Notes on

Statistical Mechanics

K. P. N. Murthy,

Adjunct Faculty, School of Basic Sciences

Indian Institute of Technology Bhubaneswar,

Bhubaneswar 752 050, Odisha, INDIA

January 1 - April 30, 2017

3

Binomial, Poisson, and Gaussian Distrubutions

3.1 Binomial Distribution

Consider a system consisting of one coin. It has two micro states : H and T. The probability for the system to be in micro state H is p and that in micro state T is $q = 1 - p$.

Consider the case with $p = 0.6$ and hence $q = 1 - p = 0.4$. A possible Maxwell ensemble of micro states is

$$\{T,\ H,\ H,\ H,\ T,\ H,\ H,\ T,\ H, T\ \}.$$

Notice that the ensemble contains ten elements. Six elements are H and four are T. This is consistent with the given probabilities: $\mathcal{P}(H) = 6/10$; and $\mathcal{P}(T) = 4/10$.

However a Gibbs ensemble is constructed by actually tossing N identical and independent coins. In the limiet $N \to \infty$, sixty percent of the coins shall be in micro state H and forty, T. To ensure this we need to take the size of the ensemble N, to be very large. How large ? You will get an answer to this question in what follows.

Let us say, we attempt to construct the ensemble by actually carrying out the experiment of tossing identical coins or by tossing the same coin several times independently. What is the probability that in the experiment there shall be n_1 'Heads' and hence $n_2 = (N - n)$ 'Tails' ? Let us denote this by the symbol $B(n_1, n_2; N)$, where N is the number of independent identical coins tossed or number of times a coin is tossed independently. It is readily seen,

$$B(n_1, n_2; N) = \frac{N!}{n_1!\ n_2!}\ p^{n_1}\ q^{n_2}\ ; \quad n_1 + n_2 = N. \tag{3.1}$$

$B(n_1, n_2; N)$ is called the Binomial distribution. Let $n_1 = n$, $n_2 = N - n$, we can write the Binomial distribution for the single random variable n as,

$$B(n; N) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

Figure (3.1) depicts Binomial distribution for $N = 10$, $p = 0.5$ (Left) and 0.35 (Right). **First moment of n** : What is average value of n ? The average, also called the mean, the first moment, the expectation value *etc.* is denoted by the symbol $\langle n \rangle$ and is given by,

$$\begin{aligned}
\langle n \rangle = \sum_{n=0}^{N} n\ B(n; N) &= \sum_{n=1}^{N} n\ \frac{N!}{n!\ (N-n)!}\ p^n\ q^{N-n}, \\
&= Np \sum_{n=1}^{N} \frac{(N-1)!}{(n-1)!\ [N-1-(n-1)]!}\ p^{n-1}\ q^{N-1-(n-1)}, \\
&= Np \sum_{n=0}^{N-1} \frac{(N-1)!}{n!(N-1-n)!}\ p^n\ q^{N-1-n}, \\
&= Np(p+q)^{N-1} = Np.
\end{aligned}$$

