MIT Açık Ders Malzemeleri

http://ocw.mit.edu

8.333 İstatistiksel Mekanik I: Parçacıkların İstatistiksel Mekaniği

2007 Güz

II.D Çok Sayıda Rassal Değişken

Birden çok rassal değişkenin varlığında, sonuçlar kümesi N-boyutlu bir uzaydır, $S_x = \{-\infty < x_1, x_2, \cdots, x_N < \infty\}$. Örneğin, bir gaz parçacığının konum ve hızını tanımlamak altı koordinat gerektirir.

- *Birleşik OYF* $p(x)$, $x = \{x_1, x_2, \ldots, x_N\}$ noktası etrafında $d^N\mathbf{x} = \prod_{i=1}^{N} dx_i$ hacim elemanında bir sonucun olasılık yoğunluğudur. Birleşik OYF şu şekilde normalleştirilir:

$$p_{\mathbf{x}}(\mathcal{S}) = \int d^N\mathbf{x}\; p(\mathbf{x}) = 1\ . \tag{II.32}$$

Sadece ve sadece N rassal değişken *bağımsızsa* bileşik OYF tekil OYF'lerin çarpımına eşittir,

$$p(\mathbf{x}) = \prod_{i=1}^{N} p_i(x_i)\quad . \tag{II.33}$$

- *Koşulsuz OYF,* rassal değişkenlerin bir altkümesinin davranışını, diğerlerinin değerlerinden bağımsız olarak tanımlar. Örneğin, bir gaz parçacığının sadece konumuyla ilgileniyorsak bir koşulsuz OYF, verili bir noktada tüm hız bileşenleri üzerinden integral alarak tanımlanabilir, $p(\vec{x}) = \int d^3\vec{v}\; p(\vec{x}, \vec{v})$; daha genel olarak

$$p(x_1, \cdots, x_m) = \int \prod_{i=m+1}^{N} dx_i\; p(x_1, \cdots, x_N)\ . \tag{II.34}$$

- *Koşullu OYF,* rassal değişkenlerin bir altkümesinin davranışını, diğerlerinin belirtilmiş değerleri için tanımlar. Örneğin, belli bir $\vec{x}$ noktasındaki parçacığın hızı için OYF, $p(\vec{v}|\vec{x})$ ile gösterelim, birleşik OYF $p(\vec{v}|\vec{x}) = p(\vec{x}, \vec{v})/\mathcal{N}$ ile orantılıdır. $p(\vec{v}|\vec{x})$'nin normalleştirilmesiyle bulunan orantı sabiti,

$$\mathcal{N} = \int d^3\vec{v}\; p(\vec{x}, \vec{v}) = p(\vec{x}), \tag{II.35}$$

$\vec{x}$ noktasındaki parçacığın koşulsuz OYF'sidir. Genel olarak, koşulsuz OYFler Bayes Teoremi'nden elde edilir:

$$p(x_1, \cdots, x_m \mid x_{m+1}, \cdots, x_N) = \frac{p(x_1, \cdots, x_N)}{p(x_{m+1}, \cdots, x_N)}\quad . \tag{II.36}$$

Dikkat edilirse, eğer rassal değişkenler bağımsızsa koşulsuz OYF koşullu OYF'ye eşittir.

- Bir $F(x)$ fonksiyonunun *beklenen değeri*, önceden olduğu gibi bulunur:

$$\langle F(\mathbf{x})\rangle = \int d^N\mathbf{x}\; p(\mathbf{x})F(\mathbf{x}) \; . \tag{II.37}$$

- *Birleşik karakteristik fonksiyon,* birleşik OYF'nin N-boyutlu Fourier dönüşümünden elde edilir:

$$\tilde{p}(\mathbf{k}) = \left\langle \exp\left(-i\sum_{j=1}^{N} k_j x_j\right)\right\rangle \quad . \tag{II.38}$$

