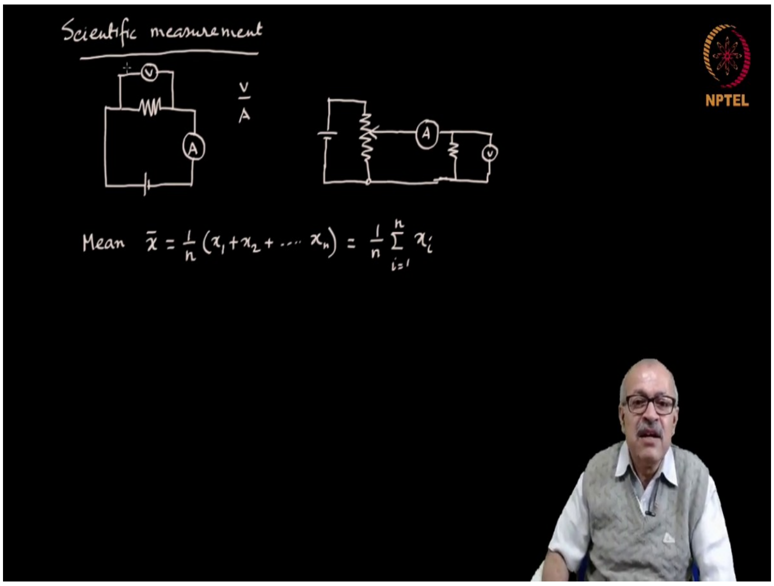
Research Methodology

Prof. Soumitro Banerjee
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

Lecture - 29
Elements of Scientific Measurement, Part 01

Today, I will be starting with the general methods of experimentation. Mostly the purpose of an experimentation is the measurement of certain quantities.

(Refer Slide Time: 00:38)



So, we will talk about scientific measurement, and as you will see later, scientific measurement is at the heart of testing hypotheses. But today let us focus on the elements of scientific measurement.

You might think that different areas, for example, physics, chemistry, biology, earth science, they have different measurements to make. Somebody might be measuring the charge of an electron. A biologist may have discovered a new species of bird or insect, and then wants to specify that specific type of insect by measuring the average body weight of that insect. An earth scientist may face the question: ‘what is the average density of the earth’s crust?’ So on and so forth. You can imagine various types of measurements.

A biologist may face the question that, in a beaker there is a solution in which some kind of a microbe is growing, and what is the density, that is, the count of the microbes per unit volume? So, these are the different types of measurements. You might think that they are entirely different. How can there be a general methodology of measurement?

Of course, the techniques, the instruments used—these are different. But there are certain general ideas flowing through the whole act of measurement. That is what we need to understand.

The reason is that, as we earlier said, one of the demands of experimental research is repeatability. So, any measurement you make, you have to get the result in such a form and you have to report the results in such a form, so that anybody anywhere in the world can repeat the experiment and can get the same result.

Now, imagine that you have measured the mass of an electron sitting in IISER Kolkata, and then you have published: it is this much. Another person sitting in Siberia should be able to get the same value. But if you do the experiment again sitting in IISER Kolkata, you will not get the same value. You will get a slightly different value, and so it is true for the person sitting is Siberia or in Mexico. How is it that we can still say that the result is repeatable? Therefore, it has to be reported in a form that should be repeatable.

These are the questions that we will face today. But in every case, we always design an experiment in such a way that we get a large number of results, and we can get some kind of an average from there. Unless this is true, the experiment is faulty.

For example, suppose you are asked to measure the resistance of a resistor. This is a very common thing that different scientific disciplines need to do, be it physics, chemistry, earth science, everywhere. In geology, we have to measure the resistance between two points on Earth. A physicist has to measure the resistance in order to find out whether something has gone into superconductivity or not. If a chemist invents a different material, you have to find its resistivity. So, resistance measurement is something very common.

Suppose you have been asked to measure the resistance of a resistor. How would you do that?

Now, I have tried this by asking this question to many students, and mostly the answer I get is something like this: that I will connect a voltage source and then I have to measure the current through it. So, I will connect an ammeter and I will have to measure the voltage across it by a voltmeter. So, to measure the resistance, all I will do is to allow a current to pass through. I will measure the voltage, I will measure the current, then V by A is the resistance. Wrong.

Wrong, because this experiment will allow you to get just one value of voltage and one value of current and therefore, one value of resistance. By the concept of experimentation, this is faulty because you always have to set up the experiment in such a way that you get a large number of readings, so that you can make an average. How should you do it? In this particular case, for example, if you have a voltage source, then that has to be connected to a potential divider, and this point can be taken. A potential divider is something like this.