Second factorial moment of n : The second factorial moment of n

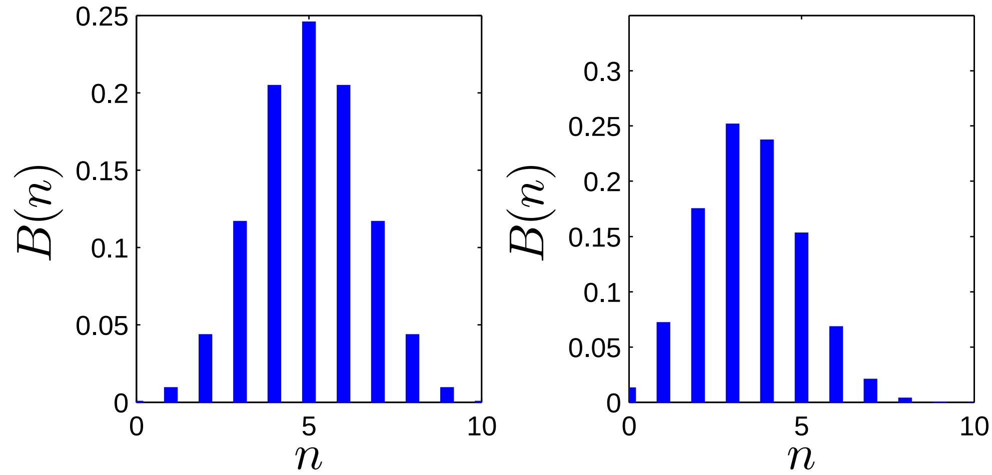

Figure 3.1: Binomial distribution : $B(n;N) = \frac{N!}{n!(N-n)!}p^n(1-p)^{N-n}$ with $N = 10$; $B(n;N)$ *versus* n; depicted as sticks; (Left) $p = 0.5$; (Right) $p = .35$.

is defined as $\langle n(n-1)\rangle$. It is calculated as follows.

$$
\begin{aligned}
\langle n(n-1)\rangle &= \sum_{n=0}^{N} n(n-1)B(n;N), \\
&= \sum_{n=2}^{N} n(n-1)\frac{N!}{n!(N-n)!}p^n q^{N-n}, \\
&= N(N-1)p^2 \sum_{n=2}^{N} \frac{(N-2)!}{(n-2)![(N-2)-(n-2)!}p^{n-2}q^{(N-2)-(n-2)}, \\
&= N(N-1)p^2 \sum_{n=0}^{N-2} \frac{(N-2)!}{n![(N-2)-n]!}p^n q^{(N-2)-n}. \\
&= N(N-1)p^2(q+p)^{N-2} = N(N-1)p^2.
\end{aligned}
$$

Moments of n : We can define higher moments. The k-th moment is defined as

$$
M_k = \langle n^k\rangle \;=\; \sum_{n=0}^{N} n^k \; B(n). \tag{3.2}
$$

Variance of n : An important property of the random variable is variance.

It is defined as,

$$\sigma_n^2 = \sum_{n=0}^{N}(n-M_1)^2 \; B(n), = \sum_{n=0}^{N} n^2 B(n) - M_1^2, = M_2 - M_1^2. \tag{3.3}$$

We have,

$$\begin{aligned} \langle n(n-1)\rangle &= N(N-1)p^2, \\ \langle n^2\rangle - \langle n\rangle &= N^2p^2 - Np^2, \\ \langle n^2\rangle &= N^2p^2 - Np^2 + Np, \\ \sigma_n^2 = \langle n^2\rangle - \langle n\rangle^2 &= Npq. \end{aligned} \tag{3.4}$$

The square-root of variance is called the standard deviation. A relevant quantity is the relative standard deviation. It is given by the ratio of the standard deviation to the mean. For the Binomial random variable, we have,

$$\frac{\sigma_n}{\langle n\rangle} = \frac{1}{\sqrt{N}}\sqrt{\frac{q}{p}}. \tag{3.5}$$

The relative standard deviation is inversely proportional to $\sqrt{N}$. It is small for large N. It is clear that the number of elements N, in a Gibbs ensemble should be large enough to ensure that the relative standard deviation is as small as desired. Let me now describe a smart way of generating the moments of a random variable.

3.2 Moment Generating Function

Let $B(n)$ denote the probability that n coins are in micro state "Heads" in an ensemble of N coins. We have shown that,

$$B(n) = \frac{N!}{n!(N-n)!}\; p^n q^{N-n}. \tag{3.6}$$

The moment generating function is defined as

$$\widetilde{B}(z) \;=\; \sum_{n=0}^{N} z^n \; B(n). \tag{3.7}$$

The first thing we notice is that $\widetilde{B}(z=1)=1$. This guarantees that the probability distribution $B(n)$ is normalised. The moment generating function is like a discrete transform of the probability distribution function. We transform the variable n to z.