- *Birleşik momentler ve birleşik kümülantlar* sırasıyla $\tilde{p}(k)$ ve $\ln\tilde{p}(k)$ ile türetilir:

$$\begin{aligned}\langle x_1^{n_1}x_2^{n_2}\cdots x_N^{n_N}\rangle &= \left[\frac{\partial}{\partial(-ik_1)}\right]^{n_1}\left[\frac{\partial}{\partial(-ik_2)}\right]^{n_2}\cdots\left[\frac{\partial}{\partial(-ik_N)}\right]^{n_N}\tilde{p}(\mathbf{k}=\mathbf{0}) \quad , \\ \langle x_1^{n_1}x_2^{n_2}\cdots x_N^{n_N}\rangle_c &= \left[\frac{\partial}{\partial(-ik_1)}\right]^{n_1}\left[\frac{\partial}{\partial(-ik_2)}\right]^{n_2}\cdots\left[\frac{\partial}{\partial(-ik_N)}\right]^{n_N}\ln\tilde{p}(\mathbf{k}=\mathbf{0}) \quad .\end{aligned} \tag{II.39}$$

Önceden tanımladığımız grafiksel bağıntılar, birleşik momentler (etiketlenmiş noktaların tüm öbekleri) ile birleşik kümülantlar (bağlantılı öbekler) arasında da uygulanabilir. Örneğin,

$$\begin{aligned}\langle x_\alpha x_\beta\rangle &= \langle x_\alpha\rangle_c\langle x_\beta\rangle_c + \langle x_\alpha x_\beta\rangle_c \quad , \quad \text{ve} \\ \langle x_\alpha^2 x_\beta\rangle &= \langle x_\alpha\rangle_c^2\langle x_\beta\rangle_c + \langle x_\alpha^2\rangle_c\langle x_\beta\rangle_c + 2\langle x_\alpha x_\beta\rangle_c\langle x_\alpha\rangle_c + \langle x_\alpha^2 x_\beta\rangle_c \quad .\end{aligned} \tag{II.40}$$

$\langle x_\alpha x_\beta\rangle_c$ bağlantılı korelasyonu, eğer x_α ve x_β bağımsız rassal değişkenler ise sıfırdır.

- *Birleşik Gauss tipi dağılım,* denklem (II.15)'in N tane rassal değişken için genelleştirilmiş halidir,

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N\det[C]}}\exp\left[-\frac{1}{2}\sum_{mn}\left(C^{-1}\right)_{mn}(x_m-\lambda_m)(x_n-\lambda_n)\right] \quad , \tag{II.41}$$

burada C bir simetrik matris, ve C^{-1} onun tersidir. Normalleştirme çarpanını bulmanın en kolay yolu, C'yi köşegenleştiren birimsel matrisi kullanarak, $y_j = x_j - \lambda_j$ değişkenlerine bir doğrusal dönüşüm uygulamaktır. Bu işlem, normalleştirmeyi, varyansları C'nin özdeğerleriyle belirlenen N tane Gauss fonksiyonunun çarpımına indirger. Özdeğerlerin çarpımı, det$[C]$ determinantıdır (Bu ayrıca, C matrisinin pozitif tanımlı olmasını gerektirir). Buna karşılık gelen birleşik karakteristik fonksiyon benzer yoldan elde edilir,

$$\tilde{p}(\mathbf{k}) = \exp\left[-ik_m\lambda_m - \frac{1}{2}C_{mn}k_m k_n\right] \quad , \tag{II.42}$$

ki bu gösterimde indisler üzerinde toplama kuralı uygulanır.

Buradan, Gauss fonksiyonunun birleşik kümülantları $\ln \tilde{p}(k)$ kullanılarak bulunur,

$$\langle x_m \rangle_c = \lambda_m \quad , \quad \langle x_m x_n \rangle_c = C_{mn} \quad , \tag{II.43}$$

daha yüksek derece kümülantların tümü ise sıfırdır. $\{\lambda_m\} = 0$ özel durumunda, dağılımın tek *momentleri* sıfır olup, momentleri kümülantlar cinsinden ifade eden genel kurallar gereğince çift momentler, ilgili rassal değişkenlerin tüm ikili gruplama yolları üzerinden toplayarak elde edilir, örneğin,

$$\langle x_a x_b x_c x_d \rangle = C_{ab}C_{cd} + C_{ac}C_{bd} + C_{ad}C_{bc}. \tag{II.44}$$

Bu sonuç bazen *Wick teoremi*. diye anılır.