And then you have to connect the unknown resistor and here you have to connect the ammeter. The unknown resistor, this is what you measure, and you can put a voltmeter across it, so that you can measure the voltage. What will happen? As you move this jockey across, when the jockey is here you get the full voltage; when the jockey is here you get zero voltage; in between you get all the possible voltages.

So, you can vary the voltage and at each value of the voltage you can measure the current, thereby tabulate the results in form of a number of voltages and the currents. Finally, you divide the voltage by the current and get a number of resistance values. Each will be different from the others slightly, but that is what the whole point is. By averaging them out, you can get the final result as some kind of a mean value that you can state.

So, what is the mean value then? The mean value, if the variable is x, I will call it x bar. It will be x1 plus x2 plus dot dot dot xn; n number of readings, so you divide that by n. So, effectively what you are doing is that 1 by n sum over i is equal to 1 to n sum over xi. So, that is the average value. That is the mean value. This is the first thing that you get.

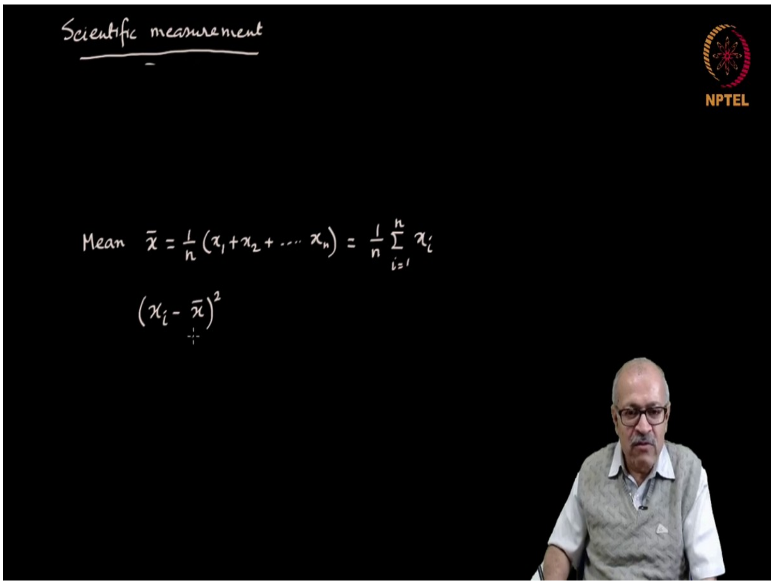
$$\text{Mean}: \ \bar{x} \quad = \quad \frac{1}{n}(x_1 + x_2 \cdots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

I initially stated a few different ways, a few different possibilities of measurement, depending on the field in which it is being done. What x represents for a physicist could be the value of the resistance measured in one run. Or it could be the value of the mass of an electron that is measured. You measure it again, you get x2, you measure it again you get x3. So, in a large number of measurements, ultimately you get the average of it.

For a biologist, if he or she wants to measure the average weight of a new insect that has been discovered, x1, x2, x3 will represent the measurement of each one's weight and so on and so forth. Finally, after having done this, you would like to obtain the mean.

I hope the point is understood. The point I am making is that, always you will have to set up the experiment in such a way that you can get a large number of readings, and then you can take an average. So, finally, you have got the mean.

(Refer Slide Time: 09:57)

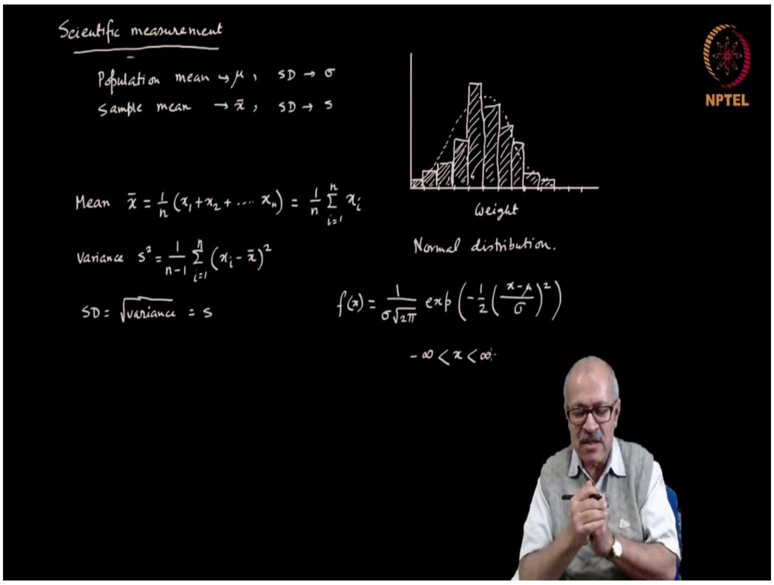


Then you would like to know how much, on an average, the data points vary from the mean. How much are they different from the mean. So, you are interested in average character of xi minus x-bar, its character. But, you see, x i minus x bar could be a positive quantity as well as a negative quantity and therefore, if you average it out, you are likely to get zero. That is not what you want. So, you square it in order to get a positive number and then take an average of it.

