BELİRSİZLİĞİN ÖLÇÜLMESİ ve ENTROPİ

Dr. Osman Zeki ÇETİNKAYA

Bu çalışmanın amacı, termodinamikteki entropi ve iletişim teorisindeki informasyon kavramlarının istatistiksel karakterini ve hepsinde ortak olan belirsizlik olgusunun ölçülmesi sorununu açıklamaya çalışmaktır. Deterministik anlayışın, özellikle Laplace'çı katı yorumun çağdaş fizikteki gelişmelerle pozitif bilimlerde aşılması, tüm bilim felsefesini de etkilemiş, ve olasılığa dayanan bir yaklaşım toplumsal bilimlere kadar her alanda hâkim olmaya başlamıştır. Şüphesiz bu olgu, her şeyi olasılığın belirsizliğine terk ederek bilimsel kanunların inkârı anlamına gelmemektedir. Sadece, kanunları katı bir mekanik determinizm anlayışıyla mutlaklaştırmak yerine, her kanunun henüz tamamlanmamış olduğunu, doğa ile ilgili bilgilerimizin göreli karakterinden hareketle kanunların sadece bir yaklaşım olduğunu kabul etmek demektir. Her kanun kendisini karşıtı ile, yani kanunlardan sapmalarla açıklar. Kanunun belirttiği genel ve zorunlu bağıntı, kendisini ancak tek başına ve çeşitli özellikleri ile meydana gelen rastlantısal olaylarla gösterir. Sonsuz çeşitlilikteki doğa ve toplumsal olaylarda rastgelelik içinde genel ve zorunlu olan kanuna ulaşmak, yani rastgelelikle determinizmi ayni anda var olan iki kategori olarak ele almak çağımızın ulaştığı bilimsel yaklaşımın temelini oluşturmaktadır.

Burada ele alınacak her üç teorinin de ortak yanı, tümdengelimci olmalarıdır. Dolayısıyla hepsi de matematiksel bir disiplindirler.

1) *Termodinamikte Entropi Kavramı*

Mantıksal ve matematiksel bilimler yanında doğa bilimlerinin tümdengelimci çıkarsamalarla oluşmuş dallarından biri de termodinamiktir. «Termodinamik enerji ve enerjinin dönüşümleri ile, ısı etkilerini içeren sistemlerde değişim ve denge eğilimlerini inceler» (1). Konusu basınç, hacim, sıcaklık vb. gibi makroskopik özelliklerdir. Termodinamiğin bu

---

(1) Tez, Z. «Termodinamik Üzerine», Doğa ve Bilim Dergisi, Sayı: 5, Mart-Nisan 1981, s. 73.

özellikleri mikroskopik yapıyı incelemeden açıklayabilmesi önemli bir özelliğidir. Termodinamiğin dört temel kanunu ve entropinin tanımlandığı istatistiksel bağıntı atom teorisinden önce, postula olarak ortaya konmuş, geçerlilikleri çok sayıda deney sonunda görülmüştür. Atom veya moleküller ile ilgili modellerden bağımsız, çok genel makroskopik önerilerdir. Makroskopik durum, mikroskopik durumların istatistiksel bir ortalamasıdır. Örneğin sıcaklık ve basınç gibi nicelikler istatistiksel bir ortalama niteliğindedirler. «Mikroskopik yapı ile termodinamik bulgular arasındaki ilişki istatistiksel termodinamik» (2) adlı disiplini meydana getirir. Tek tek atomlarm hareketlerini gözlemek mümkün olmadığı için, pratik olarak gözleyebildiğimiz tek özellik, istatistiksel görüşlerle elde edilen ortalama değerlerdir. Termodinamiğin istatistiksel teorisi bu temele dayanır. İstatistik kavramlar bir sistemdeki atom ve moleküller hakkında mikroskopik bilgilerle birleştirildiği zaman da, «İstatistik mekanik» disiplini elde edilir ve dolayısıyla entropi veya bunun sonuçlarına dayanan olasılık önerileri yapmak mümkün olur (3).