II.E Rassal Değişkenlerin Toplamı ve Merkezi Limit Teoremi

$X = \sum_{i=1}^{N} x_i$, toplamını ele alalım. Burada x_i rassal değişkenler, ve bunların birleşik OYF'si $p(\mathbf{x})$ olsun. X'in OYF'si,

$$p_X(x) = \int d^N\mathbf{x}\; p(\mathbf{x})\delta\left(x - \sum x_i\right) = \int \prod_{i=1}^{N-1} dx_i\; p\left(x_1, \cdots, x_{N-1}, x - x_1 \cdots - x_{N-1}\right), \tag{II.45}$$

yazılabilir ve ona karşılık gelen karakteristik fonksiyon (denklem .(II.38) kullanılarak) şöyledir:

$$\tilde{p}_X(k) = \left\langle \exp\left(-ik\sum_{j=1}^{N} x_j\right)\right\rangle = \tilde{p}\left(k_1 = k_2 = \cdots = k_N = k\right). \tag{II.46}$$

Toplamın kümülantları, $\ln \tilde{p}_x(k)$, ifadesi açılarak,

$$\ln \tilde{p}(k_1 = k_2 = \cdots = k_N = k) = ik \sum_{i_1=1}^{N} \langle x_{i_1} \rangle_c + \frac{(-ik)^2}{2} \sum_{i_1, i_2}^{N} \langle x_{i_1} x_{i_2} \rangle_c + \cdots, \qquad \text{(II.47)}$$

şöyle bulunur:

$$\langle X \rangle_c = \sum_{i=1}^{N} \langle x_i \rangle_c \quad , \quad \langle X^2 \rangle_c = \sum_{i,j}^{N} \langle x_i x_j \rangle_c \quad , \quad \cdots . \qquad \text{(II.48)}$$

Eğer rassal değişkenler bağımsızsa, $p(\mathbf{x}) = \prod p_i(x_i)$, ve $\tilde{p}_X(k) = \prod \tilde{p}_x(k)$. Denklem (II.48)'deki çapraz kümülantlar sıfırdır, ve X 'in $n$inci kümülantı basitçe tekil kümülantların toplamıdır, $\langle X^n \rangle_c = \sum_{i=1}^{N} \langle x_i^n \rangle_c$. N rassal değişkenin tümü bağımsız olarak aynı $p(x)$ dağılımından alındığında, $\langle X^n \rangle_c = N \langle x^n \rangle_c$ olur. Bu ise daha önce binom dağılımı için elde edilen sonucun genelleşmiş halidir. N'nin büyük değerleri için, toplamın ortalama değeri N ile orantılıdır, ortalama etrafındaki dalgalanmalar ise, standart sapmayla ölçüldüğü üzere, ancak $\sqrt{N}$ gibi büyür. $y = (X - N \langle x \rangle_c) / \sqrt{N}$, rassal değişkeninin ortalaması sıfır, kümülantları ise $\langle y^n \rangle_c \propto N^{1-m/2}$. biçimindedir. N'yi sonsuza götürdüğümüz limitte sadece ikinci kümülant sıfıra gitmekten kurtulacağından, y'nin OYF'si normal dağılıma yakınsar:

$$\lim_{N \to \infty} p\left(y = \frac{\sum_{i=1}^{N} x_i - N \langle x \rangle_c}{\sqrt{N}} \right) = \frac{1}{\sqrt{2\pi \langle x^2 \rangle_c}} \exp\left(-\frac{y^2}{2 \langle x^2 \rangle_c} \right). \qquad \text{(II.49)}$$

(Hatırlatmak gerekirse, sadece birinci ve ikinci kümülantı olan tek dağılım Gauss tipi normal dağılımdır.)

Çok sayıda rassal değişkenin toplamı için OYF'nin normal dağılıma yakınsaması, bu tip toplamların sıklıkla karşılaşıldığı istatistiksel mekanik bağlamında çok önemli bir sonuçtur. *Merkezi limit teoremi* bu sonucun daha genel bir biçimini ifade eder: Rassal değişkenlerin bağımsız olması gerek şart değildir, çünkü denklem (II.49)'un geçerliliği için $\sum_{i_1,\ldots,i_m}^{N} \langle x_{i_1} \ldots x_{i_m} \rangle_c \ll \mathcal{O}(N^{m/2})$ koşulu yeterlidir.

