MIT Açık Ders Malzemeleri
http://ocw.mit.edu

# 8.333 İstatistiksel Mekanik I: Parçacıkların İstatistiksel Mekaniği

2007 Güz

# II. Olasılık

## II.A Genel Tanımlar

Termodinamiğin yasaları, makroskobik cisimler hakkındaki gözlemlere dayanır, ve onların termal özelliklerini kapsar. Öte yandan, madde, hareketleri klasik veya kuantum mekaniğinin temel yasalarıyla yönetilen atom ve moleküllerden oluşur. İlkesel olarak, makroskobik bir cismin davranışını, bileşenleri hakkındaki bilgiden türetebilmek mümkün olmalı. Bu, önümüzdeki derslerde kinetik teori tarafından ele alınacak problemdir. Aslında, aşırı çok sayıda parçacığın tüm dinamiğini tanımlamak hayli göz korkutucu bir iştir. Göstereceğimiz gibi, bir makroskobik sistemin denge özelliklerini tartışmak için, bileşimindeki parçacıkların davranışını tamamiyle bilmek gerekmez. Gereken tek şey parçacıkların belli bir mikroskobik durumda bulunma olasılığıdır. İstatistiksel mekanik, dolayısıyla, sistemin özünde olasılıkçı bir tarifidir. Olasılıklarla çalışma aşinalığı, bu yüzden istatistiksel mekaniğin önemli önkoşullarından biridir. Buradaki amacımız, olasılık teorisinde bazı önemli sonuçları gözden geçirmek, ve dersin geri kalanında kullanılacak gösterimleri tanıtmaktır.

İnceleyeceğimiz kavram, mümkün sonuçları $S \equiv \{x_1, x_2, \ldots\}$ olan, $x$ *rassal değişken*idir. Değişkenin sonuçları, bozuk para atımında $S_{\text{para}} = \{\text{tura, yazı}\}$, veya zar atımında $S_{\text{zar}} = \{1,2,3,4,5,6\}$ olduğu gibi *ayrık*, veya bir gaz parçacığının hızında $S_{\vec{v}} = \{-\infty < v_x, v_y, v_z < \infty\}$ veya sıfır sıcaklıkta bir metalin elektronunun enerjisinde $S_\epsilon = \{0 \leq \epsilon \leq \epsilon_F\}$ olduğu gibi *sürekli* olabilir. Bir *olay* $E$, sonuçların herhangi bir alt kümesine verilen addır $(E \subset S)$, ve bir $p(E)$ olasılığı tanımlıdır; örneğin $p_{\text{zar}}(\{1\})=1/6$, veya $p_{\text{zar}}(\{1,3\}) = 1/3$. Aksiyomatik bir bakış açısıyla, olasılıkların aşağıdaki koşulları sağlaması gerekir:

(i) *Pozitiflik:* $p(E) \geq 0$, yani, olasılık negatif olmamalıdır.

(ii) *Toplanabilirlik:* $p(A$ veya $B) = p(A) + p(B)$, eğer *A* ve *B* bağımsız olaylarsa.

(iii) *Normalizasyon:* $p(S)=1$, yani rassal değişkenin $S$'nin içinden bir sonucu olmalıdır.

Kullanım açısından, çeşitli sonuçlara olasılık değerlerini nasıl atayacağımızı bilmek isteriz. Burada iki yaklaşım mümkündür:

(1) Nesnel olasılıklar, *deneysel* olarak, rassal değişkenin pek çok denemesinde bir sonucun hangi sıklıkla ortaya çıktığından elde edilir. Eğer rassal işlem $N$ kere tekrar edilirse, ve $A$ olayı $N_A$ kere ortaya çıkarsa olasılık:

$$p(A) = \lim_{N\to\infty} \frac{N_A}{N}.$$

Örneğin, bir zar $N$ = 100, 200, 300 kez atıldığında, sonucun 1 geldiği durumların sayısı $N_1$ = 19, 30, 48 olabilir. Elde edilen oranlar, .19, .15, .16, $p_{\text{zar}}(\{1\})$ için gittikçe daha güvenilir bir tahmin verir.