Sıcaklık ve ısı arasında ayrım 18'inci yüzyıla kadar yapılmamıştı. Isının enerji teorisi ancak maddenin atom teorisi kurulduktan sonra gelişebilmiştir. Termodinamiğin, enerjinin sakımımmı ifade eden birinci kanunu tümdengelimci bir düşünceyle ortaya atılmışken, ikinci kanunu hemen hemen yalnızca tümevarımcı bir düşünceyle ortaya çıkmıştır. L. S. Carnot (1796-1832) 1824'de, ısı daha bir enerji şekli olarak tanınmadan önce ısı makinaları üzerine teorik analizini yayınlamış ve bu düşüncelerin daha sonraları Kelvin (1824-1907) ve Clausius (1822-1888) tarafmdan geliştirilmesi termodinamiğin ikinci kanununun makroskopik ifadesine yol açmıştır. Carnot ısı enerjisini öteki enerji biçimlerine tam olarak dönüştürebilecek sürekli çalışan bir makina yapılamayacağını, termodinamiğin ikinci kanunu olarak ortaya atmış, Clausius ise entropi kavramını geliştirmiştir. Bu kanunla ısı ve işin enerji aktarım biçimleri olduğu ve birbirlerine dönüştürülmelerinin asimetrik olduğu gösterilmiştir. Bu dönüşümlerde korunan şeyin enerji olduğu, entropi için ise korunmanın söz konusu olmadığı ortaya çıkmıştır. Entropi enerji dönüşüm süreçlerindeki etkinliğin bir ölçüsü olarak görülmüştür (4).

Termodinamiğin ikinci kanunu olarak bilinen ifade, Q ısı ve T de sistemin mutlak sıcaklığını göstermek üzere,

---

(2) Tez, Z. s. 75.

(3) Reif, F. «İstatistik Fizik», Berkeley Fizik Dersleri, C: 5, A.Ü. Fen Fakültesi, 1977, s. 285.

(4) Tez, Z. s. 86.

$$dS = \frac{dQ}{T} \qquad (1)$$

olarak ifade edilebilir. Sistem dQ kadar ısı soğurduğu zaman, entropisindeki değişme dS kadar olmaktadır. Dolayısıyla sistemin entropisinden çok, entropisindeki değişmeden söz etmek daha doğrudur.

Her doğal işlemde, işleme giren bütün sistemler gözönüne alınırsa, bir entropi artışı vardır. Entropi ya artar ya da sabit kalır. Adyabatik olarak (ısı geçirmeksizin) yalıtılmış bir sistemin entropisi artma eğilimindedir. Her sistem yalıtılmış halde daha olası bir yapıya doğru normal bir evrim izler. Entropinin azalmasını mümkün kılan bir işlem yoktur (5).

Entropinin daima artması özelliği, enerjinin daha elverişsiz hale geleceği anlamındadır: «Pek uzak bir gelecekte bütün evrenin mutlak bir biçimlilik kazanacağını düşünebiliriz. Böyle bir hale varıldığı zaman evrenin enerjisinde bir değişiklik olmayacaksa da, bütün fiziksel, kimyasal ve belki de biyolojik işlemler sona erecektir» (6). «Isı ölümü» denen bu durum ısı kaynaklarının giderek evrensel bir ılıklığa doğru tükeneceklerini ve evrenin öleceğini öngörür.

20'nci yüzyılın başlarından itibaren gelişen kuantum mekaniği ile atom ve molekül özelliklerinin anlaşılmasında büyük gelişme kaydedilmiştir. Kuantum teorisine göre, atomlar veya moleküller yalnızca sonlu sayıdaki farklı yapılarda var olabilirler. Atom düzeyinde süreklilik yoktur, fakat yalnız kararlı yapılar ve atomik sistem enerji yutarak veya yayarak bir yapıdan diğerine aniden atlarlar. Nicel fizik sistemlerinin kesikli şekillenmelerinin her biri Planc tarafından bir «Complexion» yani, sistemin girilebilir durumları olarak adlandırılmıştır (7). Ölçemediğimiz çok sayıda mikroskopik değişkenler, bilinmeyen bütün bu nicelikler, sistemin girilebilir durumları denen nicel yapıların sonsuz çeşitliliğini ele almayı mümkün kılar. Bu mikroskopik durumların toplam sayısı P'den hareket ederek sistemin entropisi tanımlanabilir:

Bir sistemin S entropisi, sistemin girilebilir durumlarının P sayısı cinsinden,

$$S = k \ln P \qquad (2)$$

---

(5) Sears, F.W. «Fizik Prensipleri», İTÜ., 1955.

(6) Sears, F.W. s. 512.

