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

2

Maxwell's Mischief

2.1 Experiment and Outcomes

Toss a coin : You get either "Heads" or "Tails". The experiment has two outcomes. Consider tossing of two coins. Or equivalently toss a coin twice. There are four outcomes : An outcome is an ordered pair. Each entry in the pair is drawn from the set $\{H,\ T\}$.
We can consider, in general, tossing of N coins. There are 2^N outcomes. Each outcome is an ordered string of size N with entries drawn from the set $\{H, T\}$.

Roll a die : You get one of the six outcomes :

{ ⚀ ; ⚁ ; ⚂ ; ⚃ ; ⚄ ; ⚅ }

Consider throwing of N dice. There are then 6^N outcomes. Each outcome is an ordered string of N entries drawn from the basic set of six elements given above.

Select randomly an air molecule in this room and find its position and momentum : Consider the air molecule to be a point particle. In classical mechanics, a point particle is completely specified by its three position (q_1, q_2, q_3) and three momentum (p_1, p_2, p_3) coordinates. An ordered set of six numbers

$$\{q_1,\ q_2,\ q_3,\ p_1, p_2,\ p_3\}$$

is the outcome of the experiment. A point in the six dimensional phase space represents an outcome of the experiment or the microstate of the system of single molecule. We impose certain constraints *e.g.* the molecule is always confined to this room. Then all possible strings of six numbers, consistent with the constrains, are the outcomes of the experiment.

2.2 Sample space and events

The set of all possible outcomes of an experiment is called the sample space. Let us denote it by the symbol Ω.

- $\Omega = \{H,\ T\}$ for the toss of a single coin.
- $\Omega = \{HH, HT, TH, TT\}$ for the toss of two coins.

A subset of Ω is called an event. Let $A \subset \Omega$ denote an event. When you perform the experiment if you get an outcome that belongs to A then we say the event A has occured.

For example consider tossing of two coins. Let event A be described by the statement that *the first toss is H*. Then A consists of the following elements: $\{HH,\ HT\}$.

The event corresponding to the roll of an even number in a dice, is the subset

2.3 Probabilities

Probability is defined for an event. What is the probability of the event $\{H\}$ in the toss of a coin ? One-half. This would be your immediate response. The logic is simple. There are two outcomes : "Heads" and "Tails". We have no reason to believe why should the coin prefer "Heads" over "Tails" or *vice versa*. Hence we say both outcomes are equally probable. What is the probability of having at least one "H" in a toss of two coins ? The event corresponding this statement is $\{HH, HT, TH\}$ and contains three elements. The sample size contains four elements. The required probability is thus 3/4. All the four outcomes are equally probable [1]. Then the probability of an event is the number of elements in that event divided by the total number of elements in the sample space. For *e.g.*, the event A of rolling an even number in a game of dice, $P(A) = 3/6 = 0.5$. The outcome can be a continuum. For example, the angle of scattering of a neutron is a real number between zero and π. We then define an interval $(\theta_1,\ \theta_2)$, where $0 \ \le\ \theta_1\ \le\ \theta_2\ \le \pi$, as an event. A *measurable* subset of a sample space is an event.

[1] Physicists have a name for this. They call it the axiom (or hypothesis or assumption) of *Ergodicity*. Strictly ergodicity is not an assumption; it is absence of an assumption required for assigning probabilities to events.

2.4 Rules of probability

The probability p that you assign to an event, should be obey the following rules.

$$\begin{aligned} p &\geq 0 \\ p(A \cup B) &= p(A) + p(B) - p(A \cap B) \\ p(\phi) = 0 \quad & \quad p(\Omega) = 1 \end{aligned} \tag{2.1}$$

In the above ϕ denotes a null event and Ω, a sure event.

How does one assign probability to an event ?

Though this question does not bother the mathematicians, the physicists should worry about this[2]. They should find the right way to assign probabilities to get the phenomenon right. We can say that the subject of statistical mechanics mainly deals with finding the right way to characterize a micro state, the sample space, the events and the assigning of probabilities to the events, depending on the system and phenomenon under investigation.

2.5 Random variable

The next important concept in probability theory is random variable , denoted by the symbol $x = X(\omega)$ where ω denotes an outcome and x a real number. Random variable is a way of stamping an outcome with a number : Real number, for a real random variable. Integer, for an integer random variable.

