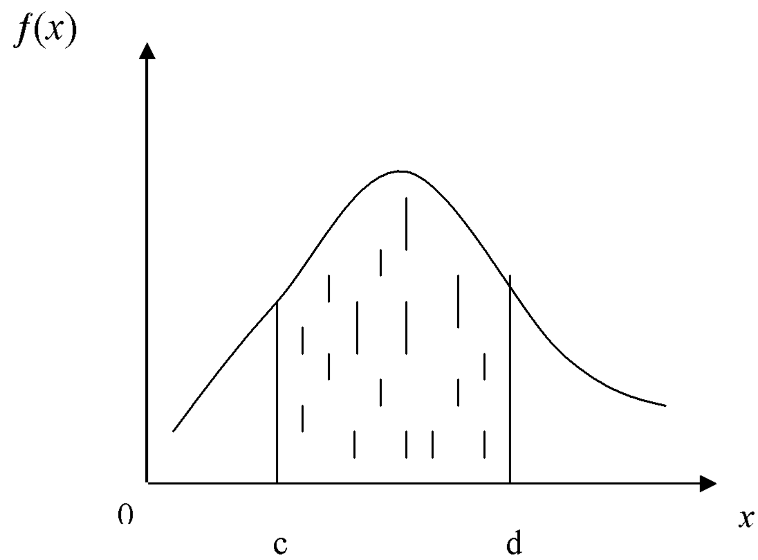
$$E(X) = x_1 f(x_1) + \ldots + x_N f(x_N) + \ldots = \sum_{i=1}^{\infty} f(x_i) \qquad (4.7)$$

dir.

Tanım 4.3.2. t herhangi bir reel sayı olmak üzere X rassal değişkeninin $\varphi(x) = e^{itx} = \cos(tx) + i\sin(tx)$ kompleks fonksiyonu verilsin. $E(\varphi(x)) = E(e^{itx})$ ortalama değeri t'nin bir fonksiyonudur. Aşağıdaki fonksiyona X rassal değişkeninin karakteristik fonksiyonu denir.

$$\varphi_x(t) = E(e^{itx}) = E[Cos(tx)] + iE[Sin(tx)] \qquad (4.8)$$

Tanım 4.3.3.



X, üstteki şekilde gösterilen $(-\infty,\infty)$ aralığında tanımlanan sürekli rastgele değişken olsun. Aşağıdaki koşulları sağlayan *f(x)* fonksiyonuna X rastgele değişkeninin olasılık yoğunluk fonksiyonu denir.

1-) $f(x)\geq 0$, $-\infty < x < \infty$

2-) $f(x)$ eğrisi altında kalan ve x-ekseni ile sınırlanan alan 1'e eşittir.

Tanım 4.3.4. Bir beklenmeyen durdurma zamanı N ile, bağımsız özdeşçe dağılmış ayrık serbest değişkenler $X_1, X_2,\ldots$ ile verildiğinde, bir rastgele durdurulmuş dizinin entropisi

$$H\left(X^N\right)=(EN)H\left(X_1\right)+H\left(N/X^{\infty}\right)H\left(X^N\right) \tag{4.9}$$

olup, X^N durdurulmuş diziyi belirtir. X^N durdurulmuş dizisi içinde rastgelelik, X^{∞} duraksız dizisi üzerinde şartlı durdurma zamanında artan rastgeleliğin artan aramaları başına X için beklenen aramaların sayısıdır (Ekroot ve Cover, 1991).