(7) Brillouin, L. «Science and information theory», Academic Press inc., second ed. 1962, s. 120.

bağıntısı ile tanımlanır. Bağıntıdaki k sabiti Boltzmann sabiti olarak adlandırılır. Entropi böylece sistemin rastgelelik derecesinin logaritmik bir ölçüsünü sağlar.

Girilebilir durumların sayısı P, parçacıkların kütle ve enerjisiyle doğru orantılıdır ve enerjinin son derece hızla artan bir fonksiyonudur[8]. Aldığı değer ise bu parametrelere bağlı olarak teorik düzeyde hesaplanabilir.

Boltzmann-Planc formülü olarak bilinen (2) bağıntısı entropinin istatistiksel tanımıdır.

II) *İstatistikte Belirsizliğin Ölçülmesi*

Matematiğin uygulamalı bir disiplini olan istatistik, determinizm anlayışının günümüzde aldığı çağdaş anlamla fizikten biyolojiye kadar bilimin her alanında yaygın olarak kullanılmaya başlamıştır. Olasılık teorisinden hareket ederek, tümdengelimci bir yöntemle matematiksel modellerin oluşturulması ile başlayan süreç, uygulama alanında bu kez tersine bir çıkarsama ile, belli varsayımlara uygunluk koşuluyla gözlenen değerlerin hangi olasılıkla ortaya çıktığını hesaplamaya çalışır. Gözlemlerin sınırlılığı ve gözlem anında araştırıcının gerçeğe müdahalesi gibi, gerçeği tam olarak gözlemeyi olanaksız kılan faktörlerin olumsuz etkilerinin bu şekilde ortadan kaldırılmasına çalışılır.

n sayıda değişken ile ilgili N sayıda gözlem yapıldığını varsayalım. Her değişken N boyutlu bir uzayda bir vektör olarak gösterilebilir. Yukarıda açıklanan nedenlerle, vektörlerin bu uzaydaki konumlarının mutlak değil, ancak olasılıkla belirlenmiş olacağı açıktır. Vektörlerin (yani değişkenlerin) uzaydaki şekillenmeleri için iki uç durum düşünelim: İlk olarak bütün vektörlerin bu uzayda hiçbir ilişki ifade etmeyecek şekilde dağıldığı tam bir düzensizlik, rastgelelik hali, ikinci olarak da bütün vektörlerin birkaç noktada yoğunlaştığı tam bir belirlenme hali. İlk durumda değişkenlerle ilgili belirsizliğin en fazla olduğu, ikinci durumda ise en aza indiği söylenebilir. Gerçekte rastlanan durumlar ise bu iki uç halin arasında yer alır. Benzer şekilde, belirsizlik n boyutlu uzayda yer alan N sayıdaki vektör için de söylenebilir. Bu kez vektörler değişkenleri değil, gözlemleri ifade edecek ve gözlemler arasında bir düzen (ya da düzensizlik) söz konusu olacaktır.

İstatistiğin bütün bilimlere uygulanabilen bir teknik olması dolayısıyla, yukarıdaki açıklamada değişken kavramı soyutlanarak kullanıl-

(8) Reif, F., s. 123.

mıştır. Fiziğe paralellik kurmak amacıyla, değişkenler bir fiziksel sistemin makroskopik değişkenleri olarak ele alınırsa, istatistik anlamdaki belirsizliğin daha önce tanımlanan entropi kavramı ile olan özdeşliği ortaya çıkar. Entropiye karşılık olarak istatistikte «Parsimoni» kavramı ile karşılaşırız. Her iki kavram arasında analojik bir ilişki kurulabilir. Parsimoni, genel olarak bir teorinin mümkün olduğu kadar az sayıda aksiyom, teorem, kanun ve varsayımlara dayanması ilkesidir ve bu anlamda bütün bilimler için geçerlidir. Özel olarak istatistikte, değişkenler arasındaki ilişkiyi en az sayıda açıklayıcı değişkenlerle ve denklemlerle ifade etmek anlamına gelir.