[2]Maxwell and Boltzmann attached probabilities to events in some way; we got Maxwell-Boltzmann statistics.

Fermi and Dirac had their own way of assigning probabilities to Fermions *e.g.* electrons, occupying quantum states. We got Fermi-Dirac statistics.

Bose and Einstein came up with their scheme of assigning probabilities to Bosons, populating quantum states; and we got Bose-Einstein statistics.

Complex number, for a complex random variable[3]. Thus the random variable $x = X(\omega)$ is a set function.

Consider a continuous random variable $x = X(\omega)$. We define a probability density function $f(x)$ by

$$f(x)dx \quad = \quad P(\omega | x \le X(\omega) \le x + dx) \tag{2.2}$$

In other words $f(x)dx$ is the probability of the event (measurable subset) that contains all the outcomes to which we have attached a real number between x and $x + dx$.

Now consider a continuous random variable defined between a to b with $a < b$. We define a quantity called the "average" of the random variable x as

$$\mu = \int_a^b dx \ x \ f(x).$$

μ is also called the mean, expectation, first moment *etc.*

Consider a discrete random variable n, taking values from say 0 to N. Let $P(n)$ define the discrete probability. We define the average of the random variable as

$$\mu = \sum_{n=0}^{N} n \ P(n).$$

But then, we are accustomed to calculating the average in a different way. For example I am interested in knowing the average marks obtained by the students in a class, in the last mid-semester examination. How do I calculate it ? I take the marks obtained by each of you, sum them up and divide by the total number of students. That is it. I do not need notions like probabilities, probability density, sum over the product of the random variable and the corresponding probability, integration of the product of the continuous

[3] In fact, we stamped dots on the faces of die; this is equivalent to implementing the idea of an integer random variable : attach an integer between one and six to each outcome.
For a coin, we stamped "Heads" on one side and "Tails" on the other. This is in the spirit of defining a random variable; we have stamped figures instead of numbers.

random variable and its probability density function *etc*.

Historically, before Boltzmann and Maxwell, physicists had no use for probability theory in their work. Newton's equations are deterministic. There is nothing chancy about a Newtonian trajectory. We do not need probabilistic description in the study of electrodynamics described by Maxwell equations; nor do we need probability to comprehend and work with Einstein's relativity - special or general.

However mathematicians had developed the theory of probability as an important and sophisticated branch of mathematics.

It was Ludwig Eduard Boltzmann who brought, for the first time, the idea of probability into physical sciences; he was championing the cause of kinetic theory of heat and atomic theory of matter. Boltzmann transport equation is the first ever equation written for describing the time evolution of a probability distribution.

2.6 Maxwell's mischief : Ensemble

However, Maxwell, had a poor opinion about a physicist's ability to comprehend mathematicians' writings on probability theory, in general, and the meaning of average as an integral over a probability density, in particular.

After all, if you ask a physicist to calculate the average age of a student in the class, he'll simply add the ages of all the students and divide by the number of students.

To be consistent with this practice, Maxwell proposed the notion of an **ensemble** of outcomes of an experiment (or an ensemble of realisations of a random variable). Let us call it Maxwell ensemble [4].

[4] Later we shall generalise the notion of Maxwell ensemble and talk of ensemble as a

Consider a collection of a certain number of independent realisations of the toss of a single coin. We call this collection a Maxwell ensemble if it it obeys certain conditions, see below.

Let N denote the number of elements in the ensemble. Let n_1 denote the number "Heads" and n_2 number of 'Tails". We have $n_1 + n_2 = N$. If $n_1 = Np$, and hence $n_2 = Nq$, then we say the collection of outcomes constitutes a Maxwell ensemble.

Thus an ensemble holds information about the outcomes of the experiment constituting the sample space; it also holds information about the probability of each outcome. The elements of an ensemble are drawn from the sample space; however each element occurs in an ensemble as often as to reflect correctly its probability.

For example consider an experiment of tossing a p-coin[5] with $p = 0,75$; then a set of four elements given by $\{H,\ H,\ H,\ T\}$ is a candidate for an ensemble underlying experiment. . A set of eight elements $\{H,\ H,\ H,\ H,\ H,\ H,\ T,\ T\}$ is also an equally valid ensemble for this experiment. Thus the size of an ensemble is somewhat arbitrary. If the number of times each outcome occurs in the ensemble is consistent with its probability it would suffice.