II.F Büyük Sayılar için Kurallar

Makroskobik sistemlerin denge özelliklerini tanımlamak için, istatistiksel mekaniğin mikroskobik serbestlik derecelerinin çok büyük sayısı N ile uğraşması gerekir. Gerçekten de, bazılarını bu bölümde göreceğimiz gibi, $N \to \infty$ *termodinamik limiti* alındığında bir takım sadeleştirmeler olur.

Termodinamik limitte tipik olarak üç tür N bağımlılığı karşımıza çıkar:

(a) *Yoğun* nicelikler, örneğin sıcaklık T, ve genelleştirilmiş kuvvetler, basınç P, ve manyetik alan $\vec{B}$ gibi, N'den bağımsızdırlar, yani $\mathcal{O}(N^0)$.

(b) *Kapsamsal* nicelikler, örneğin enerji E, entropi S, ve genelleştirilmiş yerdeğiştirmeler, hacim V, ve mıknatıslanma $\vec{M}$ gibi, N ile doğru orantılıdırlar, yani $\mathcal{O}(N^1)$.

(c) *Üstel* bağımlılık, yani $\mathcal{O}(\exp(N\varphi))$ ise ayrık mikro durumları sayarken, veya faz uzayında uygun hacimleri hesaplarken karşımıza çıkar. Diğer asimtotik bağımlılıklar da baştan gözardı edilemez. Örneğin, sabit yoğunlukta N tane iyonun Coulomb enerjisi $Q^2/R \sim N^{5/3}$ biçiminde artar. Bu tip bağımlılıklar gündelik fizikte nadiren karşımıza çıkarlar. İyonların Coulomb etkileşmesi ters yüklü diğer iyonlarla hızla perdelenir, ve toplam enerji kapsamsal bir nicelik olur. (Astrofizik problemlerinde durum böyle değildir çünkü kütle çekim enerjisi perdelenemez. Örneğin, bir kara deliğin entropisi kütlesinin karesiyle orantılıdır.)

İstatistiksel mekanikte sıklıkla üstel değişkenlerin toplamları veya integralleriyle karşılaşırız. Aşağıda verilecek sonuçlar sayesinde bu tip toplamaların termodinamik limitte yapılması oldukça basitleşir.

(1) *Üstel niceliklerin toplanması*

Aşağıdaki toplamda

$$\mathcal{S} = \sum_{i=1}^{\mathcal{N}} \mathcal{E}_i \quad , \tag{II.50}$$

her bir terim pozitif ve N'ye üstel olarak bağımlı olsun,

$$0 \le \mathcal{E}_i \sim \mathcal{O}\big(\exp(N\phi_i)\big), \tag{II.51}$$

ve terim sayısı $\mathcal{N}$ de N'nin bir kuvvetiyle orantılı olsun. Böyle bir toplam yerine, en büyük terimi $\mathcal{E}_{max}$ şu anlamda yaklaşık olarak kullanılabilir. Toplamdaki her terim için $0 \le \mathcal{E}_i \le \mathcal{E}_{max}$ olduğundan,

$$\mathcal{E}_{\max} \leq \mathcal{S} \leq \mathcal{N}\mathcal{E}_{\max} \ . \qquad \text{(II.52)}$$

Bir yoğun nicelik $\ln S/N$ teriminden aşağıdaki biçimde sınırlanmış olarak tanımlanabilir:

$$\frac{\ln \mathcal{E}_{\max}}{N} \leq \frac{\ln \mathcal{S}}{N} \leq \frac{\ln \mathcal{E}_{\max}}{N} + \frac{\ln \mathcal{N}}{N} \ . \qquad \text{(II.53)}$$

$\mathcal{N} \propto N^p$ için $\ln \mathcal{N}/N$ terimi büyük N limitinde sıfıra gider, ve

$$\lim_{N\to\infty} \frac{\ln \mathcal{S}}{N} = \frac{\ln \mathcal{E}_{\max}}{N} = \phi_{\max} \ . \qquad \text{(II.54)}$$