Let us now take the first derivative of the moment generating function with respect to z. We have,

$$\begin{aligned}\frac{d\widetilde{B}}{dz} = \widetilde{B}'(z) \;&=\; \sum_{n=0}^{N} n \; z^{n-1} \; B(n), \\ z\widetilde{B}'(z) \;&=\; \sum_{n=0}^{N} n \; z^n \; B(n). \end{aligned} \tag{3.8}$$

. Substitute in the above $z=1$. We get,

$$\widetilde{B}'(z=1) \;=\; \langle n \rangle. \tag{3.9}$$

Thus the first derivative of $\widetilde{B}$ evaluated at $z=1$ generates the first moment.

Now take the second derivative of $\widetilde{B}(z)$ to get

$$\begin{aligned}\frac{d^2\widetilde{B}}{dz^2} \;&=\; \sum_{n=0}^{N} n(n-1)z^{n-2}B(n), \\ z^2\frac{d^2\widetilde{B}}{dz^2} \;&=\; \sum_{n=0}^{N} z^n \; n(n-1) \; B(n). \end{aligned} \tag{3.10}$$

Substitute in the above $z=1$ and get,

$$\left.\frac{d^2\widetilde{B}}{dz^2}\right|_{z=1} = \langle n(n-1) \rangle, \tag{3.11}$$

For the Binomial random variable, we can derive the moment generating function :

$$\widetilde{B}(z) \;=\; \sum_{n=0}^{N} z^n B(n) = \sum_{n=0}^{N} \frac{N!}{n!\ (N-n)!}\ (zp)^n\ q^{N-n}, = (q+zp)^N. \tag{3.12}$$

The moments are generated as follows.

$$\frac{d\widetilde{B}}{dz} \;=\; N(q+zp)^{N-1}p, \tag{3.13}$$

$$\langle n \rangle \;=\; \left.\frac{d\widetilde{B}}{dz}\right|_{z=1} = Np, \tag{3.14}$$

$$\frac{d^2\widetilde{B}}{dz^2} \;=\; N(N-1)(q+zp)^{N-2}p^2, \tag{3.15}$$

$$\langle n(n-1) \rangle = \left.\frac{d^2\widetilde{B}}{dz^2}\right|_{z=1} \;=\; N(N-1)p^2. \tag{3.16}$$

3.3 Binomial → Poisson

When N is large, it is clumsy to calculate quantities employing Binomial distribution. Consider the following situation.

I have N molecules of air in this room of volume V. The molecules are distributed uniformly in the room. In other words the number density, denoted by ρ is same at all points in the room. Consider now an imaginary small volume $v < V$ completely contained in the room. Consider an experiment of choosing randomly an air molecule from this room. The probability that the molecule shall be in the small volume is $p = v/V$; the probability that it shall be out side the small volume is $q = 1 - (v/V)$. There are only two possibilities. We can use Binomial distribution to calculate the probability for n molecules to be present in v.

Consider first the problem with $V = 10M^3$, $v = 6M^3$ and $N = 10$. The value of p for the Binomial distribution is 0.6. The probability of finding n

molecules in v is then,

$$B(n; N = 10) = \frac{10!}{n!(10-n)!}(0.1)^n(0.9)^{10-n}. \tag{3.17}$$

The table below gives the probabilities calculated from the Binomial distribution. Consider the same problem with $v = 10^{-3}\ M^3$ and $N = 10^5$. We have

n	$B(n;10)$	n	$B(n;10)$
0	0.0001	6	0.2508
1	0.0016	7	0.2150
2	0.0106	8	0.1209
3	0.0425	9	0.0403
4	0.1115	10	0.0060
5	0.2007	−	−

Table 3.1: Probabilities calculated from Binomial distribution : $B(n; N = 10, p = .1)$

$p = 10^{-4}$ and $Np = 10$. Immediately we recognise that Binomial distribution is not appropriate for this problem. Calculation of the probability of finding n molecules in v involves evaluation of factorial of 100000.

What is the right distribution for this problem and problems of this kind ? To answer this question, consider what happens to the Binomial distribution in the limit of $N \to \infty$, $p \to 0$, and $Np = \mu$, a constant[1]. Note that

$$Np = Nv/V = \rho v = \text{constant}.$$

We shall show below that in this limit, Binomial goes over to Poisson distribution.