Tanım 4.3.5. Rastgele durdurulmuş dizinin entropisi, bağımsız özdeşçe dağılmış serbest değişkenlere bağlı olmayıp, dizinin uzunluğuna ve karakterine bağlıdır. $X_1, X_2,\ldots$ bağımsız özdeşçe dağılmış serbest değişkenler, *p(x)* olasılık yoğunluk fonksiyonu ve $x\in X$ olsun. $X_1, X_2,\ldots$ gözlenen süreçlerinin beklenmeyen fonksiyonları için N durdurma zamanını göz önüne alırsak, $\{N=n\}$ durdurma zamanı sonraki $\left\{X_{n+1}, X_{n+2},\ldots\right\}$ dizisinden bağımsızdır. X^N'e göre $X_1, X_2,\ldots,X_N \in X^*$ ile sonlu uzunluklu diziyi, X ile de bütün sonlu uzunluklu dizilerin kümesi tanımlansın. X^0 olmayan elemanlar dizisini gösterir. $x\in X^*$ için $q(x)=P_r\left\{X^N=x\right\}$ olduğunda $q(x)$ olasılık yoğunluk fonksiyonu bir N durdurma zamanına neden olur.

Bu kısımda X^N durdurulmuş dizisinin

$$H\left(X^N\right)=-\sum_{x\in X^*} q(x)\log q(x) \tag{4.10}$$

entropisinin nasıl hesaplandığını açıklamaya çalışacağız. $H\left(X^N\right)$ rastgele durdurulmuş dizisinin tanımlayıcı karmaşıklığıdır. Örnek olarak, kumar oynayan bir kişi, oyuna bir dolar ile başlar. $n=1,2,\ldots$ zamanda oyun oynar. Bu kişinin parası p olasılıkla artar ve q=1-p olasılıkla azalır. Oyun kumarbazın parası sıfırlanınca biter. Bu olay yani, bu kumarbazın parasının artması ya da azalması bağımsız özdeşçe dağılmış bir dizi oluşturur. $p<\frac{1}{2}$ için kumarbaz kesin olarak iflas eder. Sonucun kesin olmasına rağmen, yol rastgeledir. Yol entropisi

$$\frac{H(p)}{q-p} \tag{4.11}$$

'dir. Yani, $-[p\log p+q\log q]\frac{1}{q-p}$ ifadesine yol entropisi denir (Ekroot ve Cover,1991).

Tanım 4.3.6. (Durdurma Zamanı) $\{B_n\}$, σ--cebirlerin artan bir dizisi ve $\{B_n\}$'e uyarlanmış bağımsız özdeşçe dağılmış rastgele değişkenler $X_1, X_2, \ldots$ olsun. Eğer her $n=0,1,\ldots$ için $\{N=n\}$ olayı $\{B_n\}$'e aitse N' ye durdurma zamanı denir (Ekroot ve Cover, 1991).

Tanım 4.3.7. Eğer N sınırlı bir durdurma zamanı ise, bazı m negatif olmayan tam sayıları vardır öyle ki $P_r\{N\le m\}=1$'dir (Ekroot ve Cover, 1991).

Teorem 4.3.8. Keyfi r ve s sayıları için

$$H^{(r+s)}=H^{(r)}+H^{(s)} \tag{4.12}$$

veya denk olarak $H^{(r)}=r.H^{(1)}$'dir. Bu önermenin ispatı Khinchin (1957)'de verilmiştir. Ancak bütünlüğün sağlanması için burada ispat daha açık bir ifade ile verilecektir.

İspat. $r=1$ için $H^{(r)}=r.H^{(1)}$ olduğu aşikârdır. Kabul edelim ki bu eşitlik $r=t$ için doğru olsun. Tümevarımdan

$$H^{(t)}=t.H^{(1)} \tag{4.13}$$

sağlansın. Tümevarımdan göstermeliyiz ki (4.13) eşitliği $r=t+1$ içinde doğrudur. Kabul edilsin ki sistem A_i durumunda olsun. $t+1$ denemeden sonra sistemin sonucunu belirleyen sonlu şema iki bağımlı şemanın çarpımı olarak göz önüne alınabilir.

A) $H_i^{(1)}$ entropili bir sonraki denemeye karşılık gelen şema,

B) t deneme sonra sistemin sonucunu belirleyen şema.