2.7 Calculation of probabilities from an ensemble

Suppose we are given the following ensemble : $\{H, T, H, H, T\}$. By looking at the ensemble, we can conclude that the sample space contains two outcomes $\{H,\ T\}$.

We also find that the outcome H occurs thrice and T occurs twice. Hence

collection identical copies of a macroscopic system. We shall call it a Gibbs ensemble

[5] a p-coin is one for which the probability of 'Heads' is p and that of the 'Tails' is $q = 1 - p$.

We also conclude that the probability of H is $3/5$ and that of T is $2/5$.

2.8 Construction of ensemble from probabilities

We can also do the reverse. Given the outcomes and their probabilities, we can construct an ensemble. Let n_i denote the number of times an outcome i occurs in an ensemble. Let N denote the total number of elements of the ensemble. Choose n_i such that n_i/N equals p_i; note that we have assumed that p_i is already known.

2.9 Counting of the elements in events of the sample space : Coin tossing

Consider tossing of N identical coins or tossing of a single coin N times. Let us say the coin is fair. In other words $P(H) = P(T) = 1/2$.

Let $\Omega(N)$ denote the set of all possible outcomes of the experiment. An outcome is thus a string N entries, each entry being "H" or "T". The number of elements of this set $\Omega(N)$ is denoted by the symbol, $\widehat{\Omega}(N)$. We have $\widehat{\Omega}(N) = 2^N$.

Let $\Omega(n_1, n_2; N)$ denote a subset of $\Omega(N)$, containing only those outcomes with n_1 'Heads' and n_2 'Tails'. Note $n_1 + n_2 = N$. How many outcomes are there in the set $\Omega(n_1, n_2; N)$?

Let $\widehat{\Omega}(n_1, n_2; N)$ denote the number of elements in the event $\Omega(n_1, n_2; N)$. In what follows I shall tell you how to derive an expression[6] for $\widehat{\Omega}(n_1, n_2; N)$.

[6] I remember I have seen this method described in a book on Quantum Mechanics by

Take one outcome belonging to the event $\Omega(n_1, n_2; N)$. There will be n_1 'Heads' and n_2 'Tails" in that outcome . Imagine for a moment that all these 'Heads' are distinguishable. If you like, you can label them as H_1, H_2, $\cdots$, H_{n_1}. Carry out permutation of all the 'Heads' and produce $n_1!$ new configurations. From each of these new configurations, produce $n_2!$ configurations by carrying out the permutations of the n_2 'Tails'. Thus from one outcome belonging to $\Omega(n_1, n_2; N)$, we have produced $n_1! \times n_2!$ new configurations. Repeat the above for each element of the set $\Omega(n_1, n_2; N)$, and produce $\widehat{\Omega}(n_1, n_2; N)n_1!n_2!$ configurations. A moment of thought will tell you that this number should be the same as $N!$. Therefore

$$\widehat{\Omega}(n_1, n_2; N) \times n_1! \times n_2! = N!$$

Let us work out an example explicitly to illustrate the above. Consider tossing of five coins. There are $2^5 = 32$ outcomes/microstates listed below. The number in the brackets against each outcome denotes the number of "Heads" in that outcome.

1.	H	H	H	H	H	(5)	17.	T	H	H	H	H	(4)
2.	H	H	H	H	T	(4)	18.	T	H	H	H	T	(3)
3.	H	H	H	T	H	(4)	19.	T	H	H	T	H	(3)
4.	H	H	H	T	T	(3)	20.	T	H	H	T	T	(2)
5.	H	H	T	H	H	(4)	20.	T	H	T	H	H	(3)
6.	H	H	T	H	T	(3)	21.	T	H	T	H	T	(2)
7.	H	H	T	T	H	(3)	22.	T	H	T	T	H	(2)
8.	H	H	T	T	T	(2)	23.	T	H	T	T	T	(1)
9.	H	T	H	H	H	(4)	25.	T	T	H	H	H	(3)
10.	H	T	H	H	T	(3)	26.	T	T	H	H	T	(2)
11.	H	T	H	T	H	(3)	27.	T	T	H	T	H	(2)
12.	H	T	H	T	T	(2)	28.	T	T	H	T	T	(1)
13.	H	T	T	H	H	(3)	29.	T	T	T	H	H	(2)
14.	H	T	T	H	T	(2)	30.	T	T	T	H	T	(1)
15.	H	T	T	T	H	(2)	31.	T	T	T	T	H	(1)
16.	H	T	T	T	T	(1)	32.	T	T	T	T	T	(0)

The outcomes numbered 4, 6, 7, 10, 11, 13, 18, 19, 20, 25 are the ones with three "Heads" and two "tails"). These are the elements of the event $\Omega(n_1 = 3; n_2 = 2; N = 5)$.