(2) *Öznel* olasılıklar, gelecek sonuç hakkında kesin bilgi eksikliğiyle ilgili belirsizliklere dayanarak bir teorik tahmin sağlar. Örneğin, $p_{\text{zar}}$ ({1}) =1/6, değerlendirmesi, zar atışında altı farklı sonucun mümkün olduğu ve zarın hileli olduğuna dair bir şüphemiz yoksa her sonucun eşit derecede muhtemel olduğu bilgisine dayanır. İstatistiksel Mekanikte tüm olasılık değerlendirmeleri öznel tabanlıdır. Olasılıkların bu öznel değerlendirmesinin sonuçları ölçümlerle sınanmalı, ve sonuçlar hakkında daha fazla bilgi elde edildiğinde gerekirse yeniden değerlendirilmelidir.

## II.B Bir Rassal Değişken

Ayrık rassal değişkenin özellikleri iyi bilindiğinden, biz burada konumuzla daha ilgili olan sürekli rassal değişkenlere odaklanacağız. Sonuçları gerçel sayılar olan bir $x$ rassal değişkenini ele alalım, yani $S_x = \{-\infty < x < \infty\}$.

- *Birikimli olasılık fonksiyonu* (BOF) $P(x)$, sonucun $x$'den küçük *herhangi* bir değer olma olasılığıdır, yani $P(x) = \text{prob.}(E \subset [-\infty, x])$. $P(x)$ $x$'in tekdüze artan, ve sınır değerleri $P(-\infty) = 0$ ve $P(+\infty) = 1$ olan bir fonksiyonu olmalıdır.
- *Olasılık yoğunluğu fonksiyonu* (OYF) $p(x) \equiv dP(x)/dx$ biçiminde tanımlanır. Bu nedenle, $p(x)dx = \text{prob.}(E \subset [x, x+dx])$. Bir olasılık yoğunluğu olduğundan *pozitiftir*, ve şu şekilde normalize edilir:

$$\text{prob.}(\mathcal{S}) = \int_{-\infty}^{\infty} dx\, p(x) = 1 \ . \tag{II.1}$$

Ayrıca, $p(x)$ *olasılık yoğunluğu* olduğu için üst sınırı yoktur, yani $0 < p(x) < \infty$.

- Rassal değişkenin herhangi bir $F(x)$ fonksiyonunun *beklenen değeri* şöyle tanımlanır:

$$\langle F(x)\rangle = \int_{-\infty}^{\infty} dx \; p(x)F(x) \; . \tag{II.2}$$

$F(x)$ fonksiyonunun kendisi de, OYF'si $p_F(f)df$ = prob.$(F(x) \subset [f,f + df])$ olan, rassal bir değişkendir. $F(x)=f$ denkleminin birden fazla $x_i$ çözümü olabilir,

$$p_F(f)df = \sum_i p(x_i)dx_i \, , \quad \Longrightarrow \quad p_F(f) = \sum_i p(x_i) \left|\frac{dx}{dF}\right|_{x=x_i} \quad . \tag{II.3}$$

$|dx/dF|$ çarpanları, $x$'den $F$'ye değişken dönüşümünün *Jacobiyen*'leridir.. Örneğin, $p(x) = \lambda \exp(-\lambda|x|)/2$ ve $F(x) = x^2$ fonksiyonunu alalım. $F(x)=f$, eşitliğinin iki çözümü vardır: $x_{\pm} = \pm\sqrt{f}$, Jacobiyenleri ise $\left|\pm f^{-1/2}/2\right|$'dir. Dolayısıyla $f > 0$ olduğunda,

$$P_F(f) = \frac{\lambda}{2}\exp\left(-\lambda\sqrt{f}\right)\left(\left|\frac{1}{2\sqrt{f}}\right| + \left|\frac{-1}{2\sqrt{f}}\right|\right) = \frac{\lambda \exp\left(-\lambda\sqrt{f}\right)}{2\sqrt{f}} \; ,$$

ve $f < 0$ için $p_F(f) = 0$. Dikkat edilirse $f = 0$ noktasında $p_F(f)$ (integrali alınabilir) bir ıraksaklık sahibidir.