Bu şemanın entropisi $H_k^{(t)}$'dir. Eğer A şemasının çıkışı A_k olayında ise genel ilişkiye göre, $H(A\vee B)=H(A)+H(B/A)$'dan

$$H_i^{(t+1)} = H_i^{(1)} + \sum_{k=1}^{n} P_{ik} P_k^{(t)}, \sum_{k=1}^{n} P_k P_{k\ell} = P_\ell \quad (l = 1,2,\ldots,n)$$

olduğundan

$$H^{(t+1)} = \sum_{i=1}^{n} P_i H_i^{(t+1)} = \sum_{i=1}^{n} P_i \left(H_i^{(1)} + \sum_{k=1}^{n} P_{ik} p_k^{(t)} \right)$$

$$= \sum_{i=1}^{n} P_i H_i^{(1)} + \sum_{k=1}^{n} \sum_{i=1}^{n} P_i P_{ik} H_k^{(t)}$$

$$= \sum_{i=1}^{n} P_i H_i^{(1)} + \sum_{k=1}^{n} H_k^{(t)} \sum_{i=1}^{n} P_i P_{ik}$$

$$= \sum_{i=1}^{n} P_i H_i^{(1)} + \sum_{k=1}^{n} P_k H_k^{(t)}$$

$$= H^{(1)} + t.H^{(1)} = (t+1) H^{(1)}$$

olup böylece tümevarımdan ispat tamamlanır.

Bir rastgele durdurulmuş dizinin entropisi aşağıdaki teoremde verilmiştir. Teoremde $X_1, X_2, \ldots$ ifadeleri bağımsız özdeşçe dağılmış rastgele değişkenler, $X^N \in X^*$ rastgele durdurulmuş dizi ve X^∞'da duraksız dizi olarak kabul edilecektir.

Teorem 4.3.9. (Durdurma Zamanı Teoremi) Bağımsız özdeşçe dağılmış $X_1, X_2, \ldots$ için herhangi bir durdurma zamanı N olsun. O zaman $EN = \infty$ ve $H(X_1) = 0$ durumları olmadığında

$$H(X^N) = (EN) H(X_1) + H(N / X^\infty) \tag{4.14}$$

ispatlanabilir (Ekroot ve Cover, 1991).

İspat. İspat iki kısma ayrılacaktır. İlk olarak bir sınırlı durdurma zamanı için ispat verilecek. Sonra sınırlı olma şartı kaldırılacaktır. N durdurma zamanı, n ile sınırlı yani $P_r\{N \le n\} = 1$ olsun. X_1^n ile $(X_1, X_2, \ldots, X_n)$, X_1^0 ile olmayan elemanların dizisi X^∞ ile de $X_1, X_2, \ldots$ duraksız dizisi tanımlansın. $X_1, X_2, \ldots, X_n$ bağımsız özdeşçe dağılmış rastgele değişkenlerinin, $p(x)$ olasılık yoğunluk fonksiyonuna uygun olduğunu kabul edelim.

Şimdi $H\left(X_1^n, N\right)$ açılımını ilk olarak zincir kuralına göre, ikinci olarak ta bağımsız özdeşçe dağılmış $X_1, X_2, \ldots, X_n$ ' e göre yapalım.

$$\begin{aligned} H(X_1^n, N) &= H(X_1^n) + H(N / X_1^n) \\ &= nH(X_1) + H(N / X_1^n) \end{aligned} \tag{4.15}$$

bulunur. $N \leq n$ için;

$$\begin{aligned} H(X_1^n, N) &= H(X_1^N, X_{N+1}^n, N) \\ &= H(X_1^N) + H(N / X_1^N) + H(X_{N+1}^n / X_1^N, N) \\ &= H(X_1^N) + H(X_{N+1}^n / X_1^N, N) \\ &= H(X_1^N) + H(X_{N+1}^n / N) \end{aligned} \tag{4.16}$$

$$= H(X_1^N) + \left[P(N=0) H(X_{N+1}^n / N = 0) + \ldots + P(N = n-1) H(X_{N+1}^n / N = n-1) \right]$$

$$= H\left(X_1^N\right) + \left[P(N=0) \sum_{j=1}^{L} P\left(X_{N+1}^n / N = 0\right) \log P\left(X_{N+1}^n / N = 0\right) - \ldots \right.$$