(2) *Eyer Noktası İntegrali*

Benzer şekilde, şu biçimdeki bir integral

$$\mathcal{I} = \int dx \exp\left(N\phi(x)\right) \ , \qquad \text{(II.55)}$$

yaklaşık olarak integrandın maksimum değerine eşit alınabilir. Bu maksimum değer, $\phi(x)$ fonksiyonunun maksimum olduğu x_{max} noktasındadır. Bu nokta etrafında açarsak,

$$\mathcal{I} = \int dx \exp\left\{N\left[\phi(x_{\max}) - \frac{1}{2}|\phi''(x_{\max})|(x - x_{\max})^2 + \cdots\right]\right\} . \qquad \text{(II.56)}$$

Maksimum noktasında birinci türev $\phi'(x_{\max})$ sıfırdır, ikinci türev $\phi''(x_{\max})$ ise negatiftir. Seriyi ikinci derecede sonlandırdığımızda sonuç

$$\mathcal{I} \approx e^{N\phi(x_{\max})} \int dx \exp\left[-\frac{N}{2}|\phi''(x_{\max})|(x - x_{\max})^2\right] \approx \sqrt{\frac{2\pi}{N|\phi''(x_{\max})|}} e^{N\phi(x_{\max})} , \qquad \text{(II.57)}$$

olur. Burada integrali alınan bölge $[-\infty,\infty]$ aralığına genişletilmiştir. Yapılan makul bir yaklaşımdır çünkü x_{max}. noktasının komşuluğu dışından gelen katkı ihmal edilecek kadar küçüktür.

Yukarıdaki sonuca iki tür düzeltme yapılabilir. Birincisi, $\phi(x)$'in x_{max}. civarındaki açılımında daha yüksek dereceli terimler vardır. Bu düzeltme tedirgeme terimleri gibi görülebilir, ve $1/N$'nin kuvvetleri cinsinden bir seri verir. İkincisi, fonksiyonun başka $\mathcal{O}\big(\exp\{-N[\phi(x_{\max})-\phi(x'_{\max})]\}\big)$ yerel maksimumları olabilir. x'_{max} noktasındaki bir maksimum, benzer bir Gauss tipi integral getirir ve bu denklem (II.57)'ye eklenebilir. Tabi ki bu katkılar mertebesinde daha küçüktür. Tüm bu düzeltmeler termodinamik limitte sıfıra gittiğinden,

$$\lim_{N\to\infty}\frac{\ln\mathcal{I}}{N}=\lim_{N\to\infty}\left[\phi(x_{\max})-\frac{1}{2N}\ln\left(\frac{N|\phi''(x_{\max})|}{2\pi}\right)+\mathcal{O}\left(\frac{1}{N^2}\right)\right]=\phi(x_{\max}) \quad . \qquad \text{(II.58)}$$

İntegrallerin alınmasında *eyer noktası* metodu, yukarıdaki sonucun daha genel integrandlara ve integral yollarının kompleks düzleme genelleştirilmesidir (Kompleks düzlemde uygun ekstremum bir eyer noktasıdır.). Yukarıda sunulan basitleştirilmiş biçimi bu dersin amaçları için yeterlidir.

- N!'in büyük N değerleri için *Stirling yaklaşıklığı* ifadesi, eyer noktası integraliyle elde edilebilir. N!'i bir integral olarak temsil etmek için şu sonuçtan başlayalım,

$$\int_0^\infty dx e^{-\alpha x}=\frac{1}{\alpha}. \qquad \text{(II.59)}$$

Her iki tarafın α'ya göre üstüste türevlerini aldığımızda,

$$\int_0^\infty dx x^N e^{-\alpha x}=\frac{N!}{\alpha^{N+1}}. \qquad \text{(II.60)}$$