- OYF'nin *moment*leri, rassal değişkenin kuvvetlerinin beklenen değerleridir. $n$'inci moment şöyle yazılır:

$$m_n \equiv \langle x^n \rangle = \int dx p(x) \; x^n. \tag{II.4}$$

- *Karakteristik fonksiyon*, dağılımın momentlerini elde etmemizi sağlar. Açıkça, OYF'nin Fourier dönüşümü olup şu şekilde tanımlanır:

$$\tilde{p}(k) = \left\langle e^{-ikx}\right\rangle = \int dx p(x) \; e^{-ikx}. \tag{II.5}$$

OYF, karakteristik fonksiyonun ters Fourier dönüşümüyle tekrar elde edilebilir:

$$p(x) = \frac{1}{2\pi}\int dk \tilde{p}(k) \; e^{+ikx}. \tag{II.6}$$

Dağılımın momentleri, $\tilde{p}(k)$ fonksiyonunu $k$'nın kuvvetleri cinsinden açarak bulunur:

$$\tilde{p}(k) = \left\langle \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} x^n \right\rangle = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle \quad . \tag{II.7}$$

Herhangi bir $x_0$ noktası etrafında OYF'nin momentleri de şu seri açılımıyla bulunabilir:

$$e^{ikx_0}\tilde{p}(k) = \left\langle e^{-ik(x-x_0)} \right\rangle = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle (x-x_0)^n \rangle \quad . \tag{II.8}$$

- *Kümülant üreten fonksiyon*, karakteristik fonksiyonun logaritmasıdır. Seri açılımı dağılımın kümülantlarını şu şekilde verir:

$$\ln \tilde{p}(k) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c \quad . \tag{II.9}$$

Momentler ve kümülantlar arasındaki bağıntılar, denklem (II.7)'de verilen $\tilde{p}(k)$ 'nın logaritmasını açarak ve aşağıdaki eşitliği kullanarak bulunabilir,

$$\ln(1+\epsilon) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\epsilon^n}{n} \quad . \tag{II.10}$$

İlk dört kümülant, dağılımın, sırasıyla *ortalama*, *varyans*, "*skewness*", ve "*curtosis*" değerleri olarak adlandırılır, ve momentlerle şu şekilde ifade edilir:

$$\begin{aligned} \langle x \rangle_c &= \langle x \rangle , \\ \langle x^2 \rangle_c &= \langle x^2 \rangle - \langle x \rangle^2 , \\ \langle x^3 \rangle_c &= \langle x^3 \rangle - 3 \langle x^2 \rangle \langle x \rangle + 2 \langle x \rangle^3 , \\ \langle x^4 \rangle_c &= \langle x^4 \rangle - 4 \langle x^3 \rangle \langle x \rangle - 3 \langle x^2 \rangle^2 + 12 \langle x^2 \rangle \langle x \rangle^2 - 6 \langle x \rangle^4 . \end{aligned} \tag{II.11}$$

Kümülantlar, bir OYF'yi tanımlamanın kullanışlı ve özlü bir yolunu sunarlar.

Önemli bir teorem momentlerin kümülantlar cinsinden kolayca hesaplanmasını sağlar: $n$inci kümülantı, grafiksel olarak $n$ tane noktası bulunan, bağlantılı bir öbek olarak temsil edelim. Burada $m$inci moment, $m$ noktanın tüm olanaklı altbölümlerini, daha küçük (bağlantılı veya bağlantısız) öbeklere toplayarak elde edilir. Her altbölümün toplama katkısı, temsil ettiği bağlantılı kümülantların çarpımıdır. Bu sonucu kullanarak ilk dört moment kümülantlar cinsinden şöyle hesaplanır:

$$
\begin{aligned}
\langle x\rangle &= \langle x\rangle_c\,, \\
\langle x^2\rangle &= \langle x^2\rangle_c + \langle x\rangle_c^2\,, \\
\langle x^3\rangle &= \langle x^3\rangle_c + 3\langle x^2\rangle_c\langle x\rangle_c + \langle x\rangle_c^3\,, \\
\langle x^4\rangle &= \langle x^4\rangle_c + 4\langle x^3\rangle_c\langle x\rangle_c + 3\langle x^2\rangle_c^2 + 6\langle x^2\rangle_c\langle x\rangle_c^2 + \langle x\rangle_c^4\,.
\end{aligned}
\qquad (\mathrm{II}.12)
$$