Take outcome No. 4. Label the three heads as H_1, H_2 and H_3. Carry out

Gasiorowicz. Check it out

permutations of the three "Heads" and produce $3! = 6$ elements. These are

(4)	H	H	H	T	T					
	H_1	H_2	H_3	T	T	H_2	H_3	H_1	T	T
	H_1	H_3	H_2	T	T	H_3	H_1	H_2	T	T
	H_2	H_1	H_3	T	T	H_3	H_2	H_1	T	T

Take an element from the above set. Label the 'Tails' as T_1 and T_2. Carry out permutation of the 'Tails' and produce $2! = 2$ elements. Do this for each of the above six elements.

Thus from the outcome No, 4, we have produced $3! \times 2! = 12$ outcomes, listed below.

(4)	H	H	H	T	T					
	H_1	H_2	H_3	T_1	T_2	H_2	H_3	H_1	T_1	T_2
	H_1	H_2	H_3	T_2	T_1	H_2	H_3	H_1	T_2	T_1
	H_1	H_3	H_2	T_1	T_2	H_3	H_1	H_2	T_1	T_2
	H_1	H_3	H_2	T_2	T_1	H_3	H_1	H_2	T_2	T_1
	H_2	H_1	H_3	T_1	T_2	H_3	H_2	H_1	T_1	T_2
	H_2	H_1	H_3	T_2	T_1	H_3	H_2	H_1	T_2	T_1

Repeat the above exercise on the outcomes numbered 6, 7, 10, 11, 13, 18, 19, 20, 25. Table below depicts the results for outcome no. 6.

(6)	H	H	T	H	T					
	H_1	H_2	T_1	H_3	T_2	H_2	H_3	T_1	H_1	T_2
	H_1	H_2	T_2	H_3	T_1	H_2	H_3	T_2	H_1	T_1
	H_1	H_3	T_1	H_2	T_2	H_3	H_1	T_1	H_2	T_2
	H_1	H_3	T_2	H_2	T_1	H_3	H_1	T_2	H_2	T_1
	H_2	H_1	T_1	H_3	T_2	H_3	H_2	T_1	H_1	T_2
	H_2	H_1	T_2	H_3	T_1	H_3	H_2	T_2	H_1	T_1

Thus we produce $\widehat{\Omega}(n_1 = 3, n_2 = 2; N = 5) \times n_1! \times n_2!$ outcomes. This number is just the number of permutations of $N = 5$ objects labelled H_1, H_2, H_3, T_1, T_2 and it equals $N!$. Therefore,

$$\widehat{\Omega}(n_1 = 3, n_2 = 2; N = 5) = \frac{5!}{3!2!} = 10$$

$\widehat{\Omega}(n_1, n_2; N)$ is called the binomial coefficient[7]

2.10 Gibbs ensemble

Following Gibbs, we can think of an ensemble as consisting of large number of identical mental copies of a macroscopic system [8]. All the members of an ensemble are in the same macro state[9]. However they can be in different micro states. Let the micro states of the system under consideration, be indexed by $\{i = 1, 2, \cdots\}$. The number of elements of the ensemble in micro state j divided by the size of the ensemble equals t the probability of the system to be in micro state j. It is intuitively clear that the size of the ensemble should be large ($\to \infty$) so that it can capture exactly the probabilities of different micro states of the system[10]. Let me elaborate on this issue, see below.

2.11 Why should a Gibbs ensemble be large ?

What are the values of n_1 and n_2 for which $\widehat{\Omega}(n_1, n_2; N)$ is maximum[11] ?

[7]We have the binomial expansion given by

$$(a+b)^N = \sum_{\{n_1,n_2\}}^{\star} \frac{N!}{n_1! n_2!} a^{n_1} b^{n_2}$$

The sum runs over all possible values of $\{n_1, n_2\}$. The superscript $\star$ on the summation sign should remind us that only those values of n_1 and n_2 consistent with the constraint $n_1 + n_2 = N$ are permitted.