Yukarıdaki sonuç sadece tamsayı N için geçerli olsa da, analitik sürekliliği kullanarak her N için şöyle bir fonksiyon tanımlamak mümkündür:

$$\Gamma(N+1)\equiv N!=\int_0^\infty dx x^N e^{-x}, \qquad \text{(II.61)}$$

Denklem (II.61)'deki integral tam olarak denklem (II.55) biçiminde olmasa da, yine benzer bir yöntemle alınabilir. İntegrand, $\phi(x)=\ln x-x/N$ olmak üzere $\exp(N\phi(x))$ şeklinde yazılabilir. Üstel fonksiyonun $x_{max}=N$ noktasında bir maksimumu vardır, bu noktada $\phi(x_{max})=\ln N-1$, ve $\phi''(x_{max})=-1/N^2$. Denklem (II.61)'deki integrandı bu nokta etrafında açarak, ve integrali $[-\infty,\infty]$ aralığında alarak,

$$N! \approx \int dx \exp\left(N\ln N - N - \frac{1}{2N}(x-N)^2\right) \approx N^N e^{-N}\sqrt{2\pi N}, \qquad \text{(II.62)}$$

bulunur. Stirling formülü denklem (II.62)'in logaritması alınarak elde edilir:

$$\ln N! = N\ln N - N + \frac{1}{2}\ln(2\pi N) + \mathcal{O}(\frac{1}{N}). \qquad \text{(II.63)}$$

II.G Bilgi, Entropi, ve Tahmin

- *Bilgi:* Bir rassal değişken ele alalım $i = 1,...,M$ olmak üzere, ayrık sonuçlar kümesi $S = \{x_i\}$ ve gerçekleşme olasılıkları $\{p(i)\}$ olsun. Bilgi teorisi bağlamında, bir olasılık dağılımının *bilgi içeriği*nin belli bir anlamı vardır: Rassal değişkenin N bağımsız sonucundan bir mesaj oluşturalım. Bu mesajdaki her karakter için M farklı olanak olduğundan, mesajın $N \ln_2 M$ bitlik görünür bilgi içeriği vardır; yani bu miktarda ikili bitlik bilgi, mesajı tam olarak iletmek için aktarılmalıdır. Öte yandan, $\{p(\mathrm{i})\}$ olasılıkları olası mesajların tipini sınırlar. Örneğin, eğer $p_2 \gg p_1$ ise, içinde x_2'den çok x_1 içeren bir mesaj oluşturmak çok olası değildir. Özellikle, büyük N limitinde, mesajın her sembolden "kabaca" $\{N_i = Np_i\}$ tane içermesini bekleriz.† Dolayısıyla, tipik mesajların sayısı, $\{N_i\}$ kez kullanılan $\{x_i\}$'nin kaç değişik biçimde düzenlenebileceğinin sayısına tekabül eder, ve şu multinom katsayısıyla verilir,

$$g = \frac{N!}{\prod_{i=1}^{M} N_i!}. \qquad \text{(II.64)}$$

Bu, mesajların toplam sayısı M^n'den çok daha küçüktür. Olası *g* dizilimden bir tanesini belirtmek için

$$\ln_2 g \approx -N\sum_{i=1}^{M} p_i \ln_2 p_i \qquad (\text{for } N \to \infty), \qquad \text{(II.65)}$$

kadar bitlik bilgi gerekir. Son sonuca ulaşırken $\ln N!$ için Stirling yaklaşıklığı

† Daha açıkçası, $N \to \infty$ limitinde, $Np_i \pm \sqrt{N}$ aralığının dışında bir N_i bulma olasılığı, N'nin üstel fonksiyonu kadar küçüktür.

kullanılmıştır. Aynı sonuca aşağıdaki eşitlikler kullanılarak da ulaşılabilir:

$$1 = \left(\sum_i p_i\right)^N = \sum_{\{N_i\}} N! \prod_{i=1}^{M} \frac{p_i^{N_i}}{N_i!} \approx g \prod_{i=1}^{M} p_i^{N p_i}, \tag{II.66}$$

son aşamada, önceki bölümde açıklandığı gibi toplam yerine en büyük terimi alınmıştır.

Shannon Teoremi, N denemede hataların oranının, $N \to \infty$ limitinde sıfıra gitmesini sağlamak için gereken minimum bit sayısının $\ln_2 g$ olduğunu daha kesin şekilde ispatlar. Her tekdüze olmayan dağılım için, bu değer, bağıl olasılıklar üzerine hiç bilgi olmadığında gereken bit sayısı $N\ln_2 M$'den daha azdır. Dolayısıyla, aradaki farkın deneme başına düşen miktarı olasılık dağılımının bilgi içeriği olarak atfedilir, ve şöyle yazılır:

$$I\left[\{p_i\}\right] = \ln_2 M + \sum_{i=1}^{M} p_i \ln_2 p_i \quad . \tag{II.67}$$