So, you get a make an average of this after having done a square, so that you get a positive number and this leads to the definition of what is known as ‘variance’. Let me just write it then.

$$\text{Variance}: \; s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(Refer Slide Time: 10:50)



The variance is, I will call it s square, is in this case, you have to take the same way summation over i is equal to 1 to n and this that case we are taking a average of x i minus x bar square. But then, you have to divide it by the number of, what should I say, independent variables, the ‘degrees of freedom’ as it is called. So, that happens to be 1 by n minus 1.

Why n minus 1? Because the number of degrees of freedom. Notice that, when you are calculating the variance, the value of the x bar is already known. The value of x bar has been already calculated. And when you have calculated x bar, if you know the n minus 1 data points, then you can calculate the nth data point. So, nth data point does not remain something independent; does not represent a degree of freedom. So, you need to take 1 by n minus 1.

So, this is the representation of variance and the standard deviation is just the square root of variance. So, this is just s. We represent that by s. So, this is the first thing that you do. That means, whenever you take any reading, these are the things that you have calculate first. That is the first step.

Now, this is important because we are making a measurement. That measurement cannot be done with just one reading. We have to take a number of readings and whenever you are taken a number of readings, you are interested in stating the mean. And also how much the data points vary from the mean and that is represented by the standard deviation.

Now, the moot point. A problem in measurement, the actual problem in measurement is that, whatever you are trying to measure are out there. It is an objective quantity. The rest mass of an electron is an objective quantity, something that is a constant, and you are trying to measure it. The average mass of that particular insect is an objective quantity. It is there. You are only trying to find out the value. What is that value? That is the average of *all* possible insects.

So, in order to get that value, you have to collect all possible insects of that particular species, and measure them, and then take the average—which is practically impossible. So, if you are trying to state the value of the weight of that particular species—the male of the particular species—then you really have to collect all possible samples, all possible organisms of that particular species and measure.

But you cannot do that. So, effectively what you do is, to take *samples* from a population. So, there is a population of insects, you take a few samples, you measure it and thereby hope that what you get, that means, the value of x bar, is close to the actual value of the mean in the population.

So, there is something called *population mean*, and what you are actually measuring is a *sample mean.* And you are hoping that by obtaining the sample mean you will come close to the population mean. This is a problem in all experimentation.

Physicists might think that, no, we are free from that, because our population is not like insect population where there is a inherent variability within the population. All insects are not of the same mass. But all electrons are of the same mass. They are identical particles. So, I can get away with easier task.

No, it is not like that because every time you make a measurement of the mass of an electron, you get a slightly different value. You cannot never get exactly the same value. Therefore, in order to get the absolutely correct value of the mass, you really have to make an infinite number of measurements which is not possible. So, you are also effectively taking a *sample,* a smaller number of samples, from an ideally infinite population of measurements.

Ideally, you could have made an infinite number of measurements, taken the average. But you cannot do that and so, you are making a choice of doing a smaller number of measurements, which is also the same as sampling. The fact that you are drawing a sample is not different in different fields. It is basically the same: that you want to reach obtain the population mean, but you actually have, in your hand, the sample mean.

Let us give the names. The population mean: let us call it mu and the sample mean: let us call it x bar. So, you have x bar; you are trying to reach mu. Similarly, the population SD, standard deviation for population: let us call it sigma, and for the sample standard deviation, let us call it s, the why we have called it.

So, the problem of all measurements is that we have a handle on x bar and s, but we want to know mu and sigma. Can we do that? How accurately can we do that? How confident can we be in the statement that we make, because we have made a smaller number of samples? So, this is the essential problem in measurement.

After having done this, the next step is to find out the distribution of the values. If you are measuring the weight of all the male sparrows for example, you have collected a few samples and you have measured them. Ultimately you have got some kind of numbers, a collection of numbers.

Then what we normally do is that, suppose I want to plot a graph of whatever we are trying to measure. For example, if I am obtaining the weight of the male sparrows, then the weights will be divided into a few bins. From this weight to this weight is a bin and we will now find out how many of the sparrows fell in that bin.

Now, if that exercise is done, then for each box, each bin, there would be a number. Maybe here we have a number like this. So, I will fill it like this and then the next bin is like that, I will fill it up like that. So, in different bins you will get different numbers, and then ultimately you will get a graph that will look something like this. Because the extreme cases will be rare, something like this. It may not be exactly a graph like this, but you can see that, it will have a peak at the middle and it will taper off at the two sides, what is known as the 'bell curve'.