$$\left. - P(N = n-1) \sum_{j=1}^{L} P\left(X_{N+1}^n / n-1\right) \log P\left(X_{N+1}^n / n-1\right) \right]$$

$$= H\left(X_1^N\right) + \left[-\sum_{k=0}^{n-1} P\{N = k\} \sum_{j=1}^{L} P\left(X_{N+1}^n / N = k\right) \log P\left(X_{N+1}^n / N = k\right) \right]$$

$$= H\left(X_1^N\right) + \sum_{k=0}^{n-1} P_r \{N = k\} H\left(X_{N+1}^n / N = k\right)$$

$$= H\left(X_1^N\right) + \left[P_0 H\left(X_1^n / 0\right) + P_1 H\left(X_2^n / 1\right) + \ldots + P_{n-1} H\left(X_n^n / n-1\right) \right]$$

$$= H\left(X_1^N\right) + \sum_{k=0}^{n-1} P_r \{N = k\} H\left(X_{k+1}^n\right)$$

$$= H\left(X_1^N\right) + \sum_{k=0}^{n-1} P_r \{N = k\} (n-k) H\left(X_1\right)$$

$$= H\left(X_1^N\right) + \left[nH\left(X_1\right) - \sum_{k=0}^{n-1} P_k . k . H\left(X_1\right) \right]$$

$$= H\left(X_1^N\right) + nH\left(X_1\right) - H\left(X_1\right) \sum_{k=0}^{n-1} k P_k$$

$= H\left(X_1^N\right) + nH\left(X_1\right) - (EN) H\left(X_1\right)$ 'dir.

(4.16) eşitliği entropi için zincir kuralıdır. (4.15) ve (4.16) eşitliklerinden;

$$nH(X_1)+H(N/X_1^n)=H(X_1^n)+nH(X_1)-(EN)H(X_1)$$

veya

$$H(X_1^N)=H(N/X_1^n)+(EN)H(X_1) \qquad (4.17)$$

elde edilir. n zaman sonra dizinin bağımsızlığı N 'dir ve

$$H(N/X_1^n)=H(N/X^\infty) \qquad (4.18)$$

dur.

(4.17) ve (4.18) eşitliklerinden;

$$H(X_1^n)=(EN)H(X_1)+H(N/X^\infty) \qquad (4.19)$$

olup sınırlı durdurma zamanı için ispat tamamlanır. Şimdi, N üzerinde sınırlılık şartını kaldıralım yani N sınırsız olsun.

$$N_n=\min(n,N)=\begin{cases} N, & eğer\ N<n \\ n, & eğer\ N\geq n \end{cases}$$

göz önüne alalım. Bir sınırlı durdurma zamanı N_n'dir. (4.19)'a uygularsak,

$$H(X_1^{Nn})=H(X_1)E(N_n)+H(N_n/X^\infty) \qquad (4.20)$$

olur.

$$H(X_1^{Nn})\to H(X_1^N),\ EN_n\to EN$$

ve böylece

$$H(N_n\,|\,X^\infty)\to H(N\,|\,X^\infty)$$

olup sınırsız durdurma zamanları içinde teorem ispatlandı.

Teorem 4.3.10. Herhangi bir durumda zamanı (muhtemelen rastgele hale getirilmiş) N olsun. O zaman $X^N\in X^*$ iken rastgele durdurulmuş dizi

$$H(X^n\,|\,N)=(EN)H(X_1)-\mathrm{I}(N\,|\,X^\infty) \qquad (4.21)$$

şeklinde tanımlanır (Ekroot ve Cover, 1991).

İspat. $H(X^N\,|\,N)=H(X^N,N)-H(N)$

$$=H(X^N)+H(N\,|\,X^N)-H(N)$$

$$= H(X^N) - H(N)$$

$$= (EN)H(X_1) + H(N \mid X^\infty) - H(N)$$

$$= (EN)H(X_1) - \mathrm{I}(N; X^\infty)$$

olur.