[1]Note that for a physicist, large is infinity and small is zero.

3.4 Poisson Distribution

We start with

$$\widetilde{B}(z) = (q + zp)^N. \tag{3.18}$$

We can write the above as

$$\begin{aligned}
\widetilde{B}(z) &= q^N \left(1 + \frac{zp}{q}\right)^N, \\
&= (1-p)^N \left(1 + \frac{zp}{q}\right)^N, \qquad (3.19) \\
&= \left(1 - Np\,\frac{1}{N}\right)^N \left(1 + \frac{zNp}{q}\frac{1}{N}\right)^N. \qquad (3.20)
\end{aligned}$$

In the above replace Np by μ and q by 1 to get,

$$\widetilde{B}(z) = \left(1 - \frac{\mu}{N}\right)^N \left(1 + \frac{z\mu}{N}\right)^N.$$

In the limit of $N \to \infty$ we have by definition[2],

$$\begin{aligned}
\widetilde{B}(z) &\sim \exp(-\mu)\exp(z\mu), \\
&= \widetilde{P}(z). \qquad (3.21)
\end{aligned}$$

Thus in the limit $N \to \infty$, $p \to 0$ and $Np = \mu$, we find $\widetilde{B}(z) \to \widetilde{P}(z)$, given by

$$\widetilde{P}(z) = \exp[-\mu(1-z)]. \tag{3.22}$$

[2]exponential function is defined as

$$\exp(x) = \mathop{limit}_{N\to\infty} \left(1 + \frac{x}{N}\right)^N$$

The coefficient of z^n in the power series expansion of $\widetilde{P}(z)$ gives $P(n)$,

$$P(n) \quad = \quad \frac{\mu^n}{n!}\exp(-\mu). \tag{3.23}$$

The above is called the Poisson distribution [3]. Thus in the limit of $N \to \infty$, $p \to 0$, $Np = \mu$, the Binomial distribution goes over to Poisson distribution. Figure (3.2) depicts Poisson distribution for $\mu = 1.5$ and 9.5.

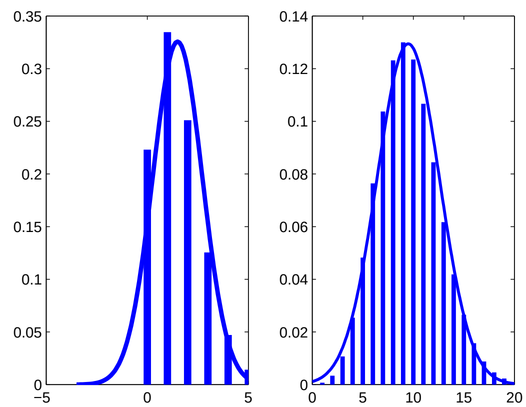

Figure 3.2: Poisson distribution with mean μ, depicted as sticks. Gaussian distribution with mean μ and variance $\sigma^2 = \mu$ depicted by continuous line. (Left) $\mu = 1.5$; (Right) $\mu = 9.5$. For large μ Poisson and Gaussian coincide

[3]We shall come across Poisson distribution in the context of Maxwell-Boltzmann statistics. Let n_k denote the number of 'indistinguishable classical' particles in a single-particle state k. The random variable n_k is Poisson-distributed.

3.4.1 Binomial → Poisson à la Feller

Following Feller [4], we have

$$\begin{aligned}
\frac{B(n;N)}{B(n-1;N)} &= \frac{N!\ p^n\ q^{N-n}}{n!\ (N-n)!}\ \frac{(n-1)!\ (N-n+1)!}{N!\ p^{n-1}\ q^{N-n+1}}, \\
&= \frac{p\ (N-n+1)}{n\ q}, \\
&= \frac{N\ p - p\ (n-1)}{n\ q}, \\
\begin{pmatrix} \overset{N\to\infty}{p \to 0} & {}_{Np=\mu} \end{pmatrix} &\overset{\sim}{} \frac{\mu}{n}.
\end{aligned} \tag{3.24}$$

Thus we get for large N, $B(n;N) = B(n-1;N)\mu/n$.