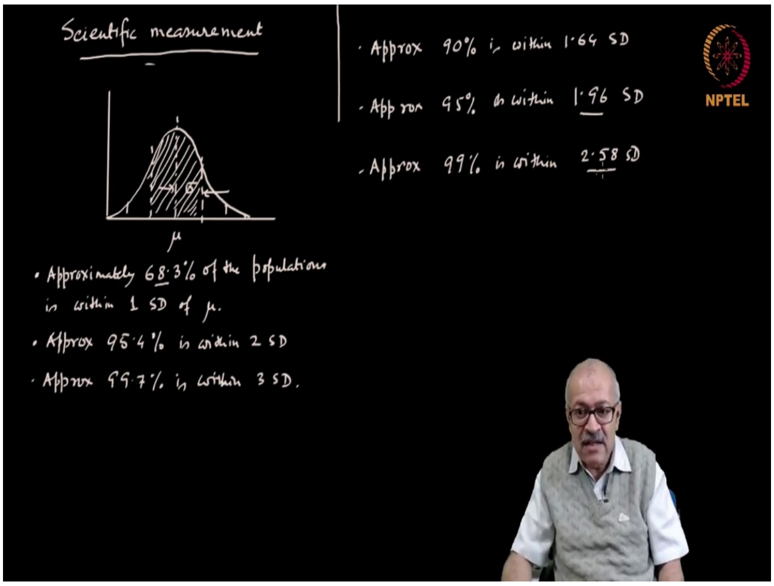
If you take a very large number of readings, then it more or less approximates the bell curve. So, the bell curve is something that is very important in measurement. It is called the 'normal distribution'. So, this is called the normal distribution. Whenever you have measured something and have done this exercise of obtaining what is known as a histogram; basically we divide whatever we are measuring into segments, into bins, and find out how many data points fell in this bin, count them, and plot the graph accordingly.

And that graph normally approximates a normal distribution. When I said 'normally', I have something in mind. I will illustrate that a little later because in some cases it could be a non-normal distribution, but the normal distribution is most common. The normal distribution, if somebody is mathematically inclined, is given by f x equal to 1 by sigma square root 2 pi exponential of minus half x minus mu by sigma square for minus infinity the whole range.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad -\infty < x < \infty$$

This is the expression for the normal distribution. You are not expected to memorize the expression for the normal distribution. That is not really needed. But having this expression, it is not difficult to integrate that over certain ranges. So, let me just show you what actually is done.

(Refer Slide Time: 23:17)



If you actually measure and plot this. Then it will be a graph like this. Some sort of an idealization of this will be a graph something like this. Here this is the mean that you have calculated, and the standard deviation will be somewhere like this. So, this difference will be the standard deviation which I represented as sigma. There is something that appears here and the mean appears here.

Now, this graph is what you get if you plot this expression. But in actual situations, from the measurement you will have the value of the mu and the value of the sigma, and from there you will have to obtain the shape of this graph. Now, what we normally do is that we ask ourselves: consider the range between mu minus sigma to mu plus sigma; that means, if I am talking about this area; how much is this area? If the area under the whole curve is 1, how much is this area? Now, these are already calculated. I will just write it.

Firstly, approximately 68.3 percent of the area or population is within 1 standard deviation of the mean. It really helps to remember just this number: 68.3 percent. So, if you consider an area within 1 standard deviation, means mu minus sigma to mu plus sigma, within that range around 68.3 percent of the population remains; the rest is outside.

Now, you can also talk about 2 standard deviations or 3 standard deviations and there I can say that 95.4 percent of the total area is within about 2 standard deviations; 99.7 percent of the total area is within 3 standard deviations. Let me write: approx 95.4

percent is within 2 standard deviations and approx 99.7 percent is within 3 standard deviations, which means practically the whole area is within 3 standard deviations. This is important. I will come to that later.

We need to get, not really memorize, but a feel for the numbers: 68.3 percent is within 1 standard deviation, 95.4 percent under 2 standard deviation, 99.7 percent or practically the whole under 3 standard deviations. What we actually want to know is slightly different. What we actually want to know is, again obtained by integrating the curve within some range, approximately 90 percent remains within what standard deviation, approximately 95 percent is within what standard deviation, and approximately 99 percent is within which standard deviation? That is what really we are interested in. I will show you later.

So, approximately 90 percent is within 1.64 standard deviations. This is more important. 95 percent is withing 1.96 standard deviation, 99 percent is within 2.58 standard deviation. This is important because as we will see, these information are actually used when we ultimately try to design an experiment. I will come to that. But these are the numbers that you need to learn by heart.

1.96 standard deviation, 2.58 standard deviation: these are more important. This is relatively less important. Within 1.96 standard deviation the area under the curve is about 95 percent. If you want to find out how much standard deviation we need to take to obtain 99 percent of the area, it is 2.58 standard deviation.