5. SONUÇLAR ve ÖNERİLER

5.1. Sonuçlar

Ergodik teori ve enformasyon teori için entropi önemli bir kavramdır. Ergodik teoride entropi değeri bulunurken yörünge kavramından, enformasyon teoride ise bilgi kaynağı kavramından faydalanılmaktadır. Bu durumdan yola çıkılarak verilen sonlu Markov yörüngelerinin ve bilgi kaynağının entropi değerleri hesaplanabilmektedir. Ancak sonlu Markov yörüngesinde tepe sayısı arttıkça ve bilgi kaynağında mertebe arttıkça entropi değerini hesaplamak zorlaşmaktadır.Ekroot ve Cover (1993) ile Ash (1965) entropi kavramı için önemli olan ve durumu net bir şekilde açıklayan örnekler vermişlerdir. Bu örneklerden yola çıkarak 1.deneyin belirsizliğinin yani entropisinin 2.deneye göre daha büyük olduğu anlaşılmaktadır.

Entropi bir sistemin düzensizliğini ve karmaşıklığını ölçmeye çalışır. Entropi bir çok bilim dalında kullanılan bir kavramdır. Özellikle enformasyon teori, Fizik, Kimya vb. bilim dallarında farklı yönleriyle incelenmiştir. Bu tezde ölçüm entropi incelenmiş ve aldığımız bir sonlu Markov yörüngesinin ve bir bilgi kaynağının ölçüm entropisini hesaplayıp, mertebeyide hesaplama işlemine katarak entropi tahminleri bulunmuş ve bunlar arasında bir bağlantı kurulmaya çalışılmıştır.

5.2. Öneriler

Çalışmamızda sonlu bir Markov yörüngesinin, bir bilgi kaynağının ve durdurulmuş bir dizinin entropisi incelenmiştir. Bir sonlu Markov yörüngesinin entropi değerinin hesaplanabilmesi için hangi özellikleri taşıması gerekir, bunlar incelenerek bir yörüngenin entropisi hesaplanmaya çalışılmıştır.

Ayrıca bir bilgi kaynağının entropi değerini hesaplarken mertebe kavramı kullanılmış ve bir yörüngenin entropisi ile bir bilgi kaynağının entropisi karşılaştırılmıştır. Tezimizde Ekroot ve Cover (1993) ile Ash (1965)'ın entropi tanımlarının birbirini karşıladığını bulduk ve genellemelere ulaştık. İnanıyoruz ki bir bilgi kaynağının entropisi hesaplanırken mertebesi arttığında kaynağın entropi değeri orijinal entropi değerine daha çok yaklaşacaktır. Aynı şekilde sonlu bir Markov yörüngesinde tepe sayısı arttıkça orijinal entropi değerine daha çok yaklaşılacaktır. Bu problemler açık problem olarak kalır.

KAYNAKLAR

AKDENİZ, F., 1988. Olasılık ve istatistik, Baki Kitabevi, Adana.

ASH, R. B., 1965. Information Theory, Dover Publications, Inc. New York.

BILLINGSLEY, P., 1965. Ergodic Theory and Information, Wiley, NewYork.

COHN, D. L., 1980. Measure Theory, Birkhäuser, Boston, 373 s.

DENKER, M., GRILLENBERG, C., and SIGMUND, K., 1976. Ergodic Theory on Compact Spaces. Springer Lecture Notes in Math. 527 s.

EKROOT, L., and COVER T. M., 1991. The entropy of a randomly stopped sequence, IEEE Trans. Inform. Theory, 37: 1641-1644.

EKROOT, L., and COVER, T. M., 1993. The entropy of Markov trajectories, IEEE Trans. Inform. Theory, 39 (4): 1418-1421.

İNAL, H. C.,1988. Olasılıksal Süreçlere Giriş (Markov Zincirleri), Hacettepe Yayınları, Ankara.

KHINCHIN, A. I., 1957. Mathematical Foundations of Information Theory Dover Publications, Inc. New York.

LIOYD, S., and PAGELS, H., 1988. Complexity as thermodynamic depth, Ann. Phys., 188: 186–213.