[8]For example the given coin is a system. Let p denote the probability of "Heads" and $q = 1 - p$ the probability of "tails". The coin can be in a micro state "Heads" or in a micro state "Tails".

[9]This means the values of p and q are the same for all the coins belonging to the ensemble.

[10]If you want to estimate the probability of Heads in the toss of a single coin experimentally then you have to toss a large number of identical coins. Larger the size of the ensemble more (statistically) accurate is your estimate .

[11]you can find this in several ways. Just guess it. I am sure you would have guessed the answer as $n_1 = n_2 = N/2$. We know that the binomial coefficient is largest when

It is readily shown that for $n_1 = n_2 = N/2$ the value of $\widehat{\Omega}(n_1, n_2; N)$ is maximum. Let us denote this number by the symbol $\widehat{\Omega}^m(N)$. We have, $\widehat{\Omega}^m(N) = \widehat{\Omega}(n_1 = N/2, N_2 = n/2; N)$.

Thus we have

$$\widehat{\Omega}(N) = \sum_{\{n_1,n_2\}}^{\star} \frac{N!}{n_1!\ n_2!} = 2^N; \quad (\star \Rightarrow n_1 + n_2 = N) \tag{2.3}$$

$$\widehat{\Omega}^m(N) = \widehat{\Omega}(n_1 = n_2 = N/2; N) = \frac{N!}{(N/2)!\ (N/2)!} \tag{2.4}$$

Let us evaluate $\widehat{\Omega}^m(N)$ for large values of N. We employ Stirling's first approximation [12] : $N! = N^N\ \exp(-N)$ and get

$$\widehat{\Omega}^m(N) = = \frac{N^N \exp(-N)}{\left[(N/2)^{(N/2)} \exp(-N/2)\right]^2} = 2^N \tag{2.5}$$

$n_1 = n_2 = N/2$ if N is even, or when n_1 and n_2 equal the two integers closest to $N/2$ for N odd. That is it.

If you are more sophisticated, take the derivative of $\widehat{\Omega}(n_1, n_2; N)$ with respect to n_1 and n_2 with the constraint $n_1 + n_2 = N$, set it zero; solve the resulting equation to get the value of N for which the function is an extremum. Take the second derivative and show that the extremum is a maximum.

You may find it useful to take the derivative of logarithm of $\widehat{\Omega}(n_1, n_2; N)$; employ Stirling approximation for the factorials : $\ln(m!) = m\ \ln(m)\ -\ m$ for large m. Stirling approximation to large factorials is described in the next section.

You can also employ any other pet method of yours to show that, for $n = N/2$, the function $\widehat{\Omega}(n; N)$ is maximum.

[12] First Stirling Approximation : $N! \approx N^N\ \exp(-N)$.
We have,

$$\begin{aligned}
N! &= N \times (N-1) \times \cdots \times 3 \times 2 \times 1 \\
\ln\ N! &= \ln\ 1 + \ln\ 2 + \ln\ 3 + \cdot + \ln\ N \\
&= \sum_{k=1}^{N} \ln(k) \approx \int_1^N \ln\ x\ dx = (x \ln\ x - x)\big|_1^N = N \ln\ N - N - 1 \\
&\approx N \ln N - N \\
N! &\approx N^N \exp(-N)
\end{aligned}$$

The above implies that almost all the outcomes of the experiment belong to the event with equal number of 'Heads' and 'Tails'. The outcomes with unequal number of 'Heads' and 'Tails' are so overwhelmingly small that the difference falls within the small error arising due to the first Stirling approximation.

For estimating the tiny difference between $\widehat{\Omega}(N)$ and $\widehat{\Omega}^m(N)$, let us employ the second Stirling approximation[13] : $N! = N^N \exp(-N)\sqrt{2\pi N}$ and get

$$\widehat{\Omega}^m(N) = \widehat{\Omega}(n_1 = n_2 = N/2; N) = 2^N \times \sqrt{\frac{2}{\pi N}}.$$

Thus we have,

$$\begin{aligned} \frac{\widehat{\Omega}^m(N)}{\widehat{\Omega}(N)} &= \sqrt{\frac{2}{\pi N}} && (2.6) \\ &\propto \frac{1}{\sqrt{N}} && (2.7) \end{aligned}$$