İstatistik anlamda belirsizliğin aksiyomatik bir teori olarak kurulması aşağıda ele alınacaktır. Bunun için yalnızca olasılık teorisi bir öncül teori olarak gerekecektir.

p olasılığı ile meydana gelen bir X olayı ele alalım. Tanımlamaya çalışacağımız belirsizlik fonksiyonu $p=1$ için sıfıra eşit (belirsizlik yok) ve p nin kesinlikle azalan bir fonksiyonu olmalıdır. q olasılığı ile meydana gelen, X den bağımsız ikinci bir Y olayı tanımlanırsa, her iki olayın birlikte belirsizliği, X ve Y olaylarının belirsizliklerinin toplamına eşit olmalıdır :

$$B(pq) = B(p) + B(q) \tag{3}$$

Zira, X ve Y bağımsız oldukları için, $P(XY) = pq$ olacak ve, Y olayının gerçekleştiğini öğrendiğimiz zaman ek belirsizlik,

$$B(pq) - B(p)$$

olacaktır. Oysa olaylar bağımsız olduklarından X in gerçekleştiğini öğrenmemiz Y nin olasılığını değiştirmeyecek, ek belirsizlik $B(q)$ olacaktır. Buradan da (3) eşitliğine ulaşılır.

Bu koşul belirsizlik fonksiyonunun tanımlanmasında önemli bir sonuca yol açar. (3) eşitliğini sağlayan fonksiyon, logaritmik fonksiyonların aşağıdaki özelliğini çağrıştırmaktadır:

$$\log(ab) = \log(a) + \log(b)$$

gerçekten de, $p=1$ için $B(p) = 0$ ve $B(p)$ nin p nin azalan bir fonksiyonu olması, yani $p>q$ için $B(p)<B(q)$ koşulları da(*) ancak logaritmik fonksiyonların özellikleri ile sağlanabilir[9].

(*) $p > q$ için $\log(p) > (q)$ olacağı açıktır. Ancak ileride görüleceği gibi, burada kastedilen logaritmik fonksiyonun (−1) ile çarpılmış şeklidir.

(9) Yaglom, A.M. - Yaglom, I.M. «İhtimaliyet ve İnformasyon», Türk Matematik Derneği, 1966, s. 43.

Belirsizliğin formelleştirilmesi, yukarıda belirtilen özelliklere dayanılarak dört aksiyomla mümkün olabilir:

Aksiyomlar (10):

1. $B(1) = 0$

2. $B(p)$ p nin kesinlikle azalan bir fonksiyonudur:

   $p>q$ için $B(p)<B(q)$

3. $B(p)$, p nin sürekli bir fonksiyonudur. Bu aksiyom, p deki küçük bir değişmenin $B(p)$ de küçük bir değişmeye karşılık olacağını, sezgisel olarak bekleyebilmemize dayanır.

4. $B(pq) = B(p) + B(q)$ $\qquad 0<p, q\leqslant 1$ (**)

Bu aksiyonlara dayanarak, belirsizlik fonksiyonunun biçimi için aşağıdaki teoreme geçilebilir:

Teorem I: $B(p)$ yukarıdaki aksiyomları sağlarsa, c herhangi bir sabit olmak üzere,

$$B(p) = -c \log_2(p) \tag{4}$$

yazılabilir (*).

Nitekim, 4'üncü aksiyoma dayanarak,

$$\begin{aligned} B(p^2) &= B(p) + B(p) \\ &= 2\,B(p) \end{aligned}$$

olacak, indüksüyon yöntemi ile de daha genel olan aşağıdaki eşitliğe ulaşılacaktır:

$$B(p^m) = m\,B(p) \tag{5}$$

Benzer şekilde, n bir tam sayı olmak üzere,

---

(10) Ross, S. «A first course in Probability», Macmillan Pub. Co. 1976, s. 276.

(**) Olasılıkların sıfır değeri için fonksiyonun tanımsız olması, logaritmik fonksiyonlara uygunluğu sağlamak için kabul edilmiştir. Bu kabul, gerçekleşmeyen olaylar için belirsizliğin tanımsız olacağı şeklinde yorumlanabilir.

(*) Logaritmanın tabanı önceden verilmezse, c katsayısı kaldırılabilir. Çünkü her zaman $b = a^{1/c}$ olmak üzere,

$$c \log_a(p) = \log_b(p) \text{ yazılabilir.}$$

$$B(p) = n\,B(p^{1/n})$$

$$B(p^{1/n}) = \frac{1}{n}\,B(p) \qquad (6)$$

ve nihayet, (5) ve (6) eşitliklerinden,

$$\begin{aligned} B(p^{m/n}) &= m\,B(p^{1/n}) \\ &= \frac{m}{n}\,B(p) \end{aligned}$$

yazılabilir, bu ifade de,

$$B(p^{u}) = \mu B(p)$$

eşitliğine denktir.

3'üncü aksiyomun bir sonucu olarak B(p) sürekli olduğu için, (7) eşitliği tüm negatif olmayan u değerleri için geçerlidir.