Start with $B(n=0;N) = q^N$. We have

$$q^N = (1-p)^N = \left(1 - \frac{Np}{N}\right)^N = \left(1 - \frac{\mu}{N}\right)^N \overset{N\to\infty}{\sim} \exp(-\mu).$$

Thus for $N \to \infty$, we get $B(n=0;N) = P(n=0;\mu) = \exp(-\mu)$. We get, $P(n=1;N) = \mu \exp(-\mu)$, $P(n=2;N) = (\mu^2/2!)\ \exp(-\mu)$, $P(n=3;N) = (\mu^3/3!)\exp(-\mu)$. Finally prove by induction $P(n;N) = (\mu^n/n!)\exp(-\mu)$.

The next item in the agenda is on Gaussian distribution. It is a continuous distribution defined for $-\infty\ \leq x\ \leq +\infty$. Before we take up the task of obtaining Gaussian from Poisson (in the limit $\mu \to \infty$), let us learn a few relevant and important things about continuous distribution.

[4] William Feller, *An Introduction to PROBABILITY : Theory and its Applications*, Third Edition Volume 1, Wiley Student Edition (1968)p.153

3.5 Characteristic Function

Let $x = X(\omega)$ be a continuous random variable, and $f(x)$ its probability density function. The Fourier transform of $f(x)$ is called the characteristic function of the random variable $x = X(\omega)$:

$$\phi_X(k) = \int_{-\infty}^{+\infty} dx\ \exp(-ikx)\ f(x).$$

Taylor expanding the exponential in the above, we get

$$\phi_X(k) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \int_{-\infty}^{\infty} dx\ x^n\ f(x) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!}\ M_n. \qquad (3.25)$$

Thus the characteristic function generates the moments.

3.6 Cumulant Generating Function

The logarithm of the characteristic function is called the cumulant generating function. $\psi_X(k) = \ln \phi_X(k)$. Let us write the above as,

$$\begin{aligned} \psi_X(k) &= \ln\left(1 + \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} M_n\right), \\ &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \left(\sum_{m=1}^{\infty} \frac{(-ik)^m}{m!}\ M_m\right)^n. \qquad (3.26) \end{aligned}$$

We now express $\psi_X(k)$ as a power series in k as follows

$$\psi_X(k) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \zeta_n. \qquad (3.27)$$

where ζ_n is called the n-th cumulant.

From the above equations we can find the relation between moments and cumulants.

3.7 The Central Limit Theorem

Let $\{X_i : i - 1, 2, \cdots, N\}$ denote independent random variables and

$$Y = \sum_{i=1}^{N} X_i \tag{3.28}$$

their sum. Also let $\phi_i(k)$ denote the characteristic function of X_i and $\phi_Y(k)$ that of Y. Then by convolution theorem,

$$\phi_Y(k) \quad = \quad \prod_{i=1}^{N} \phi_i(k) \tag{3.29}$$

If the random variable are also identical, and $\phi_X(k)$ denote the characteristic of the common distribution, then

$$\phi_Y(k) = [\phi_X(k)]^N \tag{3.30}$$

Now consider the scaled random variable defined as,

$$Y = \frac{1}{N}\sum_{i=1}^{N} X_i. \tag{3.31}$$

The characteristic function of Y is obtained by scaling k to k/N, see below.

$$\phi_Y(k) \quad = \quad \left[\phi_X\left(k \to \frac{k}{N}\right)\right]^N \tag{3.32}$$

We can write the above in terms of cumulant generating function, as follows.