İstatistiksel fizik ve alan teorisinde çeşitli diyagramatik hesaplama yöntemlerinin başlangıç noktası olan bu teorem, denklem (II.7) ve (II.9)'de $\tilde{p}(k)$ için verilen ifadelerin eşitlenmesiyle basitçe ispat edilir

$$\sum_{m=0}^{\infty}\frac{(-ik)^m}{m!}\langle x^m\rangle = \exp\left[\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}\langle x^n\rangle_c\right] = \prod_n\sum_{p_n}\left[\frac{(-ik)^{np_n}}{p_n!}\left(\frac{\langle x^n\rangle_c}{n!}\right)^{p_n}\right]. \qquad (\mathrm{II}.13)$$

İfadenin iki tarafındaki $(-ik)^m$ terimlerinin katsayıları eşitlenirse

$$\langle x^m\rangle = \sum_{\{p_n\}}'\prod_n\frac{m!}{p_n!(n!)^{p_n}}\langle x^n\rangle_c^{p_n}. \qquad (\mathrm{II}.14)$$

Toplam üzerinde bulunan $\sum np_n = m$kısıtı, yukarıda verilen grafiksel yoruma karşılık gelir: sayısal çarpan $m$ tane noktayı $n$ noktalı $\{p_n\}$ öbeklerine bölme yollarının sayısıdır.

## II.C Bazı Önemli Olasılık Dağılımları

Bu bölümde, sıklıkla karşılaşılan üç olasılık dağılımının özellikleri incelenecektir.

**(1)** *Normal (Gauss tipi) dağılım* OYF'si aşağıdaki gibi olan, sürekli, gerçel, rassal değişken $x$'i tanımlar:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\lambda)^2}{2\sigma^2}\right] \quad . \qquad (\mathrm{II}.15)$$

Buna karşılık gelen karakteristik fonksiyon da Gaussiyen biçimindedir:

$$\tilde{p}(k) = \int_{-\infty}^{\infty}dx\,\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\lambda)^2}{2\sigma^2} - ikx\right] = \exp\left[-ik\lambda - \frac{k^2\sigma^2}{2}\right] \quad . \qquad (\mathrm{II}.16)$$

Dağılımın kümülantları $\ln\tilde{p}(k) = -ik\lambda - k^2\sigma^2/2$, ifadesinde denklem (II.9) kullanılarak bulunabilir:

$$\langle x\rangle_c = \lambda \quad , \quad \langle x^2\rangle_c = \sigma^2 \quad , \quad \langle x^3\rangle_c = \langle x^4\rangle_c = \cdots = 0 \quad . \tag{II.17}$$

Dolayısıyla, normal dağılım ilk iki kümülantıyla tamamıyla tanımlanır. Bu durum, denklem .(II.12)'de verilen öbek açılımına dayalı momentlerin hesabını oldukça basitleştirir:

$$\begin{aligned}\langle x\rangle &= \lambda\,,\\ \langle x^2\rangle &= \sigma^2+\lambda^2,\\ \langle x^3\rangle &= 3\sigma^2\lambda+\lambda^3,\\ \langle x^4\rangle &= 3\sigma^4+6\sigma^2\lambda^2+\lambda^4, \quad \cdots \quad .\end{aligned} \tag{II.18}$$

Normal dağılım, alan teorisindeki tedirgemeli hesapların çoğunda başlangıç noktası olarak kullanılır. Yüksek dereceli kümülantların sıfır olması, tüm grafiksel hesapların bir noktalı ve iki noktalı öbeklerin (propagatör olarak bilinir) çarpımını içermesini gerektirir.