Her p değeri için,

$$u = -\log_2(p)$$

dönüşümü yapılırsa,

$$p = 2^{-u}$$

olacak ve (7) eşitliğinden de,

$$\begin{aligned} B(p) &= B(2^{-u}) \\ &= u\,B(2^{-1}) \end{aligned}$$

ve $B(2^{-1}) = c$ denirse,

$$B(p) = -c\,\log_2(p)$$

eşitliği ispatlanmış olur. Burada c katsayısının, ilk iki aksiyoma dayanarak pozitif olacağı açıktır:

$$c = B(2^{-1}) > B(1) = 0$$

$c = 1$ kabul edilirse, belirsizliğin bit (binary digit-iki tabanına göre) birimi ile ifadesi mümkün olur.

Bu aşamada, belirsizlik fonksiyonunun rastlantısal değişkenler için alacağı biçim incelenebilir. $x_1, x_2, \ldots x_n$ değerlerini $p_1, p_2, \ldots p_n$ olasılıklarıyla alan bir X rastlantısal değişkeni ele alalım. $-\log_2(p_i)$ X in $X = x_i$ değerini aldığı zamanki belirsizlik olacağı için, X rastlantısal değişkeni için belirsizliğin beklenen miktarı,

$$H(X) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{8}$$

olacaktır. Bu büyüklüğün X in aldığı değerlere değil de, yalnızca olasılıklarına bağlı olması önemli bir özelliktir. Olasılık dağılımları aynı, fakat değerleri farklı iki rastlantısal değişken için belirsizlik birbirine eşit olacaktır.

H(X) maksimum değerini, bütün p_i olasılıkları birbirine eşit olduğu zaman, yani maksimum belirsizlik halinde alır. Bu durumda

$p_i = \dfrac{1}{n}$ dir ve p_i lerin bu değeri için,

$$H(X)_{max} = \log(n)$$

olduğu kolaylıkla gösterilebilir. Minimum değeri ise teorik olarak sıfıra eşittir ve bir p_i değeri bire, diğer bütün p_i değerleri sıfıra eşit olduğu zaman ortaya çıkar. Bu durumda belirsizlik yoktur.

Belirsizliğin bir ölçümü olan yukarıdaki H(X) büyüklüğünün, fizikteki entropinin istatistiksel tanımı, yani (2) eşitliği ile verilen Boltzmann-planc formülü ile olan benzerliği açıktır. Her iki kavram da düzensizliği, belirsizliği veya rastgeleliği ifade etmek üzere türetilmişlerdir. Kavramların özdeşliği dolayısıyla H(X) büyüklüğüne X rastlantısal değişkeninin entropisi adı da verilebilir.

Şimdi de, X değişkeninin yanı sıra $y_1, y_2 \ldots y_m$ değerlerini alan ikinci bir Y rastlantısal değişkeni ele alalım ve genellik amacıyla her iki değişkenin bağımlı olduğunu kabul edelim. (X, Y) rastlantısal vektörünün değeri hakkında H(X, Y) ile göstereceğimiz belirsizlik, tanımı gereği,

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i,y_j)\log p(x_i,y_j) \qquad (9)$$

olacaktır. Y değişkeninin y_i değerini aldığı gözlenmiş olsun. Bu durumda X için kalan belirsizlik miktarı,

$$H_{Y=y_j}(x) = -\sum_{j=1}^{n} p(x_i|y_j)\log p(x_i|y_j) \qquad (10)$$

eşitliği ile verilecektir. Buradan da, Y gözlendiği zaman X için ortalama belirsizlik miktarı, veya Y nin gerçekleşmesine bağlı X'in entropisi,

$$H_Y(X) = \sum_{j=1}^{m} H_{Y=y_j}(X)\, p_Y(y_j) \qquad (11)$$

olacaktır. Burada $p_Y(y_j)$ Y değişkeninin marjinal olasılıklarıdır.

H(X, Y) yi H(X) ve $H_Y(X)$ e bağlayan aşağıdaki önerme X ve Y için belirsizliğin, Y nin belirsizliği ile Y gözlendiği zaman X deki ortalama belirsizliğin toplamına eşit olduğunu göstermektedir:

$$H(X,Y) = H(Y) + H_Y(X) \qquad (12)$$

Bu önermenin ispatı, H(X,Y) nin açık ifadesinden elde edilebilir[11].