SHANNON, C. E., 1948. A mathematical theory of communication. The Bell System Technical J. 27, 379–423 and 623–656.

THOMAS, M. C., and JOY, A. T., 1991. Elements of Information Theory, John Wiley & Sons.

YAGLOM, A. M., and YAGLOM , I., 1966. M. Çeviren: Lutfi Biran, İhtimaliyet ve İnformasyon, Türk Matematik Derneği Yayınları, İstanbul, Sayı:28.

WALTERS, P., 1982. An Introduction to Ergodic Theory, Springer Verlag, Berlin.

ÖZGEÇMİŞ

26.08.1983 yılında Şanlıurfa'da doğdu. İlköğretimini Bahçelievler İ.Ö.O. ve lise öğrenimini Şanlıurfa Kız Lisesi'nde tamamladı. 2001 yılında Fırat Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümünü kazandı. 2005 yılında buradan mezun oldu. 2005 yılının Eylül ayında Harran Üniversitesi Fen Bilimleri Enstitü'sünde Yüksek Lisans'a başladı.

ÖZET

Klasik ergodik teoride ve enformasyon teoride, Shannon (1948) tarafından önerilen entropi ifadesi uzun yıllardan beri pek çok yazar tarafından incelendi. Özellikle Ekroot ve Cover (1993) ve Ash (1965) tarafından verilen çalışmalar çoğu inceleme için bir yol oldu. Ekroot ve Cover (1993) sonlu durum indirgenemez Markov zincirlerinin yörüngelerinin entropisini verdiler. Genelde bir enformasyon kaynağının entropisini hesaplamak zordur. Ash (1965), Markov kaynaklarının belli bir sınıfının belirsizliğini (entropi) inceledi. Ek olarak, bir rastgele durdurulmuş dizinin entropi ifadesi, Ekroot ve Cover (1991) tarafından incelendi.

Bu tez dört bölümden oluşmaktadır. İlk bölümde genelde çok fazla detaya girmeden ihtiyaç duyduğumuz tanım ve teoremler verilmektedir. İkinci bölümde, Markov zincirleri, enformasyon kaynakları ve ölçüm entropi gibi bazı sonuçlar incelenmektedir. Üçüncü bölüm, gerekli metot ve materyallere ayrılmaktadır. Dördüncü bölüm üç alt kesimden oluşmaktadır. İlk alt kesimde, sonlu Markov zincirlerinin entropisi incelenmekte ve bazı uygulamalar yapılmaktadır. İkinci alt kesimde, enformasyon kaynakları ve bir enformasyon kaynağının entropisi incelenmekte ve grafikler ve tablolar yardımıyla bazı uygulamalar elde edilmektedir. Ayrıca, birinci ve ikinci alt kesimlerde bazı karşılaştırmalar yapılarak, bir grup sonuç verilmektedir. Son alt kesimde, bir rastgele durdurulmuş dizinin entropisi incelenmektedir.

SUMMARY

In the classical ergodic theory and information theory, the notion of entropy proposed by Shannon (1948) has been studied by many authors for long years. Specially the works given by Ekroot and Cover (1993) and Ash (1965) is a path for many studies. Ekroot and Cover (1993) have given the entropy of trajectories of finite state irreducible Markov chains. In general, the entropy of an information source is difficult to calculate. Ash (1965) has investigated the uncertainty (entropy) of a certain class of Markov sources. In addition, the notion of entropy of a randomly stopped sequence has been investigated by Ekroot and Cover (1991).

This thesis is consisting of four chapters. In the first chapter generally without going into too many details some definitions and theorems, which are needed by us, are given. In the second chapter some basic results such as Markov chains, information sources and measure entropy are studied. The third chapter is devoted to necessary method and materials.

The fourth chapter contains three subsections. In the first subsection, the entropy of finite Markov chains is investigated and some applications are done. In the second subsection, information sources and the entropy of an information source are studied and some applications are obtained by means of tables and graphs. In addition, in the first and the second subsection, by making some comparisons a group of results is given. In the last subsection, the entropy of a randomly stopped sequence is investigated.