**(2)** *Binom dağılımı:* Rassal bir değişkenin olası iki sonucu $A$ ve $B$, bunların göreli olasılıkları $p_A$ ve $p_B = 1 - p_A$ olsun (örneğin bozuk para atışı). $N$ denemede $A$ olayının $N_A$ kez olma olasılığı (örneğin 12 denemede 5 kez tura gelmesi) binom dağılımıyla verilir

$$p_N(N_A) = \binom{N}{N_A} p_A^{N_A} p_B^{N-N_A} \ . \tag{II.19}$$

Denklemdeki çarpan,

$$\binom{N}{N_A} = \frac{N!}{N_A!(N-N_A)!}\ , \tag{II.20}$$

$(p_A+p_B)^N$ ifadesinin binom açılımındaki katsayılar olup, $N_A$ tane $A$ ve $N_B = N - N_A$ tane $B$ olayının olası dizilimlerinin sayısını verir. Bu ayrık dağılım için karakteristik fonksiyon şöyle yazılır:

$$\tilde{p}_N(k) = \langle e^{-ikN_A}\rangle = \sum_{N_A=0}^{N} \frac{N!}{N_A!(N-N_A)!} p_A^{N_A} p_B^{N-N_A}\, e^{-ikN_A} = \left(p_A e^{-ik} + p_B\right)^N \quad . \tag{II.21}$$

Bunun sonucu kümülant üreten fonksiyon

$$\ln \tilde{p}_N(k) = N \ln \left( p_A e^{-ik} + p_B \right) = N \ln \tilde{p}_1(k), \tag{II.22}$$

olup, ln $\tilde{p}_1(k)$ ise tek bir deneme için kümülant üreten fonksiyondur. Dolayısıyla, $N$ denemeden sonraki kümülantlar basitçe tek bir denemedeki kümülantların $N$ katıdır. Her bir denemede $N_A$ için izin verilen değerler $p_B$ olasılığıyla 0 ve $p_A$ olasılığıyla 1 olduğundan, her $\ell$ için $\langle N_A^\ell \rangle = p_A$, elde edilir. $N$ deneme sonrasında ilk iki kümülant şöyle olur:

$$\langle N_A \rangle_c = N p_A \quad , \quad \langle N_A^2 \rangle_c = N(p_A - p_A^2) = N p_A p_B \quad . \tag{II.23}$$

Ortalama civarındaki dalgalanmaların bir ölçüsü, varyansın karekökü olan *standart sapma* ile verilir. Binom dağılımının ortalaması $N$ ile ölçeklenirken, standart sapması ancak $\sqrt{N}$ ile büyür. Dolayısıyla, bağıl belirsizlik büyük $N$ için daha küçük olur.

Binom dağılımı, bir takım $\{A,B,...,M\}$ sonuçları $\{p_A,p_B,...,p_M\}$ olasılıklarıyla ortaya çıkıyorsa, basitçe *multinom* dağılımına genelleştirilir. $\{N_A,N_B,...,N_M\}$ sonuçlarını $N = N_A+N_B +N_M$ denemede bulma olasılığı şudur:

$$p_N \left( \{N_A, N_B, \cdots, N_M\} \right) = \frac{N!}{N_A! N_B! \cdots N_M!} p_A^{N_A} p_B^{N_B} \cdots p_M^{N_M} \quad . \tag{II.24}$$

**(3)** *Poisson dağılımı:* Poisson işleyişinin klasik örneği radyoaktif bozunumdur. Bir parça radyoaktif malzemenin $T$ zaman aralığı boyunca gözlenmesi şunları gösterir:

(a) Bir ve sadece bir olayın (bozunum) $[t,t+dt]$ aralığında gerçekleşme olasılığı, $dt$ sıfıra giderken $dt$ ile orantılıdır.

(b) Farklı zaman aralıklarındaki olayların olasılıkları birbirinden bağımsızdır.