$$\begin{aligned}
\phi_Y(k) \quad &= \quad \exp\left(\ln\left[\phi_X\left(k \to \frac{k}{N}\right)\right]^N\right) \\
&= \quad \exp\left[N \ \ln \phi_X\left(k \to \frac{k}{N}\right)\right],
\end{aligned}$$

$$
\begin{aligned}
&= \quad \exp\left[N\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}\frac{\zeta_n}{N^n}\right] \\
&= \quad \exp\left[\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}\frac{\zeta_n}{N^{n-1}}\right], \\
&= \quad \exp\left[-ik\mu-\frac{k^2}{2!}\frac{\sigma^2}{N}+\sum_{n=3}^{\infty}\frac{(-ik)^n}{n!}\frac{\zeta_n}{N^{n-1}}\right], \\
&= \quad \exp\left[-ik\mu-\frac{k^2}{2!}\frac{\sigma^2}{N}+\mathcal{O}(1/N^2)\right] \\
&\underset{N\to\infty}{\sim} \quad \exp\left[-ik\mu-\frac{k^2}{2!}\frac{\sigma^2}{N}\right].
\end{aligned}
\tag{3.33}
$$

Thus in the limit of $N \to \infty$, the characteristic function of Y, is given by

$$
\phi)Y(k) \quad = \quad \exp\left(-ik\mu-\frac{k^2}{2!}\frac{\sigma^2}{N}\right) \tag{3.34}
$$

We shall show shortly that the above expression is the characteristic function of a Gaussian random variable with mean μ and variance σ^2/N. Also we see that for the entire derivation above to hold good, we require that the variance σ^2 of the common distribution must be finite *i.e.* $\sigma^2 < \infty$.

In fact it is clear that the central limit theorem obtains even if the random variables participating in the sum are not be identical. The only requirement is that they should be independent of each other and and none of them should have infinite variance.

Thus, the sum of N independent finite variance - random variables tends to have a Gaussian distribution for large N. The Gaussian distribution has a variance inversely proportional to N, and hence is small for large N. This is called the central limit theorem.

3.8 Poisson $\to$ Gaussian

Start with the moment generating function of the Poisson random variable: $\widetilde{P}(z;\mu) = \exp[-\mu(1-z)]$. Substitute $z = \exp(-ik)$ and get, $\widetilde{P}(k;\mu) = \exp\left[-\mu\left\{1-\exp(-ik)\right\}\right]$. Carry out the power series expansion of the exponential function and get,

$$\widetilde{P}(k;\mu) \;=\; \exp\left[\sum_{n=1}^{\infty}\frac{(-ik)^n}{n!}\,\mu\right]. \tag{3.35}$$

We recognise the above as the cumulant expansion of a distribution for which all the cumulants are the same μ. For large value μ it is adequate to consider only small values of k. Hence we retain only terms upto quadratic in k. Thus for k small, we have,

$$\widetilde{P}(k) \;=\; \exp\left[-ik\mu - \frac{k^2}{2!}\mu\right]. \tag{3.36}$$

The above is the Fourier transform or the characteristic function of a Gaussian random variable with mean as μ and variance also as μ.

Thus in the limit $\mu \to \infty$, Gaussian distribution with mean and variance both equal to μ is a good approximation to Poisson distribution with mean μ, see Fig. 2.

3.9 Gaussian

A Gaussian of mean μ and variance σ^2 is given by

$$G(x) \;=\; \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{3.37}$$

The characteristic function is given by the Fourier transform formally expressed as $\widetilde{G}(k) = \int_{-\infty}^{+\infty} dx\ \exp(-ikx)G(x)$. The integral can be worked out, and I leave it as an exercise. We get, $\widetilde{G}(k) = \exp\left[-ik\mu - k^2\sigma^2/2\right]$. Consider a Gaussian of mean zero and variance σ^2. It is given by

$$g(x) \quad = \quad \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{x^2}{\sigma^2}\right]. \tag{3.38}$$

The width of the Gaussian distribution is 2σ. The Fourier transform of $g(x)$ is denoted $\widetilde{g}(k)$ and is given by

$$\widetilde{g}(k) \quad = \quad \exp\left[-\frac{1}{2}k^2\sigma^2\right]. \tag{3.39}$$

The Fourier transform is also a Gaussian with zero mean and standard deviation $1/\sigma^2$. The width of $\widetilde{g}(k)$ is $2/\sigma$. The product of the width of $g(x)$ and the width of its Fourier transform $\widetilde{g}(k)$ is 4.

If $g(x)$ is sharply peaked then its Fourier transform $\widetilde{g}(k)$ will be broad and *vice versa.*