[13] A better approximation to large factorials is given by Stirling's second formula : $N! \approx N^N \ \exp(-N)\ \sqrt{2\pi N}$
We have

$$\begin{aligned} \Gamma(N+1) = N! = \int_0^\infty dx\ x^N e^{-x} &= \int_0^\infty dx\ e^{F(x)} \\ F(x) &= N \ln x - x \\ F'(x) &= \frac{N}{x} - 1\ ;\ F''(x) = -\frac{N}{x^2} \end{aligned}$$

Set $F'(x) = 0$; this gives $x^\star = N$. At $x = x^\star$ the function $F(x)$ is maximum. (Note : $F''(x = x^\star)$ is negative). Carrying out Taylor expansion and retaining only upto quadratic terms, we get,

$$F(x) = F(x^\star) + \frac{(x - x^\star)^2}{2} F''(x = x^\star) = N \ln N - N - \frac{(x-N)^2}{2N}$$

We have,

$$\begin{aligned} N! = \int_0^\infty dx\ e^{F(x)} &= N^N e^{-N} \int_0^\infty dx \exp\left[-\frac{(x-N)^2}{2N}\right] \\ &= N^N e^{-N} \sqrt{N} \int_{-\sqrt{N}}^\infty dx \exp(-x^2/2) \\ &\underset{N\to\infty}{\sim} N^N e^{-N} \sqrt{2\pi N} \end{aligned}$$

Let us define $S(\Omega(n_1, n_2; N)) = \ln \widehat{\Omega}(n_1, n_2; N)$ as entropy of the event $\Omega(n_1, n_2; N)$. For $n_1 = n_2 = N/2$ we get an event with largest entropy. Let us denote it by the symbol $S_B(N)$.

Let $S_G = S(\Omega(N)) = N \ln 2$ denote the entropy of the sure event. It is the logarithm of the number of outcomes of the experiment of tossing N independent and fair coins.

We find, for large N,

$$S_B = S(\Omega(n_1 = n_2 = N/2; N)) \quad = \quad N \ln 2 - \frac{1}{2} \ln N = S_G - \frac{1}{2} \ln N \tag{2.8}$$

The entropy of the sure event is of the order of N; the entropy of the event with $N/2$ "Heads" and $N/2$ 'Tails" is less by an extremely small quantity of the order of $\ln(N)$. Hence when you toss a very large number of coins independently you will get almost always $N/2$ 'Heads' and $N/2$ 'Tails".

For example take a typical value for $N = 10^{23}$. We have $S_G = 0.69 \times 10^{23}$ and $S_B = 0.69 \times 10^{23} - 24.48$.

Note that only when N is large, we have S_B equals S_G. It is precisely because of this, we want the number of elements to be large, while constructing a Gibbs ensemble. We should ensure that all the micro states of the system are present in the ensemble in proportions, consistent with their probabilities.

For example I can simply toss N independent fair coins just once and if N is large then I am assured that there shall be $(N/2) \pm \epsilon$ 'Heads' and $(N/2) \mp \epsilon$ 'Tails', where ϵ is negligibly small : of the order of $\sqrt{N}$. Consider the probability for random variable n to take values outside the interval

$$\left(\frac{N(1-\epsilon)}{2}, \quad \frac{N(1+\epsilon)}{2} \right)$$

where ϵ is an arbitrarily small number. Let us denote this probability as $P_{\text{out}}(N)$. One can show that in the limit $N \to \infty$, the probability $P_{\text{out}}(N)$ goes to zero.

Take $\epsilon = 1/100$ and calculate $P_{\rm out}(N)$ for $N = 10^3$, 10^4, 10^5, 10^6, and 10^8. The table below depicts the results, and we see that $P_{\rm out}$ nearly zero for large N.

N	$\epsilon = 1/100$; $P_{\rm out} = P\left(\frac{N}{2}(1-\epsilon) \le n \le \frac{N}{2}(1+\epsilon)\right)$
10^3	0.752
10^4	0.317
10^5	0.002
10^6	1.5×10^{-23}
10^7	2.7×10^{-2174}

In statistical mechanics we consider a macroscopic system which can be in any of its numerous microscopic states. These micro states are analogous to the outcomes of an experiment. In fact the system switches spontaneously from one micro state to another when in equilibrium. We can say the set of all micro states constitute a micro state space which is the analogue of the sample space. We attach a real number representing the numerical value of a property *e.g.* energy, to each micro state. This is analogous to random variable. We can define several random variables on the same micro state space to describe different macroscopic properties of the system.