$T$ aralığında tam olarak $M$ tane bozunum gözlemleme olasılığı Poisson dağılımıyla verilir. Bu dağılım, $T$ aralığını, büyüklüğü $dt$ olan $N = T/dt \gg 1$ tane dilime bölerek binom dağılımının bir limiti olarak elde edilir. Herbir dilimde $p= \alpha dt$ olasılıkla bir olay olurken, $q = 1-\alpha dt$ olasılıkla hiç olay gerçekleşmez. *dt* aralığında birden fazla olay olması olasılığı aşırı küçük olduğundan işleyiş binom dağılımına denktir. Denklem (II.21) kullanılarak, karakteristik fonksiyon şöyle yazılır:

$$\tilde{p}(k) = \left(pe^{-ik} + q\right)^n = \lim_{dt\to 0}\left[1 + \alpha dt\left(e^{-ik} - 1\right)\right]^{T/dt} = \exp\left[\alpha(e^{-ik} - 1)T\right] \quad . \qquad \text{(II.25)}$$

Poisson OYF denklem (II.6)'daki ters Fourier dönüşümü ve üstel fonksiyonun kuvvet serisi açılımı kullanılarak,

$$p(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp\left[\alpha\left(e^{-ik} - 1\right)T + ikx\right] = e^{-\alpha T} \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx} \sum_{M=0}^{\infty} \frac{(\alpha T)^M}{M!} e^{-ikM} \ , \qquad \text{(II.26)}$$

elde edilir. $k$ integrali alındığında

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ik(x-M)} = \delta(x - M) \ , \qquad \text{(II.27)}$$

olduğundan, sonuç:

$$p_{\alpha T}(x) = \sum_{m=0}^{\infty} e^{-\alpha T} \ \frac{(\alpha T)^M}{M!} \delta(x - M) \ . \qquad \text{(II.28)}$$

Dolayısıyla $M$ tane olayın olasılığı, $p_{\alpha T}(M) = \mathrm{e}^{-\alpha T}(\alpha T)^M/M!$. Dağılımın kümülantları şu ifadeden elde edilir:

$$\ln \tilde{p}_{\alpha T}(k) = \alpha T(e^{-ik} - 1) = \alpha T \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!}, \quad \Longrightarrow \quad \langle M^n \rangle_c = \alpha T \quad . \qquad \text{(II.29)}$$

Tüm kümülantlar aynı değerde olup, momentler denklem (II.12) kullanılarak bulunur:

$$\langle M \rangle = (\alpha T), \quad \langle M^2 \rangle = (\alpha T)^2 + (\alpha T), \quad \langle M^3 \rangle = (\alpha T)^3 + 3(\alpha T)^2 + (\alpha T). \qquad \text{(II.30)}$$

**Örnek:** Galaksimizdeki yıldızların $n$ yoğunluğunda rasgele dağıldıklarını varsayarsak (açıkçası gerekçesiz bir varsayım) en yakın yıldızın $R$ uzaklıkta olma olasılığı nedir? Küçük bir $dV$ hacminde bir yıldız bulma olasılığı $ndV$, ve bu olasılıklar birbirinden bağımsız olduğundan, bir $V$ hacmindeki yıldızların sayısı, denklem (II.28)'de verildiği gibi, $\alpha = n$ olan bir Poisson dağılımına sahiptir. İlk yıldızla $R$ uzaklıkta karşılaşma olasılığı $p(R)$, merkezin etrafındaki $V = 4\pi R^3/3$ hacminde sıfır yıldız bulma olasılığı $p_{nV}(0)$ ile $R$ yarıçapında $dV = 4\pi R^2 dR$ hacme sahip kabukta bir yıldız bulma olasılığı

$p_{ndV}$(1)'in çarpımına eşittir. Her iki olasılık da denklem (II.28) kullanılarak hesaplanabilir, ve dolayısıyla

$$\begin{aligned} p(R)dR = p_{nV}(0)\, p_{ndV}(1) =& e^{-4\pi R^3 n/3}\, e^{-4\pi R^2 n dR}\, 4\pi R^2 n dR, \\ \Longrightarrow \qquad p(R) =& 4\pi R^2 n \exp\left(-\frac{4\pi}{3} R^3 n\right) . \end{aligned} \tag{II.31}$$