Özel olarak X ve Y rastlantı değişkenleri bağımsız iseler,

$$H(X,Y) = H(X) + H(Y) \qquad (13)$$

olur ve bu eşitlik «Toplam entropiler kanunu» olarak adlandırılır[12].

III) *İletişim Teorisinde Entropi*

1928'de Hartley, telefon haberleşmeleriyle ilgili olarak yaptığı çalışmalarda entropi kavramını uygulamayı denemiştir. Hartley belirsizlik ölçütü olarak, n sayıda mümkün sonucu olan bir değişken için log(n) sayısını düşünmüştür. Bu yaklaşım olasılıklara yer vermediği için (ya

(11) İspat için bkz. Ross, S. s. 279 ve Yaglom, A. M. — Yaglom, I. M. s. 58.

(12) Yaglom, A. M. — Yaglom, I. M. s. 56

da değişkenin her değerine zımnen ayni olasılığı tanıdığı için) yanlış bir yaklaşımdı. Bu yanlışı Shannon düzeltecek ve X değişkeninin olasılığı p_i olan bir x_i sonucuna $-\log(p_i)$ye eşit bir belirsizlik tanıyacaktır. X değişkeninin belirsizlik derecesi olarak da, bütün sonuçların belirsizliklerinin ortalama değerini alacaktır:

$$H(X) = -\sum_{i=1}^{n} p_i \log(p_i) \qquad (14)$$

$-\log(p_i)$ belirsizliklerinin, istatistiksel bir ifade ile söylersek, beklenen değerinden başka bir şey olmayan bu ifadeye Shannon entropi adını vermiştir. Böylece, «Entropi ve informasyon arasındaki ilişki Shannon tarafından tekrar keşfedilmiş» ([13]) olmaktadır.

Bu formülasyon informasyonlarm iletilmesi sırasında seçme serbestliğinin ölçülmesi olarak da yorumlanabilir. Seçme serbestliği az ise, yani kesinliğin olasılığı fazla ise yüksek derecede bir yapılanma, eğer seçme serbestliği fazla ise, yani kesinlik olasılığı düşükse yüksek derecede düzensiz bir durum var demektir: Rastgelelik yüksektir.

Entropi kavramının bu şekilde ifade edilmesi, belirli haberleşme problemlerinin çözümü için ortaya atılmış ve önceleri yalnızca bu alanda uygulanmıştır. H(X) entropisi X değişkeninin aldığı değerlere bağlı olmadığından, bu formülasyon informasyonun süresi veya anlamı ile ilgili değildir.

İstatistikle ilgili olarak entropi için gösterilen her özellik iletişim teorisi için de geçerli olacaktır.

(12) eşitliği ile ortaya çıkan sonuç, informasyon teorisinde oldukça önemlidir. X rastlantı değişkenindeki belirsizlik miktarı, Y rastlantı değişkeni gözlendiği zaman ortalama olarak azalır. Bu sonucu sağlayan aşağıdaki özelliktir :

Teorem 2 : $H_Y(X) \leqslant (X)$

Teoremdeki eşitlik hali, ancak ve ancak X ve Y değişkenleri bağımsız iseler mümkündür([14]).

Bu teorem Y nin gerçekleşmesinin X in belirsizliğini ancak azaltacağını veya bu belirsizliği değiştirmeyeceğini (Bağımsızlık hali), fakat

(13) Brillouin, L. s. 161.
(14) İspat için bkz. Ross, S. s. 281.

hiçbir zaman bunu artıramayacağım ifade etmektedir. Şu halde, Y nin gerçekleşmesiyle X in belirsizliğinin ne kadar azaldığı,

$$I(X,Y) = H(X) - H_Y(X) \qquad (15)$$

farkı ile hesaplanabilir. Bu farka X değişkeni hakkında Y değişkeninin verdiği informasyon (bilgi) miktarı veya X hakkında Y nin içerdiği informasyon denir[15].

Entropi belirsizlik derecesini soyut bir tarzda ölçerken, bir entropi farkı ile tanımlandığından informasyonu ölçmek mümkündür. X değişkeninin H(X) entropisi, X in kendisi hakkında içerdiği informasyon veya X hakkında mümkün olan en büyük informasyon olarak tanımlanabilir.