• *Entropi:* Denklem (II.64) istatistiksel mekanikte, M farklı bileşenin karışımı bağlamında sıklıkla karşımıza çıkar; doğal logaritması *karışma entropisi* ile ilişkilidir. Daha genel olarak, *her olasılık dağılımı için bir entropi* tanımlayabiliriz:

$$S = -\sum_{i=1}^{M} p(i) \ln p(i) = -\left\langle \ln p(i) \right\rangle \quad . \tag{II.68}$$

Bu entropi, minimum değeri olan sıfırı, delta fonksiyonu dağılımı $p(i) = \delta_{i,j}$ için, bir maksimum olarak $\ln M$ değerini ise tekdüze $p(i) = 1/M$ dağılımı için alır. Dolayısıyla, S dağılımın yaygınlığının (düzensizliğinin) bir ölçüsüdür, ve rassal değişkenler $\{x_i\}$'lerin değerlerine bağlı değildir. $f_i = F(x_i)$ biçiminde bire-bir eşleme entropiyi değiştirmezken, bire-çok eşleme dağılımı daha düzenli yapacağından S azalır. Örneğin, iki değer x_1 ve x_2 aynı f_i'ye eşlendiğinde entropideki değişim şöyledir:

$$\Delta S(x_1, x_2 \to f) = \left[p_1 \ln \frac{p_1}{p_1 + p_2} + p_2 \ln \frac{p_2}{p_1 + p_2}\right] < 0. \tag{II.69}$$

• *Tahmin:* Entropi S, aynı zamanda olasılıkların öznel tahminini sayıya dökmekte kullanılabilir. Hiç bilgi olmadığında, en iyi önyargısız tahmin M tane sonucun eşit derecede olası olmasıdır. Bu, entropisi maksimum olan dağılımdır. Eğer ek bilgi varsa, önyargısız tahmin, bu bilginin getirdiği kısıtlar altında entropiyi maksimize ederek elde edilir. Örneğin, eğer $\langle F(x) \rangle = f$ olduğu biliniyorsa, şu ifadeyi maksimize edebiliriz:

$$S(\alpha, \beta, \{p_i\}) = -\sum_i p(i) \ln p(i) - \alpha \left(\sum_i p(i) - 1 \right) - \beta \left(\sum_i p(i) F(x_i) - f \right), \quad \text{(II.70)}$$

burada Lagrange çarpanları α ve β sırasıyla, normalleşme ve $\langle F(x) \rangle = f$ kısıtlarını hesaba katmak üzere alınmıştır. Optimizasyonun sonucu bir $p_{\mathrm{i}} \propto \exp(-\beta F(x_i))$, dağılımı olup β'nın değeri kısıt tarafından belirlenir. Bu işlem keyfi sayıda kısıt için genelleştirilebilir. Kolayca görülebilir ki, eğer bir dağılımın ilk n momenti (ve dolayısıyla n kümülantı) belirlenirse, önyargısız tahmin $n$inci derece bir polinomun üstel fonksiyonudur

Denklem (II.68)'in bir benzeri olarak, sürekli bir rassal değişken ($S_x = \{-\infty < x < \infty\}$) için bir entropi tanımlayabiliriz:

$$S = -\int dx\, p(x) \ \ln p(x) = -\langle \ln p(x) \rangle \quad . \quad \text{(II.71)}$$

Ancak, bu tanımlamada bazı problemler vardır, S'nin birebir eşleme altında değişmez olmaması gibi ($f = F(x)$ değişken dönüşümü sonrasında entropi $\langle |F'(x)| \rangle$ kadar değişir) Bir kanonik dönüşümün Jacobiyeni bire eşit olduğundan, kanonik eşlenik çiftler, klasik istatistiksel mekanikte koordinatlar için uygun bir seçenek sunarlar. Sürekli değişken ayrıklaştırılırsa belirsizlikler giderilir. Bu kuantum istatistiksel mekaniğinde oldukça doğal olur, çünkü orada durumların bir ayrık merdiveniyle çalışmak genellikle mümkündür. Faz uzayının ayrıklaştırılması için uygun hacim Planck sabiti $\hbar$ tarafından belirlenir.