(15) denkleminin,

$$I(X,Y) = H(Y) - H_X(Y) \qquad (16)$$

şeklinde de yazılabileceği açıktır. Buradan, X değişkeni hakkında bir Y değişkeninde bulunan I(X,Y) informasyonunun Y değişkeninin H(Y) entropisinden küçük veya buna eşit olacağı sonucu çıkar. Zaten, Y değişkeninde, bir diğer X değişkeni hakkında bulunan informasyon, Y de kendisi hakkında bulunan informasyondan, yani H(Y) den daha büyük olamaz.

Shannon'un informasyonu yukarıda özetlenen şekilde formelleştirmesi, informasyonun nicel olarak ölçülebilmesini de mümkün kılmıştır. Her informasyon, 0 ve 1 sembollerine, yani ikili sisteme dayanarak bir vektörle ifade edilebilir. X kesikli rastlantısal vektörünün değerinin A mekânında gözlendiğini ve 0 ve 1 işaretlerinden oluşan bir haberleşme devresi ile B mekânına gönderileceğini kabul edelim. X değişkeninin sadece dört değer aldığını ve bunları aşağıdaki iki farklı şekilde kodladığımızı düşünelim :

(I)		(II)	
$x_1 \rightarrow$	00	$x_1 \rightarrow$	0
$x_2 \rightarrow$	01	$x_2 \rightarrow$	10
$x_3 \rightarrow$	10	$x_3 \rightarrow$	110
$x_4 \rightarrow$	11	$x_4 \rightarrow$	111

(15) Yaglom, A. M. — Yaglom, I. M. s. 73.

Kodlama işleminde dikkat edilecek nokta, bir kodun diğer bir kodun genişletilmesi şeklinde yapılmamasıdır. Örneğin II inci durumda hiçbir değer sıfır ile başlamamalıdır. Aksi takdirde x_1 in genişletilmiş bir şekli olacaktır.

Kodlamanın temel amaçlarından biri, beklenen bit sayısının minimum kılınmasıdır. X in aldığı değerlerin olasılıklarının sıra ile,

$\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ ve $\frac{1}{8}$ olduğunu varsayalım. Bu durumda ikinci kodlamanın beklenen bit sayısı,

$$\frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(3) + \frac{1}{8}(3) = 1.75 \text{ bit}$$

olacaktır. Birinci durumda ise bu değer 2 bit'tir. Dolayısıyla ikinci kodlamanın daha etkin olduğu söylenebilir.

Kodlamanın bir diğer amacı ise etkinliktir. Veri bir rastlantısal X vektörü için kodlamada ulaşılabilecek maksimum etkinlik nedir? Bu soru bizi entropi kavramına götürür: Gönderilecek bit'lerin ortalama sayısı en az X in entropisi kadar büyük olacaktır. İspatını vermeden, informasyon teorisindeki bu önemli teoremi ifade edelim:

X, $x_1, x_2, \ldots x_N$ değerlerini $p(x_i)$ olasılıklarıyla alsın. X in, x_i ye n_i bit karşılığı veren herhangi bir kodlaması için,

$$H(X) \leqslant \sum_{i=1}^{N} n_i \, p(x_i) \qquad (17)$$

eşitsizliği geçerlidir(18).

Yukarıda verilen ikinci kodlama için H(X) entropisi hesaplanırsa,

$$H(X) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{4}\log\frac{1}{8}$$

$$= 1.75$$

bulunur ki, bu da kodlamanın beklenen bit sayısına eşittir. Sonuç olarak, X değişkeni için en etkin kodlamanın ikinci kodlama olduğu ortaya çıkar.

(16) Ross, S. s. 283.

YARARLANILAN KAYNAKLAR

1. Abramson, N. «Information theory and coding» McGraw-Hill Book co., 1963.
2. Brillouin, L. «Science and information theory» Academic press inc., second ed. 1962.
3. Feinstein, A. «Foundations of information theory» McGraw-Hill Book co., 1958.
4. Reif, F. «İstatistik fizik» Berkeley fizik dersleri, c. 5 A. Ü. Fen Fakültesi, 1977.
5. Ross, S. «A first course in Probability» Macmillan Pub. co. 1976.
6. Sears, F. W. «Fizik Prensipleri» c. 1, İ.T.Ü. 1955.
7. Tez, Z. «Termodinamik Üzerine» Doğa ve Bilim dergisi, s. 5, Mart-Nisan 1981.
8. Yaglom, A. M. — Yaglom, I. M. «İhtimaliyet ve İnformasyon» Türk Matematik Derneği, 1966.